



HAL
open science

Vers un outil d'aide à la décision en évaluation des systèmes interactifs et la prise en compte conjointe de critères techniques et socio-culturels

André Nendjo Ella

► **To cite this version:**

André Nendjo Ella. Vers un outil d'aide à la décision en évaluation des systèmes interactifs et la prise en compte conjointe de critères techniques et socio-culturels. Interface homme-machine [cs.HC]. Université de Valenciennes et du Hainaut-Cambrésis, 1999. Français. NNT: . tel-01665759

HAL Id: tel-01665759

<https://theses.hal.science/tel-01665759>

Submitted on 16 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numéro d'ordre : 99-37

THESE

présentée à

**L'Université de Valenciennes
et du Hainaut - Cambrésis**

en vue de l'obtention du

DOCTORAT

Spécialité :

INFORMATIQUE

par

André NENDJO ELLA
Ingénieur ESIE

**VERS UN OUTIL D'AIDE A LA DECISION EN EVALUATION
DES SYSTEMES INTERACTIFS ET LA PRISE EN COMPTE
CONJOINTE DE CRITERES TECHNIQUES ET SOCIO-CULTURELS**

Soutenue le 22/12/1999 devant la commission d'examen :

Mme : Anne-Marie JOLLY-DESODT	ENSAIT	Rapporteur
Mr : Jean VANDERDONCKT	Université Catholique de Louvain - Belgique	Rapporteur
Mr : Christophe KOLSKI	Université de Valenciennes	Directeur
Mr : Noël MALVACHE	Université de Valenciennes	Examinateur
Mr : Frédéric SEMET	Université de Valenciennes	Président
Mr : Pascal YIM	Ecole Centrale de Lille	Examinateur

*« Le Monde est un tout,
c'est-à-dire qu'il faut prendre en compte à la fois,
non seulement les actions logiques (basées sur le raisonnement),
mais aussi les actions non logiques (basées sur les sentiments et l'émotion) »
(W. Pareto, 1919)*

REMERCIEMENTS

Le travail présenté dans ce mémoire a été réalisé au Laboratoire d'Automatique et de Mécanique Industrielles et Humaines (LAMIH-UMR CNRS 8530) de l'Université de Valenciennes et du Hainaut-Cambrésis, dans le groupe de Travail « Raisonnement Automatique et Interaction Homme – Machine » (GT RAIHM), dirigé par le Professeur Christophe KOLSKI. Je le remercie pour m'avoir suivi, guidé et conseillé avec enthousiasme tout au long de mes travaux. Je lui exprime ma profonde reconnaissance pour avoir enrichi ma recherche et ce mémoire par le biais de ses nombreuses remarques et de ses nombreux conseils.

Je tiens à remercier Madame Anne-Marie JOLLY-DESODT, Professeur à l'ENSAIT, pour m'avoir conseillé dans le cadre de mes travaux et pour avoir accepté d'être rapporteur de ce travail.

Mes remerciements vont aussi au Professeur Jean VANDERDONCKT de l'Université Catholique de Louvain, rapporteur de cette thèse, pour l'intérêt porté à cette recherche.

Je remercie également le Professeur Noël MALVACHE pour avoir accepté d'être membre de mon jury et pour m'avoir aidé dans le cadre de ma thèse en m'orientant vers le GRAISYHM, ce qui m'a valu des contacts fructueux avec une socio-technicienne du groupe de travail pour la partie de mes travaux concernant l'acceptabilité socio-culturelle.

Un grand merci également au Professeur Frédéric SEMET pour avoir accepté d'examiner ce travail.

Je tiens particulièrement à remercier Monsieur Pascal YIM, Maître de Conférences à l'École Centrale de Lille, pour avoir assuré efficacement un encadrement lors du développement en PROLOG d'une partie du système à la décision en évaluation.

A Catherine JACQUES, spécialiste de la socio-technique et Chercheur à l'IFRESI (Institut Fédératif de Recherche sur les Économies et les Sociétés Industrielles), je suis reconnaissant de l'intérêt manifesté pour ce travail et pour avoir accepté de réaliser une étude complémentaire dans le cadre de mes travaux relativement au problème d'acceptabilité sociale des Interfaces Homme(s) - Machine (s). Ces travaux ont été présentés au Congrès international de Reims sur « l'éthique et la qualité dans l'évaluation ».

Mes pensées vont à Fabrice WAWAK de l'Université de Madrid pour m'avoir aidé dans la modélisation des ensembles flous, en corrigeant les étapes de la modélisation depuis l'Espagne. Ses corrections des rapports d'étape et ses propositions m'ont permis d'élaborer une modélisation des critères de décision plus fiable.

Que mes collègues, avec qui j'ai eu l'occasion de travailler, trouvent ici toute ma reconnaissance. Je pense particulièrement à Serge DEBERNARD, Igor CREVITS, Thierry POULAIN, Elyes CHERIFI, Dimitri TABARY et Manuel LAMBERT. Un grand merci en particulier à Houcine EZZEDINE pour son aide relative à une des études de cas.

Je remercie enfin mes parents, mes beaux-parents, mes enfants et ma femme pour m'avoir soutenu tout au long de ma thèse ainsi que pour leur soutien et leur patience à mon égard. Que ce mémoire leur soit dédié, particulièrement à mon père et à ma mère, décédés, qui n'ont malheureusement pas eu la chance de voir et d'assister à la phase finale de ma thèse.

SOMMAIRE

Glossaire	9
Introduction générale	11
 Chapitre I : Evaluation des systèmes interactifs : problématique, diversité des critères	
Introduction	15
I. Principe de l'évaluation	15
I.1. Synthèse préalable à l'aide des questions de la méthode quintilienne concernant l'évaluation de la qualité des IHMs.....	15
I.2. Les intervenants de l'évaluation.....	18
I.3. Traitement des données recueillies.....	20
I.4. Conclusion et besoins pour l'évaluation des IHMs.....	21
II. Dimensions et critères d'acceptabilité pratique	21
II.1. Dimension et critères d'utilité.....	21
II.2. Dimension et critères d'utilisabilité.....	22
II.3. Autres dimensions et critères	24
III. Dimensions et critères d'acceptabilité sociale	25
III.1. Importance des critères d'acceptabilité sociale dans la conception et l'évaluation des Interfaces Homme(s) - Machine(s)	25
III.2. Autres exemples tirés des Pays en Voie de Développement (PVD).....	32
III.3. Degré d'exigence socio-culturelle selon les pays.....	36
III.4. Résumé des dimensions et critères à prendre en compte.....	39
III.4.1. La dimension et les critères d'éthique.....	39
III.4.2. La dimension sociale et ses critères.....	39
III.4.3. La dimension culturelle et ses critères.....	41
III.5. Positionnement des critères d'évaluation par rapport au cycle de développement.....	42
IV. Contraintes liées aux choix des méthodes et techniques d'évaluation dans un contexte d'évaluation globale	43
Conclusion	45

Chapitre II : Des méthodes et techniques aux outils logiciels d'aide à l'évaluation des systèmes interactifs

Introduction	47
I. Approches centrées sur les utilisateurs	47
I.1. Approches empiriques de diagnostic d'usage.....	49
I.2. Estimation de la charge de travail.....	62
I.2.1. Echelles subjectives.....	63
I.2.2. Méthodes objectives.....	66
I.3. Tests de conception.....	68
I.4. Conclusion sur les approches centrées sur les utilisateurs et positionnement dans le cycle de développement.....	72
II. Approches centrées sur une expertise	72
III Méthodes analytiques	75
III.1. Approches fondées sur une théorie.....	75
III.2. Modèles formels prédictifs.....	79
III.2.1. Modèles de tâches.....	79
III.2.2. Modèles linguistiques.....	83
III.3. Les modèles formels dits de qualité de l'IHM : approches logicielles.....	85
III.4. Conclusion sur les modèles analytiques.....	91
IV. Synthèse sur les méthodes et techniques usuelles	93
V. Vers de nouveaux outils d'aide à l'évaluation	94
V.1. Aides logicielles FACE.....	94
V.2. Plan d'aide pour l'évaluation (Denley et Long).....	96
V.3. Discussions sur FACE et l'outil d'aide de Denley et Long.....	97
Conclusion	102

Chapitre III : Principes pour une aide à la décision en évaluation

Introduction	101
I. Principe et structure d'un système d'aide à la décision	102
I.1. Problématique.....	102
I.2. Alternatives.....	103
I.3. Les critères et leur évaluation.....	105
I.4. Stratégies et préférences.....	105

I.5.	Les types de résultats.....	106
I.5.	Conclusion.....	106
II.	Approches possibles en aide à la décision et justification de l'approche suivie.....	106
II.1.	Analyse monocritère.....	107
II.2.	L'aide multicritère.....	108
II.2.1.	Approches opérationnelles pour l'agrégation des performances : exemples représentatifs d'approches multicritères conduisant à un critère unique de synthèse.....	109
II.2.2.	Surclassement de synthèse.....	109
II.3.	Approche utilisant les techniques de l'intelligence artificielle.....	119
II.4.	Approches basées sur les ensembles flous.....	119
II.5.	Synthèse et positionnement des différentes méthodes d'aide à la décision	123
II.6.	Conclusion.....	126
III.	Démarche globale basée sur les ensembles flous.....	127
III.1.	Evaluation des critères et préférences de l'expert.....	128
III.2.	Comparaison préférences-critères (Degré de Compatibilité Flou).....	130
III.3.	Règles floues.....	131
III.4.	Méthode de surclassement répondant à notre problématique (ELECTRE I).....	132
III.5.	Conditions d'arrêt.....	134
	Conclusion.....	134

Chapitre IV : Première maquette de faisabilité d'un système d'Aide à la Décision en évaluation : ADHESION

	Introduction.....	138
I.	Problème à traiter.....	138
II.	Critères de décision, actions potentielles et choix stratégiques envisagés par la démarche du système.....	140
II.1.	Critères et actions potentielles.....	140
II.2.	Choix stratégiques et schéma fonctionnel du système.....	142
III.	Description globale du système ADHESION.....	144
III.1.	Justification du choix du support d'implémentation.....	145
III.1.1.	Présentation et principes.....	145
III.1.2.	Aspects de la programmation et primitives dans Prolog III.....	145

III.2.	Architecture du système d'aide à la décision ADHESION et fonctionnalités.....	149
III.3.	Explications liées à la maquette développée.....	150
III.4.	Conclusion.....	156
IV.	Module de raisonnement flou implanté en PROLOG III.....	156
IV.1.	Constitution de la base des faits et algorithme.....	156
IV.2.	Algorithme de calcul des DCFs.....	159
	Conclusion.....	163

Chapitre V : Mise en œuvre et validation technique de la démarche et de la maquette sur deux cas d'étude

	Introduction.....	166
I.	Mise en œuvre et validation sur un premier cas d'étude : le projet ASTREE.....	166
I.1.	Contexte du projet ASTREE.....	166
I.1.1.	Les méthodes et techniques d'évaluation utilisées initialement	167
I.1.2.	Le protocole expérimental mis en œuvre initialement.....	167
I.1.3.	Dispositif expérimental initialement mis en place.....	167
I.1.4.	Présentation des résultats d'évaluation initialement obtenus.....	169
I.2.	Mise en œuvre du système d'aide ADHESION.....	170
I.2.1.	Préparation des données (contraintes).....	171
I.2.2.	Tableau de synthèse de la phase de préparation des données.....	175
I.3.	Saisie des contraintes.....	175
I.4.	Présentation des propositions de solution fournies par ADHESION pour le cas étudié (projet ASTREE).....	177
I.5.	Comparaison entre les méthodes utilisées pour l'évaluation d'ASTREE et celles proposées par ADHESION.....	179
I.5.1.	Elaboration des critères de comparaison.....	180
I.5.2.	Tableau de synthèse des critères de comparaison.....	181
II.	Mise en œuvre et validation sur un second cas d'étude : le projet MOGADOR.....	185
II.1.	Contexte du projet MOGADOR.....	185
II.1.1.	La technique d'évaluation utilisée initialement.....	186
II.1.2.	Le protocole expérimental mis en œuvre initialement.....	186
II.1.3.	Dispositif expérimental initialement mis en place.....	186
II.1.4.	Présentation des résultats initialement obtenus.....	187
II.2.	Mise en œuvre du système d'aide ADHESION.....	189
II.2.1.	Tableau de synthèse de la phase de préparation des données.....	189
II.2.2.	Saisie des contraintes (données) par l'utilisateur de l'outil d'aide.....	190
II.2.3.	Présentation des propositions de solution fournies par ADHESION	

pour le cas étudié (projet MOGADOR).....	191
II.3. Comparaison entre la méthode utilisée pour l'évaluation de MOGADOR et celles proposées par ADHESION.....	194
Conclusion	194
Conclusion générale et perspectives	197
REFERENCES BIBLIOGRAPHIQUES	202
ANNEXES	217
ANNEXE I : Evaluation du système ADHESION	
ANNEXE II : Ensembles et raisonnement flous	
ANNEXE III : Grille d'évaluation socio-technique pour l'évaluation des IHMs en terme d'Acceptabilité Sociale	
ANNEXE IV : Eléments de modélisation des données et des connaissances pour l'aide à la décision en évaluation	
INDEX DES FIGURES ET TABLEAUX	230

GLOSSAIRE

C	Ensemble des critères c_i pour l'aide à la décision
G_i	Ensemble des valeurs possibles pour l'évaluation des actions vis-à-vis du critère c_i
$g_i(a_j)$	Evaluation dans l'ensemble G_i de l'action a_j pour le critère c_i . Lorsqu'il n'y a pas ambiguïté concernant l'action, nous remplacerons l'écriture $g_i(a_j)$ par g_i
G	Espace des critères dans lequel sont exprimées les évaluations des actions de A. Cet espace se définit par $G = G_1 \times G_2 \times \dots \times G_n$
$a_j \text{Pak}$	Expression symbolisant le fait que le décideur préfère l'action a_j à l'action a_k
U	Fonction d'utilité en Théorie de l'Utilité
U_i	Fonction partielle, suivant le critère c_i , d'une fonction d'utilité additive : $U = U_1 + \dots + U_n$
p_i	Poids attribué au critère c_i pour une méthode multicritère pondérée
$c(a_j, a_k)$	Concordance entre les deux actions a_j et a_k dans les méthodes Electre I et PROMETHEE
N	Sous-ensemble des actions résultant de la méthode Electre I
$F_i(a_j, a_k)$	Fonction de préférence entre les actions a_j et a_k suivant le critère c_i pour la méthode Prométhée suivante dans la méthode du point de mire
$\mu_A(x)$	Degré d'appartenance de l'élément x à l'ensemble flou A
$\tilde{\mu}_A$	Fonction d'appartenance de l'ensemble flou continu A
$\Pi(A)$	Mesure de possibilité du sous-ensemble A
π	Distribution de possibilité
$\tilde{\pi}_A$	Distribution de possibilité associée à l'événement flou A
G_{ij}	Ensemble flou correspondant à l'évaluation floue du critère c_i pour l'action a_j
\tilde{P}_i	Ensemble flou correspondant au poids flou associé au critère c_i
r_m	Règle floue m
L	Variable linguistique avec V_T ensemble des termes liés à la variable X univers de discours, Gr sa règle syntaxique et R_S sa règle sémantique
Gr	Grammaire d'une variable linguistique avec V_T ensemble des termes de la variable, T_A ensemble des termes auxiliaires, S terme de départ et P ensemble des règles de production
V_D	Ensemble des termes descripteurs d
V_M	Ensemble des termes modificateurs m
MAT_{Pref}	Matrice de dimension $n \times p$ dont les éléments sont des ensembles flous évaluant les préférences du décideur
MAT_{Crit}	Matrice de dimension n contenant des distributions de possibilité évaluant les critères
MAT_{Act}	Matrice de dimension p contenant les DCFs de chaque action a_j provenant de l'agrégation des critères dans la matrice MAT_{Comp}

Q	Quantité totale d'information requise dans un histogramme brut
MAT_{Comp}	Matrice de dimension $n \times p$ contenant les degrés de compatibilité flous évaluant chaque action vis-à-vis de chaque critère
\leftrightarrow	
\tilde{A}	Support de l'ensemble flou A
$ A $	Cardinalité de l'ensemble flou A
\tilde{A}_α	α -coupe de l'ensemble flou A suivant le niveau α
\tilde{Q}	Quantité floue
\tilde{I}	Intervalle flou
\tilde{N}	Nombre flou
$M(a,b,c,d)$	Ensemble flou trapézoïdal

Introduction générale

L'origine des Interfaces Homme(s)-Machine(s) date de la fin des années 40. Ce fut d'abord la naissance des interfaces « BATCH » non interactives (les pionnières), suivies des interfaces unidimensionnelles. Dans les années 80, une collaboration entre les éléments du couple Homme-Machine est préconisée : la machine prépare et trie l'information, l'Homme prend les décisions. L'évolution et le développement croissant d'interfaces évoluées prenant en compte l'évolution de la société conduisent à des systèmes interactifs présents à tous les niveaux, dans les entreprises et la vie de tous les jours. Ces derniers possèdent de plus en plus de fonctionnalités et nécessitent des méthodes et techniques d'évaluation de plus en plus robustes pour leur validation.

L'évaluation consiste en général à comparer un modèle de l'objet évalué à un modèle de référence permettant d'établir des conclusions [SENACH, 1990]. D'après WHITEFIELD *et al.* [1991], l'évaluation est une estimation de la conformité entre les performances réelles du système et ses performances souhaitées. Cette estimation englobe une méthode (les procédures utilisées pour réaliser l'estimation) et un rapport (le produit résultant).

Le thème de ce travail concerne une des phases importantes de l'évaluation des Interfaces Homme(s)-Machine (s). En effet, le choix, dans une démarche globale d'évaluation (aussi bien sous l'angle de l'acceptabilité pratique que de l'acceptabilité sociale), de la méthode ou technique appropriée conditionne la qualité du résultat obtenu. La complexité des paramètres influençant le choix des méthodes et leur nature (informations objectives et subjectives) nous a orienté vers une démarche d'aide à la décision en évaluation basée sur la théorie des sous-ensembles flous et des possibilités. Cette étape, généralement dévolue à un évaluateur ou à un groupe d'évaluateurs, est délicate à franchir dans la mesure où à la complexité de la tâche d'évaluation est associée la difficulté, pour les évaluateurs eux-mêmes, de choisir la/les méthode(s) ou technique(s) d'évaluation appropriée(s) au système à évaluer. Les utilisateurs de notre système d'aide à l'évaluation peuvent être de plusieurs catégories : il peut s'agir de chefs de projets qui s'interrogent sur les méthodes et critères d'évaluation disponibles au regard de ses propres critères, d'évaluateurs novices (voire expérimentés), ou d'apprenants en interaction Homme(s)-Machine(s).

Dans le cadre de cette recherche pluridisciplinaire, nous avons bénéficié de la collaboration d'une socio-technicienne de l'IFRESI (Institut Fédératif de Recherche sur les Economies et les Sociétés Industrielles) et du CLERSE (Centre Lillois d'Etudes et de Recherches Sociologiques et Economiques) pour l'identification de critères socio-culturels en évaluation, d'un informaticien de l'Ecole Centrale de Lille, spécialiste en particulier de modélisation logique à l'aide de Prolog III, et d'un groupe d'automaticiens de l'université de Lille 1 et de l'université de Madrid, coutumiers de modélisations par sous-ensembles « flous ».

Cette réflexion a conduit à la première version d'un système ayant pour entrée différentes contraintes (moyens humains, localisation, forme du produit, type d'application, etc.) pouvant influencer sur le choix d'une méthode donnée. Ces contraintes, exprimées d'abord par des réponses à des questions à l'utilisateur du système, permettent à l'outil de sélectionner automatiquement

la/les méthode(s) ou technique(s) adaptées au système à évaluer. La liste de ces facteurs ou contraintes qui influencent le choix d'une méthode d'évaluation a été esquissée dans le cadre des travaux de thèse de BALBO [1994]. Nous avons renforcé ces contraintes et introduit également des contraintes socio-culturelles. Cette démarche est dite globale, c'est-à-dire qu'elle permet d'évaluer un système en termes d'acceptabilité pratique et sociale [NIELSEN, 1993].

D'après [NIELSEN, 1993], pour parvenir à des systèmes interactifs dits acceptables, il s'agit de les considérer selon deux angles : l'acceptabilité sociale et l'acceptabilité pratique (figure 1). Dans l'acceptabilité pratique, on retrouve les notions d'utilité et d'utilisabilité largement étudiées dans la littérature vis-à-vis de la conception et de l'évaluation de systèmes interactifs. Par contre, la notion d'acceptabilité sociale revêt souvent un aspect que l'on pourrait presque qualifier de « mystérieux » et est largement négligée par les équipes de développement de systèmes interactifs, malgré le besoin croissant de systèmes utilisés dans un contexte international [NIELSEN, 1990] [RUSSO et BOOR, 1993] [NENDJO ELLA *et al.*, 1996].

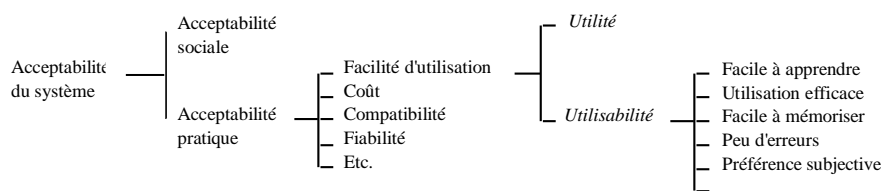


Figure 1 : Principales variables cibles pour l'évaluation [NIELSEN, 1990]

Lors du développement de systèmes interactifs dans les Pays en Voie de Développement (PVD), la prise en compte de critères socio-culturels est d'une importance capitale. Ces travaux le démontreront sur quelques exemples représentatifs issus d'une expérience de travail de plusieurs années dans les PVD et entrant dans le cadre de projets d'informatisation, d'automatisation ou d'observation dans lesquels l'auteur de cette thèse était directement ou non impliqué. Pour chaque exemple, des problèmes potentiels et solutions possibles seront discutés.

Dans cette thèse, nous nous intéressons à l'élaboration d'un système d'aide à la décision basé sur les sous-ensembles flous et les possibilités, en mettant un accent sur l'activité décisionnelle développée lors de situations de non satisfaction de l'utilisateur du système lorsque le système d'aide propose des solutions (méthodes et techniques) non satisfaisantes, c'est-à-dire non adaptées aux caractéristiques de la situation. Dans ce contexte, ce mémoire est structuré en cinq chapitres.

Le *premier chapitre* aborde en premier lieu les problèmes relatifs à l'évaluation des Systèmes interactifs. Dans un premier temps, on précisera le principe de l'évaluation et les difficultés liées au choix des méthodes d'évaluation. On discutera ensuite de l'importance des critères socio-culturels.

Après avoir présenté dans le premier chapitre le contour de l'évaluation, nous aborderons dans le *deuxième chapitre* la notion d'aide à l'évaluation, en précisant différentes approches

existantes. C'est ainsi que seront présentées une classification des différents outils d'aide à l'évaluation et une synthèse des méthodes d'évaluation.

Après avoir discuté de l'évaluation proprement dite et des différentes limites des outils d'aide à l'évaluation existants, le **troisième chapitre** traitera des problèmes liés à l'aide à la décision en évaluation. On présentera également les différentes stratégies et les préférences qui ont été établies à partir des connaissances expertes (données du décideur) et modélisées à l'aide des sous-ensembles flous (par des modèles trapézoïdaux). Après avoir effectué l'agrégation des critères et exposé les règles utilisées, nous présenterons en détail la démarche proposée pour l'aide à la décision en évaluation.

Dans le **quatrième chapitre**, une description de l'algorithme utilisé et une présentation des pages-écrans de la première maquette seront fournis. On verra notamment les pages-écrans de saisie de contraintes et des propositions de solutions issues d'un scénario fictif.

Le **cinquième chapitre** est consacré à la validation de la maquette sur deux scénarios d'évaluation provenant de cas réels ; l'un de ces deux scénarios concernera une situation où la prise en compte des critères socio-culturels s'avère particulièrement cruciale.

Ce mémoire se terminera par des perspectives de recherche. Les perspectives sont nombreuses, liées aussi bien au développement du système et à ses tests qu'à des scénarios d'évaluation lors de projets réels ou simulés ou à l'expertise à intégrer dans le système.

Chapitre I : Evaluation des systèmes interactifs : problématique, diversité des critères

Introduction

- I. Principe de l'évaluation**
- II. Dimensions et critères d'acceptabilité pratique**
- III. Dimensions et critères d'acceptabilité sociale**
- IV. Contraintes liées aux choix des méthodes et techniques d'évaluation dans un contexte d'évaluation globale**

Conclusion

Introduction

Avant d'aborder les problèmes relatifs à l'aide à la décision en évaluation, ce chapitre s'intéressera à la problématique de l'évaluation des systèmes interactifs.

Nous présenterons d'abord le principe de l'évaluation, en évoquant notamment les intervenants de l'évaluation, les données à recueillir et leur traitement. Dans un second temps, nous parlerons des dimensions et critères d'acceptabilités pratique et sociale. Nous ferons enfin une synthèse des contraintes liées au choix des méthodes et techniques d'évaluation.

I. Principe de l'évaluation

L'IHM doit être considérée en tant que composant du système Homme(s)-Machine(s) global. Dans ces conditions, son évaluation consiste à le *vérifier* et à chercher à le *valider*. Le système Homme(s)-Machine(s) est *vérifié* s'il correspond aux spécifications issues de la définition des besoins. Il est *validé* s'il correspond aux besoins en respectant les contraintes du domaine d'application (sécurité, production, ergonomie,...) ; dans le cas contraire, ses insuffisances par rapport à des critères identifiés *a priori* sont mises en évidence et doivent être corrigées [GRISLIN, 1995]. L'évaluation ergonomique d'une IHM consiste à s'assurer que l'utilisateur est capable de réaliser sa tâche au moyen du système de communication qui lui est proposé. Pour [SENACH, 1990], « toute évaluation consiste à comparer un modèle de l'objet évalué à un modèle de référence permettant d'établir des conclusions ». Nous expliquerons plus en détail ce principe d'évaluation au cours des pages qui suivent.

Lors de l'évaluation de logiciels interactifs, le modèle que l'on peut qualifier d'observé (ou d'analysé) est donc comparé à un modèle de référence (voir figure I.1). Une évaluation basée sur une comparaison n'est pas simple. Ce modèle dit de référence (nous l'appellerons plutôt « modèle de l'évaluateur ») correspond en fait à un modèle que va se construire l'équipe d'évaluation en prenant en compte les propriétés ou dimensions d'évaluation que doit respecter l'interface. Ces propriétés se composent, dans un contexte d'acceptabilité globale ou internationale (acceptabilité sociale et pratique), de plusieurs dimensions : d'*utilité*, d'*utilisabilité*, *économique*, d'*environnement*, d'*éthique*, *sociale* et *culturelle* ; elles doivent être « introduites » dans le modèle ou prises en compte dans la démarche. La construction et la validation des modèles utilisés constituent les principales sources de difficultés pour une bonne évaluation.

Il faut distinguer le modèle conceptuel de la machine, créé par le concepteur d'interface, et la représentation mentale que se fait l'utilisateur du système, construite progressivement par la pratique de ses acquis antérieurs et de ses interactions. Il existe une différence souvent importante entre la théorie formalisée et la réalité puisque l'interaction Homme(s)-Machine(s) est complexe à décrire. Cette difficulté permet de poser la question de la fiabilité des résultats d'évaluation qui sont dépendants des méthodes d'évaluation utilisées.

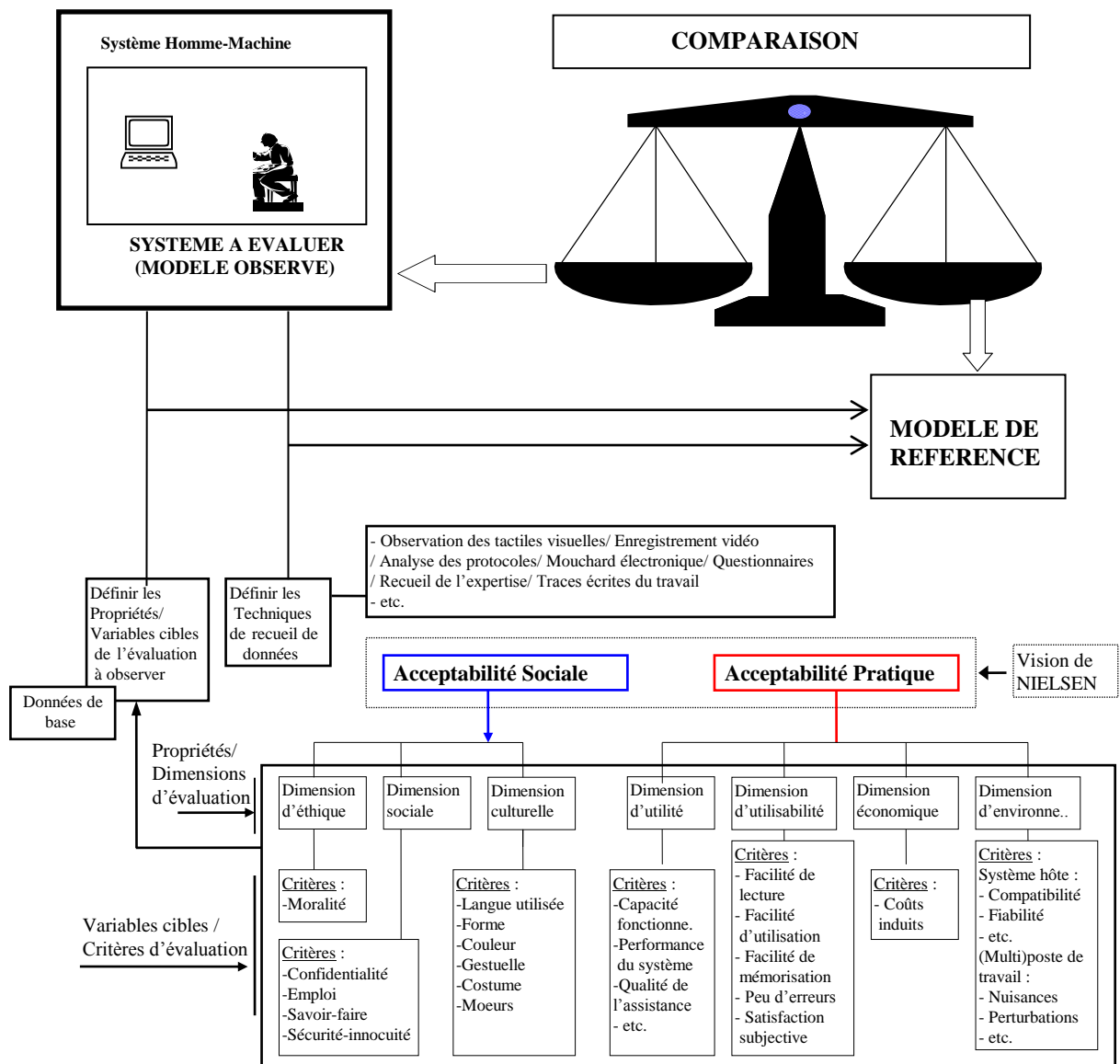


Figure I.1 : Principe de l'évaluation des IHMs en termes d'acceptabilité pratique et sociale, complété, inspiré et précisé à partir des informations issues de [NIELSEN, 1993][SENACH, 1990][GRISLIN, 1995][NENDJO ELLA et JACQUES, 1999].

I.1. Synthèse préalable à l'aide des questions de la méthode quintilienne concernant l'évaluation de la qualité des IHMs

Concernant l'évaluation, il s'agit dans un premier temps de répondre aux six questions de la méthode quintilienne¹ adoptée en technique de communication :

- Quoi ? De quoi s'agit-il ? (objet, opération, nature, quantité, etc.)
- Qui ? Qui est concerné ? (exécutants, nombre, qualification, etc.)
- Où ? Où cela se produit-il ? (situer le problème dans l'espace)

¹ Marcus Fabius Quintilianus enseigna la rhétorique à Rome vers les années 100 après JC.

- Quand ? Quand cela survient-il ? (situer le problème dans le temps)
- Comment ? Comment procède-t-on ? (de quelle manière cela se passe-t-il ?)
- Pourquoi ? Pourquoi cela se passe-t-il ainsi ? (causes, etc.)

De quoi s'agit-il ?

Il s'agit ici de contrôler la qualité aussi bien technique, économique, ergonomique, socio-culturelle, etc., des IHMs. Cette qualité intègre des notions de performance et de coûts mesurables et devrait s'inscrire dans un plan d'assurance qualité vue sous l'angle du génie logiciel.

Qui effectue l'évaluation ?

L'évaluation de la qualité ergonomique des logiciels est le plus souvent le fait de spécialistes de la communication Homme(s)–Machine(s). Ceux-ci travaillent en relation étroite avec les concepteurs et utilisateurs représentatifs (s'ils existent). En fait, comme nous le préciserons plus loin, d'autres catégories d'intervenants sont potentiellement concernées.

Quand s'effectue l'évaluation ?

L'évaluation peut se pratiquer tout au long du cycle de développement d'un système interactif. C'est ce que souligne l'esprit des méthodes [BARTHET, 1988]. Dans cette optique, on note deux types d'évaluation : l'évaluation *a priori* et l'évaluation *a posteriori*. L'évaluation *a priori* peut commencer à partir des phases de spécification fonctionnelle (voir figure I.2 ci-dessous), de conception préliminaire et de conception détaillée.

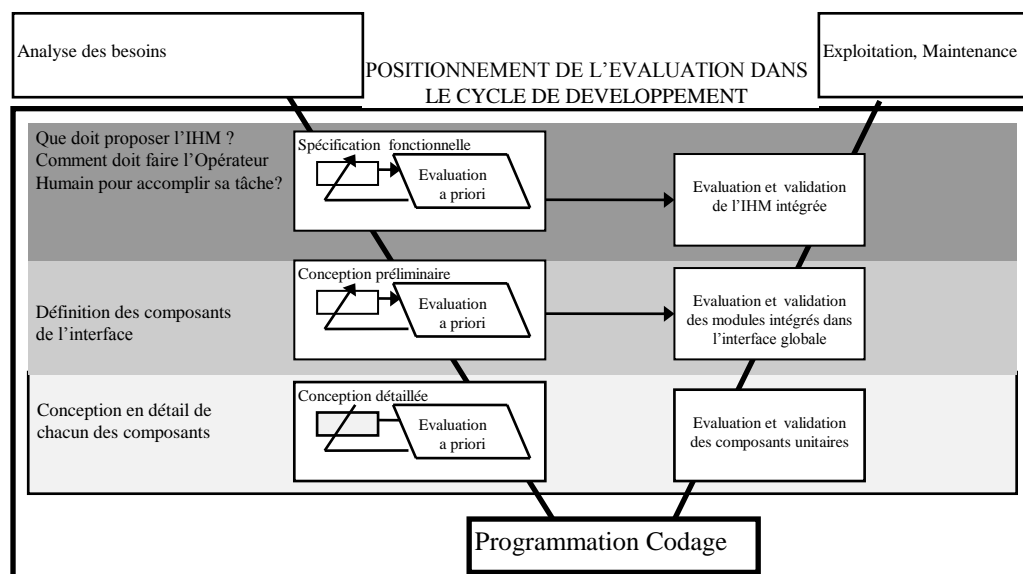


Figure I.2 : Positionnement de l'évaluation dans le processus de développement, inspirée de [GRISLIN, 1995]

L'évaluation *a posteriori* se déroule une fois que l'IHM est réalisée. Dans cette optique, les utilisateurs finaux ou représentatifs sont fortement souhaités.

Où s'effectue l'évaluation ?

L'évaluation peut se pratiquer sur le terrain, c'est-à-dire dans un contexte réel d'utilisation, ou en laboratoire ; l'évaluation peut éventuellement être menée sans outillage spécialisé (seule intervient ici la compétence de l'évaluateur : heuristiques et théories prédictives). Elle peut aussi se pratiquer dans un laboratoire d'utilisabilité, local doté d'un équipement spécialisé (régie audio-visuelle, reconstitution du contexte réel d'utilisation, etc.).

Comment procède-t-on pour évaluer les interfaces ?

Comme l'expliquent BALBO et COUTAZ [1992], la réponse dépend de l'approche (artisanale, science appliquée ou ingénierie). Généralement, le processus commence par la définition d'une situation de référence. Il inclut ensuite une phase d'expérimentation puis une analyse comparative des performances attendues et effectives. Il s'agit ensuite d'expliquer les phénomènes et d'identifier les remèdes. Il existe d'autres techniques, notamment les approches prédictives qui ne nécessitent pas de phase expérimentale.

L'approche artisanale est caractérisée par le fait que le savoir découle de l'expérience, de la créativité, voire de l'intuition [BALBO et COUTAZ, 1992]. Ce savoir empirique peut s'appuyer sur des principes et des théories mais il est tout de même heuristique, c'est-à-dire implicite, informel et difficilement généralisable. L'approche sciences appliquées repose quant à elle sur la connaissance scientifique de théories, modèles et lois développés pour expliquer ou prédire des phénomènes.

Pourquoi évalue-t-on les interfaces Homme(s)-Machine(s) ?

L'évaluation des Interfaces Homme(s)-Machine(s) vise à s'assurer que celles-ci respectent les objectifs recherchés. Le but de l'évaluation peut ainsi s'avérer très varié :

- vérifier le respect du cahier des charges,
- vérifier le respect des critères ergonomiques,
- vérifier le respect des critères socio-culturels,
- vérifier le respect des principes du droit et informatique supervisés pour le cas de la France par la CNIL (Commission Nationale Informatique et Liberté),
- vérifier que l'interface pourra évoluer dans les conditions naturelles du site hôte,
- vérifier les principes de l'assurance qualité des logiciels,
- etc.

I.2. Les intervenants de l'évaluation [GRISLIN et KOLSKI, 1996]

L'évaluation peut être effectuée par un ergonome, le concepteur lui-même, voire un utilisateur [RAVDEN et JOHNSON, 1989]. Il semble que le choix de l'évaluateur soit fonction

de la tâche, de ses compétences, de la phase de développement au cours de laquelle se déroule l'évaluation. L'évaluation demande des compétences d'ergonomes connaissant la tâche, d'informaticiens (ou concepteurs), mais on préférera l'intervention d'une équipe pluridisciplinaire d'évaluateurs [NORMAN, 1986][MEINADIER, 1991][ADREIT *et al.*, 1992][POULAIN, 1994].

En fait, on considère maintenant de plus en plus que l'évaluation n'est véritablement efficace que si la démarche adoptée pour le processus de développement de logiciel est participative [LANNING, 1991][KOLSKI *et al.*, 1993][CACM, 1993]. Par exemple, dans les applications industrielles complexes, la notion de groupe de travail est primordiale [ABED, 1990]. Ces groupes qui composent l'équipe de développement s'organisent autour de quatre types principaux de connaissances : celles de l'ergonome, de l'ingénieur, de l'utilisateur et du concepteur.

- ❖ **Les connaissances de l'ergonome** : celui-ci doit être familier des exigences physiques et cognitives, ainsi que des limites et des ressources théoriques d'un utilisateur réalisant une tâche donnée dans un environnement socio-technique ; il doit aussi être largement sensibilisé au fonctionnement du système technique.
- ❖ **Les connaissances de l'ingénieur** : l'ingénieur est censé connaître l'application, son comportement physique, son modèle et l'interaction entre ses variables. Il est aussi capable de recenser les informations utiles pour le maintien du fonctionnement normal du système et de contribuer à l'analyse des modes opératoires qu'effectue l'utilisateur sur le terrain.
- ❖ **Les connaissances de l'utilisateur** : elles se basent sur la pratique et l'expérience de différentes situations. L'utilisateur connaît l'efficacité de certaines interventions ; dans le meilleur des cas, il a déjà utilisé différents outils informatiques. Il est donc théoriquement en mesure de recenser une grande partie des tâches qui lui ont été assignées et d'en décrire les contraintes.
- ❖ **Les connaissances du concepteur de l'IHM** (informaticien, automaticien) : il apporte ses connaissances liées aux méthodes et outils de développement. Il doit en outre faciliter la liaison entre les données recueillies et leur modélisation.

Ainsi, l'évaluation devrait être suivie et réalisée par un personnel qualifié dans la mise en œuvre d'expérimentations et qui possède des connaissances ergonomiques appropriées au produit concerné. Dans la littérature, on considère que ce type de personnel doit inclure des personnes spécialistes en ergonomie, ou encore en physiologie et psychologie appliquée. Cependant, si l'on considère la grande variété des IHMs, il est clair que certains autres groupes professionnels peuvent réussir à mener à bien des tests utilisateurs. Certaines des tâches incluses dans la conduite de l'évaluation peuvent, à l'occasion, être prises en charge par d'autres personnes, mais un « entraînement » spécifique sera souvent requis préalablement pour assurer des résultats satisfaisants [McCLELLAND, 1990]. Dans ce cas, la collecte des données doit être réalisée de façon bien définie, ce qui est souvent difficile dans un système déjà conçu et pas forcément « ouvert ».

I.3. Traitement des données recueillies [GRISLIN, 1995][PHILIPPE, 1993]

Les mesures effectuées sur le système Homme(s)-Machine(s) regroupent la totalité des variables observées par rapport aux deux propriétés les plus communément étudiées (l'utilité et l'utilisabilité). Elles forment en quelque sorte une base de données expérimentales avec des informations aussi bien objectives (par exemple des taux d'erreurs de manipulation de l'IHM enregistrées automatiquement) que subjectives (par exemple issues des questionnaires ou d'interviews). En fonction de leur nature et dans le but d'élaborer le modèle observé, différents traitements spécifiques sont appliqués.

Le traitement des données repose sur l'utilisation de méthodes d'analyse de données issues des statistiques descriptives [VOLLE, 1981]. L'analyse des données s'applique à des résultats statistiques bruts et permet de faciliter la description et l'interprétation des phénomènes où le nombre des observations est important. L'application de ces méthodes numériques nécessite des connaissances approfondies sur leur fondement théorique, ainsi qu'une bonne pratique pour interpréter les résultats. Sans entrer dans le détail [LOSLEVER, 1988], nous dirons simplement que l'analyse des données issues de l'évaluation repose sur trois orientations :

- ❑ **La mise en évidence de l'effet des facteurs** est basée principalement sur l'utilisation de l'analyse de la variance pour reconnaître les données qui interviennent réellement sur le système homme-machine et modifient le résultat de l'évaluation.
- ❑ **Le traitement des données objectives** fait appel aux méthodes d'analyse multidimensionnelles. Après un codage adapté, elles consistent à faire apparaître les relations existant entre les variables observées, avec par exemple les méthodes factorielles et de classification. SIMON [1993] propose par exemple d'utiliser différents codages pour étudier l'influence de différents facteurs sur le chemin oculaire d'opérateurs de supervision.
- ❑ **L'analyse des données subjectives**, principalement pour le dépouillement des questionnaires, comme par exemple pour l'analyse des données recueillies avec l'échelle de Cooper-Harper, fait également appel à certaines des méthodes multidimensionnelles telles que l'analyse factorielle ou les sous-ensembles flous [GUERRA, 1991].

Pour l'analyse des données recueillies sur l'IHM, on pourra avoir recours à *la méthode expérimentale* de WALLISER [1977], détaillée dans [PHILIPPE, 1993]. Elle peut se représenter comme un cycle itératif d'évaluation qui passe par : (i) le champ théorique d'investigation, point de départ et définition des connaissances du domaine ; (ii) le modèle hypothétique, hypothèse de recherche, définition des variables à contrôler et à mesurer ; (iii) le champ empirique, lieu des observations, et (iv) le modèle observé ou identifié, à partir duquel une conclusion est apportée sur les résultats obtenus. Cette démarche rejoint le principe proposé de comparaison des deux modèles de l'interaction Homme(s)-Machine(s). L'évaluation d'une IHM, en passant par la définition des données de base à recueillir, le recueil puis le traitement des données, nécessite une compétence pluridisciplinaire que se partagent un certain nombre d'intervenants. Enfin, l'évaluation de l'IHM fait partie de l'évaluation du système global ; aussi l'analyse de l'ensemble de la documentation, notamment les aides utilisateur, doit-elle être prise en compte de façon

explicite dans les critères. Ceci nous paraît important, même si elles ne seront en principe quasiment jamais consultées pour une interface qui respecte toutes les recommandations ergonomiques et qui s'avère adaptée à l'utilisateur (facile d'utilisation, rapide à apprendre, intuitive,...), ou pour laquelle l'utilisateur est parfaitement formé. Après avoir collecté les données, il est souvent nécessaire, vu leur nombre, d'effectuer des traitements afin de faciliter l'interprétation pour l'évaluation.

I.4. Conclusion et besoins pour l'évaluation des IHMs

Toute évaluation consiste à identifier ou à prévoir les difficultés que rencontrent les utilisateurs, à détecter les points forts et les points faibles du système, à en comprendre les raisons pour ensuite trouver des solutions correctives. Comme nous venons de le voir, avec toujours un même objectif, les techniques employées sont très dépendantes du contexte. Aussi, avant toute analyse, est-il nécessaire de bien choisir la méthode à employer, liant à la fois le critère à évaluer à la situation concrète [GRISLIN *et al.*, 1995a]. C'est l'un des points sur lesquels nous insistons dans la suite de ce travail, en nous basant sur les classifications actuelles. Dans la majeure partie des cas de développement d'applications informatiques interactives, les interactions Homme(s)-Machine(s) et l'évaluation de leur qualité ne sont pas explicitement et systématiquement prises en compte lors du cycle de vie. Pourtant, il est évident que la qualité des IHMs est cruciale. Aussi, l'utilisation de méthodes et de techniques d'évaluation appropriée s'impose pour l'équipe chargée du développement.

II. Dimensions et critères d'acceptabilité pratique

Pour les évaluateurs, la construction du modèle de l'IHM est dirigée par le contexte, c'est-à-dire que la sélection des propriétés pertinentes pour le diagnostic de la qualité ergonomique dépend non seulement des objectifs de l'évaluation, mais aussi des caractéristiques de la population et des exigences des tâches. Les différentes propriétés ou dimensions d'évaluation sont définies successivement. Elles proviennent d'un examen de la littérature à ce sujet. Notons que certaines telles que celles concernant l'utilisabilité, ont fait l'objet de nombreuses validations. D'autres, telles que celles concernant les aspects socio-culturels nécessitent encore de nombreuses études et expérimentations complémentaires.

II.1. Dimension et critères d'utilité

Cette dimension détermine si l'interface permet à l'utilisateur d'atteindre ses objectifs de travail. Elle correspond par exemple aux capacités fonctionnelles, aux performances et à la qualité de l'assistance technique fournie à l'utilisateur par le système. Il nous semble pourtant pertinent de s'intéresser aussi à l'**utilité** de l'interface en observant *l'adéquation à la tâche et la répartition du travail*. *L'adéquation à la tâche* consiste à vérifier si les procédures cognitives développées par l'utilisateur sont similaires à celles initialement développées par le concepteur, et donc à estimer si la tâche réelle redéfinie par l'utilisateur est en accord avec la tâche prescrite à effectuer [ABED, 1990]. Concernant *la répartition du travail*, il s'agit d'étudier l'adaptation du poste de travail à la tâche, en s'intéressant à la façon dont est organisé le travail en terme de répartition entre l'homme et la machine [MILLOT, 1988 ; DEBERNARD, 1993]. Notons que

l'ensemble de ces critères est rassemblé par [COTE *et al.*, 1993] sous la dénomination de *conformité fonctionnelle*.

De plus, *la performance du système Homme(s)-Machine(s)* n'est évidemment pas à délaissier lors de l'évaluation. L'interface est conçue pour assurer la communication entre l'homme et la machine certes, mais le système Homme(s)-Machine(s) doit produire des résultats. C'est pourquoi un (ou plusieurs) critère(s) de performance du système global est indispensable pour déterminer si les objectifs sont atteints. La difficulté réside dans le choix d'une (ou plusieurs) donnée(s) représentative(s), pertinente(s) et surtout mesurable(s). Différents critères représentatifs liés à l'utilité de l'IHM sont listés ci-dessous :

Critères de la dimension d'utilité :

- **Capacité fonctionnelle**
- **Performance du système**
- **Qualité de l'assistance**
- **etc.**

II.2. Dimension et critères d'utilisabilité

La dimension d'utilisabilité rend compte de la qualité de l'interaction Homme(s)-Machine(s), particulièrement en termes de facilité d'apprentissage et d'utilisation, ainsi que de qualité de la documentation. Le concept d'utilisabilité a été proposé par EASON [1984] et affiné par NIELSEN [1993] pour différencier l'utilité potentielle d'un produit de son utilité réelle. Avant d'évaluer, il convient de définir le plus tôt possible quelles seront les variables à observer et bien sûr la manière de les recueillir. L'analyse de l'IHM doit mettre à jour un ensemble de variables cibles, c'est-à-dire les données de base à recueillir. Ce sont en effet le système à évaluer et son utilisation qui définissent le comportement que l'on mesure lors d'une évaluation.

La littérature propose un très grand nombre de recommandations ergonomiques sous forme de guides [SMITH et MOSIER, 1986b][SCAPIN, 1987][BROWN, 1988][VANDERDONCKT, 1994] ou de standards (BSI 91/40677, ISO 9241). Pour l'évaluation, il s'agit soit de les utiliser comme propriétés à observer ou mesurer, soit d'en extraire des **critères ergonomiques d'évaluation**. En se basant sur plusieurs travaux de recherche [SMITH et MOSIER, 1986b][SCAPIN, 1990][WHITEFIELD *et al.*, 1991], une classification de ces informations a été proposée par BASTIEN et SCAPIN [1993] pour aboutir à un jeu de critères et sous-critères qui prennent en compte ces recommandations. Dix-huit critères élémentaires sont ainsi énumérés et explicités ci-dessous. Plusieurs études ont montré qu'ils pouvaient servir de descripteurs pour l'évaluation des IHMs et ont validé leur définition [BASTIEN, 1996][POLLIER, 1991][BONNARDEL, 1992][GRISLIN, 1995]. Ces critères ou variables cibles ergonomiques d'évaluation sont listés ci-dessous et décrits plus en détail dans le tableau I.1.

Critères de la dimension d'utilisabilité :

- **Le guidage**
 - * Incitation
 - * Groupement par localisation

Chapitre I : Evaluation des systèmes interactifs : problématique, diversité des critères

Critères	Définitions
<p>Le guidage</p> <ul style="list-style-type: none"> • Incitation • Groupement par localisation • Groupement par Format • Retour informatif immédiat • Clarté 	<p>Moyens mis en œuvre pour conseiller, orienter, informer et conduire l'utilisateur lors de l'interaction avec la machine (localisation dans les séquences, actions permises et leurs conséquences, information sur demande). Les critères associés sont :</p> <ul style="list-style-type: none"> • Informations fournies à l'utilisateur sur l'état dans lequel il se trouve, actions possibles ou attendues et moyens de les mettre en œuvre. Une bonne incitation guide et évite par exemple d'avoir à apprendre une série de commandes, permet aussi de connaître le mode et la position dans le dialogue (aides disponibles, format adéquat d'entrée de données, unités de mesures,...). • Positionnement des items les uns par rapport aux autres pour indiquer leur appartenance ou non à une même classe, ou leur distinction : facilité à repérer (puis à mémoriser), informations organisées logiquement (listes hiérarchiques, groupement des options dans des menus en fonction des objets sur lesquels elles s'appliquent,...). • Indices graphiques (format, couleur) des items les uns par rapport aux autres pour indiquer leur appartenance ou non à une même classe, ou leur distinction (facilité à reconnaître les liens entre informations, distinguer les aires de fonctions, ou les labels et champs d'entrée,...). • Réponses de la machine après des actions de l'utilisateur (temps le plus court possible). Le retour d'information engendre la confiance et la satisfaction de l'utilisateur comme la compréhension du dialogue. • Caractéristiques lexicales de présentation des informations : facilité de lecture (textes en sombre sur fond clair plus faciles à lire que l'inverse, textes en majuscules et minuscules plus rapidement lus qu'en majuscules, titres centrés, curseur repérable,...).
<p>La charge de travail</p> <ul style="list-style-type: none"> • Concision • Actions minimales • Charge mentale 	<p>Coût du travail pour l'utilisateur, quantité de travail (mentale et physique) investie. Les surcharges comme les sous-charges (génératrices de baisse de vigilance) ne sont pas acceptables car elles sont la cause, à plus ou moins long terme, d'une défaillance de l'utilisateur et donc du système homme-machine. Les critères associés sont :</p> <ul style="list-style-type: none"> • Charge de travail perceptive et mnésique en rapport avec des éléments individuels d'entrée / sortie : capacités de la mémoire à court terme limitées (pas de codes supérieurs à 4 ou 5 caractères, ou alors utiliser des mnémoniques ou abréviations,...). • Charge de travail au niveau des options ou moyens utilisés pour atteindre un but (minimiser le nombre d'étapes dans la sélection de menus, ne pas entrer plusieurs fois des données, éviter la ponctuation dans des entrées de commande,...). • Densité informationnelle : Charge de travail perceptive et mnésique pour des ensembles d'éléments, éviter d'imposer la mémorisation de longues et nombreuses informations, procédures ou activités nécessitant une activité cognitive complexe (limiter la densité des informations, n'afficher que ce qui est nécessaire, l'information ne doit pas nécessiter de traduction d'unités,...).
<p>Le contrôle explicite</p> <ul style="list-style-type: none"> • Actions explicites • Contrôle utilisateur 	<p>Le contrôle explicite : correspond à la fois au contrôle qu'a l'utilisateur sur l'interface et au caractère explicite de ses actions. Le contrôle du dialogue par l'utilisateur est un facteur d'acceptation du système. Les critères associés sont :</p> <ul style="list-style-type: none"> • Relation entre le fonctionnement de l'application et les actions des utilisateurs : le système doit exécuter seulement les opérations demandées par l'utilisateur, et ce au moment où il les demande (action explicite de validation après sélection d'une option par pointage, l'entrée de commande doit se terminer par <ENTER> avec des possibilités d'édition au préalable,...). • L'utilisateur doit toujours « avoir la main » et donc pouvoir contrôler le déroulement des traitements en cours. Ses actions doivent être anticipées et des options appropriées fournies pour chaque cas ; l'utilisateur doit toujours pouvoir intervenir.
<p>L'adaptabilité</p> <ul style="list-style-type: none"> • Flexibilité • Prise en compte de l'expérience de l'utilisateur 	<p>Capacité du système à réagir selon le contexte, les besoins et préférences de l'utilisateur. Une interface ne peut convenir à la fois à tous les utilisateurs potentiels. Pour qu'elle n'ait pas d'effets négatifs, elle doit s'adapter. Les critères associés sont :</p> <ul style="list-style-type: none"> • Moyens mis à la disposition de l'utilisateur pour la personnalisation de l'interface afin d'adopter ses propres stratégies de travail • Moyens mis en œuvre pour permettre au système de respecter le niveau d'expertise de l'utilisateur. Des utilisateurs expérimentés ont besoin de dialogues moins explicites que les novices (prévoir des raccourcis, des choix d'entrées simples ou multiples selon l'expérience, autoriser différents modes de dialogue,...).
<p>La gestion des erreurs</p> <ul style="list-style-type: none"> • Protection • Qualité des messages • Correction 	<p>Moyens permettant d'éviter ou de réduire les erreurs et de les corriger. Les interruptions provoquées par celles-ci ont des conséquences sur l'activité de l'utilisateur et peuvent causer des accidents graves.</p> <ul style="list-style-type: none"> • Moyens mis en place pour détecter les erreurs d'entrée de données ou de commandes (demande de confirmation avant de quitter une session, protection des labels de champs, accès impossible aux aires non nécessaires à l'utilisateur,...). • Pertinence et exactitude de l'information fournie à l'utilisateur sur la nature des erreurs commises et sur les actions à entreprendre pour corriger (aucune action ne doit suivre une action invalide, fournir des messages d'erreurs orientés tâches,...). • Moyens mis à la disposition de l'utilisateur pour corriger les erreurs. La facilité de correction diminue les perturbations dues aux erreurs (retour arrière ou annulation d'une action possible).
<p>Signification des codes</p>	<p>Adéquation entre l'objet ou l'information affichée ou demandée et son référent. Des codes et dénominations « signifiants » disposent d'une relation sémantique forte avec leur référent (règles de contraction ou d'abréviation explicitées, distinction des labels,...).</p>
<p>Compatibilité</p>	<p>Accord existant entre les caractéristiques de l'utilisateur (mémoire, perception, habitudes, etc.) et l'organisation de l'interface (entrées, sorties, dialogue, architecture, etc.) : procédures et tâches respectant les attentes ou habitudes de l'utilisateur, traductions, transpositions, références à la documentation minimisées.</p>
<p>Homogénéité et Consistance</p>	<p>Conservation des choix d'objets de l'interface (codes, procédures) pour des contextes identiques et non conservation pour des contextes différents (standardisation et homogénéité des codes, procédures, moyens, localisation, format d'écran, procédure,...).</p>

Tableau I.1 : Critères ergonomiques, selon BASTIEN et SCAPIN [1993]

- * Groupement par format
 - * Retour informatif immédiat
 - * Clarté
 - **La charge de travail**
 - * Concision
 - * Actions minimales
 - * Charge mentale
 - **Le contrôle explicite**
 - * Actions explicites
 - Contrôle utilisateur
 - * **L'adaptabilité**
 - * Flexibilité
 - * Prise en compte de l'expérience de l'utilisateur
 - **La gestion des erreurs**
 - Protection
 - Qualité des messages
 - Correction
 - **Signification des codes**
 - * **Compatibilité**
- Homogénéité et Consistance**

L'ensemble de ces critères vise à aider l'évaluateur à estimer la qualité ergonomique de l'interface et à prendre si nécessaire des décisions de modification et/ou d'amélioration. Il existe des relations de causalité entre ceux-ci (par exemple une baisse de la performance peut être la conséquence d'une surcharge de travail) ; c'est pourquoi ils ne doivent pas être étudiés séparément. En fait, les bonnes interfaces réalisent des compromis entre des caractéristiques aussi diverses et souvent contradictoires que la prise en compte de l'intuition et l'expérience de l'utilisateur, la consistance et l'adaptabilité, la flexibilité et les actions explicites, en tenant compte fondamentalement des caractéristiques des utilisateurs potentiels et de leurs attentes.

Les critères d'évaluation en terme d'ergonomie concernent principalement la qualité de l'interface au sens de l'**utilisabilité**. La facilité d'utilisation, d'apprentissage et l'efficacité sont des critères fondamentaux de qualité. Pour l'évaluation, il faudra donc se construire un modèle *a priori* de l'interaction Homme(s) - Machine(s) en prenant en compte l'utilité et l'utilisabilité de l'inter-face. Il est en outre important de bien délimiter le niveau de précision du modèle afin de pouvoir observer et faire ressortir les aspects pertinents vis-à-vis des objectifs du système. Il s'agit donc de déterminer ce qui permettra de porter un jugement sur l'IHM concernée.

II.3. Autres dimensions et critères

La dimension économique est la propriété qui permet de prendre en compte les coûts induits (formation des opérateurs, stages à l'étranger pour les Pays en Voie de Développement, coûts de transports pour les experts devant installer le système dans les pays du Sud, etc.). Dans

certaines Pays en Voie de Développement, ces coûts peuvent remettre en cause toute rentabilité recherchée en informatisant ou en automatisant les installations.

Le manque de main-d'œuvre est parfois un handicap majeur pour les pays du Sud. C'est pourquoi le développement des IHMs doit tenir compte des réalités économiques des pays dans le cadre d'un développement international. Cela revient à dire que pour certains pays du Sud, il est préférable de concevoir des IHMs très simples d'utilisation, moins coûteuses et respectant les fonctionnalités recherchées, en lieu et place de systèmes complexes et coûteux de même fonctionnalité qui finiront par être abandonnés en raison de coûts induits non supportables ou de main-d'œuvre non qualifiée. Le critère d'évaluation de cette dimension est présenté ci-dessous.

Critère de la dimension économique :

- **Coûts induits**

- **La dimension d'environnement** permet de prendre en compte la qualité de l'environnement du/des poste(s) de travail, notamment le niveau de nuisance et les perturbations. Dans certains pays d'Afrique par exemple, les salles de contrôle se trouvent dans des milieux très bruyants, de telle sorte que les opérateurs ont du mal à se concentrer lors de la surveillance des installations.

Critères de la dimension d'environnement :

- **Perturbation**
- **Nuisances sonores et vibratoires**
- **Accessibilité**
- **Conditions climatiques, etc.**

III. Dimensions et critères d'acceptabilité sociale

Les dimensions et critères d'acceptabilité socio-culturels ont fait l'objet d'une attention particulière dans le cadre d'une collaboration avec des sociologues [NENDJO ELLA et JACQUES, 1999].

III.1. Importance des critères d'acceptabilité sociale dans la conception et l'évaluation des Interfaces Homme(s) - Machine(s)

Concernant les exigences sociales, on constate que les mentalités ont beaucoup évolué : les considérations sociales deviennent dans certains pays quasiment aussi importantes que les contraintes d'utilité et d'« utilisabilité ». Les contraintes de la **dimension culturelle** (qui caractérise une partie de l'acceptabilité socio-culturelle des IHMs) renvoient plus particulièrement aux critères suivants : **langue, formes, couleurs, gestuelle, symboles, signes, costumes et moeurs** qui transparaissent au niveau de l'interface logicielle. Chaque pays ou groupe de pays possède en effet une culture propre, des valeurs particulières et des traditions précises, qui peuvent, selon les cas, revêtir une importance considérable lors de la mise en place d'une nouvelle IHM. Cette dernière, en mettant en scène la vie quotidienne, doit être compatible avec la culture qui

caractérise le pays, ce qui signifie qu'elle doit respecter les valeurs, traditions et coutumes locales.

Les figures I.3 et I.4 sont relatives à un exemple de non prise en compte des critères d'acceptabilité culturelle (problèmes de **mœurs**, de **costumes** et de **gestuelle** non respectés dans un pays en voie de développement) et montrent comment celle-ci a entraîné la modification en profondeur d'une interface, entraînant ainsi une baisse importante de la productivité. Il s'agit d'une interface graphique conçue et réalisée aux Etats-Unis pour un programme d'hygiène en Egypte.



Figure I.3 : Composant graphique de l'interface réalisée
Source : *How Fluent is Your Interface ?*

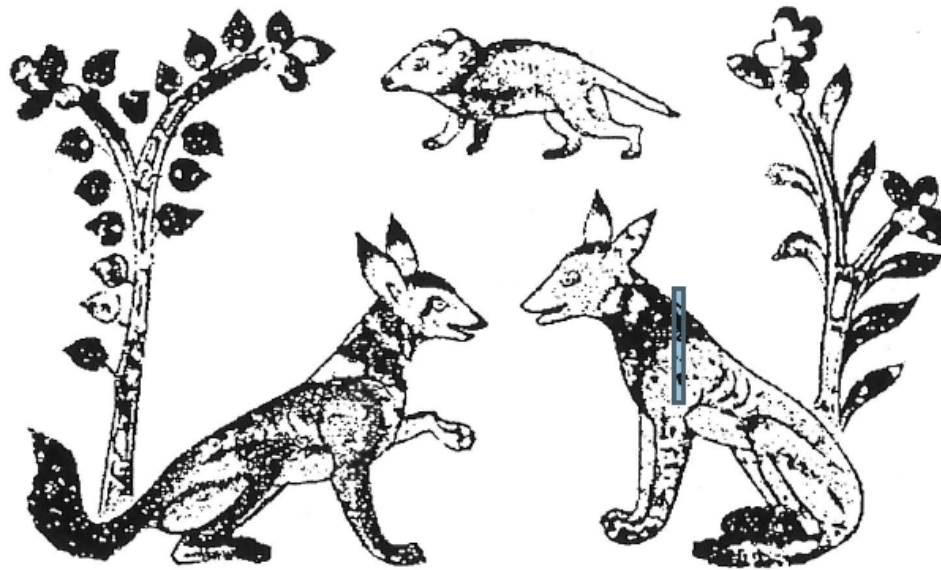


Figure I.4 : Composant graphique modifié de l'interface
Source : *How Fluent is Your Interface ?*
[RUSSO et BOOR, 1993]

Lorsque l'image visible en figure I.3, qui fait apparaître deux femmes égyptiennes dont l'une est en train de subir une contraception, avait été présentée en Egypte, des femmes égyptiennes l'avaient trouvée inacceptable (problème de dignité et de culture). Certaines l'avaient commentée en ces termes : « nous ne ressemblons pas aux femmes occidentales » et « pourquoi laisse-t-elle ses doigts nus ? ». Décision fut donc prise de modifier l'interface, avec les conséquences de coûts que ceci impose. Une série de modifications déboucha sur des graphiques socialement acceptables, dont celui de la figure I.4. Les concepteurs furent donc obligés de revoir ou d'éliminer les images contenant des éléments sujets à controverse.

Un effort est par contre fait pour le critère de **langue** (autre critère appartenant à la dimension culturelle). En effet, les interfaces sont de plus en plus traduites en langue locale. Après deux ans en Egypte et dans d'autres pays arabes, l'auteur de cette thèse a pu constater que les logiciels et claviers d'ordinateurs étaient traduits en langue arabe. La figure I.5 montre un exemple de prise en compte de la **langue** (un des critères socio-culturels) dans le développement des logiciels.

L'exemple de la figure I.5, extrait des composants graphiques de Viewpoint Xerox, montre l'importance grandissante que les grandes compagnies accordent à la dimension culturelle. Sur cette figure, la traduction en trois langues (anglais, arabe et chinois) des composants graphiques témoigne des grandes tendances observées actuellement dans le développement des interfaces Homme(s)-Machine(s). Mais malgré cette prise de conscience des aspects socio-culturels, la langue est le seul critère à être de plus en plus pris en compte.



The Enchanted Mouse

Once there was an ascetic whose wishes always came true. One day as he was sitting on a river bank, a hawk passed by with a little mouse in its talons, and dropped the mouse near the ascetic. Taking pity on it, he wrapped it in his sleeve to take home with him. But then he worried that his wife might object to raising a mouse, so he made a wish for it to change into a little girl. His wish came true, and the mouse changed into a little girl endowed with grace and beauty, and the ascetic hurried home with her and said to his wife:

"Here is my daughter, treat her like your own child."

الفأرة المسحورة

زعموا ان ناسكا كان مستجاب الدعوة؛ فبينما هو ذات يوم قاعد على شاطئ نهر اذ مرت به حداة في رجلها درصة فوقعت منها عند الناسك. فادركه لها رحمة، فأخذها ولفها في رده، واراد ان يذهب بها الى منزله. ثم خاف ان يشن على امرأته تربيتها، فدعا ربه ان يحولها جارية. فتحولت جارية واعطيت حسنا وجمالا. فانطلق بها الناسك الى منزله، وقال لامرأته:

هذه ابنتي فاصنعي بها صنيعك بولدك.

鼠精

從前有一個苦行者，他的願望總是能實現。有一天，當他坐在河堤上時，一隻老鷹從旁邊飛過，將它爪中捉著的小老鼠，掉落在苦行者的附近。

那位苦行者懷著憐憫之心，將這隻小老鼠包藏在他的袖子裏，要帶回家去。但是他又擔心他的妻子可能會反對撫養這隻小老鼠，因此他就下了一個願，祈望將這隻小老鼠變成一個美麗大方的小姑娘。他的願望馬上實現了，小老鼠變成了一個漂亮的小姑娘。苦行者急忙帶著小姑娘回家並對他的妻子說：

「這是我的女兒，請您像親生女兒一樣的對待她。」

Figure I.5 : Composant graphique issu de Viewpoint Xerox
Source : *Designing User Interfaces for International Use* [NIELSEN, 1990]

Les autres critères socio-culturels devraient suivre cette tendance dans les années à venir. Notons que les problèmes culturels font déjà l'objet d'études importantes actuellement dans le monde, avec pour objectif de se rapprocher de plus en plus des réalités locales. Les figures I.6a à I.6f illustrent un autre exemple de prise en compte des critères culturels (ici la langue) dans le développement des IHMs pour les pays suivants : Grande-Bretagne, Chine, Corée, pays arabes, Thaïlande et Urdu .

Certains laboratoires se spécialisent de plus en plus dans la traduction automatique des logiciels. On notera le réseau francophone de l'ingénierie linguistique FRANCIL qui s'agrandit

et rassemble déjà plusieurs pays, dont certains en voie de développement (Egypte, Tunisie, Maroc, Belgique, Canada, France, Suisse).

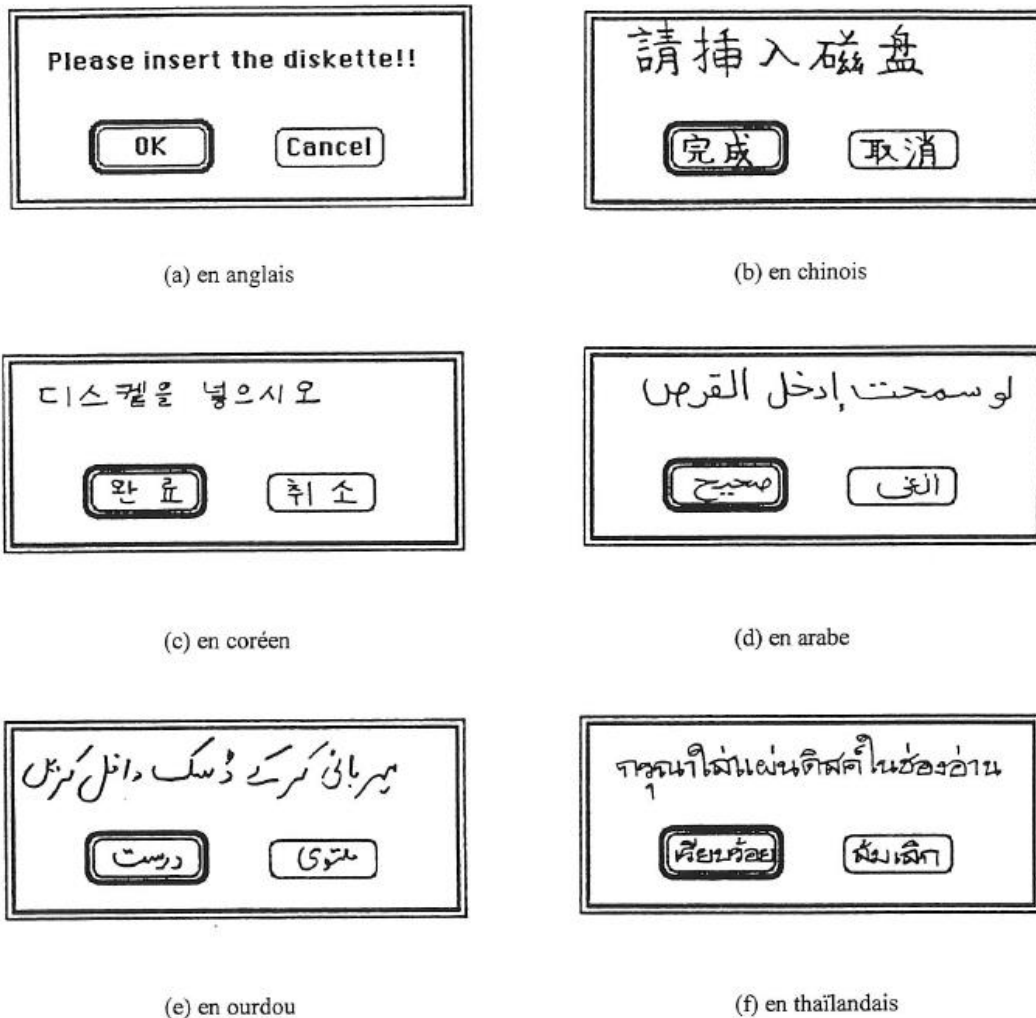


Figure I.6 : Exemples de boîte de dialogue disponible en plusieurs langues

Source : *Designing User Interfaces for International Use* [NIELSEN, 1990]

Les aspects socio-culturels, qui constituent des contraintes pour les concepteurs, ne sont pas toujours pris en compte dans le développement des IHMs. De plus, les évaluateurs ne disposent que très rarement de méthodes ou techniques d'évaluation pour évaluer les IHMs en terme d'acceptabilité socio-culturelle.

C'est ainsi que dans le cadre de nos travaux, des experts en socio-technique ont mis au point une technique (Annexe 3) basée sur un questionnaire et contribuant à évaluer sous un angle socio-culturel des IHMs à partir de la conception préliminaire, c'est-à-dire lorsque les premiers composants graphiques sont déjà disponibles sur papier. En cas de non respect de facteurs socio-culturels, l'interface est donc susceptible de provoquer une gêne, un malaise ou un rejet de la part de ses utilisateurs, auquel cas elle est dans l'incapacité d'être utilisée de façon efficace, et ceci

particulièrement dans certains pays. C'est pour éviter ce type de réaction que l'évaluation doit tenir compte des spécificités culturelles de chaque nation, en veillant particulièrement aux formes et aux couleurs de l'image qu'elle met en scène, de même qu'à la gestuelle, aux symboles, aux costumes et aux mœurs qu'elle reproduit.

Les problèmes de **symboles** et de **signes** sont également des critères culturels à prendre en compte. L'exemple suivant montre comment la prise en compte des symboles et signes est nécessaire pour faire passer une information compréhensible par toutes les sociétés (figure I.7.).

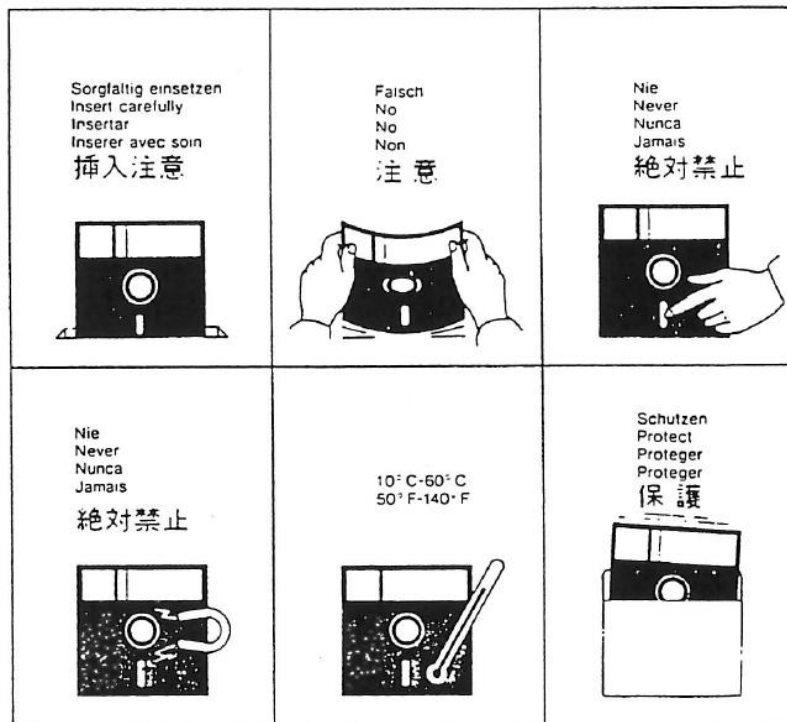


Figure I.7 : Exemple de clarifications d'usage de protections de disquette
Source : *Designing User Interfaces for International Use* [NIELSEN, 1990]

Le respect des mœurs et croyances locales intervient de la même manière pour expliquer l'acceptation ou le rejet des systèmes informatiques : toute IHM en contradiction avec celles-ci risque d'être rejetée pour non conformité avec l'environnement culturel dans lequel elle est appelée à fonctionner. L'intervention de l'informatique dans le domaine de l'hygiène alimentaire, par exemple, ne peut être acceptée par tous les pays du monde.

Les problèmes culturels relatifs aux couleurs sont également importants. C'est ainsi que les concepteurs d'interfaces doivent par exemple veiller attentivement aux **couleurs présentes** dans l'imagerie graphique : dans de nombreuses nations, chaque couleur possède une signification particulière. Ainsi le blanc symbolise-t-il la pureté au Ghana et est porté pour célébrer les naissances, tandis que le blanc au Japon symbolise la mort et est porté en cas de décès ; de même le rouge signifie-t-il danger aux Etats-Unis (voir Tableau I.2) où il servira à signaler toute situation à risque, notamment par l'intermédiaire des feux de route. En France, le bleu signifie

liberté et paix. Cette couleur est notamment utilisée pour la reconnaissance des conduites d'eau dans les centres de traitement d'eau, de même que pour les conduites d'oxygène. Toutes ces matières symbolisent la paix, la neutralité et la source de vie.

Dans certains pays, l'utilisation irréfléchie des couleurs pourrait dans ces conditions être de nature à provoquer le rejet du nouveau système, par crainte des retombées possibles issues de la confrontation avec des couleurs chargées de sens, et créer des confusions dans certains cas.

Aspects culturels/ couleurs	rouge	bleu	vert	jaune	blanc
USA	danger	masculinité	sécurité	lâcheté	pureté
France	aristocratie	liberté paix	criminalité	temporaire	neutralité
Egypte	mort	mérite confiance vérité	fertilité solidité	joie prospérité	joie
Inde	vie créativité		prospérité fertilité	succès	mort pureté
Japon	colère danger	banditisme	futur jeunesse énergie	grâce noblesse	mort
Chine	joie	cieux nuages	dynastie cieux nuages	naissance ressource puissance	mort pureté

Tableau I.2 : Exemples de caractéristiques culturelles de couleurs
[RUSSO et BOOR, 1993]

Lors du développement des IHMs de supervision (par exemple pour le contrôle de la production d'énergie électrique dans les centrales ou pour les procédés chimiques), on constate parfois que certains concepteurs ignorent la signification des couleurs et les utilisent à mauvais escient [KOLSKI *et al.*, 1993]. Dans le cadre du développement international où l'acceptabilité socio-culturelle est aussi importante que l'acceptabilité pratique, ces IHMs peuvent devenir sources de malaise pour certains opérateurs.

Les autres dimensions de l'acceptabilité socio-culturelle, telle que la **dimension sociale**, renvoient plus particulièrement aux caractéristiques propres à la société : **confidentialité, emploi, savoir-faire, sécurité-innocuité**, qui surgissent lors de l'usage des IHMs. Chaque pays ou groupe de pays possède, on l'a dit, ses caractéristiques sociales qui lui sont propres et qui peuvent, selon les cas, revêtir une importance considérable lors de la mise en place d'une nouvelle interface. Cette dernière, en remettant par exemple en cause le savoir-faire manuel des opérateurs ou les emplois déjà acquis, peut provoquer de la frustration. Les opérateurs peuvent alors refuser l'usage d'une telle interface.

Chacune des trois dimensions de l'acceptabilité sociale renvoie à des critères d'évaluation précis. La **dimension d'éthique** fait référence à un critère, essentiel, qui est celui de **moralité**. Une IHM, pour être acceptée d'emblée ou pour ne pas être contestée après sa commercialisation, doit absolument être en accord avec les principes éthiques qui donnent à la société son sens, en inculquant à ses enfants les fondements de la morale. Aussi le respect de ce critère d'éthique doit-il dans certains contextes faire l'objet d'une vérification et d'une validation avant la mise sur le marché de l'IHM.

Un exemple de système interactif dans lequel la dimension d'éthique a été complètement négligée (ce qui a entraîné une vive protestation de la part du public) est celui d'un système écossais intitulé « Grand Theft Auto » (Grand Vol de Voitures) [MIZIO, 1998]. Reposant sur un scénario particulièrement amoral (commettre des braquages, voler des voitures, écraser des piétons, vendre de la drogue, forcer les barrages de police et tuer des policiers), ce système peut inciter les (jeunes) utilisateurs à commettre des actes non seulement répréhensibles par la loi, mais aussi en contradiction totale avec la morale en vigueur. Ce cas a été révélé grâce aux observations d'un policier français qui a constaté que son fils se plaisait à tuer des policiers à travers ce système. Estimant que ce jeu était moralement inacceptable et qu'il pouvait inciter son fils à la violence, ce policier a décidé de contacter le syndicat de la police. Ce dernier, après avoir visualisé le contenu global du système, a finalement déposé plainte, estimant que ce système posait effectivement un problème d'éthique : on peut de fait s'interroger sur les conséquences éventuelles du jeu GTA dont l'objectif est de vendre de la drogue et de tuer des policiers pour le compte de la mafia.

A ce sujet, notons que WOOD-HARPER et ses collègues [1996] proposent un système d'analyse éthique dans le développement des IHMs (voir tableau I.3 ci-dessous). Cette approche comprend 3 grandes étapes à prendre en compte au cours du processus de développement des interfaces Homme(s) – Machine(s). La première étape consiste à identifier des indicateurs ou des signes ayant un rapport avec les problèmes d'éthique, en d'autres termes à mettre en évidence les critères d'éthique que nous verrons dans le chapitre II. Ces critères sont appelés ici des indicateurs ou des signes en rapport avec les problèmes d'éthique.

Etapes	Activité
1	Identifier les indicateurs ou les signes en rapport avec des problèmes d'éthique
2	Identifier leurs perspectives dominantes. Si les perspectives dominantes n'existent pas, identifier celle que vous estimez être la plus représentative
3	Dresser une carte des différentes perspectives

Tableau I.3 : Analyse d'éthique pour le développement des IHMs [WOOD HARPER *et al.*, 1996]

Dans une seconde étape, il faut identifier ou mettre en évidence les critères d'éthique les plus importants (ou dominants) du système à évaluer. Si aucun des critères n'apparaît plus

important ou dominant par rapport aux autres, il faut identifier celui qui semble le plus représentatif en terme d'éthique par rapport au système à évaluer. Par exemple, en ce qui concerne le jeu vidéo GTA que nous venons de voir, le critère dominant ou représentatif est la *moralité*. Enfin, il faut dresser une carte ou tableau des autres critères et de leur impact sur le système à évaluer. L'impact est la conséquence possible sur le plan social par exemple, que nous estimons ou que nous prévoyons en cas de non prise en compte des critères d'éthique dominants ou représentatifs ; en d'autres termes, nous dressons un tableau de « perspectives ».

III.2. Autres exemples tirés de Pays en Voie de Développement (PVD)

Lors du développement de systèmes interactifs dans les Pays en Voie de Développement, la prise en compte des critères socio-culturels est d'une importance capitale. Notre travail le démontrera sur quelques exemples représentatifs. D'après NIELSEN [1993], pour parvenir à des systèmes interactifs dits acceptables, il s'agit, on l'a dit, de les considérer selon deux angles : l'acceptabilité sociale et l'acceptabilité pratique. L'acceptabilité pratique a déjà été étudiée. Par contre, la notion d'acceptabilité sociale est rendue complexe car des facteurs exogènes difficiles à maîtriser influencent de façon considérable la vie sociale des individus. En particulier, certaines spécificités observées dans les Pays en Voie de Développement doivent attirer plus que jamais l'attention des concepteurs provenant des pays industrialisés, particulièrement lors de développements technologiques en général et de la conception de systèmes interactifs en particulier. En effet, et sans doute beaucoup plus que dans les pays industrialisés, en raison du nombre élevé d'échecs ou de difficultés graves (systèmes rapidement abandonnés, modifications profondes apportées après livraison, mauvaise utilisation,...) lors de projets de conception de systèmes interactifs pour certains pays en voie de développement, il y a matière à réflexion. Il s'avère que certaines raisons sont largement dues à des critères socio-culturels (en relation éventuellement avec des critères économiques) non ou mal pris en compte.

Cette section a pour objectif de discuter, sans souci d'exhaustivité, de problèmes représentatifs pouvant survenir lors d'un projet de développement de systèmes interactifs destinés à certains PVD. Certains exemples cités dans ce chapitre proviennent de l'expérience de travail et des stages effectués par l'auteur de cette thèse dans de nombreux Pays en Voie Développement.

Selon WISNER [1985], dans les PVD, le développement technologique doit s'adapter à certaines caractéristiques de la société, et particulièrement à ses valeurs socio-culturelles. Il faut noter que de telles valeurs sont très souvent ignorées ou négligées par les équipes de développement. Ainsi, généralement, dans les PVD, les entreprises sont limitées par les moyens financiers et le peu de main-d'œuvre qualifiée. Elles sont parfois obligées de faire appel aux fournisseurs étrangers pour la mise à jour des systèmes commandés, en assumant les conséquences financières qui en découlent (transports internationaux, hébergement, frais de mise à jour, etc.). Certaines entreprises sont par voie de conséquence obligées d'abandonner leurs interfaces après un certain temps, lorsqu'elles sont convaincues que les frais de mise à jour deviennent insupportables pour leur trésorerie, sans compter les frais d'amortissement du matériel informatique.

Les échecs sont également liés au fait que la prise en compte des considérations socio-culturelles, parfois incontournable, a un coût. Considérons l'exemple suivant concernant une entreprise angolaise qui recevait des logiciels interactifs en provenance d'un fournisseur américain. Pour des raisons d'économie sur le travail de traduction, l'entreprise angolaise avait demandé à ce que seuls les logiciels soient traduits en portugais, le manuel d'utilisation restant en anglais [WISNER, 1985]. Ce fut un échec lors de l'utilisation effective du système car l'anglais n'était pas une langue (*critère de la dimension culturelle*) accessible à tous.

L'exemple suivant se passe aux Philippines : un logiciel de gestion conçu et réalisé en Occident était bien maîtrisé par les informaticiens philippins en termes d'utilisabilité. Mais sur le plan socio-culturel, il y avait des modifications à apporter. Les informaticiens auraient été capables de les effectuer, mais pour des raisons d'intérêt financier et de rétention de savoir, le fournisseur interdisait toute modification dans le contrat de vente. Le fournisseur américain voulait lui-même s'occuper de la maintenance et des modifications, et ceci à la charge financière de l'entreprise philippine. Ceci est un exemple de transfert inachevé par rétention du savoir [WISNER, 1984].

Sous cet angle, le développement des systèmes destinés aux PVD devrait dans l'idéal se baser sur un état d'esprit totalement différent, en visant à prévoir au maximum toute forme d'« aide » aux différentes personnes qui auront à prendre en main le système et à l'utiliser, et aussi à celles qui auront éventuellement à le maintenir et à le faire évoluer : supports papier et/ou aides plus ou moins interactives (manuel de référence et d'utilisation). Ces aides devraient non seulement faciliter l'utilisation du système, mais aussi permettre une formation continue des agents sur le site.

Cet aspect d'aide fait partie des critères classiques d'utilisabilité d'un logiciel [NIELSEN, 1993][BASTIEN et SCAPIN, 1993], qui seront dans le contexte des PVD d'une importance capitale. Les développeurs occidentaux devraient donc prévoir, dans la mesure du possible, des supports (le plus souvent écrits) fournissant à l'entreprise locale un maximum d'informations indiquant par exemple les procédures de mise à jour pouvant permettre aux informaticiens de l'entreprise, pas toujours qualifiés pour ce genre de travail, d'être les plus autonomes possibles. Nous devons éviter, dans ce contexte, ce que WISNER [1985] qualifie de transfert incomplet par rétention du savoir. Nous devons à chaque fois que possible nous rapprocher du transfert dit complet, c'est-à-dire qu'il ne s'agit pas seulement de vendre le système interactif à ces pays, mais également l'organisation du travail, qui va inclure également les supports d'aide à l'utilisation et à la mise à jour.

Une telle démarche n'entraînerait pas seulement des avantages en termes d'utilisabilité d'une part et d'économies que ces entreprises du Tiers-Monde pourraient réaliser d'autre part : elle devrait permettre d'éviter les problèmes de communication que connaissent certains pays avec le reste du monde. Par exemple, une entreprise du Zaïre qui veut contacter son fournisseur pour une mise à jour urgente d'un système ne saura pas toujours le faire par téléphone ou par fax car les liaisons hertziennes ne sont pas favorables avec le reste du monde (sans compter que les connexions à Internet sont encore loin d'être généralisées dans les PVD). Il faut parfois attendre des semaines, voire des mois, car le courrier passe par les pays limitrophes. C'est le cas de

plusieurs pays du Tiers-Monde dont certains sont en guerre et/ou ne disposent pas de moyens de communication viables.

De plus, les contrats de maintenance et/ou d'évolution du système interactif destiné aux PVD doivent être rédigés différemment de la façon dont ces contrats le sont habituellement en Occident. En effet, au lieu que le contrat stipule que l'entretien et la mise à jour du système interactif sont du ressort du fournisseur (qui se trouve généralement sur un autre continent), il faut viser à permettre à l'entreprise utilisatrice d'assumer ces responsabilités si elle dispose d'un personnel qualifié pour de telles tâches.

Aux Pays-Bas, en 1972, sous l'égide de l'OTAN, s'étaient réunis 44 chercheurs pour déterminer les causes d'échec du transfert technologique en direction des Pays en Voie de Développement (PVD). Parmi les 44 chercheurs, 6 venaient des PVD, dont l'Afrique du Sud et l'Inde. Les chercheurs occidentaux avançaient des raisons liées aux différents conflits militaires qui désorganisaient toute forme d'organisation dans les PVD. Les chercheurs du Tiers-Monde exposaient plutôt des facteurs liés à la qualité de vie (nutrition, santé, éducation, problèmes familiaux,...) et aux données anthropométriques (par exemple, la taille moyenne des Pygmées de l'Afrique centrale avoisine 1,50 m, ce qui rend difficile l'utilisation de certaines machines) [WISNER, 1984].

Les conditions socio-économiques dans lesquelles vivent de nombreux travailleurs des PVD sont également l'une des causes d'échec. En effet, dans de nombreux cas, une alimentation mauvaise ou déséquilibrée ainsi que des maladies endémiques largement plus courantes que dans les pays occidentaux, entraînent des fatigues généralisées non adaptées par exemple pour des tâches humaines de supervision et de résolution de problèmes (parfois très complexes) en interaction avec des systèmes interactifs en salle de contrôle. Un tel cas fut constaté en 1988 pendant l'un de nos voyages d'étude dans une centrale électrique sur le continent africain. Tous les jours à partir de 11 heures, on constatait des signes de fatigue généralisée chez la plupart des opérateurs, avec des conséquences notables sur la performance du système homme-machine. Certains opérateurs avouaient qu'ils n'avaient pas pris de petit-déjeuner faute de moyens financiers. Mais il faut également souligner que les interfaces Homme(s)-Machine(s) étaient compliquées, le plus souvent alphanumériques, dans une langue non maîtrisée par la plupart des opérateurs. Ce problème pourrait en fait être résolu à différents niveaux, de l'organisation du travail (collation offerte aux opérateurs par exemple ; îles anthropotechnologiques : cf. plus loin) jusqu'à l'amélioration de la présentation de l'information.

Dans les PVD en général et en Afrique en particulier, l'opérateur humain est un acteur de l'entreprise, tout en possédant des revenus limités. Il est souvent chef d'une famille très nombreuse (parfois 10 à 15 enfants). La femme est généralement analphabète dans des pays où le taux de chômage peut dépasser 40% de la population active. Le rôle de la femme est souvent limité aux tâches ménagères. Les problèmes administratifs, comme l'inscription des enfants à l'école, relève des attributs du père. C'est ainsi que la société admettra de fait que le père opérateur humain ait la possibilité de traiter certains problèmes sociaux depuis le site de travail. C'est pourquoi on constate souvent que les opérateurs humains passent beaucoup de temps au téléphone, pas forcément dans le cadre professionnel, mais plutôt pour répondre aux appels

téléphoniques en provenance de leur domicile, ce qui a souvent des incidences néfastes sur les performances du système Homme(s)-Machine(s). Il arrive que l'opérateur soit informé plusieurs fois dans la journée d'un ou plusieurs cas de maladie d'enfants ou de problèmes d'intendance. L'opérateur peut également appeler son domicile pour s'assurer que tout va bien. Il arrive que, lorsque le téléphone n'existe pas (250 postes téléphoniques pour 1000 habitations : c'est la moyenne dans plusieurs pays d'Afrique), l'on puisse voir débarquer en salle de contrôle la femme et les enfants, pour informer le chef de la famille d'un problème urgent à régler (une maladie grave d'un enfant par exemple).

Ce cas a été constaté entre autres en 1992 au Gabon lors de la mise en œuvre d'un système interactif de gestion de l'eau : on pouvait constater que les appels reçus par les opérateurs avaient une fréquence de nature à perturber le travail et à dégrader la performance recherchée. Dans ce contexte très perturbateur, il est évident que l'opérateur, face à ses écrans de supervision, ne réalise pas ses tâches dans des conditions normales.

Cette possibilité qu'ont de nombreux opérateurs de régler les problèmes familiaux (plus ou moins urgents) dans le cadre professionnel ne relève pas d'une autorisation officielle de l'entreprise : c'est souvent une acceptation de fait. On semble en effet ignorer le problème et on « ferme les yeux » pour ne pas compliquer la situation des opérateurs qui ont déjà trop de soucis et du mal à nourrir leurs enfants convenablement.

L'expérience tentée par certaines entreprises, visant à interdire les membres de la famille ou la résolution de problèmes familiaux depuis le milieu professionnel, s'est le plus souvent soldée par un échec avec des conséquences notables pour le rendement du système et/ou l'ambiance dans l'entreprise. Il existe tout de même des cas plus ou moins réussis, notamment des entreprises à capitaux étrangers dirigées par des Occidentaux (par exemple Elf Congo ou Elf Gabon), dont les chefs de service et les contremaîtres sont souvent des expatriés. Dans ces entreprises, des moyens de compensation ont été fournis en contrepartie de règles draconiennes sur l'utilisation du téléphone, comme la prise en charge complète par l'entreprise des problèmes de santé et de scolarité des enfants, et ceci par la construction proche des logements des opérateurs, d'écoles et de dispensaires. Bref, les opérateurs sont éloignés des soucis familiaux et peuvent se consacrer à leur travail.

L'organisation est également présente car les travaux sont quotidiennement encadrés par les Occidentaux. Ces opérateurs africains bénéficient de programmes réguliers de formation. Notons aussi que dans ces entreprises étrangères, les opérateurs bénéficient de salaires conséquents et supérieurs à ceux de leurs compatriotes dans les entreprises nationales. Ces entreprises expatriées correspondent à ce que WISNER [1984] appelle des « îles anthropotechnologiques », où l'on ne transfère pas seulement le matériel et le logiciel, mais aussi l'organisation et les conditions de travail.

Un exemple marquant de critère social à considérer lors de la conception d'un système interactif est le suivant. S'il est précisé que le logiciel est destiné à un PVD où le taux d'alphabétisation est très faible (le Tchad par exemple), et plus précisément dans une entreprise nationale où il n'existe pas d'utilisateurs qualifiés pour la tâche visée et dont les moyens

financiers ne permettent pas de faire appel à des experts occidentaux, alors la prise en compte du critère social est fondamentale. Il s'agit véritablement de concevoir un logiciel le plus simple possible d'utilisation, tout en remplissant les contraintes du cahier des charges. Un effort ergonomique très important de la part de l'équipe de développement est ici incontournable. Il s'agit en fait d'arriver à un système interactif le plus convivial possible, prenant en compte l'expérience et les connaissances des utilisateurs.

Dans le domaine particulier de la supervision des systèmes industriels, on peut souligner l'importance d'un courant de recherche portant actuellement sur des interfaces Homme(s)-Machine(s) qualifiées d'écologiques [RASMUSSEN et VICENTE, 1989][VICENTE, 1995][MORAY *et al.*, 1995], à partir d'une théorie issue de GIBSON [1979]. La proposition de ces auteurs consiste à concevoir des interfaces Homme(s)-Machine(s) qui tirent parti des ressources les plus puissantes que les hommes possèdent pour affronter des situations complexes. Parce que l'utilisateur est plus efficace dans ses processus mentaux automatisés de bas niveau, ces auteurs déclarent qu'il est nécessaire de présenter l'information de telle sorte qu'elle ne doive que très rarement sortir de ces processus de bas niveau. Une telle interface vise à donner une vue « directe » de l'état du système, sans que l'opérateur ait à réfléchir pour comprendre et interpréter celui-ci. Il serait extrêmement intéressant de tester de telles interfaces « écologiques » dans le contexte d'applications destinées aux PVD.

Nous concluons, au vu des données précédentes, sur le fait que la réussite d'une informatisation ou d'une automatisation dans les PVD passe nécessairement par une approche socio-technique. Dans ce cadre bien précis, comme l'expliquait déjà PARETO en 1919, nous dirons que contrairement à la société occidentale, dans les PVD, la société est un mélange d'actions logiques et non-logiques : il est vraiment nécessaire d'aborder l'intégration de la société comme **un tout** et ceci quelle que soit l'application visée ; il faut prendre en compte tout ce qui la compose. Il est important que les équipes de développement occidentales changent leur état d'esprit, souvent calqué sur les besoins et les caractéristiques des applications destinées aux pays occidentaux.

III.3. Degré d'exigence socio-culturelle selon les pays

Pour éviter les problèmes liés à la non prise en compte des facteurs socio-culturels, dont l'importance varie largement d'un pays à un autre, un chercheur spécialiste de la socio-technique (JACQUES C.) de l'IFRESI (Institut Fédératif de Recherche sur les Economies et les Sociétés Industrielles) a mené dans le cadre de nos travaux une étude, validée auprès de 170 chercheurs et étudiants, afin de déterminer les pays où la prise en compte des critères d'acceptabilité socio-culturelle est primordiale ou absolument nécessaire [NENDJO ELLA et JACQUES, 1999].

Parmi les divers facteurs influençant le choix des méthodes d'évaluation se trouve la contrainte de localisation, c'est-à-dire la contrainte du lieu (site, laboratoire, pays) dans lequel l'IHM sera évaluée et mobilisée. La prise en compte de ce lieu, et notamment du pays d'accueil de l'interface, a une importance considérable dans la mesure où c'est elle qui conditionne son acceptabilité sociale : la spécificité des problèmes propres à chaque société nécessite des démarches spécifiques, qui vont de la conception jusqu'au stade de la réalisation. Toute IHM doit

de fait être acceptée par la société dans laquelle elle est destinée à fonctionner, ce qui signifie que la conception et l'évaluation devront tenir compte non seulement des dimensions d'acceptabilité pratique (utilité, utilisabilité, etc.), mais aussi des critères d'acceptabilité sociale.

Le terme « sociale » renvoie en fait à trois réalités distinctes : l'éthique, le culturel et le social, que l'évaluation doit absolument intégrer pour assurer une insertion sans risque de l'IHM dans un pays donné. Pour être socialement acceptable, une interface doit être compatible avec l'éthique et le bien-être social que chaque nation respecte et se doit d'assurer à ses citoyens, de même qu'elle doit être en accord avec la culture particulière qui la caractérise. Ainsi les dimensions sociale et éthique concernent-elles l'ensemble des pays du monde, tandis que la dimension culturelle est relative au pays considéré.

Chaque pays possède de fait une culture, qu'il partage souvent avec un groupe plus ou moins restreint de pays voisins. Divers pays peuvent ainsi être rassemblés sur la base de leur proximité géographique et culturelle, et les groupes ainsi formés comparés entre eux en fonction de leur attachement à leur culture, à leurs valeurs et à leurs traditions. 21 groupes de pays ont été définis sur cette base et soumis à un test auprès de 170 chercheurs et étudiants en sciences humaines et sociales (sociologie, ethnologie, art et culture, histoire et sociétés, démographie). Cette étude a permis d'effectuer un classement des 21 groupes (soit 209 pays) par ordre décroissant en fonction de l'importance accordée (respect, attachement) aux critères socio-culturels (voir tableaux I.4, I.5, I.6 et I.7, ci-après).

Les critères estimés dans le cadre de cette étude sont regroupés en trois dimensions d'évaluation : la dimension d'*éthique* (moralité), la dimension *sociale* (confidentialité, emploi, sécurité-innocuité, savoir-faire) et la dimension *culturelle* (langue d'usage, formes, couleurs, symboles, croyances, gestuelle, costumes, moeurs).

Le tableau I.7 présente le résultat de l'étude qui a permis d'effectuer un découpage de l'ensemble des pays du monde en deux grands ensembles, dont le premier réunit les pays dans lesquels la culture et les traditions ont un poids très important dans tous les domaines de la vie quotidienne, tandis que le second groupe rassemble à l'inverse les pays dans lesquels la prise en compte des facteurs socio-culturels est peu, voire pas importante lors de la mise en place d'une IHM. L'examen des 170 découpages proposés a permis d'établir la coupure entre ces deux ensembles de pays au niveau du groupe N (tableau I.7), [NENDJO ELLA et JACQUES, 1999].

Les résultats nous montrent que lors du développement des IHMs, dans un pays comme l'AFGHANISTAN (groupe G), les critères d'acceptabilité socio-culturelle sont très importants et exigés. Ce pays se trouve dans un groupe où le degré d'exigence des critères socio-culturels est très élevé : dans le tableau I.5, on peut ainsi lire : critères culturels : **absolument nécessaires** ; critères d'éthique : **nécessaires** ; critères sociaux : **absolument nécessaires**.

De même, si l'on considère les PAYS-BAS, on constate qu'ils se situent dans le groupe I, ce qui permet de lire, dans le tableau I.5 : critères culturels : **pas nécessaires** ; critères d'éthique : **absolument nécessaires** ; critères sociaux : **nécessaires**.

Ces informations sont utiles pour les évaluateurs et les concepteurs car elles permettent d'éviter la prise en compte systématique des critères d'acceptabilité socio-culturelle, ce qui reviendrait très coûteux dans certains cas. En effet, alors qu'ils s'avèrent d'une nécessité absolue ailleurs, ils se révèlent moins importants dans certains pays. C'est une information capitale pour le développement des systèmes interactifs dans un contexte international (cf. par exemple les applications développées et distribuées mondialement par MICROSOFT).

Ces informations sont prises en compte par l'outil d'aide à la décision développé dans le cadre de cette recherche pour le choix des méthodes et techniques d'évaluation (cf. à partir du chapitre III), méthodes et techniques qui sont accompagnées des critères d'évaluation correspondants. Lorsque le système demande le pays de destination de l'IHM développée, il sait, grâce à cette étude, s'il est nécessaire ou non de prendre en compte les critères socio-culturels.

III.4. Résumé des dimensions et critères à prendre en compte

Chacune des dimensions d'éthique, sociale et culturelle renvoie à des critères d'évaluation précis dont il s'agit ici de justifier la nécessité.

III.4.1. La dimension et les critères d'éthique

La dimension d'*éthique* fait référence à un critère, essentiel, qui est celui de *moralité*. Une IHM, pour être acceptée d'emblée ou pour ne pas être contestée après sa commercialisation, doit absolument être en accord avec les principes éthiques qui donnent à la société son sens. Aussi le respect de ce critère d'éthique doit-il dans certains contextes faire l'objet d'une vérification et d'une validation avant la mise sur le marché de l'IHM. C'est pourquoi on recommande, dans certains cas, la prise en compte des critères qui permettront au système évalué d'être socialement acceptable. Une difficulté réside toutefois dans le fait que le rôle joué par le spécialiste en éthique est généralement biaisé par sa vision personnelle sur un système donné [WOOD-HARPER *et al.*, 1996].

Critère de la dimension d'*éthique* :

- Moralité

III.4.2. La dimension sociale et ses critères

La seconde dimension d'évaluation d'une IHM est sociale et renvoie à quatre critères. Le premier d'entre eux est le critère de *confidentialité* : pour être socialement acceptable, un système, relevant du domaine médical par exemple, doit garantir la confidentialité des informations dont il est porteur (cf. Commission Nationale Informatique et Liberté). Une illustration de l'importance de ce critère peut être trouvée en France, pays qui expérimente un nouveau système de gestion informatique des données sur l'état des patients. Ce système informatique, à la disposition des médecins, permettra d'accéder à l'ensemble des données relatives à l'état de santé des malades, qui s'interrogent sur ses garanties en matière de confidentialité. Les risques éventuels de divulgation des données sont en effet de nature à rendre l'informatisation

inacceptable dans la mesure où tout ce qui relève de la personne humaine, et particulièrement de son état physique et mental, est un domaine réservé qui ne concerne que le patient et son médecin. Aussi pourrait-il être contesté s'il ne garantit pas à chacun une *confidentialité* totale.

Le second critère à prendre en compte est relatif à l'*emploi*. L'automatisation des activités manuelles se traduit souvent par une diminution des emplois nécessaires à la réalisation des tâches, qui sont désormais prises en charge par des machines. L'augmentation du taux de chômage dans certains pays développés ou en développement n'est certes pas de nature à faciliter l'acceptation, par les travailleurs en particulier et par la société en général, d'un tel effet pervers : la rentabilité économique des systèmes automatisés est rarement compatible avec les besoins en matière d'emploi. L'introduction de l'informatique pose de plus le problème de la reconversion de travailleurs qui ne peuvent être licenciés pour des raisons de droit. L'automatisation partielle peut dans ces circonstances et selon les cas apparaître comme un élément de réponse : à côté des activités automatisées sont maintenues des tâches réalisées grâce au savoir-faire manuel des travailleurs.

L'exemple suivant est issu d'une enquête menée par une équipe de chercheurs en socio-technique (dont l'un, Catherine JACQUES, a participé à la partie concernant l'acceptabilité sociale de nos travaux) et réalisée à la demande des caisses mutuelles de l'entreprise EDF (Electricité De France). Cette étude avait pour but d'étudier les conséquences en termes d'emploi de l'informatisation des feuilles de soins et de définir de nouvelles activités pour les salariés libérés par l'introduction de l'informatisation. Les besoins en personnel du traitement informatique des prestations médicales étant moindres que lors du traitement manuel des feuilles de soins, il a été proposé que certains conservent leur travail de traitement (après avoir suivi une formation pour utiliser le nouveau logiciel) et que d'autres participent à la redéfinition ou au développement des autres activités proposées par les caisses mutuelles, et principalement les activités sociales et culturelles [JACQUES et RABIER, 1997]. Cet exemple montre que l'introduction des systèmes automatisés n'est pas neutre et qu'elle pose un réel problème d'emploi, notamment en ce qui concerne la reconversion, la formation, l'organisation et le licenciement éventuel des salariés, ce qui invite, avant l'automatisation elle-même, à appréhender le contexte dans lequel se fait cette introduction. La conception et l'évaluation des interfaces Homme(s)-Machine(s) doivent donc intégrer ces paramètres et proposer, par exemple, une limitation des tâches à automatiser pour assurer une évolution progressive vers l'automatisation, c'est-à-dire une évolution qui ne crée pas d'emblée des modifications structurelles majeures et qui permet ainsi son acceptation par les travailleurs.

Cette réflexion peut être étendue au troisième critère d'évaluation retenu, le *savoir-faire*. L'introduction des machines-outils à commande numérique dans les entreprises est parfois accusée de réduire le travailleur à un rôle de presse-bouton : face à son écran de travail, l'ouvrier n'a plus besoin de mobiliser son savoir-faire, devant simplement préparer et contrôler des tâches qu'il devait auparavant réaliser de ses propres mains. Aussi l'introduction de machines nouvelles est-elle souvent accusée de provoquer la suppression des savoir-faire au niveau de l'usage, c'est-à-dire, pour les travailleurs concernés, d'être socialement inacceptable.

Le manque de maîtrise des automatismes peut de plus être de nature à provoquer la ruine des Etats en développement : l'automatisation pousse en effet à abandonner les anciens savoir-faire et à ne pas prévoir un éventuel retour aux activités manuelles en cas d'échec. Aussi semble-t-il nécessaire de conserver certaines tâches manuelles qui permettront à l'entreprise de se relever si l'automatisation ne peut être intégrée par les travailleurs : l'automatisation gagne plus, dans ce cas, à être partielle plutôt que totale.

Le quatrième et dernier critère retenu est celui de *sécurité*, associé à l'*innocuité*, c'est-à-dire au caractère non nuisible de l'interface : pour être socialement acceptable, la dernière « condition » à remplir pour un système est d'assurer la sécurité de ceux qui l'utilisent d'une part et de ceux qui en bénéficient d'autre part. La catastrophe de la centrale nucléaire de Tchernobyl en 1986 est un exemple bien connu, qui montre les risques qu'il peut y avoir, pour des centaines de milliers d'individus, de ne pas s'assurer de la capacité d'un système à respecter les règles de sécurité élémentaires. La centrale comportait au moment de sa conception (incluant les procédures de travail) les limites de sa propre survie dans la mesure où elle n'avait aucun moyen de pallier et de limiter les dégâts de certaines défaillances humaines [REASON, 1990], dont la conception et l'évaluation n'ont pas tenu compte. On peut de ce fait dire que la centrale, par certains de ses aspects, était inacceptable socialement, pour la seule raison qu'elle ne garantissait pas une sécurité optimale à son environnement humain. L'évaluateur de l'IHM doit donc s'assurer, notamment lorsque le domaine d'application est dangereux ou délicat, que l'interface réalisée est en mesure de tolérer (ou d'éviter) les erreurs humaines.

Les critères attachés à cette dimension sont les suivants :

Critères de la dimension sociale :

- **Confidentialité, légalité**
- **Savoir-faire**
- **Sécurité-innocuité**
- **etc.**

III.4.3. La dimension culturelle et ses critères

La troisième et dernière dimension d'évaluation des IHMs qui caractérise l'acceptabilité sociale est la dimension *culturelle*, qui renvoie plus particulièrement aux caractéristiques de l'imagerie, c'est-à-dire aux formes, aux couleurs, à la gestuelle, aux symboles, aux costumes, aux mœurs et aux croyances qui transparaissent au niveau de l'IHM. Chaque pays ou groupe de pays possède, on l'a dit, une culture propre, des valeurs particulières et des traditions précises, qui peuvent, selon les cas, revêtir une importance considérable lors de la mise en place d'une nouvelle interface. Cette dernière, en mettant en scène la vie quotidienne, doit être compatible avec la culture qui caractérise le pays, ce qui signifie qu'elle doit respecter les valeurs, traditions et coutumes locales. Les critères rattachés à cette dimension sont les suivants :

Critères de la dimension culturelle :

- **Langue utilisée**
- **Forme**
- **Couleur**

- Gestuelle
- Costume
- Mœurs

III.5. Positionnement des critères d'évaluation par rapport au cycle de développement

Après avoir identifié les critères d'évaluation, il est possible de les positionner selon les étapes d'un cycle de développement en V en étendant la proposition de GRISLIN [1995] aux critères d'utilisabilité (voir figure I.8).

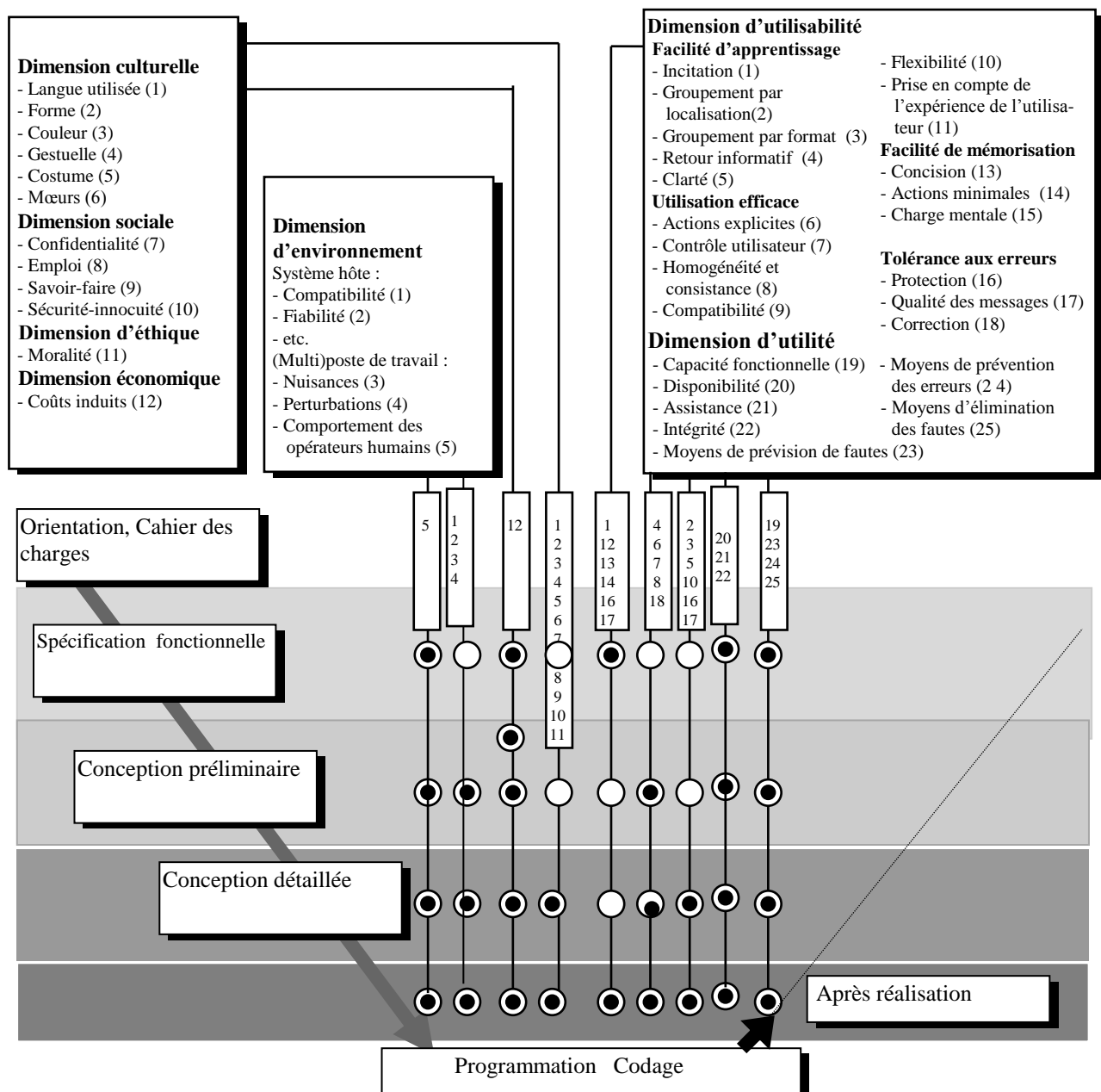


Figure I.8 : Première proposition à valider des phases et critères d'évaluation associés dans un cycle de développement en V

Le but de ce positionnement est de montrer dans quelles étapes du cycle de développement un critère d'évaluation peut être pris en compte.

Après avoir mis en évidence les critères d'évaluation, il s'agit de dresser un inventaire des **contraintes** qui influencent le **choix** des **méthodes** et **critères** d'évaluation (voir tableau I.8).

IV. Contraintes liées aux choix des méthodes et techniques d'évaluation dans un contexte d'évaluation globale

Selon certains auteurs à l'image de STANTON et BABER [1995], BALBO et COUTAZ [1992], le choix des méthodes est influencé par les paramètres suivants : les moyens humains, les ressources matérielles, les connaissances requises, les facteurs situationnels et la nature des résultats fournis. Ces informations sont nécessaires pour mettre au point un système d'aide à la décision pour le choix des méthodes et techniques d'évaluation appropriées pour chaque type d'application.

Les moyens humains désignent l'ensemble des acteurs impliqués directement ou indirectement dans l'activité d'évaluation. Nous distinguons trois classes d'intervenants : les sujets, les évaluateurs spécialistes (ergonomes, socio-techniciens, informaticiens, etc.) et une catégorie qui regroupe tous les autres acteurs (les développeurs, les concepteurs d'interface, les clients, les agents commerciaux, etc.).¹

Les facteurs situationnels définissent le contexte de l'évaluation. Comme le montre le tableau I.8, nous proposons de distinguer, dans l'étape considérée dans le processus de développement, le lieu dans lequel l'évaluation peut être pratiquée, la typologie des applications et des techniques d'interaction, l'enveloppe budgétaire et les contraintes de planning. Les ressources matérielles recouvrent tous les moyens physiques impliqués dans l'activité d'évaluation. Nous convenons de distinguer les ressources utilisées ou disponibles comme instruments de capture et les ressources matérielles qui font l'objet de l'évaluation.

Les connaissances concernent les descriptions nécessaires en entrée au bon fonctionnement de la méthode : modèle de l'utilisateur, modèle de la tâche, spécifications externes de l'interface, définition de scénarios, etc. Plus le spectre des informations d'entrée est large, plus la méthode est coûteuse à mettre en œuvre. En revanche, on peut espérer obtenir en sortie des résultats robustes et précis. Les connaissances requises peuvent provenir de disciplines variées, telles que l'ergonomie, la psychologie cognitive, la sociologie, le droit appliqué à l'informatique et les connaissances en assurance qualité.

L'objet de l'évaluation peut être caractérisé par les produits du processus de développement. Ce peut être le modèle de tâche. L'objet de l'évaluation peut être décrit de manière formelle, semi-formelle ou informelle et peut admettre plusieurs formes de support (papier, maquette, carton, ordinateur). A chacune de ces formes de présentation correspond un coût de développement et des compétences.

¹ Il serait éventuellement possible de distinguer les évaluateurs à compétence simple (spécialistes en ergonomie en général) des évaluateurs à compétence double (aussi bien spécialistes de l'ergonomie que d'un autre domaine).

De même, les ressources déjà étudiées peuvent être caractérisées par le *niveau d'abstraction* des données qu'elles permettent de recueillir. Pour illustrer l'intérêt de cette dimension, nous allons prendre comme exemple le cas de la capture automatique par ordinateur. Dans l'état actuel de cette technique, seules sont captées les actions physiques de type clavier-souris et, au mieux, la manipulation d'objets liés à l'accomplissement de tâche élémentaire.

La localisation est également un facteur déterminant qui influence le choix des méthodes et techniques d'évaluation. En effet, selon que l'évaluation est effectuée en laboratoire, sur site, ou selon que l'IHM à évaluer est destinée à un pays donné, les méthodes et critères diffèrent pour la même IHM.

Contraintes	Valeurs possibles
Description en entrée [BALBO ET COUTAZ, 1992]	<ul style="list-style-type: none"> Description du modèle de l'utilisateur et/ou de la tâche Description des spécifications externes/des scénarios Etc.
Spécialités requises [BALBO, 1994] [NENDJO ELLA et JACQUES, 1999]	<ul style="list-style-type: none"> Psychologie cognitive/Ergonomie / Sociologie Droit et informatique/Assurance qualité Etc.
Objet de l'évaluation [STANTON et BABER, 1995]	<ul style="list-style-type: none"> Modèle de l'évaluation/Modèle de la tâche Spécifications fonctionnelles Prototype/Maquette/Produit final Etc.
Support pour la capture des données [BALBO, 1994]	<ul style="list-style-type: none"> Laboratoire d'utilisabilité Ordinateur/Papier et crayon Etc.
Niveau d'abstraction [BALBO et COUTAZ, 1992]	<ul style="list-style-type: none"> Actions physiques/Tâches Etc.
Étapes du cycle de développement [STANTON et BABER, 1995]	<ul style="list-style-type: none"> Analyse des besoins, spécification Conception préliminaire/détaillée/Réalisation Etc.
Types d'application [BALBO, 1994]	<ul style="list-style-type: none"> CAO/Contrôle de procédés/Jeux vidéo/Billetterie Etc.
Techniques d'interaction [BALBO, 1994]	<ul style="list-style-type: none"> Manipulation directe Mono/Multi-utilisateur(s) Multimedia/multimodale Etc.
Budget de l'évaluation [BALBO, 1994]	<ul style="list-style-type: none"> Suffisance Possibilité d'extension
Origine des experts [BALBO, 1994]	<ul style="list-style-type: none"> Locale Externe
Localisation [BALBO, 1994] [NENDJO ELLA <i>et al.</i> , 1999b]	<ul style="list-style-type: none"> Laboratoire/Site Pays
Présence des utilisateurs [STANTON et BABER, 1995]	<ul style="list-style-type: none"> Utilisateurs finaux Utilisateurs non finaux mais représentatifs Utilisateurs non finaux et représentatifs
Nature des données [NENDJO ELLA <i>et al.</i> , 1999b]	<ul style="list-style-type: none"> Quantitative Qualitative
Temps alloué [STANTON et STANTON, 1995]	<ul style="list-style-type: none"> Heures/jours/semaines/mois
Environnement où évoluera l'IHM à évaluer [NENDJO ELLA <i>et al.</i> , 1999a]	<ul style="list-style-type: none"> Perturbation/Nuisances sonores Vibrations
Expérience des utilisateurs [NENDJO ELLA <i>et al.</i> , 1999a]	<ul style="list-style-type: none"> Novice Première expérience Experte

Tableau I.8 : Contraintes influençant le choix des méthodes et techniques d'évaluation

Conclusion

En définitive, pour parvenir à des systèmes interactifs dits acceptables, il faut les considérer non seulement sous les angles de fonctionnalité ou d'ergonomie, mais aussi selon les aspects qui influencent directement ou indirectement leur performance. Après avoir dressé un inventaire des contraintes qui influencent le choix des méthodes et critères adéquats pour l'évaluation des systèmes interactifs, nous constatons finalement que les aspects techniques ne suffisent plus, surtout dans un contexte international. L'environnement des IHMs est en effet constitué de facteurs, tels que les nuisances, de caractéristiques qui lui sont propres (et qui sont liées au cahier des charges) et de facteurs socio-culturels qui sont complexes et qui peuvent influencer négativement la performance du système Homme(s) - Machine(s).

Après avoir décrit la problématique et la diversité des critères en évaluation, il va désormais s'agir de faire une synthèse des différentes méthodes d'évaluation existantes, en les classant selon différentes approches. Cette synthèse fait l'objet du chapitre suivant.

Chapitre II : Des méthodes et techniques aux outils logiciels d'aide à l'évaluation des systèmes interactifs

Introduction

- I. Approches centrées sur les utilisateurs**
- II. Approches centrées sur une expertise**
- III. Méthodes analytiques**
- IV. Synthèse sur les méthodes et techniques usuelles**
- V. Vers de nouveaux outils d'aide à l'évaluation**

Conclusion

Introduction

L'évaluation des systèmes interactifs a fait l'objet du chapitre précédent. A cet effet, nous avons constaté que les critères d'évaluation pouvaient s'avérer extrêmement variés. Le recueil de ces données de base entraîne l'usage de méthodes et techniques d'évaluation. Mais le choix des méthodes appropriées pour une IHM donnée pose problème. En effet, différents paramètres identifiés par BALBO [1994] influencent le choix de ces méthodes. Dans cette optique, nous allons d'abord faire une classification et une synthèse des méthodes d'évaluation usuelles. C'est l'objectif de ce chapitre 2.

Plusieurs classifications existent dans la littérature. Comme l'expliquent GRISLIN et KOLSKI [1996], on trouve souvent la distinction entre des approches dites prédictives et des approches dites expérimentales : c'est le cas par exemple de la classification de NIELSEN et MOLICH [1990]. Proche de ces auteurs, SENACH [1990], qui a constitué pour nous une source importante d'inspiration, différencie les techniques en des approches analytiques et d'autres empiriques. MAGUIRE et SWEENEY [1989] et SWEENEY *et al.* [1993] proposent quant à eux trois grandes familles de méthodes où l'évaluation est basée sur l'utilisateur, la théorie ou l'expertise. WHITEFIELD *et al.* [1991] organisent leur classification en fonction de la présence ou non de l'utilisateur d'une part, et du système interactif à évaluer d'autre part. BALBO et COUTAZ [1992] classifient enfin les techniques en fonction de plusieurs paramètres, tels leur degré d'automatisation, la prise en compte de l'utilisateur, ou encore le savoir utilisé pour l'évaluation. Dans un souci de clarté, il nous a été nécessaire d'effectuer des regroupements de méthodes et techniques. Nous en avons défini trois que l'on retrouve d'ailleurs explicitement ou implicitement dans la plupart des classifications existantes, particulièrement celle de SWEENEY *et al.* (tout en rajoutant des sous-catégories, par exemple l'estimation de la charge de travail) : *centrées sur l'utilisateur, centrées sur une expertise, centrées sur des méthodes analytiques*. Bien entendu, ce recensement ne se veut pas exhaustif mais plutôt représentatif des méthodes et techniques existantes.

I. Approches centrées sur les utilisateurs

Ce type d'approche consiste à recueillir des données représentatives de l'interaction à l'aide de mesures ou d'observations provenant de l'utilisation des interfaces par des utilisateurs représentatifs de la population finale, puis à analyser les traces de leurs activités. L'évaluation doit être effectuée dans un environnement le plus proche possible de la réalité et (si possible) sans gêner l'utilisateur. On distingue les techniques associées au *diagnostic d'usage*, à *l'estimation de la charge de travail* et aux *tests de conception*, successivement présentées dans les paragraphes suivants.

I.1. Approches empiriques de diagnostic d'usage

Interviews :

Principe :

Cette technique consiste à s'entretenir avec des sujets représentatifs afin d'obtenir une

première vue d'ensemble sur les problèmes d'utilité et d'utilisabilité. En fait, les interviews viennent généralement en complément d'autres méthodes et techniques, tout en s'avérant souvent très utiles lors d'auto-confrontations suite à l'analyse des premières données recueillies.

Avantages et inconvénients principaux :

L'un des inconvénients réside dans le fait que, à part pour des interactions Homme(s)-Machine(s) simples, cette méthode ne peut être utilisée seule. De plus, sachant que ces interviews peuvent être effectuées sur le site de travail de l'utilisateur et/ou hors site, les risques sont grands de manquer de structuration et de rigueur lors des interviews. On peut en effet oublier des fonctionnalités et ne pas considérer tous les critères d'évaluation. La figure II.1 montre le principe de l'interview.

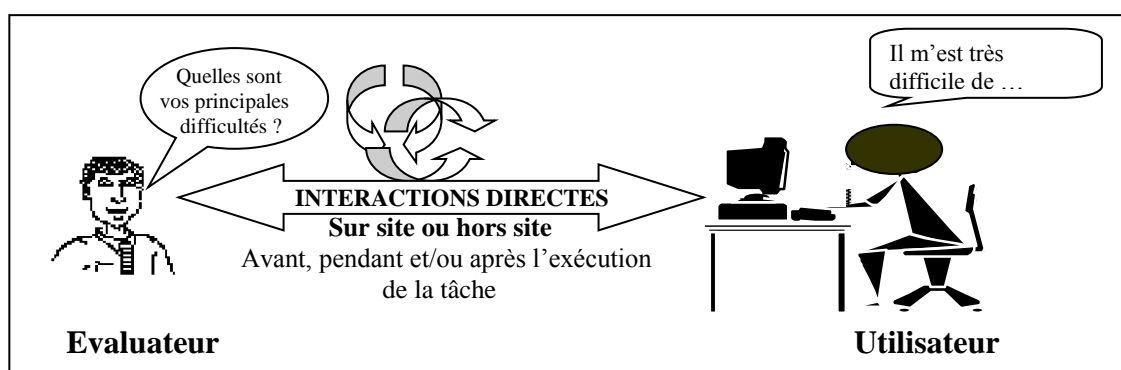


Figure II.1 : Interaction directe par interview entre utilisateur(s) et évaluateur(s)

Référence :

[OLSON et RUETER, 1987].

Questionnaires :

Principe :

Cette technique permet de recueillir par écrit un ensemble d'appréciations le plus souvent subjectives, qui peuvent être représentatives des problèmes rencontrés par les utilisateurs, sous une forme sûre et structurée, donc propice à l'analyse. Cette technique vient en complément d'autres techniques tels que les entretiens, les observations, etc. On trouve différents types de questions : les *questions fermées* (de type QCM), les *questions ouvertes* (qui permettent à un utilisateur de s'exprimer librement en utilisant son propre langage), les *questions semi-ouvertes* (ou semi-fermées) consistant à proposer à l'utilisateur un éventail de réponses et à lui donner la possibilité de s'exprimer comme il le désire, etc.

Avantages et inconvénients principaux :

Cette technique présente l'avantage d'être économique et utilisable dans la plupart des situations [RUSHINEK et RUSHINEK, 1986] et de permettre à l'utilisateur de travailler "à

tête reposée", même si elle est susceptible, d'après ROOT et DRAPER [1983], d'introduire un biais dans les réponses puisque l'utilisateur n'est pas toujours en mesure de se souvenir parfaitement des difficultés rencontrées lors des différentes situations de travail.

Référence :

[SINCLAIR, 1975].

Quick and Dirty :

Principe :

Cette méthode est basée sur un questionnaire muni d'échelles que les évaluateurs soumettent aux sujets. Le questionnaire est renforcé par une interview et les sujets sont ici essentiellement constitués par le personnel marketing de l'entreprise et non par des sujets représentatifs. Habituellement, les sujets d'évaluation sont des utilisateurs finaux, représentatifs ou non ; mais ici, la démarche consiste à choisir comme sujets ceux qui sont en contact avec les clients de la société (c'est-à-dire le personnel marketing dans le cas de la société PHILIPS, comme l'explique THOMAS [1996]). Durant l'évaluation, les questions auront pour but de détecter et de collecter des informations relatives aux mécontentements des clients (qui sont les utilisateurs finaux) et observations sur des systèmes antérieurs et similaires. Pour mieux comprendre, prenons l'exemple d'un produit comme Word 5. Comme l'explique THOMAS, lors de son utilisation par des utilisateurs finaux, des clients (utilisateurs finaux) se plaignent de certains dysfonctionnements et le signalent à l'entreprise mère ; ces informations sont alors collectées et analysées par le service marketing après que le service technique ait réglé le problème. Ces informations enregistrées par le personnel du service marketing serviront ainsi pour l'évaluation du produit similaire qu'est Word 6 et ainsi de suite.

La figure II.2 ci-dessous montre la salle d'évaluation lors de l'utilisation de la méthode « Quick and Dirty » chez « PHILIPS ». On peut y voir l'emplacement des 7 évaluateurs, le 8^{ème} fauteuil étant réservé au sujet (tous les sujets d'évaluation ne passent pas en même temps, mais un à la fois).

Cette démarche d'évaluation est enrichissante car les sujets qui font partie du service marketing ont une bonne connaissance des problèmes types que rencontrent les utilisateurs finaux (clients de l'entreprise). Cette démarche permet d'éviter la recherche des utilisateurs finaux. Les sujets font partie du personnel de l'entreprise et sont sélectionnés selon des critères tels que : la connaissance du système à évaluer, l'âge et l'intérêt éprouvé pour le produit.

Avantages et inconvénients principaux :

La technique d'évaluation « Quick and Dirty » présente l'avantage d'être rapide et peu coûteuse pour les entreprises [THOMAS, 1996]. Selon THOMAS, les méthodes formelles sont coûteuses pour les entreprises qui ont généralement un budget limité et allouent un temps bien défini pour l'évaluation afin que les coûts de revient ne remettent pas en cause la compétitivité des IHMs. Selon lui, le temps alloué pour cette technique chez

« PHILIPS » avoisine 6 jours (recueil des données et analyse) avec une équipe de 7 évaluateurs. THOMAS évoque le compromis entre l'économie sur la consultation d'un expert et les dépenses dues à l'usage d'une méthode formelle d'utilisabilité. L'autre inconvénient est que cette méthode privilégie plus les coûts de revient des produits de l'entreprise ; en d'autres termes, plus l'évaluation est longue et complexe, plus le coût de revient du logiciel à évaluer sera élevé, ce qui peut faire perdre de l'« efficacité » à une analyse approfondie des données (si le critère économique est considéré comme le plus prioritaire).

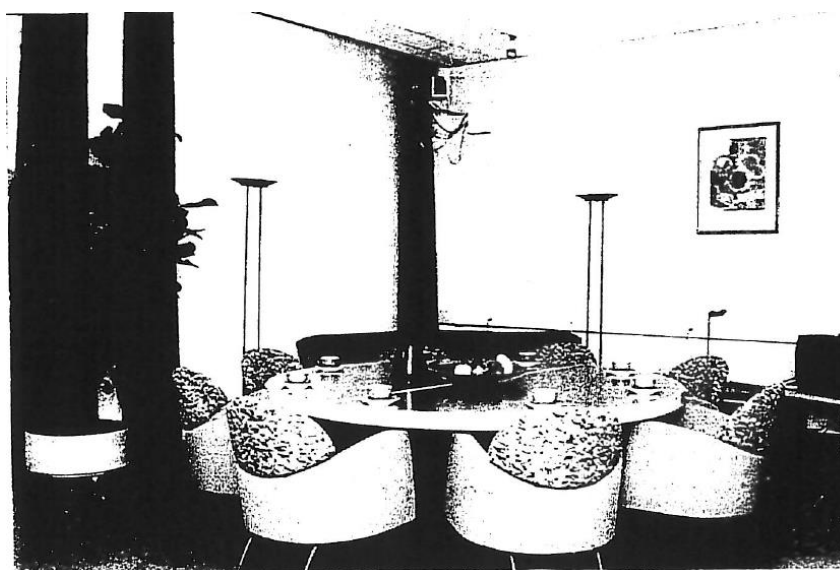


Figure II.2 : Salle d'évaluation à l'aide de la méthode Quick and Dirty chez PHILIPS [THOMAS, 1996]

Référence :

[THOMAS, 1996].

Private camera conversation :

Principe :

Cette technique consiste à placer dans une cabine un sujet représentatif. Ce dernier est seul devant une caméra vidéo et un micro, et doit s'exprimer sur tous les aspects du système (à évaluer) qu'il a utilisé. Il parle librement et exprime des données objectives et subjectives. Il n'y a pas d'interview et les sujets décident de commencer ou d'arrêter lorsqu'ils le souhaitent. Les sujets sont de tout âge et les données sont exploitées ultérieurement. La figure II.3 illustre ce principe.

Avantages et inconvénients principaux :

L'un des avantages de la méthode est que les sujets ne subissent aucune influence de la part des évaluateurs. Ils ne sont pas censurés et sont libres de parler de tous les problèmes rencontrés lors de l'utilisation du système à évaluer. Cette technique permet donc d'avoir un bon retour de la part des utilisateurs représentatifs. Le temps d'une session est très court, mais l'inconvénient est qu'il faut beaucoup de sujets pour avoir un retour significatif, exploitable et intéressant et non centré sur le point de vue d'un sujet unique.

Référence :

[DE VRIES *et al.*, 1996].

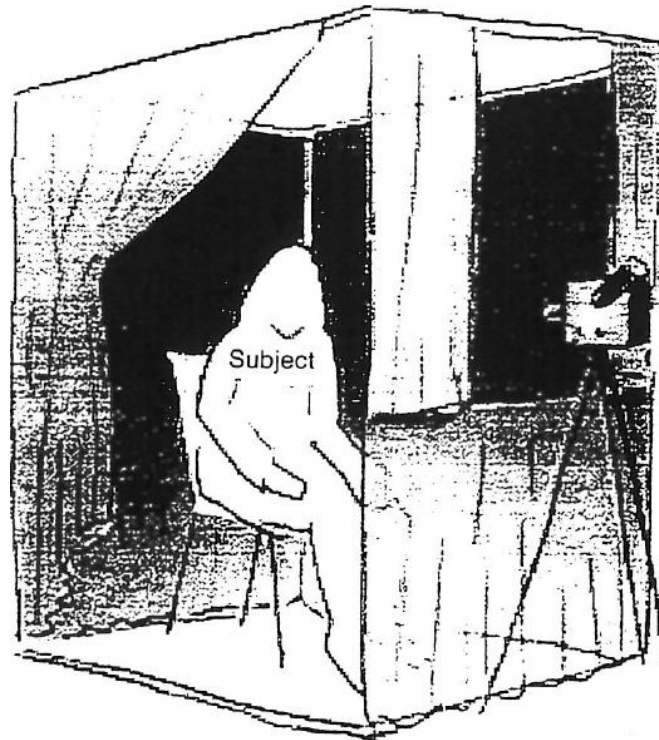


Figure II.3 : Illustration de la cabine vidéo [DE VRIES *et al.*, 1996]

Arbre des causes :

Principe :

L'arbre des causes est une technique basée sur la description graphique d'un enchaînement causal sous la forme d'une arborescence. Elle est applicable pour la majorité des systèmes interactifs et permet de mettre en évidence les causes humaines et les procédures conduisant à une défaillance. Cette technique permet de comparer des situations incidentelles ou accidentelles à celles qui auraient pu avoir lieu en l'absence de problèmes d'utilisabilité.

Avantages et inconvénients principaux :

Cette technique permet de décrire facilement et graphiquement les enchaînements des causes conduisant à un problème d'utilisabilité. L'inconvénient est qu'elle s'avère inefficace par rapport aux systèmes complexes. L'évolution temporelle des événements du système n'est cependant pas bien prise en compte et l'identification des séquences de défaillances s'avère difficile pour les systèmes complexes [VILLEMEUR, 1988].

Référence :

[VILLEMEUR, 1988].

Incidents Critiques :

Principe :

Le principe de cette technique est de faire des observations en situation et de s'entretenir avec les utilisateurs. Au cours de ces observations et entretiens, on recueille systématiquement des dysfonctionnements de l'IHM. Ce recueil est suivi d'un diagnostic des principaux dysfonctionnements du système (difficultés d'utilisation, détection des inadéquations des outils ou de la logique de travail, dysfonctionnements, etc.). Les recueils font l'objet d'une classification hiérarchique selon les types d'incidents, d'une identification des classes anormales et de propositions d'amélioration. Les informations issues du diagnostic peuvent être utilisées pour définir les premières fonctionnalités d'un système, réaliser des aménagements et préciser les objectifs de formation. La méthode des incidents critiques a été utilisée pour la première fois par FLANAGAN [1954] pour étudier des accidents d'avions évités de justesse.

Avantages et inconvénients principaux :

Cette technique est caractérisée par sa rapidité et permet d'obtenir un diagnostic global des principales classes de dysfonctionnements du système. La difficulté demeure au niveau de son utilisation, qui peut s'avérer lourde à mettre en œuvre dans le cas de systèmes complexes.

Référence :

[SENACH et ALENGRY, 1985].

Traces écrites du travail :

Principe :

Cette technique est basée sur l'analyse des traces écrites du travail des utilisateurs. Ces traces peuvent être des notes prises pendant le travail, des documents de conduite, des cahiers de doléances, etc. Grâce à cette technique, on peut mettre indirectement en évidence les lacunes de l'IHM ou les problèmes affectant l'exécution d'une tâche (voir figure II.4).

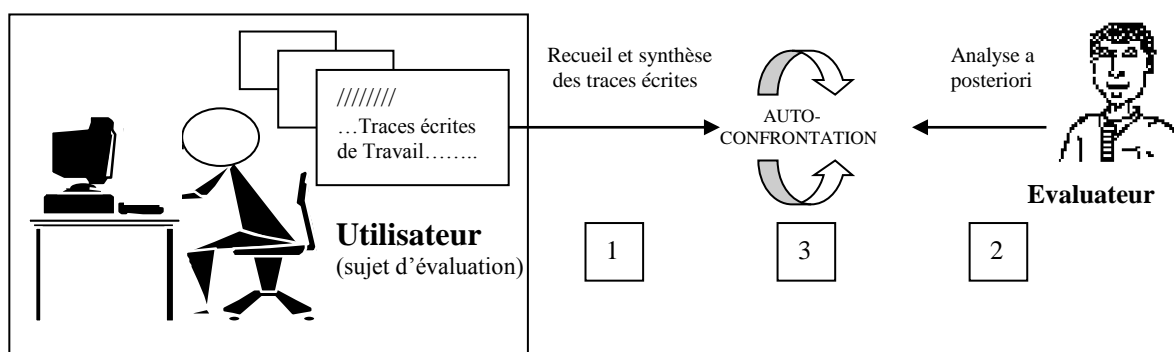


Figure II.4 : Analyse des traces écrites en tant que support à l'évaluation

Avantages et inconvénients principaux :

Ce qui caractérise et distingue cette technique des autres est qu'elle est facile à mettre en œuvre mais qu'elle reste toutefois succincte. Dans le cas des cahiers de doléances, qui peuvent s'avérer d'une contribution considérable pour l'amélioration d'une interface, on constate que les utilisateurs la délaissent très vite lorsque leurs remarques ne sont pas prises en compte.

Référence :

[SPERANDIO, 1991].

SUS (System Usability Scale) :*Principe :*

Cette technique permet d'évaluer un système selon la dimension d'utilisabilité. Elle est basée sur 10 questions sur une échelle à cinq valeurs (voir figure II.5). SUS permet d'obtenir une évaluation globale et subjective de l'utilisabilité.

pour	Fortement contre	Fortement
1. Je pense que j'aimerais utiliser ce système fréquemment	<input type="checkbox"/>	<input type="checkbox"/>
2. Je ne trouve pas le système particulièrement complexe	<input type="checkbox"/>	<input type="checkbox"/>
3. J'ai trouvé que le système était facile à utiliser	<input type="checkbox"/>	<input type="checkbox"/>
4. Je pense que j'aurais besoin de l'aide d'un technicien pour l'utilisation	<input type="checkbox"/>	<input type="checkbox"/>
5. J'ai trouvé que les différentes fonctions du système étaient bien intégrées	<input type="checkbox"/>	<input type="checkbox"/>
6. J'ai trouvé qu'il y avait trop d'inconsistances dans le système	<input type="checkbox"/>	<input type="checkbox"/>
7. J'imagine que plusieurs personnes apprendront vite à utiliser le système	<input type="checkbox"/>	<input type="checkbox"/>
8. J'ai trouvé le système encombrant à utiliser	<input type="checkbox"/>	<input type="checkbox"/>
9. Je me sens en confiance lors de l'utilisation du système	<input type="checkbox"/>	<input type="checkbox"/>
10. J'ai besoin d'apprendre beaucoup de choses avant d'utiliser le système	<input type="checkbox"/>	<input type="checkbox"/>

Figure II.5 : La technique SUS

Avantages et inconvénients principaux :

Cette technique est peu coûteuse et BROOKE la recommande pour les entreprises, qui sont d'ailleurs, d'après lui, de plus en plus nombreuses à l'utiliser, à l'image de « Digital

Equipement Co.Ltd. ». Selon BROOKE [1996], l'inconvénient est qu'elle soumet les utilisateurs à des pressions psychologiques importantes car, dans les entreprises où cette technique est utilisée, les utilisateurs des systèmes à évaluer n'ont en moyenne que 20 minutes à 1 heure pour répondre aux questions. L'inconvénient de cette méthode est qu'elle permet de n'obtenir que des résultats globaux et de ne mettre en évidence que les problèmes les plus importants, sans entrer dans les détails.

Référence :
[BROOKE, 1996].

Recueil de l'expertise :

Principe :

Cette technique consiste à rendre compte de l'activité et des problèmes de l'utilisateur en langue naturelle (entretiens, verbalisations spontanées ou provoquées) ou à l'aide de graphiques. Le Recueil de l'expertise permet de mettre en évidence des procédures (comment procéder), des connaissances sous-jacentes aux procédures, de l'organisation et des stratégies de raisonnement, les problèmes rencontrés, etc. Rappelons que cette technique est utilisée en ingénierie de la connaissance. La figure II.6 ci-dessous illustre cette technique.



Figure II.6 : Technique de recueil de l'expertise

Avantages et inconvénients principaux :

L'avantage de cette technique réside dans sa simplicité de mise en œuvre. L'un des inconvénients est qu'elle pose le problème de biais de verbalisation [DIENG, 1990].

Référence :
[DIENG, 1990].

Subjectivité dans l'évaluation (EMPI) :

Principe :

EMPI est une méthode d'évaluation d'applications multimédias dont le principe consiste à prendre en compte l'opinion subjective des utilisateurs *non spécialistes* sur les

caractéristiques logicielles des IHMs. L'objectif de la méthode est de trouver un équilibre entre l'évaluation prenant en compte des données objectives et celle orientée vers les critères subjectifs. L'évaluateur est ici un enseignant non spécialiste, un manager ou toute personne voulant apprendre l'évaluation. La méthode est présentée sous forme d'un questionnaire structuré en plusieurs thèmes : (1) qualité technique, (2) aspects ergonomiques, (3) documents multimédias et (4) outils pédagogiques. Ces thèmes sont ensuite subdivisés en critères, sous-critères et questions.

Le questionnaire dispose de 5 types de notations (voir figure II.7) :

- Une note calculée, issue de l'évaluation des questions.
- Une note instinctive initiale (rappel) donnée par l'évaluateur.
- Un indice de corrélation qui mesure les similitudes entre la note instinctive et la note calculée.
- Un indice de cohérence qui mesure la cohérence de la note instinctive entre un niveau et ses sous-niveaux.
- Une note finale qui est une combinaison entre la note calculée, la note instinctive globale et l'indice de cohérence.

Avantages et inconvénients principaux :

L'un des avantages de cette technique est qu'elle est prévue pour les non spécialistes et un effort est d'ailleurs fait pour la rendre facile à utiliser. Par contre, l'un des inconvénients majeurs de la méthode est que la fiabilité des résultats dépend de la précision des questions : moins la question est vague, plus le résultat est fiable.

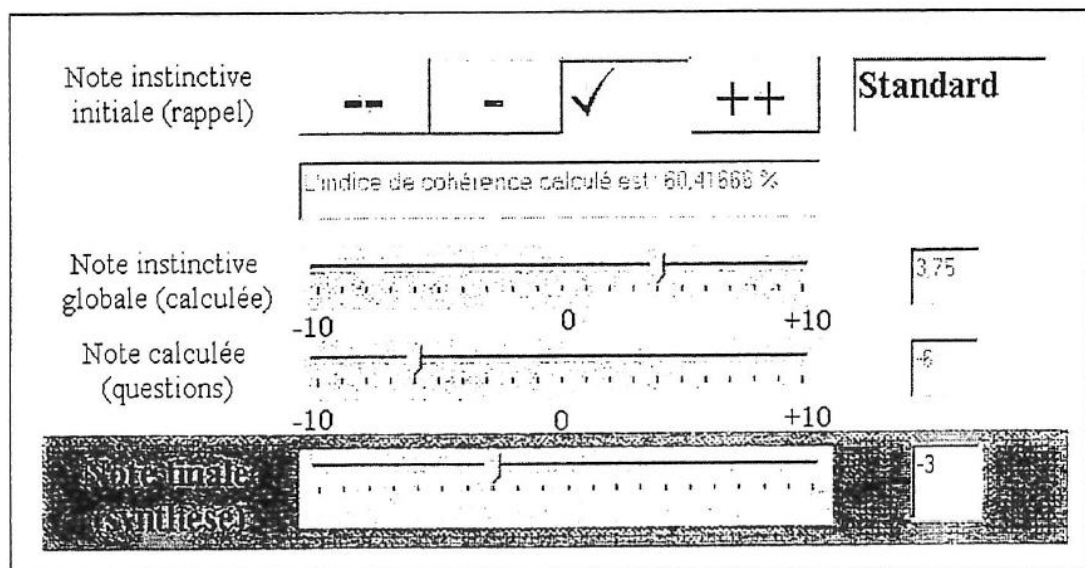


Figure II.7 : Exemple de note finale issue d'une évaluation [HÛ *et al.*, 1999]

Référence :

[HÛ *et al.*, 1999].

Mouchard électronique :

Principe :

Cette technique repose sur le recueil automatique, en situation naturelle, des actions des utilisateurs et de leurs répercussions sur le système. Elle permet de recueillir automatiquement des données objectives à l'aide de l'ordinateur, à une fin généralement d'analyse ultérieure. En plus de ce recueil automatique, les sujets sont soumis à une séance de questionnaires et de verbalisations. La figure II.8 montre le principe du mouchard électronique. Les systèmes décrits ci-après contribuent plus ou moins directement au principe du mouchard électronique.

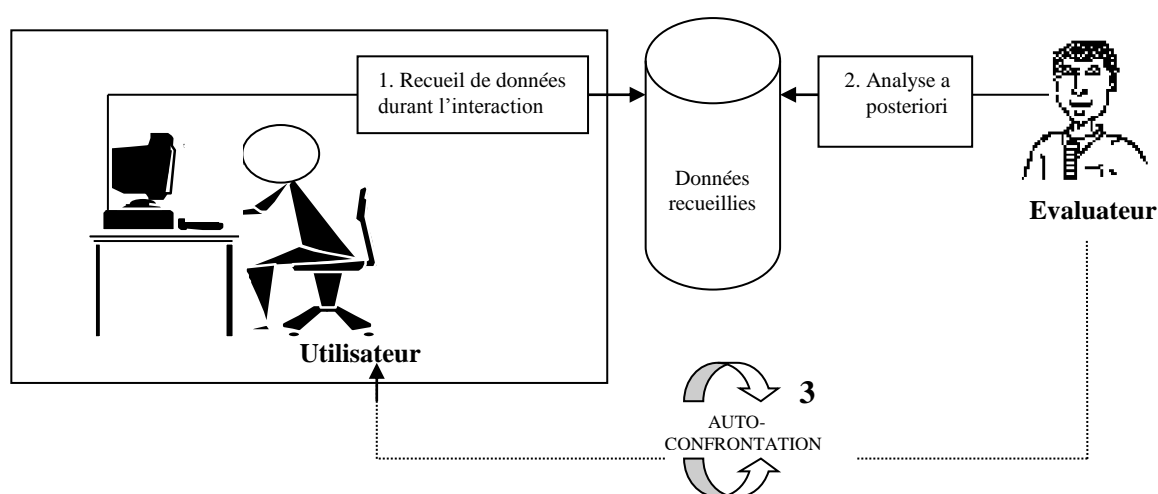


Figure II.8 : Principe du mouchard électronique

Avantages et inconvénients principaux :

L'un des avantages du mouchard électronique est que des données temporelles peuvent être recueillies avec la possibilité de rejouer des séquences de travail. Cette technique se distingue des autres méthodes et techniques d'évaluation par le fait qu'elle est objective et non perturbante pour l'opérateur. C'est une technique qui donne une appréciation fondée de la performance de l'interaction Homme(s)-Machine(s) [NEAL et SIMON, 1983].

Références :

[NEAL et SIMON, 1983] [HANSON *et al.*, 1984] [BIAS *et al.*, 1993].

Questionnaire socio-technique :

Principe :

Le questionnaire socio-technique permet d'évaluer les IHMs en terme d'acceptabilité socio-culturelle. Le principe consiste à répondre à un questionnaire sur les aspects socio-culturels. Deux stratégies sont possibles : soit l'évaluateur est un expert des aspects socio-culturels et doit vérifier la prise en compte des critères socio-culturels dans une IHM à partir du questionnaire, soit l'évaluateur n'est pas un spécialiste des aspects socio-

culturels : dans ce cas, il doit choisir un échantillon de sujets représentatifs du pays de destination de l'IHM à évaluer. A ces sujets, l'évaluateur posera des questions issues du questionnaire (voir figure II.9 et annexe 3). et pourra ainsi vérifier si les aspects socio-culturels sont pris en compte. Les réponses ont une notation allant de 1 (fortement négative) à 5 (fortement positive).

QUESTIONNAIRE SOCIO-TECHNIQUE POUR L'EVALUATION DES INTERFACES H-M	NOTE SUR 5				
	1	2	3	4	5
DIMENSION SOCIALE					
Confidentialité					
- L'interface respecte-t-elle l'anonymat de l'opérateur ?.....					
- L'interface permet-elle d'obtenir des informations relatives à la vie privée de l'opérateur ?.....	X				
Sécurité-innocuité					
- Les dysfonctionnements de l'interface peuvent-ils être nuisibles à l'environnement et donc à la société ?.....					
- L'interface permet-elle l'intervention de l'homme en cas de dysfonctionnement ?.....					
Emploi/Travail					
- La mise en place de l'interface risque-t-elle de provoquer des licenciements ?.....					
- La mise en place de l'interface peut-elle entraîner une diminution du salaire des opérateurs ?.....					
- La mise en place de l'interface va-t-elle modifier l'organisation du travail ?.....					
- La mise en place de l'interface va-t-elle opérer une redéfinition des postes des travailleurs ?.....					
- L'utilisation de l'interface est-elle frustrante pour les opérateurs ?.....					
- L'interface laisse-t-elle une autonomie aux opérateurs dans l'exécution de leurs tâches ?.....					
Savoir-faire					
- Les opérateurs ont-ils déjà été en présence de ce type d'interface ?.....					
- Les opérateurs disposent-ils des « outils » intellectuels nécessaires à la mise en œuvre de l'interface ?.....					
- Les opérateurs ont-ils déjà travaillé avec ce type d'interface ?.....					
- L'interface remet-elle en cause le savoir-faire des opérateurs ?.....					
Compétence					
- Les compétences nécessaires à la compréhension de l'interface sont-elles présentes chez les opérateurs ?.....					
- L'interface est-elle adaptée au niveau de formation des opérateurs ?.....					
- Les opérateurs disposent-ils des compétences nécessaires à l'utilisation de l'interface ?.....					
DIMENSION D'ETHIQUE					
Moralité					
- L'interface met-elle en scène des actions immorales ?.....					
- L'interface est-elle compatible avec les principes éthiques de la population locale ?.....					
Violence					
- Les actions proposées sur l'interface peuvent-elles être considérées comme violentes par la population locale ?.....					
- L'interface met-elle en scène des objets susceptibles d'être utilisés dans le cadre d'actes violents ?.....					
- L'interface présente-t-elle des scènes conflictuelles ou laisse-t-elle supposer l'existence d'une certaine tension ?.....					
DIMENSION CULTURELLE					
Langue					
- La langue d'usage de l'interface est-elle la même que celle des utilisateurs ou de la population locale?.....					
- Les opérateurs connaissent-ils la langue d'usage de l'interface?.....					
- Les opérateurs comprennent-ils la langue d'usage de l'interface?.....					
Forme					
- Les formes présentes sur l'interface peuvent-elles heurter la sensibilité de la population locale ?.....					
- Les symboles représentés sur l'interface ont-ils des significations particulières ?.....					
Si oui, cette signification est-elle la même que celle que la population locale leur attribue?.....					
- La proximité de certains symboles peut-elle être choquante pour la population locale ?.....					
Couleur					
- Les couleurs présentes sur l'interface ont-elles une signification particulière ?.....					
Si oui, cette signification est-elle la même que celle que la population locale leur attribue?.....					
- L'assemblage des couleurs est-il de nature à heurter la population locale ?.....					
- Certaines des couleurs présentes sur l'interface sont-elles en désaccord avec son contexte ?.....					
Gestuelle					

Figure II.9 : Extrait d'un Questionnaire socio-technique [NENDJO ELLA et JACQUES, 1999]

Avantages et inconvénients principaux :

L'inconvénient d'une telle méthode se situe au niveau de la deuxième stratégie d'évaluation, c'est-à-dire dans le cas où l'évaluateur n'est pas spécialiste des aspects socio-culturels. En effet, il faut dans ce cas un très grand nombre de sujets pour que le résultat soit fiable. Les sujets doivent de plus être très représentatifs du pays de destination du système à évaluer sinon les résultats seront biaisés et les conclusions ne refléteront pas la réalité socio-culturelle de l'IHM. L'avantage réside dans la simplicité d'utilisation de cette méthode qui peut être utilisée par les spécialistes de la socio-technique et des non spécialistes qui doivent par contre rechercher un nombre très important de sujets représentatifs, ce qui n'est pas chose facile.

Référence :

[NENDJO ELLA et JACQUES, 1999].

Analyse des protocoles :*Principe :*

Utilisée depuis longtemps en ergonomie, cette technique consiste à observer un utilisateur pendant la réalisation d'une tâche interactive. L'utilisateur doit penser à « haute voix », en précisant ses buts et sous-but, ses méthodes de travail, ses difficultés, ses actions, etc. L'utilisateur explique ce qu'il fait lors de l'interaction avec la machine : problèmes types, verbalisations provoquées, informations à la demande, étude de la mémorisation. La séance de travail est généralement enregistrée à l'aide de techniques audio-visuelles et/ou simplement par l'intermédiaire de prises de notes (voir figure II.10). Ces données sont ensuite analysées par l'évaluateur. Il est possible d'effectuer une partie de l'analyse du protocole avec l'utilisateur et de lui demander des compléments d'information.

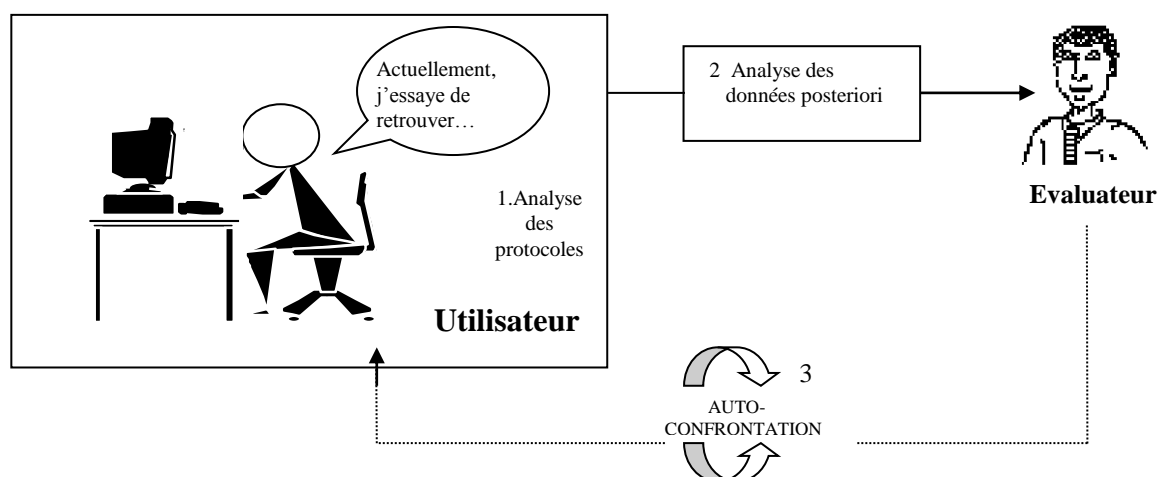


Figure II.10 : Analyse de protocoles en tant que méthode d'évaluation d'interface

Avantages et inconvénients principaux :

L'un des inconvénients est que le dépouillement est souvent très long lorsque les techniques audio et vidéo ont été utilisées, en particulier lorsque les tâches sont complexes. De

plus se pose le problème de la pertinence et/ou de l'exhaustivité des données recueillies. La variété des situations normales et anormales (dont certaines peuvent être rares) fait qu'une analyse exhaustive rigoureuse s'avère parfois impossible. Il faut toutefois noter l'intérêt de cette méthode pour des systèmes relevant des domaines où les opérateurs humains sont confrontés à des choix stratégiques ou des enchaînements de tâches [GRISLIN, 1995].

Références :

[ERICSSON et SIMON, 1984] [BELKIN *et al.*, 1987] [KOUBEK et SALVENDY, 1991].

L'analyse des mouvements oculaires :

Dans certains domaines (par exemple la conduite de procédés industriels), la perception visuelle des informations s'avère capitale. La direction absolue du regard représente un élément utilisable en parallèle avec d'autres moyens pour analyser l'activité de l'utilisateur face à un écran ou un ensemble de supports d'informations. La figure II.11 est un exemple de système oculométrique ASL qui est un système qui mesure la direction relative du regard, c'est-à-dire par rapport à la tête.

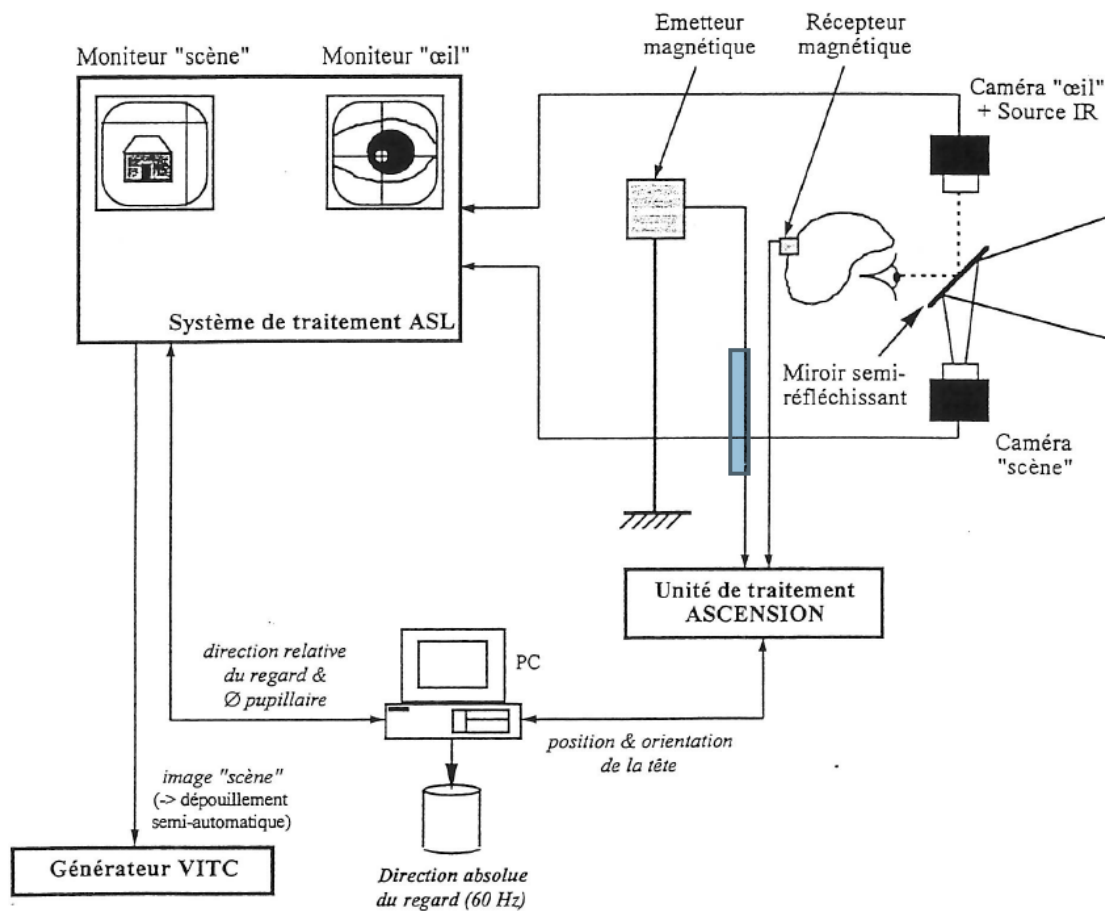


Figure II.11 : Principe d'évaluation par analyse des mouvements oculaires, cas ASL-ASCENSION [KADDOURI *et al.*, 1995]

Avantages et inconvénients principaux :

L'un des avantages réside dans le fait que les mouvements oculaires enregistrés lors de sessions de travail peuvent être rejoués en différé sur écran et commentés avec (et par) des utilisateurs. L'un des inconvénients est que cinq groupes de facteurs peuvent influencer sur la durée des fixations oculaires : la nature de l'information prélevée, la familiarité du sujet, le processus de traitement, la difficulté d'identification et le type de tâche [ACIER, 1990]. Ces durées auront alors une influence sur le temps que mettra un opérateur humain pour effectuer une tâche en interaction avec l'interface graphique. Un autre inconvénient réside dans le fait que des difficultés peuvent être rencontrées lors de l'exploitation des résultats dans la mesure où la fixation du regard ne correspond pas toujours à une prise d'information et qu'il n'est pas possible d'évaluer l'influence de la vision périphérique.

Références :

[MORAY, 1986] [ABED, 1990] [BERNARD, 1994] [RIERA *et al.*, 1994].

MRP (Maximal Repeating Pattern) :*Principe :*

Le principe de cette méthode, basée sur la capture et l'analyse automatique, consiste à détecter des répétitions de schémas d'actions [SIOCHI et HIX, 1991]. Ici, la capture nécessite d'instrumenter le programme source pour capter les actions de l'utilisateur sur les dispositifs d'entrée. Ces actions sont repérées à un bas niveau d'abstraction : par exemple la sélection d'un bouton ou la saisie d'un champ textuel. La répétition d'actions peut être considérée comme un révélateur de problème. L'usage fréquent d'une même suite d'actions peut par exemple indiquer l'inadéquation de la granularité des commandes et suggérer le besoin d'une macro-commande.

Comme l'indiquent SIOCHI et HIX [1991], une analyse MRP fournit des enseignements plus riches lorsqu'il s'agit de sujets experts : les intentions sont plus faciles à identifier que dans le cas d'utilisateurs novices dont le comportement au niveau « Keystroke » est exploratoire (cf. à ce sujet la méthode « Keystroke » en III.2.1.).

Avantages et inconvénients principaux :

L'inconvénient dans cette méthode est que lorsque l'utilisateur est novice, le résultat de l'évaluation est très moyen car les renseignements fournis par les utilisateurs novices sont moins riches. Par rapport à d'autres approches automatiques, la méthode MRP présente l'avantage de combiner à la fois la capture et l'analyse automatiques, ce qui lui confère un degré d'automatisation plus important que certaines méthodes à l'image d'ÉMA dont l'automatisation se limite au niveau de la capture, sans analyse automatique (cf. plus loin).

Référence :

[SIOCHI et HIX, 1991].

PLAYBACK

Principe :

Le principe de la méthode consiste à capturer les actions clavier lorsque l'utilisateur exécute sa tâche, et ceci au moyen d'un dispositif physique appelé « lab computer ». Ce dispositif est situé entre le clavier et son hôte. En mode enregistrement, le « lab computer » détecte, transmet, date et sauvegarde sur disque toutes les actions de l'utilisateur sur le clavier alors qu'en « replay », il les transmet au serveur pour reproduire la session de travail originale. Rappelons que PLAYBACK a été développé au début des années 80 dans le service « Human Factors » des laboratoires.

Avantages et inconvénients principaux :

Ce dispositif a l'avantage de supprimer l'intrusion de la caméra et fournit un fichier de données fiables. Ce fichier peut ensuite alimenter un analyseur statistique simple sur les séquences d'actions : par exemple le nombre de sollicitations de la touche « aide » ou « détruire ». Playback est un système pionnier mais conçu pour des applications limitées à des entrées clavier. L'inconvénient de ce système est qu'il éprouve des limites notoires, notamment dans le domaine des interfaces graphiques à manipulation directe qu'il est incapable de gérer.

Référence :

[NEAL et SIMON, 1983].

Système d'HAMMONTREE :

Principe :

Le principe de cette méthode consiste à enregistrer les actions clavier-souris et à compléter ces données numériques par un enregistrement audiovisuel. Développé sur Macintosh, ce système est constitué par : (1) un programme de capture d'événements estampillés d'une date. Il s'agit des événements fournis par la boîte à outils du Macintosh ; (2) un programme de filtrage d'événements qui permet de construire des fichiers de capture personnalisés (par exemple, un filtre qui extrait toutes les sélections de commande permet d'élaborer un fichier plus concis) ; (3) un analyseur multimédia de données qui effectue le lien entre événements et enregistrements audiovisuels par le biais de l'estampille. L'évaluateur localise et sélectionne un événement et demande à visionner la séquence qui s'y réfère ; (4) un enregistreur d'annotations verbales utilisable avec l'analyseur multimédia ; il numérise la parole et l'associe à des segments de la bande audiovisuelle et à l'événement correspondant.

Avantages et inconvénients principaux :

L'avantage de ce système est qu'il fournit une source de données, de traitement et d'analyse pour comparer et évaluer des applications ainsi que des prototypes. Ces sources de données permettent également d'évaluer la documentation des logiciels et le matériel d'instruction.

Référence :

[HAMMONTREE *et al.*, 1992].

ÉMA (Evaluation par Mécanisme Automatique) :

Principe :

ÉMA (Evaluation par Mécanisme Automatique) est un outil qui permet la capture de données comportementales issues de l'utilisation d'une application. Ce type d'évaluation nécessite donc la participation d'utilisateurs. ÉMA utilise des graphes de séquences, c'est-à-dire des graphes orientés, dont les nœuds sont des tâches élémentaires et les arcs des relations d'enchaînement entre tâches. Ces graphes traduisent la logique de fonctionnement du système et les tâches dont il est ici question sont en fait les tâches telles qu'elles sont implantées dans l'application. En d'autres termes, les graphes de séquences modélisent les tâches prévues pour une application donnée. C'est l'accomplissement des tâches utilisateurs par le biais des tâches prévues dans l'application, plus particulièrement les événements utilisateurs, que capture ÉMA. A partir des traces d'utilisation et des graphes de séquences, une détection automatique d'anomalies est effectuée.

Avantages et inconvénients principaux :

Comme l'explique BALBO [1994], l'inconvénient de cette méthode est que diverses anomalies peuvent apparaître lors de sa mise en œuvre. Les erreurs d'incitation peuvent en effet entraîner la rupture dans l'espace de tâche et l'utilisateur peut alors s'engager à tort dans une tâche qu'il ne désire pas effectuer parce que des informations le guident incorrectement ; de même lors des erreurs liées à la *Compatibilité* : l'utilisateur est correctement incité à entreprendre une tâche donnée pour se rendre compte, plus loin dans le dialogue, que la tâche du système ne correspond pas à celle qu'il désire effectuer. L'annulation immédiate traduit, selon BALBO [1994], un problème de navigation, lequel peut éventuellement être causé par des « erreurs » de conception relatives à la *Charge de travail*, à la *Cohérence*, à la *Signification des codes* et à la *Compatibilité*.

L'avantage qu'offre cette méthode réside dans le fait que contrairement aux autres méthodes automatiques (en fait celles qui seront décrites dans la partie III.3), ÉMA nécessite la présence des utilisateurs, ce qui permet de mieux apprécier certains problèmes réels d'utilisabilité que n'offrent pas d'autres outils automatiques.

Référence :

[BALBO, 1994].

I.2. Estimation de la charge de travail

L'estimation qualitative ou quantitative du niveau d'activité d'un utilisateur effectuant une tâche peut être pertinente pour évaluer la difficulté à utiliser l'IHM [WIERWILLE et CASALI, 1983]. Plusieurs spécialités contribuent à l'évaluation de la charge de travail : physiologie, psychologie, ergonomie, sciences pour l'ingénieur, etc. Les méthodes les plus connues sont les échelles subjectives, les mesures physiologiques et les mesures de la performance. On considère

généralement que la charge de travail est une grandeur indispensable pour l'évaluation des postes de travail.

I.2.1. Echelles subjectives

Les Echelles subjectives sont des techniques issues de la psychologie du travail. Elles sont basées sur une estimation subjective du travail cognitif afin d'évaluer le niveau d'activité mentale et physique par des questionnaires d'opinion. Elles trouvent leur fondement dans la théorie de l'information. En *psychologie du travail*, on utilise également des questionnaires spécifiques pour obtenir une auto-évaluation *a posteriori* par l'opérateur humain. Les plus connus sont les suivants : SWAT, TLX, l'Echelle de Cooper-Harper.

SWAT (Subjective Workload Assessment Technique) :

Principe :

Le principe de cette méthode consiste en une indication par l'opérateur humain de son niveau de difficulté sur trois dimensions d'exigence temporelle, d'effort mental et de stress. Les 3 dimensions sont représentées sur des échelles à 3 niveaux (faible, moyen, élevé), soit 27 états possibles correspondant à 27 cartes différentes (voir figure II.12). Toute exécution de tâche par l'opérateur humain est précédée d'un classement des cartes selon une progression qu'il considère comme allant du plus bas niveau de charge au plus haut niveau. A la fin du travail, l'opérateur doit évaluer celui-ci au moyen des 3 dimensions et des trois niveaux. L'évaluateur détermine ainsi son niveau de charge par rapport à la classification initiale.

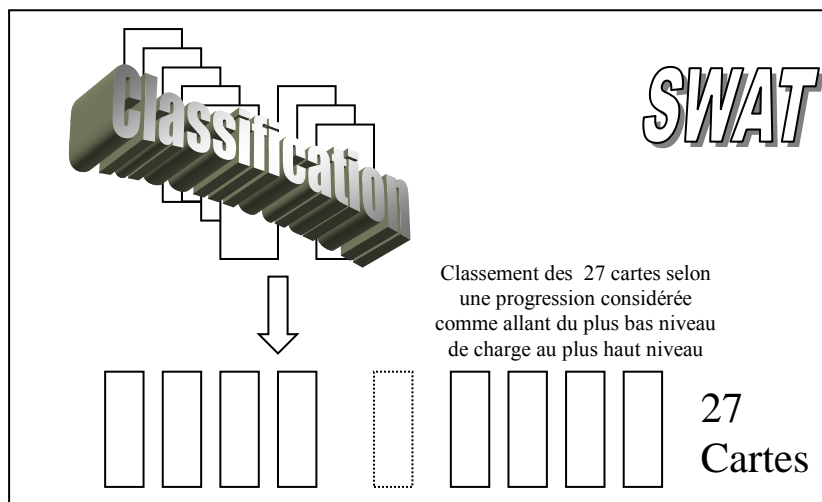


Figure II.12 : Principe de la méthode SWAT

Avantages et inconvénients principaux :

Cette méthode présente l'avantage de prendre en compte 3 dimensions (temporelle, stress, effort mental) pertinentes, suffisant pour estimer globalement le sentiment de charge d'un opérateur humain.

Référence :

[REID et NYGREN, 1988].

TLX (Task Load Index) :

Principe :

Le principe de cette méthode provenant de la NASA consiste à calculer la charge de travail par pondération sur six échelles correspondant à 6 différentiateurs sémantiques : les exigences mentales, physiques et temporelles, ainsi que la performance, l'effort et le stress (voir le principe à la figure II.13). Chaque opérateur doit ici remettre une feuille d'estimation à l'expérimentateur après chaque évaluation de charge de travail due à une situation particulière. Avant de remettre cette feuille, l'opérateur exprime sa réponse sous forme d'une croix par rapport à deux limites inférieure et supérieure.

La charge de travail a pour expression :

$$W_L = \sum_{i=1}^6 \alpha_i R_i$$

avec R_i la réponse de l'opérateur et α_i un coefficient de pondération de la réponse R_i à la $i^{\text{ème}}$ échelle.

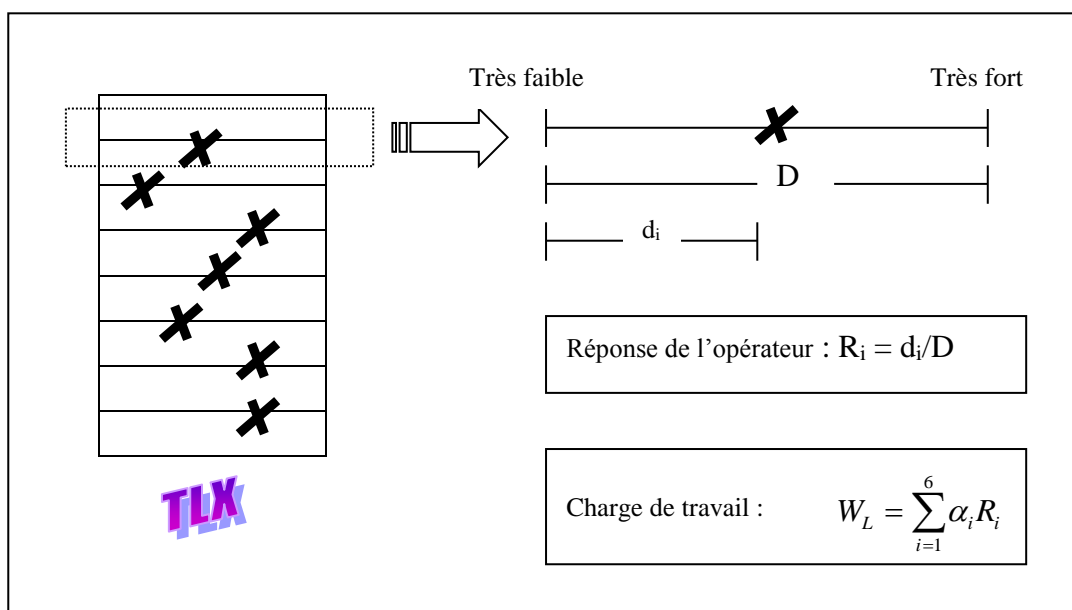


Figure II.13 : Principe de la méthode de TLX

Avantages et inconvénients principaux :

L'avantage de cette méthode est qu'elle prend en compte un nombre de différenciateurs suffisamment représentatifs des exigences du travail, ce qui permet d'obtenir une première estimation de la charge de travail, même si celle-ci est très globale.

Référence :

[HART et STAVELAND, 1988].

L'échelle de Cooper-Harper :*Principe :*

Cette méthode est structurée comme un arbre de décision à dix possibilités, basé sur les caractéristiques suivantes : probabilité d'erreur, charge de travail mental, difficulté, effort exigé compte tenu de la performance. L'opérateur humain, tout en étant guidé par des explications textuelles, doit choisir une des dix possibilités. Le niveau 1 correspond à une difficulté nulle ou extrêmement faible, alors que le niveau 10 correspond à l'impossibilité d'exécuter la tâche.

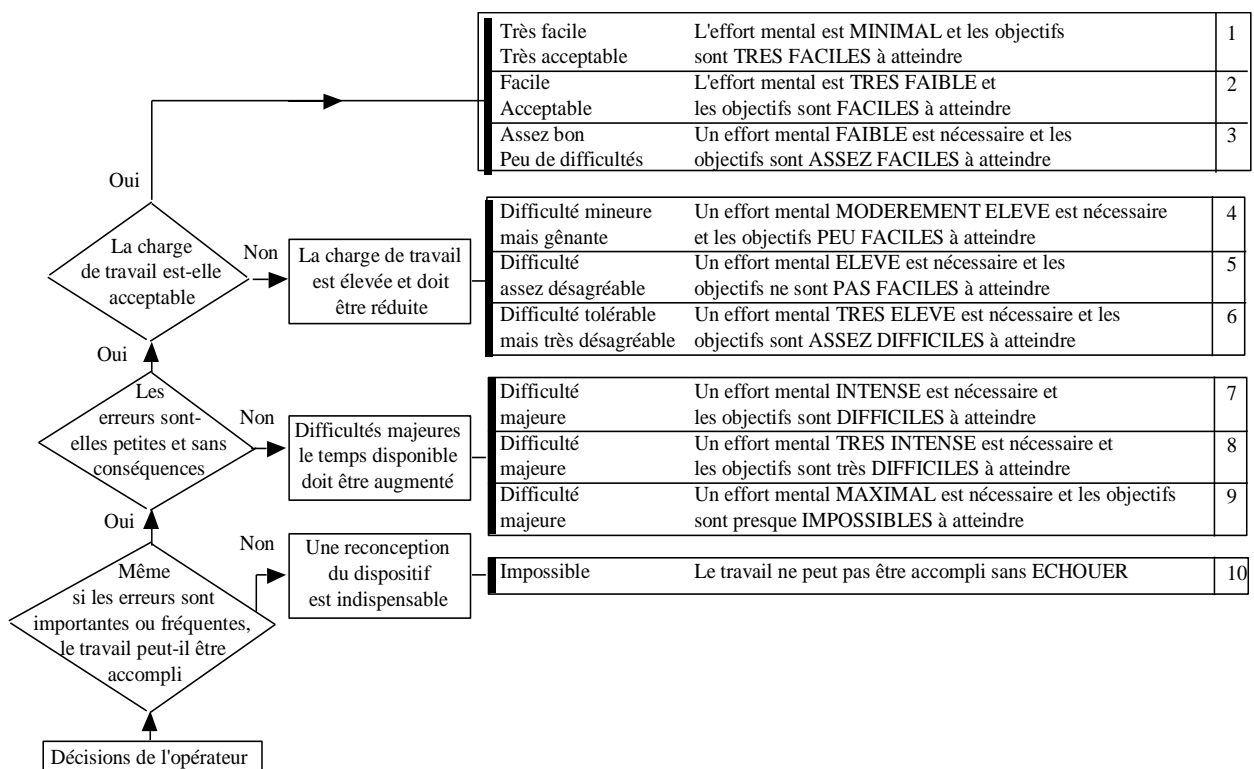


Figure II.14 : Echelle de Cooper-Harper
(version proposée aux sujets d'une expérimentation par BERGER [1991])

Avantages et inconvénients principaux :

L'inconvénient avec cette méthode est que les explications textuelles qui guident l'opérateur humain lors du choix de possibilités sont parfois difficiles à mettre en relation

avec le travail réel. Selon HART [1988], l'échelle est explicitement unidimensionnelle mais l'unique évaluation numérique obtenue est en fait le résultat d'une évaluation multidimensionnelle implicite de la charge de travail (voir figure II.14).

Référence :

[COOPER et HARPER, 1969] [WIERWILLE et CASALI, 1983].

I.2.2. Méthodes objectives

Mesures physiologiques :

Principe :

Cette démarche consiste à évaluer la charge de travail en quantifiant les changements au niveau de certaines fonctions physiologiques (activité électrique du cerveau, des muscles, clignement de l'œil, fréquence cardiaque ou respiratoire, variations de la résistance électrique de la peau, etc.).

Avantages et inconvénients :

L'inconvénient de ces méthodes est le "manque de discrimination vis-à-vis de la charge mentale, d'autres facteurs influant sur les mesures, notamment la charge physique, le stress, la fatigue, l'environnement physique du poste, etc." [MILLOT, 1988].

Référence :

[WIERWILLE, 1979] [WILSON et O'DONELL, 1988] [SPERANDIO, 1991].

Mesure de la performance :

Principe :

L'estimation de la charge de travail peut s'effectuer en mesurant une donnée représentative du niveau de difficulté de l'utilisateur lorsqu'il réalise sa tâche principale. Cette donnée peut être une erreur par rapport à une consigne, un temps de réponse, etc. Toute augmentation de la charge de travail engendre une dégradation de sa performance sur la tâche principale.

Avantages et inconvénients :

Cette technique est limitée par le fait qu'il est difficile de définir la variable ou l'ensemble des variables dépendante(s) de la tâche principale qui sera à même de refléter la charge de travail de l'utilisateur [ODGEN *et al.*, 1979].

Référence :

[LEPLAT et SPERANDIO, 1967] [ODGEN *et al.*, 1979].

Approche temporelle issue de l'Automatique Humaine :

Les approches temporelles sont utilisées par les automaticiens humains pour définir des

modèles permettant d'estimer la charge de travail. On peut citer quelques travaux comme ceux de MILLOT [1988], SIMON [1993] ou RIERA *et al.* [1994].

Principe :

Dans cette catégorie de méthodes, la charge est définie comme étant le rapport entre le temps nécessaire à l'opérateur humain pour effectuer la tâche (avec ce temps requis $TR=TR_{détection}+TR_{cognition}+TR_{action}$) et le temps dont il dispose effectivement pour l'exécuter (appelé temps disponible TD). En résumant l'approche de MILLOT [1988], on peut dire que les exigences à la fois temporelles et fonctionnelles de la tâche sont définies comme le produit des exigences temporelles et d'une fonction de gravité $G(t)$ associée à la tâche :

$$\frac{1}{TD_S(t)} = \frac{1}{TD(t)} * G(t)$$

avec : $TD(t) = TD - t$: temps disponible pour exécuter la tâche,
diminuant au fur et à mesure que le temps s'écoule.
 $G(t)$: gravité de la situation rendant compte des exigences
fonctionnelles et pondérant les exigences temporelles.
 $TD_S(t)$: temps disponible subjectif pour exécuter la tâche.

Lorsque l'opérateur traite une tâche, il est supposé investir dans son exécution la puissance de travail nécessaire pour satisfaire les exigences de la tâche, cette puissance étant définie par $1/TD_S(t)$. L'exécution de la tâche pendant un temps dt , avec la puissance $1/TD_S(t)$, correspond alors à la charge de travail instantanée suivante :

$$dWL(t) = \frac{dt}{TD_S(t)} ; \forall t / TD_S(t)$$

Avantages et inconvénients :

Les méthodes temporelles présentent l'avantage d'un classement des cartes selon une progression qu'il considère comme allant du plus bas niveau de charge au plus haut niveau. e permettre une *évaluation en ligne* de la charge si le temps disponible et le temps requis sont mesurables et ceci pour certaines tâches. Cette méthode a de plus fait l'objet de plusieurs expérimentations dans des situations monotâches et multitâches.

Références :

[MILLOT, 1988] [SIMON *et al.*, 1991].

Métriques de Whiteside :

Principe :

Cette technique consiste à paramétrer les caractéristiques d'une tâche réalisée par un utilisateur. Les différents paramètres sont les suivants :

$S = (C / T) \times P$, où

- S = taux d'achèvement de la tâche (pourcentage de réussite de l'utilisateur par rapport à l'expert) ;
- T = temps passé sur la tâche par l'utilisateur ;
- P = pourcentage effectué de la tâche (par l'utilisateur) ;
- C = temps minimal nécessaire pour exécuter la tâche par un expert.

L'évaluateur, via des enregistrements audiovisuels, des interviews et des questionnaires, vient compléter, par une évaluation qualitative, l'évaluation quantitative précédente.

Avantages et inconvénients principaux :

L'avantage de cette technique est qu'elle permet une mesure objective de la performance de l'utilisateur. Malgré tout, cette technique d'évaluation reste, pour une grande part, artisanale (au sens de LONG et DOWELL [1989]).

Références :

[WHITESIDE *et al.*, 1988] [LONG et DOWELL, 1989].

I.3. Tests de conception

Les tests de conception, bien connus en génie logiciel et trouvant leur inspiration dans le modèle Spirale [BOEHM, 1988], ont pour objet d'évaluer et de valider un produit selon des critères aussi bien techniques qu'ergonomiques. L'évaluation est effectuée selon le cycle itératif tout au long du processus de développement avec des utilisateurs si possible représentatifs. Le principe consiste à recueillir des données au moyen d'une batterie de tests qui engendrent des aménagements mis à l'épreuve à leur tour dans une nouvelle version, et ainsi de suite jusqu'à satisfaction. Les tests de conception peuvent être mis en œuvre lorsqu'il n'existe pas encore d'expérience d'utilisation du système. Trois démarches complémentaires seront successivement décrites (figure II.15) : la sélection d'alternatives, les évaluations itératives et le banc d'essai final. Une quatrième (les groupes de discussion) sera exposée par la suite.

Sélection d'alternatives :

Principe :

La méthode des tests de conception par sélection d'alternatives est utilisable lorsqu'il n'existe pas de critère de choix évident entre plusieurs possibilités ; le recueil des données empiriques doit alors permettre de *hiérarchiser les solutions envisagées au départ*. Cette méthode est la plupart du temps conduite en amont du prototypage [SENACH, 1990].

Avantages et inconvénients :

L'inconvénient d'une telle approche est qu'elle est souvent effectuée en dehors du contexte réel de la tâche, de ses contraintes, de son environnement. Une alternative aboutissant à de médiocres résultats en phase d'évaluation peut en effet s'avérer meilleure dans une situation de travail en se rattachant à d'autres informations suivant la logique de l'utilisateur pour la tâche courante.

Références :

[SENACH, 1990].

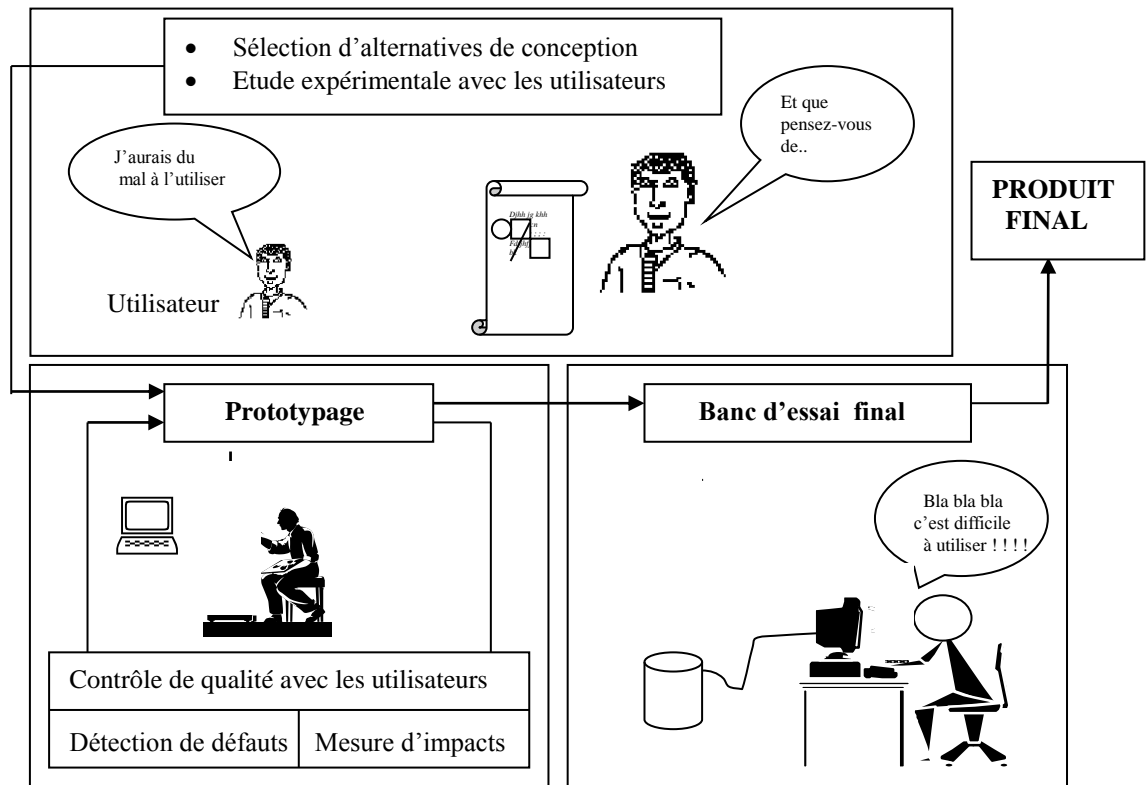


Figure II.15 : Trois approches complémentaires pour les tests de conception (inspirées de [SENACH, 1990] et [GRISLIN, 1995])

Prototypage :*Principe :*

Le prototypage permet l'évaluation précoce, avec les utilisateurs, de certains aspects de l'IHM. On évalue les performances au moyen d'une batterie de tests (diagramme d'usage, mesures,...). Le prototypage a pour objectif de minimiser les coûts de développement de l'interface tout en visant l'optimisation progressive de sa qualité selon l'état d'esprit introduit par le modèle Spirale. Les critères considérés sont principalement : (1) des critères ergonomiques généraux tels que ceux décrits dans les guides ergonomiques et guides de style ; (2) des critères spécifiques à l'application (qualité et quantité de produits, sécurité des travailleurs et des installations, environnement,...). La rentabilité du prototypage est totale si cette dernière version peut être utilisée en tant que produit final. La démarche du prototypage est résumée en figure II.16.

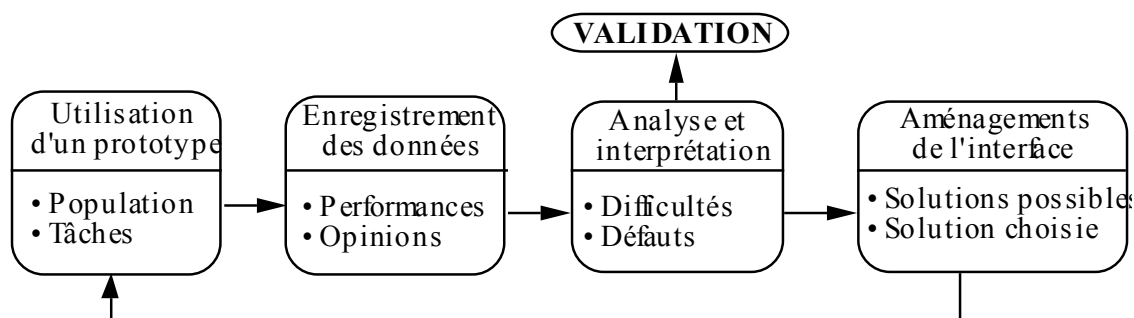


Figure II.16 : Démarche du prototypage [SENACH, 1990]

Avantages et inconvénients principaux :

Le prototypage permet des évaluations précoces lors du développement, ce qui a plusieurs avantages : confrontation des choix de conception à la réalité du terrain, l'utilisateur n'est plus devant le fait accompli, sensibilisation du concepteur à la logique de l'utilisateur et inversement. Le prototypage rend nécessaire la participation des utilisateurs à la démarche de conception, tout en servant de support de dialogue entre ces derniers et les concepteurs. C'est un avantage important compte tenu des nombreux spécialistes de l'interaction Homme(s)-Machine(s) s'accordant à dire que "les utilisateurs potentiels sont d'autant plus enclins à accepter un système qu'ils ont été associés à l'établissement du cahier des charges et qu'ils constatent qu'un certain nombre de leurs suggestions ont été implémentées" [SCAPIN, 1988]. Cette démarche doit cependant être minutieusement mise en œuvre. SENACH [1990] recense en effet plusieurs difficultés provenant particulièrement de ce que les évaluations sont souvent effectuées en situations de laboratoire où certaines conditions de travail ne peuvent être reproduites.

Références :

[BOEHM, 1988] [BODKER et GRONBAEK, 1989] [JAMES, 1991] [POULAIN, 1994].

Groupes de discussion :*Principe :*

Cette approche consiste à rassembler plusieurs corps de métier afin de discuter des différents aspects de l'interface (voir figure II.17). Ces entrevues permettent de se focaliser sur l'interaction Homme(s)-Machine(s) et notamment sur ses problèmes potentiels [CACM, 1993].

Avantages et inconvénients principaux :

On constate que pour ces tests de conception, les 'Groupes de discussions' qui font intervenir plusieurs experts (ou acteurs plus ou moins expérimentés) de spécialités différentes admettent des résultats plus intéressants que le prototypage 'intuitif'. L'un des inconvénients réside dans le fait que les compromis sont parfois difficiles à trouver entre acteurs de spécialités différentes.

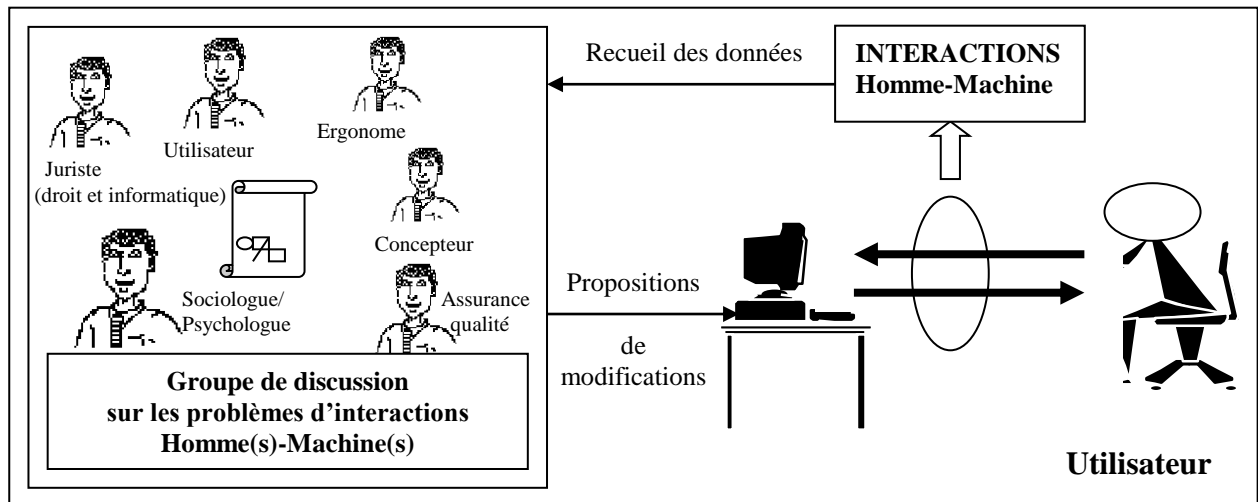


Figure II.17 : Principe du groupe de discussion

Référence :
[CACM, 1993].

Le banc d'essai final - station d'évaluation :

Principe :

En général, toute commercialisation de logiciel est précédée d'une évaluation globale de chacun de ses composants pour tenter d'en faire ressortir les points forts et les points faibles. Dans ce but, des *stations d'évaluation*, faisant office de *banc d'essai final*, font leur apparition dans certains centres de recherche et de développement. Cette technique permet de contrôler la qualité du produit final par une mesure d'utilisabilité globale du système. Les bancs d'essai interviennent une fois l'interface implémentée. On recueille de manière empirique des données objectives et subjectives sur des stations d'évaluation spécialement aménagées : microphones, magnétophones, glaces sans tain, caméras, etc.. Beaucoup d'entreprises disposent des stations de bancs d'essais ; c'est la cas par exemple chez 'LOTUS', 'MICROSOFT' ou 'WORDPERFECT'.

Avantages et inconvénients principaux :

Un intérêt majeur réside dans le fait que l'utilisateur est placé dans des conditions très proches d'un environnement réel. Selon GRISLIN [1995], ces approches sont exploratoires dans de nombreux cas et peuvent ne pas s'appuyer sur une véritable méthode.

Références :

[NEAL et SIMON, 1983] [NOVARA *et al.*, 1987] [ROWLEY, 1994].

I.4. Conclusion sur les approches centrées sur les utilisateurs et positionnement dans le cycle de développement

La plupart des approches empiriques, en particulier les techniques de diagnostic d'usage ainsi que celles permettant l'estimation de la charge de travail, nécessitent une expérience d'utilisation et un site de référence existant, ce qui n'est pas toujours possible pour la conception d'un système complètement nouveau.

Notons toutefois que le développement d'un système démarre rarement de rien puisqu'il vise généralement l'amélioration d'une situation existante, déjà informatisée ou non. Dans ce cas, toutes les techniques de diagnostic d'usage peuvent être envisagées dès les étapes de spécification et de conception et introduire un état d'esprit d'évaluation. Bien entendu, on les utilisera aussi lors des phases de tests finaux et d'exploitation lorsque l'interface existe concrètement [GRISLIN et KOLSKI, 1996].

II. Approches centrées sur une expertise

Lorsque les données relatives à l'utilisation ne peuvent être enregistrées, ne sont pas disponibles ou lorsque l'interface n'existe pas, le recueil des données du modèle de l'interaction Homme(s)-Machine(s) peut se faire par les approches expertes, basées le plus souvent sur l'intervention de spécialistes de l'interaction Homme(s)-Machine(s).

Pour ces approches expertes, on distinguera d'abord celle prenant la forme d'une expertise humaine lors de l'évaluation de l'interface Homme(s)-Machine(s) par un spécialiste. Seront ensuite présentés le cas où le spécialiste prend pour un temps la place de l'utilisateur afin de mieux cerner les difficultés rencontrées (Expérience de l'analyse) et les grilles d'évaluation ergonomique.

Expertise humaine :

Principe :

Cette technique consiste à faire intervenir un ou plusieurs experts pour juger de la qualité ergonomique d'une interface et proposer des améliorations. Cette méthode est sans doute celle qui donne les meilleurs résultats, même si elle est actuellement sous-employée, ce que démontre l'analyse de MAIRE et BRUMENT [1988] des réalisations en France qui met en évidence que le recours à un ergonomiste ne concerne qu'une faible proportion des projets industriels analysés : lorsqu'un ergonomiste est consulté, c'est souvent à une étape avancée des études de détail, principalement sur des questions d'aménagement physique de la salle de contrôle (éclairage, mobilier, etc.).

Avantages et inconvénients principaux :

Cette technique est simple à mettre en œuvre pour un diagnostic rapide des erreurs de conception. L'inconvénient est que l'évaluation est basée sur un modèle individuel acquis par la pratique et qui risque donc d'être trop personnel [HAMMOND *et al.*, 1984]. POLLIER [1991] a fait de telles observations dans une étude récente : quatre ingénieurs ergonomes expérimentés (entre 5 et 12 ans d'expérience en ergonomie de l'informatique)

avaient à évaluer la qualité ergonomique de l'interface d'une application de gestion de communication multimédia. POLLIER a constaté que cinq stratégies très imbriquées, et correspondant à des examens de l'interface selon des points de vue différents, ont contribué à des degrés et des moments divers à la démarche d'évaluation de chacun des ergonomes, et que des variations interindividuelles importantes sont apparues au niveau de la manière de procéder et des priorités adoptées par les ergonomes pour chacune des cinq stratégies¹.

De plus, en considérant que chacun des ergonomes avait trouvé en moyenne moins de la moitié (42%) des problèmes considérés par l'ensemble des quatre ergonomes, il fallait faire la synthèse des diagnostics d'au moins trois d'entre eux pour couvrir l'ensemble des problèmes. En moyenne, deux ergonomes en couvraient deux tiers. Enfin, il s'est avéré que les domaines d'expérience et les connaissances préalables concernant le produit à évaluer ont joué un rôle déterminant dans la diversité des diagnostics. POLLIER en conclut que même si les évaluations d'experts sont efficaces, il existe un biais de focalisation important.

Références :

[HAMMOND *et al.*, 1984] [POLLIER, 1991].

Expérience de l'analyste :

Principe :

Cette technique consiste à se mettre à la place des utilisateurs et à apprendre leur métier. Selon SPERANDIO [1991], elle est caractérisée par le fait qu'elle n'est pas utile pour les tâches complexes. Issue de l'ergonomie, cette méthode a été suggérée par plusieurs auteurs, tels OMBREDANE et FAVERGE [1955], pour des tâches gestuelles ou à dominance sensori-motrice, en faisant l'hypothèse que cet apprentissage permet dans certains cas à l'analyste de mieux appréhender certaines difficultés rencontrées par les opérateurs. Pour des tâches informatiques bien spécifiques (par exemple en bureautique) où l'analyste devient un utilisateur potentiel, la mise en œuvre de cette méthode paraît entraîner moins de difficultés et en tout cas permettre au moins l'évaluation de la prise en main du système. Par exemple, cette méthode a été utilisée lors de l'évaluation ergonomique des interfaces expert et utilisateur du système SYNOP : partant de l'hypothèse que ce système était susceptible d'apporter une aide à un concepteur ergonomiste, il a été demandé à deux analystes d'utiliser le système et de l'évaluer. Cette évaluation a débouché sur plusieurs dizaines de recommandations qui ont ensuite été prises en compte par l'équipe de développement [KOLSKI, 1989].

Avantages et inconvénients :

Une particularité affectant sa performance par rapport à d'autres approches est que l'analyste a souvent tendance à se prendre comme modèle et à pratiquer l'introspection sur

¹ Les stratégies concernaient respectivement l'évaluation de l'interface (1) selon des objectifs d'utilisation de celle-ci (correspondant aux tâches et sous-tâches à réaliser par les utilisateurs), (2) selon la structure, correspondant à la séquence d'examen des différents scénarios ou fonctionnalités qui ont été passés en revue par les ergonomes, (3) selon des niveaux d'abstraction et des priorités, (4) selon les objets présents sur l'interface, (5) selon des critères particuliers d'évaluation.

lui-même. La plupart des auteurs ne recommandent pas cette méthode : elle est à exclure lorsque les tâches sont compliquées et elle est de plus rarement utile. L'analyste ne peut ici que devenir un opérateur sans expérience et donc éprouver des difficultés à analyser objectivement le travail [SPERANDIO, 1991].

Références :

[OMBREDANE et FAVERGE, 1955] [SPERANDIO, 1991].

Grilles d'évaluation :

Principe :

La Grille d'évaluation permet d'évaluer la qualité par notation systématique par rapport à une liste complète des propriétés d'une interface (voir figure II.18). L'évaluateur note systématiquement l'interface selon une échelle comportant généralement de 3 à 5 points, parfois 10. Les recommandations de la littérature peuvent composer les entrées de la grille: guide de 944 conseils ou principes [SMITH et MOSIER, 1986a], aide-mémoire des questions à se poser [SCAPIN, 1987], « check-list » pour l'évaluation [RAVDEN et JOHNSON, 1989].

Une méthode d'évaluation par "check-list" a par exemple été mise au point par RAVDEN et JOHNSON [1989] dans le cadre d'un projet ESPRIT, selon le principe visible en figure II.18. Elle est utilisable par des concepteurs, des spécialistes des facteurs humains et les utilisateurs finaux du système. Elle a, d'après JOHNSON *et al.* [1989], obtenu d'excellents résultats selon des démarches d'évaluation différentes lors d'évaluations comparatives d'interfaces et d'évaluations de prototypes.

GRILLE D'ÉVALUATION		Note				
		1	2	3	4	5
Prise en compte de l'expérience de l'utilisateur, respect du niveau d'expertise						
1	L'information est-elle présentée dans le même état d'esprit que l'utilisateur s'en fait ?			X		
2	L'affichage graphique est-il compatible avec la représentation que se fait l'utilisateur du système ?	X				
3	La structure du système est-elle adaptée à la perception de la tâche de l'utilisateur ?			X		
4	Le déroulement des activités exigées pour effectuer une tâche est-il en accord avec les attentes de l'utilisateur ?		X			
5	Le système travaille-t-il dans la même voie que la pensée de l'utilisateur ?					X
6	Estimation globale du système en terme de prise en compte de l'expérience de l'utilisateur ?					X
Clarté visuelle <i>Caractéristiques lexicales de présentation des informations (cognitif, perceptif)</i>						
Groupement par localisation, positionnement des items, appartenance-distinction.						
1	Chaque écran est-il clairement identifié (titre informatif ou description) ?					X
2	L'organisation des informations sur l'écran est-elle toujours logique ?			X		
3	La séparation entre les différents informations est-elle...					X

Figure II.18 : Extrait d'une grille d'évaluation (inspirée de [RAVDEN et JOHNSON, 1989])

Avantages et inconvénients :

Malgré les efforts accomplis dans la définition et l'optimisation de telles grilles, elles peuvent être d'un apport très bénéfique pour l'évaluateur puisqu'elles l'obligent à n'oublier aucun des aspects de l'interface. Cette technique présente quelques limitations que SENACH [1990] met particulièrement en avant : les problèmes méthodologiques associés aux échelles de mesure, l'adéquation fonctionnelle des items de la grille d'évaluation pour

le logiciel considéré, l'énoncé des aspects qui doivent être analysés (celui-ci est souvent trop général, trop subjectif, certains aspects s'avérant parfois contradictoires), et la notation des propriétés identifiées faite souvent de manière très personnelle par l'évaluateur.

Références :

[JOHNSON *et al.*, 1989] [GILMORE *et al.*, 1989].

Conclusion sur les Approches centrées sur une expertise et leur positionnement dans le cycle de développement

Les approches expertes autorisent l'évaluation dans des situations d'existence ou non du système interactif en question. L'expérience de l'analyste n'est envisageable qu'une fois l'interface réalisée alors que l'expertise, les grilles d'évaluation et les groupes de discussions sont utilisables en parallèle à la conception [GRISLIN, 1995]. Ces méthodes contribuent à améliorer la définition de l'architecture de l'interface, des enchaînements graphiques, la présentation de l'information et des traitements d'affichage. Rappelons encore que, bien que l'expertise humaine soit sans doute la méthode qui donne les meilleurs résultats, le recours à un ergonomiste ne concerne qu'une faible proportion des projets industriels analysés.

III. Méthodes analytiques

Ces approches se basent sur des modèles formels, informatisés ou non, de l'interface et/ou de l'interaction Homme(s)-Machine(s), ainsi que sur la mise en œuvre progressive de métriques objectives. L'évaluation est réalisée en comparant une représentation de l'interface à une référence décrivant les qualités d'une « bonne interface » selon plusieurs critères ergonomiques prédéfinis.

Les modèles sont utilisés pour prédire, au moyen de spécifications, certains aspects liés à l'interaction Homme(s)-Machine(s) (hiérarchie des tâches, cheminement des actions de l'utilisateur, temps requis pour réaliser une tâche,...). Les métriques sont mises en place pour mesurer de façon objective certains aspects associés à la qualité de l'interaction (cohérence, compatibilité avec l'image mentale de l'utilisateur, qualité des écrans,...). Le recours à ces représentations abstraites autorise des prédictions relatives aux performances qui ne peuvent être établies avec une approche empirique, lorsqu'il n'y a pas encore d'expérience d'utilisation du système interactif [SENACH, 1990], [GRISLIN, 1995].

Cette partie est consacrée aux modèles fondés sur une théorie, puis aux modèles formels prédictifs et de qualité de l'IHM.

III.1. Approches fondées sur une théorie

Les méthodes qui vont suivre sont basées sur une théorie. Les plus usuelles d'entre elles sont « Cognitive Walkthrough » et les modèles CCT (Cognitive Complexity Theory), PUM (Programmable User Model), ICS (Interacting Cognitive Subsystems).

Cognitive Walkthrough :

Principe :

Cette méthode, consistant en l'intervention d'un expert humain, permet de prédire les aspects de l'IHM facilitant la résolution des problèmes rencontrés par l'utilisateur. A chaque étape dans le déroulement de la tâche, un questionnaire spécifique est rempli par l'évaluateur. Ce questionnaire comprend des questions concernant les buts en cours, le choix et l'exécution des actions, les retours informatifs du système, etc. L'analyse des résultats permet en principe de souligner les problèmes rencontrés par l'utilisateur lors de l'apprentissage des tâches (problèmes liés aux buts et aux actions) ; l'interface doit donc être modifiée en conséquence, la méthode préconisant de recommencer l'étude avec l'interface améliorée. Après une première phase de préparation, l'évaluateur choisit les tâches humaines à analyser, chaque tâche devant être décrite et associée à une séquence d'actions.

Avantages et inconvénients principaux :

Cette technique s'appuie sur une théorie cognitive de l'apprentissage [LEWIS *et al.*, 1990]. La méthode Cognitive Walkthrough a été utilisée à de nombreuses reprises, d'abord avec des tâches relativement simples et bien identifiées : édition de textes, envoi de mails par exemple [POLSON *et al.*, 1992], vis-à-vis desquelles elle a permis de prédire de nombreux problèmes d'utilisabilité. Notons que pour des tâches complexes, la méthode s'avère lourde à utiliser (en raison des nombreuses fiches à remplir par l'évaluateur).

Références : [LEWIS *et al.*, 1990] [POLSON *et al.*, 1992].

CCT (Cognitive Complexity Theory) :

Principe :

La méthode CCT se situe dans la lignée de GOMS. Mais contrairement à GOMS qui ne modélise que l'activité experte, CCT permet de comparer les mérites relatifs de plusieurs alternatives de conception en termes de transfert de connaissances et de difficulté d'apprentissage [KIERAS et POLSON, 1985]. Pour ce faire, KIERAS et POLSON utilisent, pour une interface donnée, un modèle de la tâche et un modèle formel du processus mental auquel il est supposé que l'utilisateur fait appel dans l'utilisation de cette interface. Dans un but d'évaluation, le modèle CCT vise à formaliser la notion de complexité du point de vue de l'utilisateur du système informatique. D'après ces auteurs, celle-ci dépend particulièrement de la complexité de la représentation de la tâche, des fonctions du dispositif et de la facilité de compréhension du système. Notons que le modèle du processus mental reprend celui de NEWELL et SIMON [1972] qui s'appuie sur un système de production utilisant des règles de la forme :
IF (condition) ALORS (action)

Avantages et inconvénients principaux :

Dans sa critique du modèle, SENACH [1990] reste réservé quant à l'interprétation de la complexité que font KIERAS et POLSON [1985], pour deux raisons essentielles : (1) une

tâche ne sera sans doute pas décrite de la même manière par des analystes différents, engendrant ainsi un nombre d'inférences différentes, donc un taux de complexité qui s'avérera variable selon la description qui aura été faite de la tâche ; (2) les différences interindividuelles, et particulièrement la notion d'expérience, ne sont pas traitées par le modèle, bien qu'il soit admis que l'efficacité de l'utilisateur dépende de son expérience, de sa connaissance des tâches à effectuer et des moyens mis à sa disposition.

Références :

[KIERAS et POLSON, 1985] [SENACH, 1990].

PUM (Programmable User Model) :

Principe :

Le modèle PUM relève du même esprit que la Théorie de la Complexité Cognitive. PUM inclut une architecture cognitive qui permet de prédire l'utilisabilité d'un système en terme de résolution de plan [YOUNG et WHITTINGTON, 1990]. Pour utiliser PUM, l'évaluateur spécifie ce que l'utilisateur doit savoir sur la tâche et sur l'interface. Cette description est ensuite compilée en règles exécutables par PUM. L'exécution des règles fournit une simulation du comportement cognitif de l'utilisateur en train d'exécuter la tâche avec l'interface spécifiée. Si le simulateur ne trouve pas de plan de résolution pour la tâche donnée, l'évaluateur doit alors déduire l'existence potentielle d'un défaut d'utilisabilité ou une difficulté d'apprentissage.

Avantages et inconvénients principaux :

L'un des inconvénients de cette méthode est d'être complexe dans sa mise en œuvre. L'évaluateur doit en effet avoir une bonne connaissance dans le domaine de la description des tâches. La démarche proposée par PUM constitue toutefois une contribution considérable dans le domaine de la modélisation du comportement cognitif d'un utilisateur.

Référence :

[YOUNG et WHITTINGTON, 1990].

ICS (Interacting Cognitive Subsystems) :

Principe :

Le modèle ICS peut se voir comme un affinement du Modèle du Processeur Humain [CARD *et al.*, 1983] qui serait axé sur l'étude du comportement cognitif [BARNARD, 1987]. Dans ICS, l'utilisabilité se mesure, pour l'essentiel, en terme de coût cognitif (alors que les métriques issues du Modèle du Processeur Humain visent plutôt des performances de bas niveau telles la capacité des mémoires des processeurs, la persistance des informations en mémoire et les cycles de base des processeurs). Dans ICS, le sujet humain est modélisé sous forme de neuf sous-systèmes qui fonctionnent en parallèle et qui communiquent via un bus commun de données (voir figure II.19).

Chaque sous-système est spécialisé dans le traitement d'un type d'information et tous répondent à la même organisation : une mémoire locale dans laquelle est recopiée l'information reçue en entrée et qui peut être réutilisée par le sous-système comme donnée d'entrée (ce tampon sert en particulier à la temporisation des informations), des procédures de traitement qui produisent en sortie des informations de type données et qui sont transmises sur le bus commun vers d'autres sous-systèmes spécialisés. Il convient de distinguer les sous-systèmes périphériques des sous-systèmes centraux. Les premiers sont de nature sensorielle (acoustique, visuel, tactile) ou bien moteurs (articulatoire, viscéral, corporel). Les sous-systèmes centraux reflètent des niveaux de traitement cognitif : les sous-systèmes « objet » et « morphonolexical » modélisent respectivement la structuration de scènes visuelles et d'expressions langagières. Les sous-systèmes propositionnel et implicationnel raisonnent au plus haut niveau d'abstraction (représentation de la connaissance et résolution de problème).

Le principe est le suivant : identifier les sous-systèmes impliqués en notant le suivi des informations dans l'architecture ICS. La lecture d'une phrase sur l'écran met par exemple en jeu les sous-systèmes visuel, objet, morphonolexical, propositionnel et implicationnel.

Avantages et inconvénients principaux :

Cette méthode est intéressante dans la mesure où plusieurs recherches interdisciplinaires sont actuellement menées sur celle-ci. L'un des inconvénients de cette méthode est que lors de la lecture d'une phrase sur l'écran, si les caractères sont trop petits (plusieurs boucles sont peut-être nécessaires pour lire la phrase), le chemin se complexifie. La même remarque peut d'ailleurs être faite si la structure syntaxique est erronée et si la phrase n'a pas de sens. Dès lors surgissent des boucles de traitement entre les sous-systèmes. Un nombre important de boucles est une façon de prédire une surcharge cognitive [BARNARD, 1987].

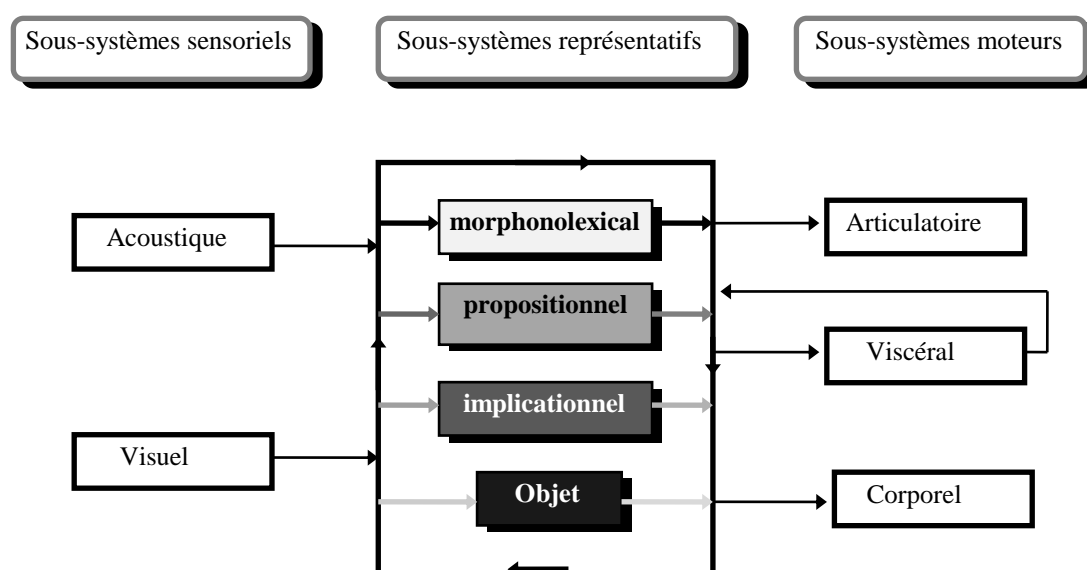


Figure II.19 : Schéma du fonctionnement cognitif de l'individu selon ICS

Référence :

[BARNARD, 1987].

III.2. Modèles formels prédictifs

La modélisation des tâches humaines est particulièrement importante dans la mesure où elle facilite la communication entre les intervenants de l'équipe de développement. D'après de nombreux auteurs [BARTHET, 1988] [SENACH, 1990] [ABED, 1990][KOLSKI, 1997], les besoins en assistance peuvent être mieux mis en évidence et considérés lors du choix des différentes fonctions d'aide aux tâches humaines ; de même, on considère que les besoins informationnels de l'utilisateur ne peuvent qu'être mieux pris en compte lors de la spécification des IHMs si ceux-ci sont situés par rapport à une modélisation des tâches humaines.

Cette modélisation peut être effectuée au moyen de différents modes de description généralement inspirés d'une décomposition en buts et sous-but. Ces décompositions peuvent respecter plusieurs formalismes issus des modèles formels prédictifs. Dans ceux-ci, on distingue les modèles de tâches et les modèles linguistiques.

III.2.1. Modèles de tâches

Les modèles de tâches consistent en une décomposition hiérarchique des tâches pour représenter l'activité cognitive d'un utilisateur. Les modèles les plus connus parmi ceux applicables aux applications interactives sont présentés ci-dessous. Dans ce type de description, la tâche est souvent décomposée par raffinement successif en partant de la tâche globale ou, au contraire, à partir des constituants élémentaires (analyse descendante ou ascendante). Elle fournit un support formel pour l'évaluation prédictive de la performance puisqu'elle propose une description mesurable du comportement de l'utilisateur et de ses performances. Sans souci d'exhaustivité, un ensemble de modèles très connus sont décrits ci-après : HTA, GOMS, Keystroke, MAD, SADT-Petri. Il aurait par exemple été possible de rajouter DIANE [BARTHET, 1988] et DIANE+ [TARBY, 1993].

HTA :

Principe :

Le principe consiste à décomposer hiérarchiquement la tâche globale en sous-tâches élémentaires pour mettre en évidence les compétences et informations nécessaires à l'utilisateur pour chaque but ou sous-but, compte tenu de la structure du système. La décomposition de la tâche s'arrête lorsque le rapport entre le coût et la probabilité de défaillance est inférieur à un coefficient prédéfini. De même, moins l'arbre est parcouru, moins le coût est important ; mais dans ce dernier cas, les chances de réussite s'affaiblissent.

Avantages et inconvénients principaux :

Le modèle de tâche HTA a l'avantage d'être l'ancêtre des modèles et a été utilisé avec succès dans l'industrie pour la formation des utilisateurs. Il a également servi de moyen

d'analyse de la fiabilité humaine pour un ensemble de tâches à réaliser dans des contextes industriels. L'inconvénient de ce modèle est que les coûts de mise en œuvre sont proportionnels à la descente dans l'arbre de description.

Références :

Les informations relatives à cette technique sont disponibles dans [DUNCAN, 1981].
[DUNCAN, 1981] [STAMMERS *et al.*, 1990] [SHEPHERD, 1993].

GOMS (Goal Operator Method Selection rules) :

Principe :

Cette méthode consiste à décrire l'activité cognitive d'un utilisateur (expert) engagé dans la réalisation d'une tâche de routine sans erreur [CARD *et al.*, 1983]. GOMS se décompose en *Buts* (« Goals ») organisés en une hiérarchie, en *Opérateurs* correspondant aux actions élémentaires (« Operators »), en *Méthodes* ou procédures de réalisation d'un but (« Methods ») et en *Règles de sélection* (« Selection rules ») des méthodes lorsque plusieurs solutions permettent d'atteindre un même but. Le modèle CPM-GOMS [GRAY *et al.*, 1992] introduit dans GOMS est une méthode étendue au calcul du chemin pour déterminer un temps critique de réalisation d'une tâche. Rappelons qu'elle est orientée vers la conception et peut servir pour l'évaluation.

Avantages et inconvénients principaux :

Le modèle GOMS est réducteur puisqu'il ne permet que de prédire les performances d'un utilisateur expert ne commettant pas d'erreur, au contraire de Keystroke décrit ci-après [COUTAZ, 1990]. Il a de plus été validé au départ pour des tâches bien définies, telles que celles nécessaires pour l'édition de texte, puis sur des tâches de plus grande complexité, par exemple [JOHN et VERA, 1992]. Par son approche formelle, GOMS a ouvert des perspectives intéressantes et a conduit à la définition d'autres modèles (tel Keystroke présenté ci-dessous).

Références :

[CARD *et al.*, 1983] [JOHN, 1988] [JOHN et VERA, 1992] [GRAY *et al.*, 1992]
[IRVING *et al.*, 1994].

Keystroke :

Principe :

Ce modèle est dérivé de GOMS (cf. ci-dessus). Cette approche permet de prédire les performances d'un utilisateur expert lorsqu'il effectue une tâche à partir de quatre descriptions : (1) la description préalable de la tâche, (2) celle du langage de commande du système, (3) celle de paramètres caractéristiques des capacités motrices de l'utilisateur et des temps de réponse du système, et (4) celle de la méthode de réalisation de la tâche.

La prédiction s'effectue selon les calculs suivants :

- Le temps d'exécution est la somme des temps passés à exécuter chacune des six classes d'opérateurs ("Keystroking" : frappe de touches du clavier ou clic de la souris, "Pointing" : désignation, "Homing" : rapatriement de la main, "Drawing" : action de dessiner, "Mental activity" : activité mentale, et "Response time" : temps de réponse du système) :

$$T_{\text{exécution}} = T_K + T_P + T_H + T_D + T_M + T_R$$

- La prédiction du temps d'exécution d'une tâche comme la somme des durées des sous-tâches la constituant :

$$T_{\text{tâche}} = T_{t1} + T_{t2} + \dots + T_{tn}$$

- Le temps de réalisation d'une tâche élémentaire est la somme du temps nécessaire à l'utilisateur pour se constituer une représentation mentale de la tâche (ce temps est supposé constant pour un type de tâche) et du temps nécessaire pour l'exécuter :

$$T_{ti} = T_{\text{acquisition}} + T_{\text{exécution}}$$

Avantages et inconvénients principaux :

L'avantage de cette méthode est d'être facilement compréhensible par chaque concepteur et de fournir des valeurs numériques simples, bien que sa mise en œuvre pose rapidement de nombreux problèmes [COUTAZ, 1990]. Validé comme GOMS pour des tâches de traitement de texte, le fait que ce modèle ne tienne pas compte des erreurs de l'opérateur et ne prévoit pas de phase de résolution de problème ne permet pas d'envisager son utilisation directe pour la prédiction de l'efficacité de l'opérateur lorsque celui-ci effectue des tâches complexes.

Références :

[CARD *et al.*, 1983] [COUTAZ, 1990].

MAD (Méthode Analytique de Description) :

Principe :

Le principe de la méthode est de décrire des tâches à partir de données issues de l'analyse du travail d'un utilisateur (logiques de traitement de l'information). La décomposition des tâches se fait de façon structurée et la description de façon déclarative (état du monde) et procédurale (façon d'arriver à un état).

Dans MAD, on trouve plusieurs opérateurs reliant les sous-tâches : (1) la séquence (les sous-tâches devant être exécutées les unes après les autres) ; (2) le parallélisme (les sous-tâches pouvant être exécutées dans n'importe quel ordre, voire en parallèle) ; (3) la boucle (pour les tâches itératives) ; (4) l'alternative (pour exprimer différentes manières d'exécuter une tâche) ; et (5) l'option (pour les tâches non obligatoires).

MAD est complété par un formalisme textuel qui permet de préciser, pour chaque tâche, les éléments suivants : (1) l'état initial (sous-ensemble de l'état du monde constitué de la tâche de la liste des arguments d'entrée de la tâche) ; (2) l'état final pour les arguments de sortie ; (3) le but ; (4) les pré-conditions (ensemble des prédicats qui doivent être satisfaits pour l'exécution de la tâche) et les post-conditions (ensemble des prédicats qui doivent être

satisfaits après l'exécution de la tâche). Cette approche facilite la description des tâches des utilisateurs et le modèle de tâche obtenu permet de maîtriser la différence entre la logique fonctionnelle et la logique d'utilisation (caractéristiques de l'utilisateur et de sa tâche), plutôt que la logique fonctionnelle de l'application et des contraintes imposées par l'ordinateur.

Avantages et inconvénients principaux :

MAD prend en compte le parallélisme et le séquentiel (synchronisation des tâches), tout en proposant une description facile à implémenter. L'inconvénient est que le rendu de l'interruption de l'activité (abandon, changement de but, erreur, modification de chemin dans l'arbre) et les erreurs humaines et/ou techniques, ne sont pas bien prises en compte dans la description. Toutefois, grâce à l'apport de certains travaux, ces manquements tendent à être comblés.

Références :

[SCAPIN, 1988] [SCAPIN et GOLBREICH, 1989] [SEBILLOTTE, 1994] [GAMBOA-ROGRIGUEZ, 1998].

SADT-Petri ou TOOD :

Principe :

Issue du LAMIH de l'Université de Valenciennes, la méthode SADT/Petri avait été prévue initialement pour modéliser des tâches de contrôle de la navigation aérienne. Elle correspond à une combinaison de réseaux de Petri, permettant dynamiquement dans le temps le déroulement des opérations, et de modèles SADT servant à décomposer un système ou une activité de manière fonctionnelle. Cette combinaison (voir figure II.20) est adaptée à la description des tâches humaines. Le modèle naît de cette combinaison et permet, au cours d'une évaluation ergonomique, d'analyser les activités de l'utilisateur lors de l'exécution de ses tâches. SADT suit une logique de décomposition fonctionnelle du système de façon descendante, modulaire, hiérarchique et structurée, à l'aide d'actigrammes, alors qu'une description détaillée de la dynamique des tâches peut être effectuée par réseaux de Petri.

Notons que l'un des objectifs de la méthode est de modéliser aussi bien les tâches prescrites théoriques que les tâches réelles, afin de les confronter formellement [ABED, 1990]. La méthode a évolué vers TOOD (Task Object-Oriented Design) en introduisant une description orientée objet des tâches. Le modèle de tâche résultant se présente sous la forme d'un ensemble d'entités coopératives concernant l'application, l'interface et l'opérateur humain [MAHFOUDI, 1997] [TABARY et ABED, 1998]. L'objectif de TOOD est d'aller jusqu'à la génération automatique de parties de l'IHM.

Avantages et inconvénients principaux :

Cette méthode est adaptée à la vérification de l'adéquation des activités des utilisateurs à celles préconisées par le concepteur puisque le mode de représentation peut être le même pour les deux analyses (tâches théoriques et réelles). Les réseaux de PETRI fournissent une

description fine du comportement des interfaces, en particulier en ce qui concerne leurs réactions suite aux opérations et commandes de l'utilisateur. Pour l'évaluation, cette description permet d'évaluer les parcours entre les différents écrans et d'avoir une idée précise des caractéristiques de l'espace théorique du problème dans lequel va naviguer l'utilisateur.

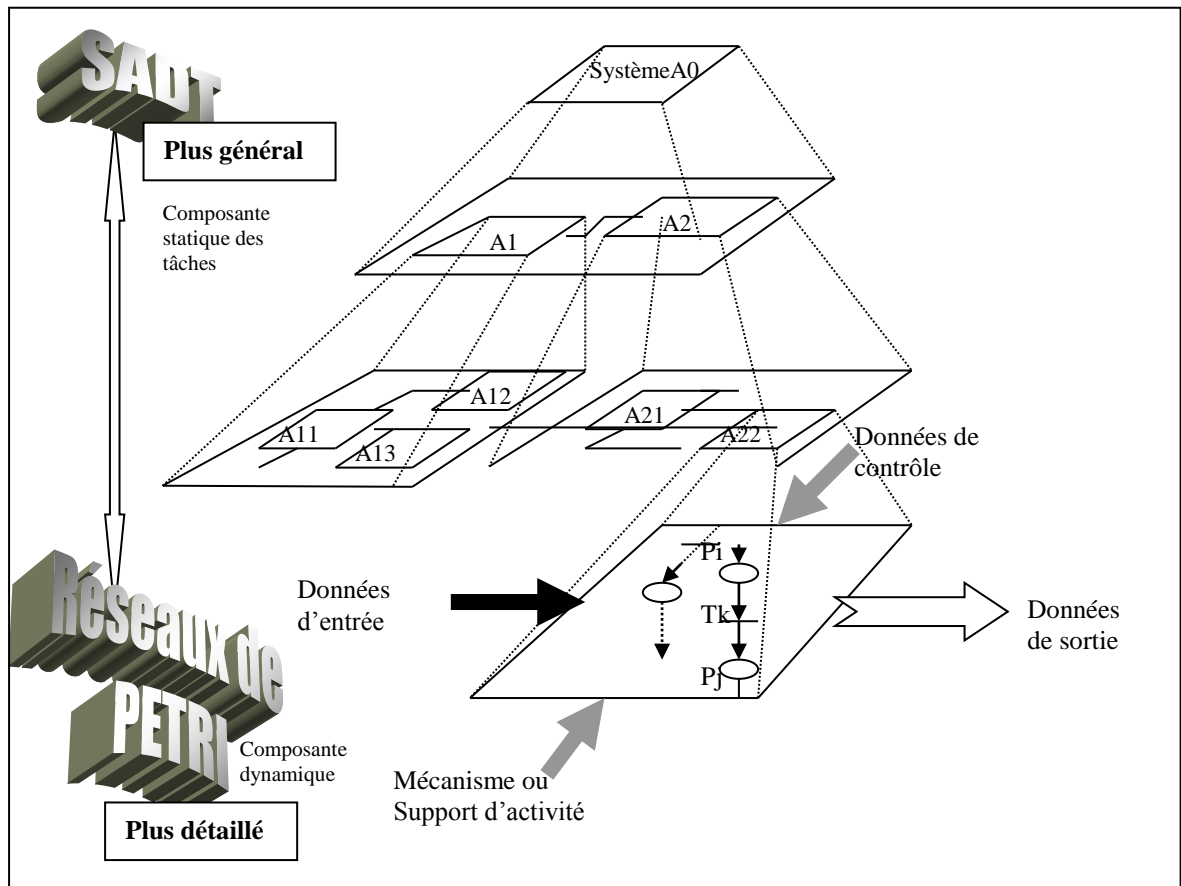


Figure II.20 : Description de la tâche par SADT et réseaux de Petri (inspirée de ABED [1990])

Références :

[ABED, 1990] [BENAÏSSA, 1993] [ABED et ANGUE, 1994] [MAHFOUDI, 1997] [TABARY et ABED, 1998].

III.2.2. Modèles linguistiques

Les modèles linguistiques explicitent la structure de l'interaction Homme(s)-Machine(s) au moyen de grammaires. Ces modèles sont utilisables pour décrire les tâches d'interaction entre l'utilisateur et l'interface, dans un but principalement d'évaluation du langage de commande et de sa cohérence. Ces grammaires formelles constituent des modèles à base d'items à un niveau qui peut être bas, c'est-à-dire parfois souvent sans considération des aspects conceptuels (définition de la tâche globale et des sous-tâches). POITRENAUD [1993] parle à ce sujet de

grammaires d'action, en les situant comme des formalismes qui mettent en avant les liens existant entre le domaine de l'action et celui du langage. Quelques méthodes représentatives sont décrites ci-dessous.

ALG :

Principe :

Le principe est de construire un modèle des actions mises en jeu par l'utilisateur, sous la forme de règles de production : il s'agit d'une décomposition récursive des buts des utilisateurs (symbole initial pour le but général, phrases pour les procédures et mots pour les actions élémentaires).

Les règles de production sont du type : *POUR effectuer-telle-action FAIRE telles-opérations*. Ce modèle a un double objectif : permettre la comparaison entre des alternatives de comparaison selon un critère de facilité d'utilisation, et faciliter l'identification des choix de conception susceptibles d'engendrer des erreurs d'utilisation.

Avantages et inconvénients principaux :

Même si cette méthode présente un intérêt pour la comparaison entre plusieurs versions d'une interface, les trois indices quantitatifs considérés sont insuffisants pour caractériser un langage de commande puisque, par exemple, des tâches faciles à décrire peuvent être difficiles à réaliser [SENACH, 1990]. En outre, ce langage ne prend pas en compte la dynamique et les types de modes d'entrée/sortie, ceux-ci ayant une influence indéniable sur la complexité des tâches en fonction du contexte opérationnel de l'environnement. Toutefois, les indices du modèle pourraient sans doute être utiles pour expliquer aux opérateurs novices les différences entre leurs performances pour une tâche donnée et celles d'opérateurs expérimentés utilisant d'autres stratégies et ceci dans un but de formation [KOLSKI, 1997].

Références :

[REISNER, 1984] [SENACH, 1990].

CLG :

Principe :

Cette méthode permet de décrire des tâches à l'aide d'une série de modèles suivant six niveaux d'abstraction : tâches et sémantique pour les buts et sous-buts, syntaxe et interaction pour la communication entre l'utilisateur et le système, présentation et unités d'entrée-sortie pour la composante physique. Bien qu'orienté vers la conception, le modèle CLG (Command Language Grammar) proposé par MORAN [1981] permet une évaluation prédictive de la facilité d'utilisation de l'interface.

D'après MORAN, le modèle CLG permet de localiser les décisions d'évaluation qui seront critiques pour l'utilisateur. Il propose d'ailleurs différentes métriques permettant de prédire l'efficacité de l'interface. Les idées de MORAN ont été exploitées par SHARATT [1987] qui a mis au point trois métriques relatives à la complexité des tâches à accomplir, à l'optimalité des composants du système décrit en CLG, et enfin à différentes classes d'erreurs susceptibles d'être commises par les utilisateurs.

Avantages et inconvénients principaux :

Cette approche n'est pas directement exploitable pour évaluer un produit fini développé avec une autre méthode : il faudrait traduire l'interface en CLG pour en faire l'analyse. Dans la mesure où l'évaluation selon les principes de CLG n'est possible que si le système est conçu en CLG, ce modèle peut pour l'instant difficilement être exploité pour des applications complexes.

Références :

[MORAN, 1981] [COUTAZ, 1990].

TAG et ETAG :*Principe :*

Le modèle TAG a pour but de modéliser la représentation mentale que se fait un utilisateur d'un langage d'interaction. Les tâches se décomposent en sous-tâches simples routinières pouvant être exécutées par l'utilisateur sans prise de décision (organisation des tâches en actions sous la forme d'un méta-langage). TAUBER [1990] propose donc de décrire les procédures d'utilisation et le modèle de fonctionnement en unités ETAG qui sont des événements de changement de localisation des objets de l'interface ajoutés à TAG. L'idée est de mettre en évidence les connaissances que l'utilisateur doit avoir sur la structure du dispositif qu'il utilise pour réaliser ses tâches. Sur le plan de l'évaluation, cette approche permet d'énumérer les propriétés de consistance conceptuelle de l'interface.

Avantages et inconvénients principaux :

TAG aide à juger la structure de tâches décrites à l'aide d'un langage, mais celui-ci se limite aux tâches simples, c'est-à-dire à la description des fonctions spécifiques du dispositif. Dans ce formalisme, le lien entre les objets mis en jeu dans les actions et les manipulations auxquelles on peut les soumettre est très faible.

Références :

[PAYNE et GREEN, 1989] [SCHIELE et GREEN, 1990] [TAUBER, 1990].

III.3. Les modèles formels dits de qualité de l'IHM : approches logicielles

Les modèles formels dits de qualité de l'IHM sont, dans l'état actuel de la recherche, complémentaires des modèles formels dits prédictifs décrits précédemment. Ils s'intéressent en effet aux propriétés mesurables de l'interface, et ceci selon des critères d'utilisabilité formalisés (le plus souvent sous forme de règles). On se focalisera dans cette partie essentiellement sur des approches consistant en outils informatiques capables d'effectuer automatiquement des mesures ergonomiques sur des pages-écrans, le plus souvent indépendamment du contexte d'utilisation¹. Pour l'évaluation d'un système interactif, de telles approches partent du principe que certains problèmes ergonomiques liés à l'utilisation de l'interface peuvent être évités *a priori*. Mais elles doivent s'intégrer dans une démarche globale d'évaluation selon un ensemble de critères interconnectés dont il s'agit de gérer la cohérence par rapport aux différents contextes

opérationnels d'utilisation. Plusieurs approches démontrent la possibilité de rendre opérationnelles les trop nombreuses recommandations ergonomiques actuellement proposées dans une forme souvent rébarbative (manuels et guides). Notons que les précurseurs dans ce domaine sont STREVELER et VASSERMAN [1984].

Display Analysis Program de Tullis

Principe :

Cette approche consiste à mesurer la qualité des affichages en fonction de six paramètres : la densité globale d'information, la densité locale, le nombre de groupes d'informations distincts, la taille moyenne des groupes, le nombre d'items, la complexité de l'affichage. Après de nombreuses expériences d'évaluations ergonomiques de pages-écrans graphiques et alphanumériques, cette méthode s'est orientée vers une démarche automatique d'évaluation d'écran. Les six paramètres précédents ont été choisis car ils représentent les critères les plus significatifs de la facilité de lecture selon une expérimentation qu'a menée TULLIS [1988] sur 10 sujets ayant à utiliser 52 pages-écrans différentes. Initialement, cette approche s'appuie sur l'hypothèse que la qualité d'un affichage dépend particulièrement des temps de recherche sur l'écran et de l'évaluation subjective que se fait l'utilisateur de l'affichage.

Avantages et inconvénients principaux :

Ce système constitue un pionnier dans cette catégorie. Il possède une bonne capacité à analyser les différentes versions d'un écran et à proposer des aménagements, prédire les temps de recherche et les évaluations subjectives qui découleront de son utilisation, et enfin à fournir des recommandations concernant son amélioration. Le système est toutefois limité aux pages-écrans alphanumériques.

Référence :

[TULLIS, 1988].

SYNOP :

Principe :

SYNOP est un système expert d'évaluation automatique de la présentation statique de synoptiques industriels à partir de recommandations relatives à la présentation des informations à l'écran. L'évaluation procède par pages-écrans créées à l'aide d'un logiciel graphique donné (IMAGIN). Ces pages-écrans sont en fait converties en fichiers binaires traités par SYNOP. C'est donc à partir d'une description de l'interface que l'évaluation s'effectue.

¹ Les approches de génération automatique des affichages ou des spécifications d'une IHM ne sont pas considérées ici, même si elles contribuent d'une certaine manière à l'évaluation *a priori* d'IHM. Le lecteur intéressé se référera à SENACH [1990], GRISLIN et KOLSKI [1996] et KOLSKI [1997]. Notons aussi que l'objectif de cette partie III.3 est la représentativité des approches proposées et non l'exhaustivité. Il aurait été possible de décrire plusieurs autres outils, tels AIDE [SEARS, 1993] [SEARS, 1994], SHERLOCK [STEPHANIDIS et al., 2000] ou UIDE [BYRNE et al., 1994].

Ce système permet aussi d'effectuer des modifications automatiques, et lorsque ces dernières s'avèrent impossibles, des recommandations sont proposées. SYNOP (pour SYNOPTiques) est parti de l'hypothèse qu'il est possible de décomposer la connaissance ergonomique liée à la perception de l'information en deux ensembles distincts : d'une part les connaissances générales issues des règles classiques d'ergonomie et, d'autre part, les connaissances spécifiques à certaines applications. Ces connaissances ont été décomposées en différents thèmes ergonomiques (la lisibilité des caractères, le format d'écran, la présentation des courbes, etc.) prenant la forme de sous-bases de connaissances d'un système expert écrit en LISP où chaque règle correspond à une recommandation ergonomique.

Le principe d'évaluation et d'amélioration ergonomique d'une vue graphique par SYNOP se décompose en trois phases : (1) interprétation des fichiers graphiques de la vue à traiter et création d'un réseau sémantique d'objets structurés LISP descriptif ; (2) évaluation ergonomique des objets du réseau (à cet effet, les règles contenues dans les sous-bases sélectionnées sont activées), par exemple, une sous-base concernant l'ergonomie de présentation des histogrammes regroupe l'ensemble des règles relatives à ce sous-problème. De même, une autre sous-base peut concerner la lisibilité des caractères, les contrastes colorés, etc. ; (3) reconstruction des fichiers graphiques correspondant à la vue améliorée et affichage à l'écran de la nouvelle vue.

Avantages et inconvénients principaux :

SYNOP ne permet pas une étude globale d'un ensemble de pages-écrans. Il est limité par son principe d'évaluation statique et écran par écran. Bien que possédant des limitations, ce système expert est représentatif des outils d'ingénierie pour l'évaluation d'interfaces Homme(s)-Machine(s) que les techniques actuelles permettent de réaliser.

Références : [KOLSKI, 1989] [KOLSKI et MILLOT, 1991].

KRI/AG :

Principe :

KRI/AG est un système à base de connaissances capable d'évaluer automatiquement une interface de dialogue. Il permet de créer des écrans à l'aide d'un outil graphique sous X-Windows et de les traduire dans un fichier sous la forme d'un pseudo-langage (UIL).

Ces fichiers sont analysés à l'aide de règles contenues dans une base de connaissances (voir figure II.21). Ces connaissances traduisent des règles issues de l'ergonomie des logiciels et de guides de style Motif. Ce système, qui est en état de prototype, démontre bien, dans la lignée des systèmes précédents, la possibilité d'opérationnaliser de nombreuses recommandations ergonomiques.

Avantages et inconvénients principaux :

Cet outil est prometteur pour une évaluation automatique, mais essentiellement pour des interfaces de dialogue. L'évaluation est effectuée hors du contexte d'utilisation.

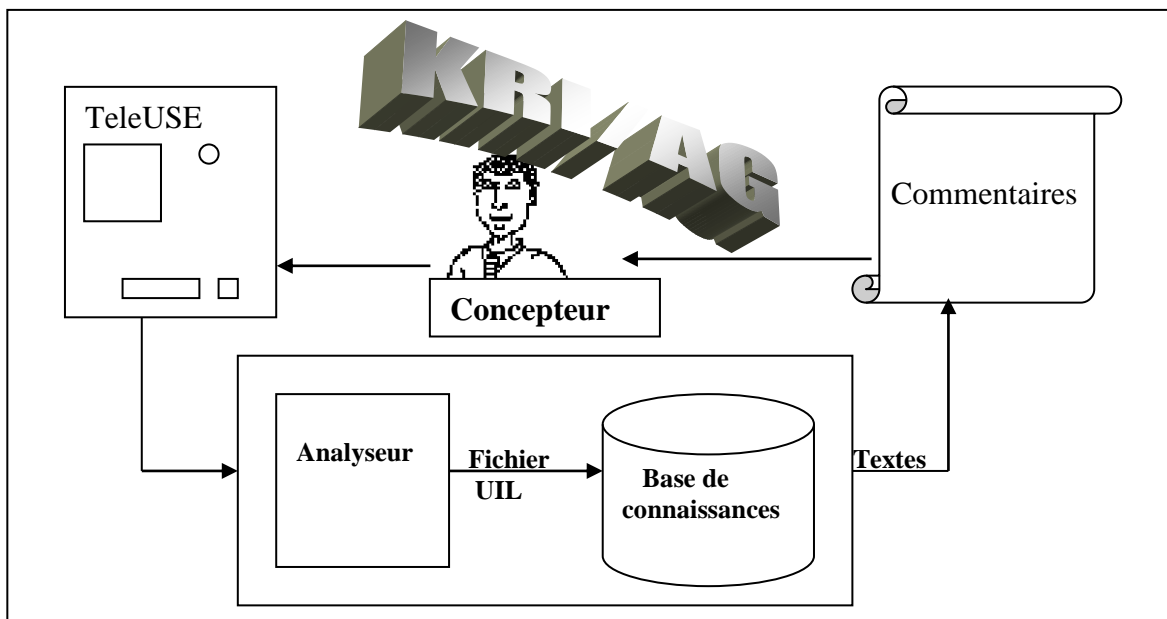


Figure II.21 : Principe d'utilisation de KRI/AG

Référence :

[LÖWGREN et NORDQVIST, 1992].

SHERLOCK :*Principe :*

Développé par MAHAJAN et SCHNEIDERMAN [1995], cet outil d'évaluation de la cohérence de l'interface convertit l'interface créée à l'aide de Visual Basic sous forme d'un fichier de description d'objets. Ce type de fichier est indépendant des plate-formes et peut être créé avec un programme de traduction approprié à partir d'outils tels que Power Builder, Galaxy et XVT.

L'outil permet ensuite d'évaluer la cohérence de l'interface. Plus précisément, il évalue le style et la taille des polices de caractères utilisées dans les boîtes de dialogue afin d'y détecter les incohérences, par exemple des couleurs inadaptées utilisées pour les fonds d'écran. Il permet également de détecter l'incohérence des majuscules dans les mots apparaissant dans les boutons, labels, titres, menus, etc., et ce dans toutes les boîtes de dialogue ; la cohérence des caractéristiques des boutons de commande, c'est-à-dire leur titre, l'emploi des majuscules, leur localisation relative, leur taille (hauteur, largeur) ; et, enfin, l'orthographe des mots utilisés dans tous les objets, de même qu'une évaluation des synonymes. Le résultat de l'évaluation est un fichier de tableaux indiquant les objets évalués et les incohérences détectées. L'évaluateur analyse ensuite les tableaux qui sont en relation avec les écrans créés avec Visual Basic.

Avantages et inconvénients principaux :

Pour l'instant, les aspects de la *Qualité ergonomique* que permettent d'évaluer cet outil sont relativement restreints, du moins dans l'état actuel des choses. Cet outil d'aide à l'évaluation doit donc être considéré comme une technique complémentaire à l'inspection de la *Qualité ergonomique* des systèmes interactifs et aux tests d'utilisation. Cet outil, bien qu'il ne fasse aucunement référence aux tâches et aux caractéristiques des utilisateurs, n'en demeure pas moins utiles.

Rappelons, par exemple, que la seule évaluation d'une interface du point de vue de la cohérence, et à plus forte raison si l'interface est complexe, est une tâche extrêmement difficile à réaliser. Selon MAHAJAN et SCHNEIDERMAN [1995], cette évaluation nécessite une comparaison de la valeur des paramètres de tous les objets à travers l'interface. De plus, pour que cette évaluation soit exhaustive, l'évaluateur doit avoir une bonne représentation du dialogue, ce qui n'est pas non plus simple à obtenir. Ainsi, tout outil facilitant ce type d'évaluation permet à l'évaluateur de consacrer plus de temps aux aspects liés aux tâches.

Référence :

[MAHAJAN et SCHNEIDERMAN, 1995].

ERGOVAL :

Principe :

Cette approche a pour buts de structurer des critères ergonomiques afin de les rendre plus utilisables par des développeurs et de permettre d'évaluer des interfaces graphiques utilisateurs dans l'environnement Windows.

ERGOVAL est un système d'évaluation à base de connaissances. Les règles ergo-miques intégrées à la base de connaissances sont issues de SMITH et MOSIER [1986b], RAVDEN et JOHNSON [1989], VANDERDONCKT [1994] et sont classifiées dans des catégories qui s'apparentent aux *Critères ergonomiques*.

ERGOVAL comporte une décomposition structurelle des objets de l'interface. Ainsi une fenêtre est-elle composée de zones de travail elles-mêmes composées de champs de données, eux-mêmes composés de libellés de champs, de zones de saisie, formats de saisie, etc.

L'exemple de la figure II.22 illustre l'utilisation d'ERGOVAL. Le cas choisi est un éditeur destiné à identifier un client. Cet éditeur permet d'ajouter des clients dans la base de données. Il permet également de sélectionner ou de supprimer un client de la base.

Avantages et inconvénients principaux :

Cet outil est prometteur pour une évaluation automatique, mais essentiellement pour des interfaces de dialogue. L'évaluation est effectuée hors du contexte d'utilisation.

L'évaluation proprement dite de la boîte de dialogue consiste à évaluer tous les objets graphiques contenus dans la boîte de dialogue. Pour chaque objet graphique, différentes règles doivent être prises en considération [FARENC et PALANQUE, 1999]. Le résultat fournit par ERGOVAL consiste en un rapport d'évaluation (liste de règles non respectées et/ou rapport qualimétrique)

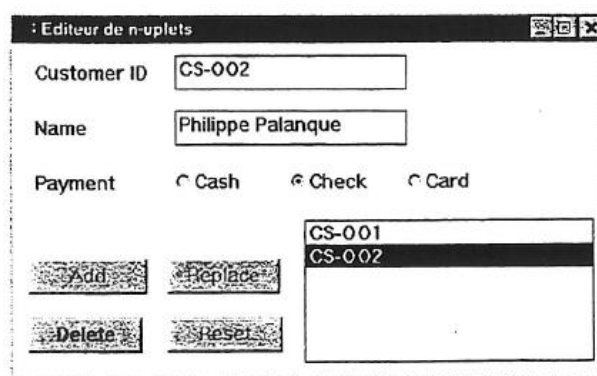


Figure II.22 : Fenêtre d'édition client [FARENC et PALANQUE, 1999]

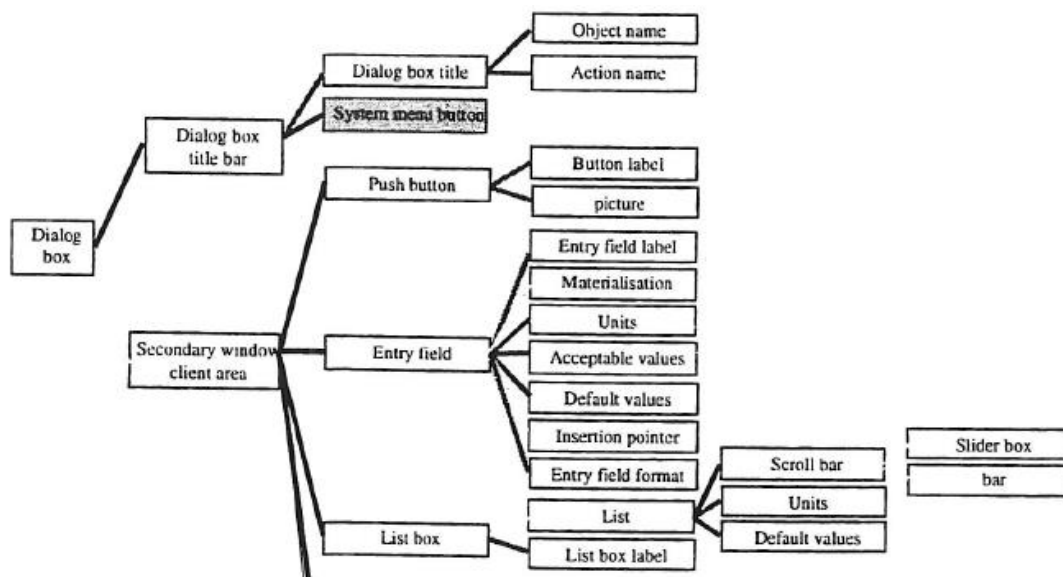


Figure II.23 : Décomposition de la boîte de dialogue [FARENC et PALANQUE, 1999]

Avantages et inconvénients principaux :

L'avantage qu'offre ERGOVAL est qu'elle effectue une évaluation précise et permet de gagner du temps pour des grandes applications. L'inconvénient majeur est qu'elle ne permet pas la saisie automatique des valeurs des objets graphiques [FARENC et PALANQUE, 1999]. De plus, selon FARENC [1997], les informations de natures sémantique et pragmatique portant sur des éléments hors du système ne pourront jamais être récupérées automatiquement et devront donc nécessairement être décrites par un opérateur humain.

Références : [FARENC, 1997] [FARENC et PALANQUE, 1999].

Bobby :

Principe :

La méthode Bobby, plus particulièrement destinée aux sites Internet, évalue, fournit une aide à la conception et propose des améliorations. Le principe consiste à tester chaque page du site en mettant en évidence les éventuelles « insatisfactions » des utilisateurs. Bobby explique tous les problèmes d'accès rencontrés en analysant les différentes pages du site. Cette méthode d'évaluation des sites Web propose des modes de présentation facilement accessibles aux utilisateurs. Mais Bobby n'évalue pas seulement les pages Web : il fournit également aux concepteurs des outils pour l'aide à la conception des sites. Les figures II.24a et II.24b ci-dessous montrent un aperçu d'une page évaluée par Bobby.

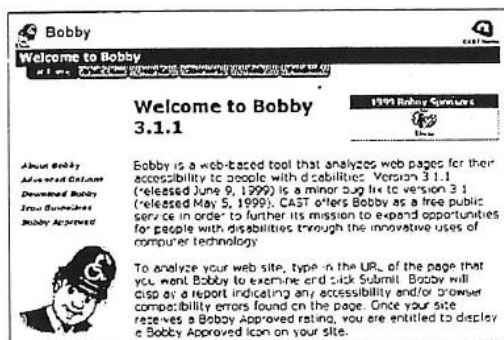


Figure II.24a : La page initiale
[COOPER, 1999]

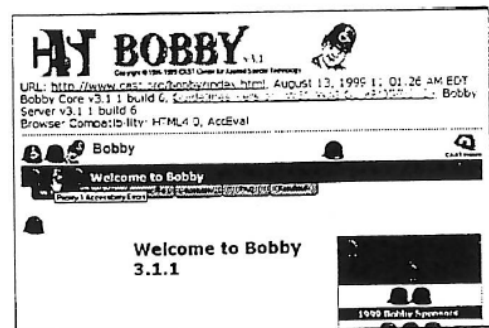


Figure II.24b : La page indiquant les erreurs d'accessibilité auxquelles on accède en cliquant sur « hat »

Avantages et inconvénients principaux:

L'avantage qu'offre Bobby est qu'elle est très simple d'utilisation et qu'elle offre des outils aux concepteurs. L'inconvénient de cette méthode est que dans la plupart des cas, les problèmes rencontrés ne sont pas traités dans la mesure où elle ne propose que des indications vagues pour résoudre le problème.

Référence :

[COOPER, 1999].

III.4 Conclusion sur les modèles analytiques

La description de l'interaction Homme(s)-Machine(s) au moyen des modèles basés sur une théorie des modèles de tâche et linguistiques fournit un support - que l'on qualifie dans la littérature de formel - pour l'évaluation prédictive de la performance puisqu'elle propose une description mesurable du comportement de l'utilisateur. Le modèle obtenu est ainsi utilisé pour l'évaluation ergonomique en analysant la représentation des activités probables des utilisateurs lors de l'exécution de leurs tâches.

De nombreux paramètres doivent être recueillis, souvent au moyen des techniques empiriques précédemment citées : informations présentées sur les interfaces, stratégies oculaires (obtenues par des techniques oculométriques), explications subjectives complémentaires (obtenues par verbalisations et questionnaires). Ces paramètres sont ensuite utilisés pour affiner le modèle théorique des tâches selon l'activité réelle des utilisateurs. Cette description permet

d'évaluer les parcours entre les différents écrans et d'avoir une idée précise des caractéristiques de l'espace théorique du problème dans lequel va naviguer l'utilisateur. On peut aussi évaluer les possibilités d'actions offertes à l'utilisateur, ainsi que la lourdeur des transactions en termes de nombre d'états. Ce type d'approche analytique permet de mettre en évidence des lacunes dans l'interaction Homme(s)-Machine(s) et des besoins en fonctions et en informations. Ces modèles sont encore actuellement à l'état de recherche et sont pour la plupart insuffisamment développés ou validés pour constituer des outils d'évaluation à part entière. Leur utilisation nécessite un travail de formalisation assez long et peu d'informations sont données pour l'instant concernant la manière d'effectuer des tests en cours de conception.

Les modèles prédictifs énumérés ci-dessus s'intéressent à l'interaction Homme(s)-Machine(s) au sens des tâches et de leur organisation. Ils sont de plus très importants en tant que support de dialogue entre chaque intervenant du projet. On oublie aussi trop souvent que de tels modèles peuvent en plus être utilisés une fois l'interface réalisée pour modéliser les activités des utilisateurs et comparer cette modélisation à celle effectuée vis-à-vis des tâches prévues par les concepteurs, afin d'en tirer les conséquences en termes d'aménagement des IHMs et/ou de formation [KOLSKI, 1997]. L'ensemble des modèles formels dits prédictifs peut être utilisé pour prédire certains aspects liés à l'interaction Homme(s)-Machine(s) (hiérarchie des tâches, cheminement des actions de l'utilisateur, alternatives, temps requis pour réaliser une tâche,...) et faciliter la mise en évidence de problèmes potentiels. Des métriques sont progressivement mises en place par les chercheurs pour mesurer objectivement des aspects associés à la qualité de l'interaction (cohérence, compatibilité avec l'image mentale de l'utilisateur, qualité des écrans,...).

Tous les modèles de tâches énumérés s'intéressent à l'interaction Homme(s)-Machine(s) au sens des tâches et de leur organisation. On peut donc les utiliser dès les étapes *d'analyse et de spécification*. Notons que les extensions KLM et Petri, associées respectivement aux modèles GOMS et SADT, permettent de formaliser l'activité de l'utilisateur. Ils sont utilisables lors de la conception détaillée, lorsque l'activité de l'utilisateur est définie par rapport à ces fonctions, et la description obtenue peut être analysée afin d'estimer l'utilité et l'utilisabilité de l'interface. Le modèle CLG apporte une description qui fait apparaître différents niveaux de l'interaction (sémantique et syntaxique). Une définition des objets manipulés et des fonctions est en effet réalisée, des modules sont mis en évidence ainsi que les liens entre fonctions et modules. De plus, la syntaxe exacte de ces fonctions est décrite en mettant en évidence l'activité de l'utilisateur. Enfin, seul CLG s'intéresse à la présentation de l'information et aux entités physiques manipulées. Ce modèle linguistique est donc utilisable en principe dès l'étape de *spécification* de l'interface jusqu'à la conception détaillée [GRISLIN, 1995].

Quant aux modèles formels dits de qualité de l'IHM, il faut considérer que la plupart de ces systèmes sont à l'état de prototypes de recherche, même s'ils ont été validés sur des cas concrets. Ils sont censés intervenir sur des pages-écrans réalisées ou en cours de prototypage afin d'apporter une aide à l'évaluateur dans sa démarche globale, en se focalisant sur des critères d'utilisabilité. De tels systèmes devraient être à terme intégrés dans des environnements graphiques de développement.

IV. Synthèse sur les méthodes et techniques usuelles, vers de nouveaux outils d'aide à l'évaluation

Après avoir vu les méthodes et techniques d'évaluation usuelles, nous allons maintenant faire une synthèse. Cette synthèse se rapproche de la classification proposée par WHITEFIELD et ses collègues [1991]. Nous l'avons mise à jour en étudiant les méthodes et techniques vis-à-vis de l'évaluation de systèmes interactifs. La figure II.25 reprend les méthodes et techniques présentées dans ce travail en les localisant dans leur approche respective.

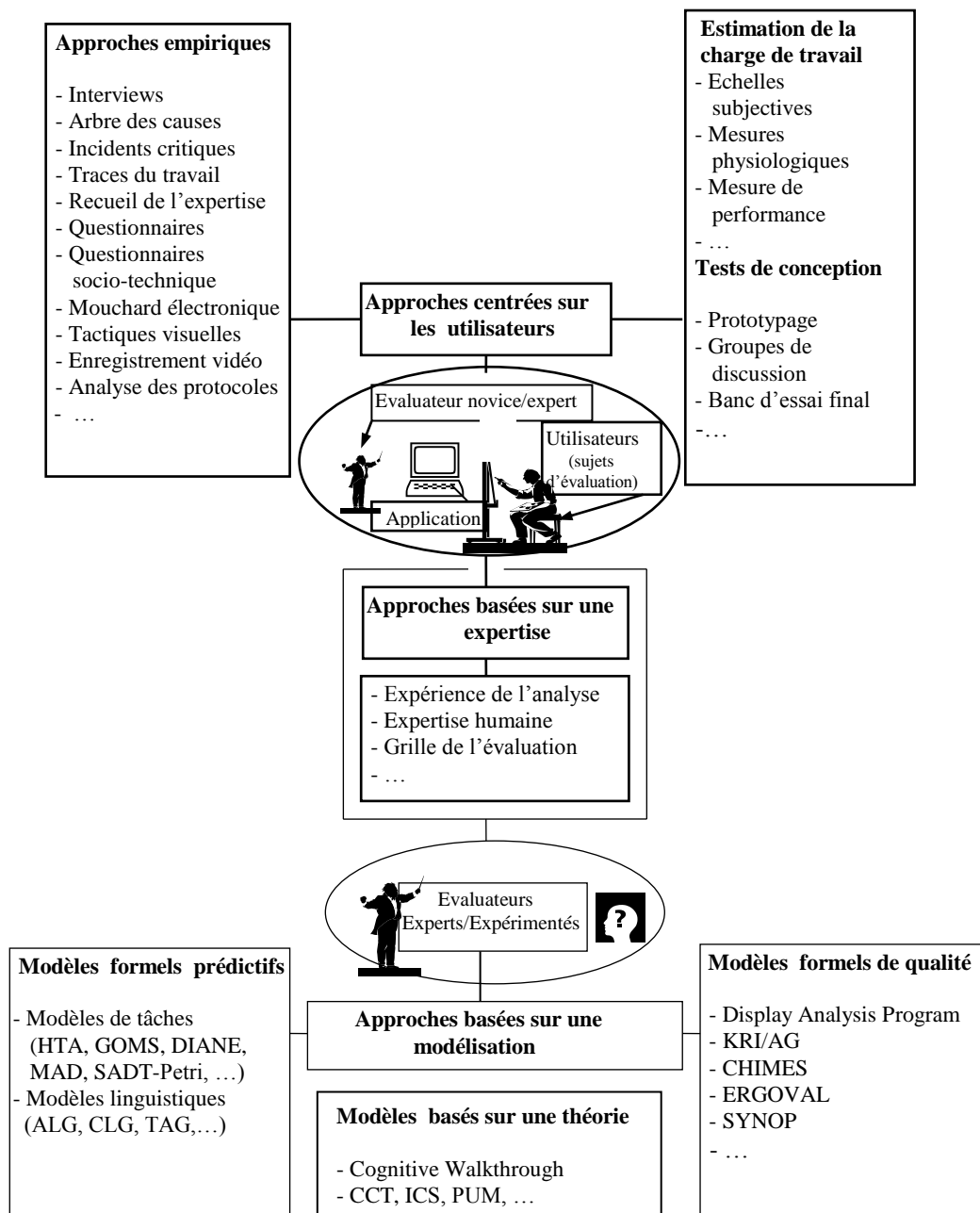


Figure II.25 : Synthèse et classification des méthodes et techniques d'évaluation des IHMs

Les méthodes de la figure II.25 ont été détaillées au cours des pages précédentes en présentant leur principe, leurs avantages et inconvénients principaux et plusieurs références bibliographiques.

Avantages et inconvénients principaux :

Les méthodes et techniques d'évaluation que nous venons d'étudier permettent aux évaluateurs d'effectuer l'évaluation partielle d'une interface Homme(s)-Machine(s) dans la mesure où toute méthode ou technique ne peut vérifier de manière efficace tous les critères d'évaluation, d'où la notion de complémentarité entre ces méthodes pour tenter de prendre en compte le maximum de critères possibles. C'est une limitation par rapport aux nouvelles approches telles que FACE et le Plan de Denley et Long qui proposent plutôt une démarche d'aide à l'évaluation qui va du choix des méthodes pouvant évaluer globalement l'IHM à la démarche et à l'organisation proprement dite d'une évaluation. Ces nouvelles approches constituent réellement une aide à l'évaluation d'une IHM sous tous ces aspects alors que les méthodes que nous venons de voir sont spécifiques à des critères donnés.

V. Vers de nouveaux outils d'aide à l'évaluation

De nouvelles approches d'aide à l'évaluation voient actuellement le jour, apportant une aide aux évaluateurs dans la démarche à suivre et leur organisation. Nous exposons deux outils représentatifs : FACE et le Plan d'aide de Denley et Long pour l'aide à l'évaluation.

V.1 Aides logicielles FACE :

Principe :

Le système FACE (Fast Audit based on Cognitive Ergonomics) de HULZEBOSCH et JAMESON [1993] a été développé sous HyperCard. Il permet à l'évaluateur de cerner la situation d'évaluation, de choisir la technique qui sera appliquée, de suivre son application, etc. Cette pile fournit aussi des check-lists et permet la production de rapports. En d'autres termes, FACE permet à l'évaluateur de préparer l'évaluation, de conduire celle-ci et d'en présenter les résultats.

Lors de la phase de préparation, l'évaluateur est incité, par des questionnaires, à fournir de l'information sur le système devant être évalué, sur les caractéristiques des utilisateurs et sur la structure organisationnelle de l'entreprise. FACE fournit, sur la base des informations relatives au système, des dimensions à prendre en compte lors de l'évaluation. Lorsque l'évaluateur passe à la phase d'évaluation, FACE l'aide à choisir la méthode qui sera en principe la plus adaptée au contexte (voir Tableau II.1), indique les aspects sur lesquels une attention particulière devra être portée, suggère des façons de conduire cette dernière et d'analyser les résultats.

Le tableau II.1 montre un extrait d'une aide (sachant que d'autres possibilités d'aide existent dans FACE) apportée à l'évaluateur pour le choix des méthodes d'évaluation :

FACE prend en compte 4 facteurs situationnels : le nombre d'évaluateurs disponibles, la complexité du problème, la disponibilité des utilisateurs et les limites des différentes fonction-

nalités du système à évaluer. Ces facteurs doivent être mis en évidence par les évaluateurs avant l'évaluation proprement dite et, à l'issue de cette étape, les évaluateurs présentent des « suggestions » de recherche accompagnées des points spécifiques liés aux quatre facteurs situationnels (voir figure II.26).

Techniques d'évaluation	Considérations
Système Walk through	<ul style="list-style-type: none"> • Le système doit être disponible pendant l'évaluation • Le système ne doit pas être très complexe • Appliquer s'il n'y a aucune autre technique orientée sur une expertise • L'introduction d'un autre test basé sur les utilisateurs est nécessaire • Les critères d'utilisabilité doivent être appropriés
Evaluation Heuristique	<ul style="list-style-type: none"> • 3 à 5 évaluateurs devraient être disponibles pendant l'évaluation • L'introduction d'un autre test est nécessaire • Cette technique est appropriée lorsqu'aucune préparation n'est possible • Temps alloué suffisant
Questionnaire et check-list	<ul style="list-style-type: none"> • Présence d'utilisateurs obligatoire • Le questionnaire doit être exhaustif • Les évaluateurs doivent connaître les utilisateurs
« Pensée à haute voix »	<ul style="list-style-type: none"> • Le système doit être disponible durant l'évaluation • Les utilisateurs doivent être suffisamment disponibles durant l'évaluation • Nécessité du matériel audio • Nécessité d'une technique complémentaire

Tableau II.1 : Aide au choix de techniques d'évaluation dans FACE (extrait)

Avantages et inconvénients principaux :

FACE ne fait donc pas d'évaluation en tant que telle. Cependant, cet outil aide l'évaluateur à sélectionner la technique d'évaluation en fonction des informations fournies, à mettre en place l'évaluation et à présenter les résultats de l'évaluation. Il est important de signaler que FACE est orienté vers une évaluation industrielle et n'aide pas au choix de méthodes analytiques. De plus, l'aide au choix des techniques d'évaluation proposée par FACE n'est pas automatique, ni adaptée aux évaluateurs non expérimentés et certaines informations et recommandations proposées sont vagues (cf. tableau II.1).

Référence :

[HULZEBOSCH et JAMESON, 1993].

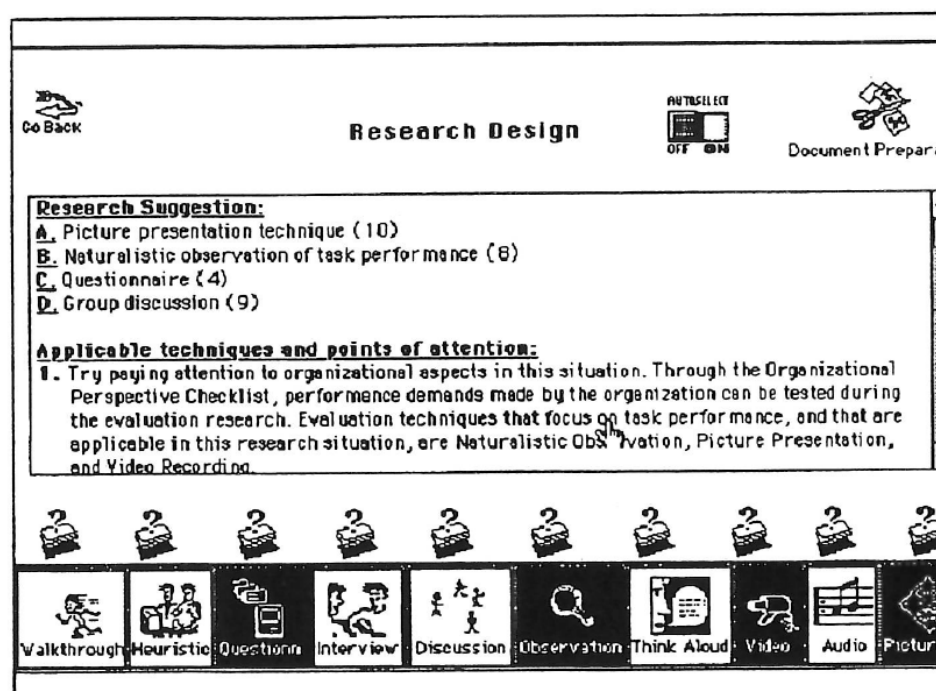


Figure II.26 : Décomposition de la boîte de dialogue [BROOKE J., 1996]

V.2. Plan d'aide pour l'évaluation (Denley et Long)

Considéré comme une aide à l'évaluation, ce support, réalisé par DENLEY et LONG [1997], est surtout destiné aux évaluateurs novices en facteurs humains. L'outil d'aide fournit une structure et une démarche pour le choix des techniques en évaluation des Systèmes Interactifs en termes de facteurs humains. Il offre également des informations relatives à la planification de l'évaluation des données sous forme heuristique à propos de l'utilisation des méthodes et de leur sélection.

Principe :

Dans un premier temps, l'évaluateur interagit avec un outil d'aide pour produire un plan d'évaluation. Dans un second temps, l'évaluateur va interagir avec le plan produit pour rendre performante l'évaluation d'un Système Interactif. Dans ce contexte, la performance est définie comme une fonction à deux variables : la *qualité du produit* et les *coûts induits* (ex : coûts des ressources nécessaires pour l'évaluation). On peut faire une distinction entre *performance réalisée* (niveau de performance atteint) et *performance désirée* (niveau de performance à atteindre ou recherché).

Les trois phases de l'outil d'aide

L'outil d'aide à l'évaluation conçu par DENLEY et LONG [1990] est composé de trois phases et chaque phase se subdivise en niveaux et les niveaux en étapes. Les étapes sont constituées d'activités :

Phase 1 : Elaboration des indicateurs de performance.

Durant cette phase, on identifiera les facteurs de performance et l'organisation nécessaire pour leur mise en œuvre. Pour y parvenir, on utilisera des guides de recommandations et autres conseils. Cette démarche permettra d'établir ce que l'on qualifiera de *performance désirée* (Pd) ou recherchée (voir figure II.27).

Niveau 1.1 : Identification des facteurs de performance

Niveau 1.2 : Identification des facteurs organisationnels

Niveau 1.3 : Spécification de la performance désirée en évaluation

Tâches :

- Identification matérielle et logicielle
- Identification des utilisateurs
- Identification des tâches
- Identification des facteurs d'environnement
- Identification de toute information incluant même celles issues du marketing

Techniques :

- Interviews
- Questionnaires
- Check-lists
- Documents de recommandations
- etc.

Phase 2 : Sélection des méthodes

Cette phase est caractérisée par le choix des méthodes d'évaluation et ceci par les évaluateurs. Ce choix se fait à partir d'une classification des méthodes d'évaluation. L'évaluateur se servira, dans cette optique, de guides d'aide conçus à cet effet.

Niveau 2.1 : Sélection des classes de méthodes

Niveau 2.2 : Spécification des classes sélectionnées

Niveau 2.3 : Spécification des méthodes sélectionnées

Phase 3 : Configuration des méthodes

Dans cette phase, chaque méthode sélectionnée est configurée grâce aux guides, et ceci afin d'établir d'autres besoins nécessaires pour le plan d'évaluation.

Niveau 3.1 : Création du plan préliminaire d'évaluation

Niveau 3.2 : Spécification du plan préliminaire d'évaluation.

V.3 Discussions sur FACE et l'outil d'aide de Long et Denley

De ces outils d'aide pour évaluer les systèmes interactifs, nous dirons qu'ils présentent beaucoup d'informations enrichissantes que tout évaluateur aurait intérêt à prendre en compte. Ces informations vont de l'organisation générale des évaluateurs à celle des clients désirant

acquérir ces systèmes interactifs. C'est une aide non négligeable à tout évaluateur novice en facteurs humains.

Les faiblesses de ces outils résident à plusieurs niveaux : tout d'abord, on notera la difficulté à prendre en compte l'organisation des évaluateurs, qui peut être différente d'un évaluateur à l'autre. Une quantité énorme d'informations parfois redondantes ne facilite pas l'exploitation de ces dernières. Les phases de la démarche d'aide ont aussi un maillon faible : le choix de la méthode d'évaluation par l'évaluateur à partir d'une classification des méthodes. En effet, ces démarches, qui relèvent plutôt des problèmes de planification que des problèmes de l'évaluation proprement dits, possèdent dans leurs structures une classification des méthodes qui ne tient compte que d'un ensemble de paramètres très limité, tels que la qualité du produit et les coûts induits.

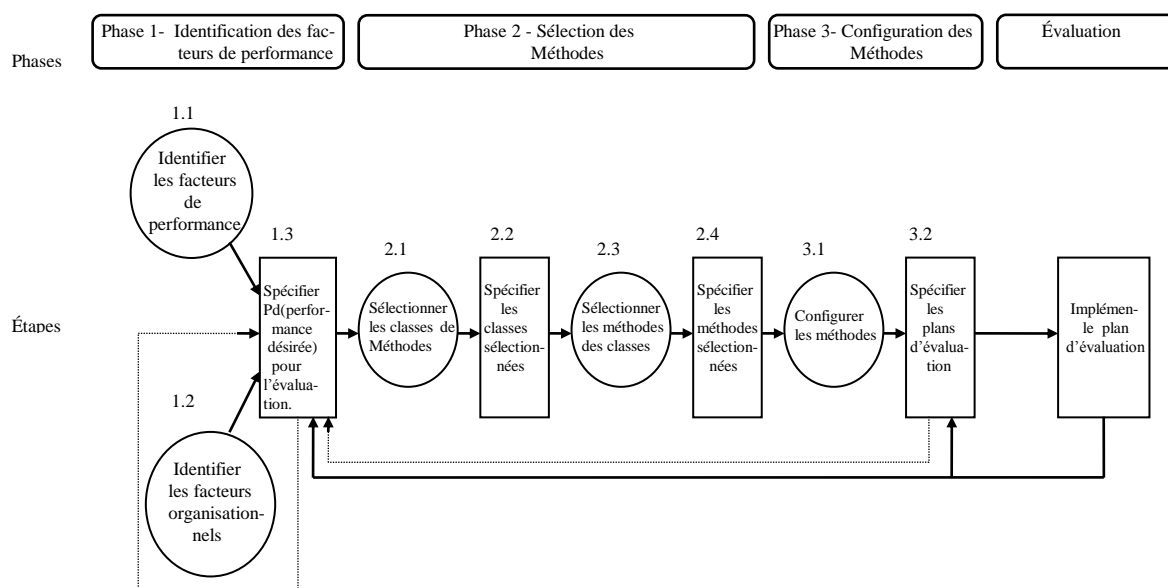


Figure II.27 : Plan d'aide à l'évaluation des IHMs [LONG et DENLEY, 1997]

En fait, l'évaluateur inexpérimenté en facteurs humains ne peut choisir la méthode appropriée à partir uniquement d'une classification des méthodes car dans chaque classe, on peut trouver près de 20 méthodes d'évaluation. Dans l'une des classes de méthodes telle que celle nécessitant la présence des utilisateurs, on peut par exemple comptabiliser un grand nombre de méthodes d'évaluation. Cette phase de sélection de la méthode d'évaluation est l'une des principales phases de ces démarches. Or, cette phase présente des faiblesses majeures, ce qui nous permet d'émettre des réserves quant à la portée de l'aide à apporter aux évaluateurs non expérimentés.

Le choix de méthodes ou techniques appropriées est l'un des problèmes majeurs en évaluation car le résultat final en dépend ; à cet effet, toute aide en évaluation qui laisserait le choix de la méthode ou technique d'évaluation au sort de l'évaluateur passerait à côté de la tâche la plus complexe en évaluation en dehors de l'analyse et du traitement des résultats. Or, les

démarches dont il est ici question ne fournissent pas de méthode ou technique appropriée à l'évaluateur ; elles ne donnent pas non plus d'information sur les techniques de traitement des données recueillies. Ces démarches relèvent plutôt des problèmes de l'organisation nécessaire pour l'évaluation et non des problèmes techniques proprement dits. Pour le cas de DENLEY et LONG, il existe toutefois une phase de configuration des méthodes après la sélection de la méthode par le novice, mais cette configuration aura une portée limitée dans la mesure où la sélection de la méthode faite au départ n'est pas fiable, c'est-à-dire que la méthode choisie ne permet pas d'atteindre les objectifs recherchés.

La sélection d'une méthode d'évaluation doit dépendre des facteurs situationnels qui sont en fait les contraintes qui se présentent aux évaluateurs, tels le budget disponible, les moyens humains, le type d'application, le lieu d'évaluation, etc. Or, il n'en est pas question ici, ce qui peut nous laisser sceptique quant au résultat final. L'un des objectifs de nos travaux est de pouvoir pallier ce genre de faiblesse en développant un outil d'aide qui fournirait automatiquement la méthode ou technique d'évaluation appropriée pour toute application. Notre démarche sera sans doute un complément aux travaux de HULZEBOSCH et JAMESON [1993] et DENLEY et LONG [1990] dans la mesure où elle permettra de renforcer le maillon faible (sélection de la méthode par l'évaluateur) de la chaîne d'aide à l'évaluation.

Conclusion

Actuellement, les méthodes utilisées sont principalement *empiriques*, basées sur la mesure et/ou l'observation de données représentatives des actions des utilisateurs, et *expertes*. Les méthodes basées sur une modélisation sont plus utilisées en laboratoire qu'en entreprise pour plusieurs raisons : le coût de revient, le temps de mise en œuvre, les connaissances requises, la fiabilité des résultats, etc. Le point commun de toutes ces méthodes et techniques d'évaluation demeure la problématique du choix de la méthode appropriée pour l'évaluation d'un système donné. L'objectif de nos travaux est justement d'étudier et de mettre au point la première version d'un système d'aide à l'évaluation qui permet de fournir automatiquement aux évaluateurs les méthodes appropriées au système qu'ils veulent évaluer. Avant de présenter les contours de notre système d'aide, nous allons d'abord expliquer, dans le chapitre 3 suivant, les différentes approches d'aide à la décision existantes.

Chapitre III : Principes pour une aide à la décision en évaluation

Introduction

- I. Principes et structure d'un système d'aide à la décision**
- II. Approches possibles en aide à la décision et justification de l'approche suivie**
- III. Démarche globale basée sur les ensembles flous**

Conclusion

Introduction

Ce chapitre a pour objectif d'expliquer les principes choisis pour l'aide à la décision pour le choix automatique des méthodes et techniques d'évaluation nécessaires aux évaluateurs dans leur processus d'évaluation. Nous allons pour cela d'abord étudier différents concepts de l'aide à la décision, puis nous dresserons un bilan des différentes méthodes d'aide à la décision existantes.

L'aide à la décision est l'activité durant laquelle on s'appuie sur des modèles clairement explicités, mais non nécessairement complètement formalisés, pour aider à obtenir des réponses aux questions que se pose un *intervenant* dans le processus de décision. Selon ROY [1985], les éléments de réponses visent à éclairer la décision et à recommander, ou simplement à favoriser, un comportement de nature à accroître la *cohérence* entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels cet *intervenant* se trouve placé d'autre part.

La décision est souvent présentée comme le fait d'un individu isolé (le *décideur*) qui choisit librement entre plusieurs *actions* à un moment donné dans le temps. Bien que la responsabilité d'une décision relève du ressort d'un individu clairement identifié, celle-ci est souvent la conséquence d'une interaction entre de multiples *acteurs* au cours d'un *processus de décision*. On distingue plusieurs types d'acteurs : les *intervenants*, les *administrés*, les *contribuables*, etc. Pendant leur intervention, les intervenants conditionnent directement la décision en fonction du système de valeurs dont ils sont porteurs, alors que les autres (administrés, contribuables, etc.) se distinguent par le fait qu'ils subissent de façon passive les conséquences de la décision, laquelle est seulement censée tenir compte de leurs préférences. SFEZ [1992] appelle cette catégorie d'acteurs les *agis*.

L'intervenant le plus important dans le processus de décision est le *décideur*. C'est celui que les modèles mis en œuvre cherchent à éclairer. Le décideur ne doit pas être confondu avec le *demandeur d'étude* au sens de ROY [1985]. Dans le même temps, nous devons distinguer le *décideur* de l'*homme d'étude*. En effet, l'*homme d'étude* est celui qui prend en charge l'aide à la décision en mettant en œuvre des modèles dans le processus de décision ; il contribue à l'orienter et à le transformer. Cet *homme d'étude* va être un *acteur* particulier du processus de décision qu'il analyse et influence par sa présence et son action, par la « neutralité », « l'objectivité » et la « validité » de son travail.

Sans remettre en cause toute pertinence à cette conception de la décision, il faut admettre qu'elle ne rend que très imparfaitement compte de la façon dont sont prises les décisions dans la plupart des organisations. Dans cette optique, *décider* ou, de façon plus large, *intervenir* dans un processus de décision n'est qu'exceptionnellement trouver une solution à un problème. Cela revient parfois à imaginer des compromis ou à faire accepter un arbitrage dans une situation de conflit [ZOPOUNIDIS et DIMITRAS, 1998].

L'aide à la décision est une recherche partielle de la vérité. Les méthodologies, les concepts, les modèles, les techniques et les théories sur lesquels elle s'appuie ont le plus souvent un enjeu différent : raisonner sur le changement que prépare un processus de décision de façon à accroître sa cohérence avec les objectifs et le système de valeurs de celui au nom de qui l'aide à la décision s'exerce. Dans cette optique, ROY et SKALKA [1984] pensent que l'aide à la

décision s'éloigne de la *Recherche Opérationnelle* classique qui cherche à s'appuyer sur un modèle de la réalité pour optimiser un critère de façon objective. Ainsi, les problèmes de décision dans ce cadre doivent se réduire à un problème monocritère (un critère de type profit ou rendement est généralement choisi). Or, cela s'avère rapidement limité dès que l'on veut s'intéresser à plusieurs aspects simultanément. A l'opposé des problèmes monocritères, ceux à plusieurs critères sont caractérisés par le fait que la solution optimale n'existe plus. Dans cette optique, SCHÄRLIG [1985] considère que l'objet du problème n'est plus la recherche de cette solution mais le suivi d'une stratégie. De plus, selon ROY [1985], l'aide à la décision contribue à construire, à asseoir et à faire partager des convictions. ROY met en avant l'idée de subjectivité présente dans les théories de la Décision, distinguant définitivement ces théories de la Recherche Opérationnelle.

Selon BOUYSSOU [1993], la méthodologie multicritère d'aide à la décision est caractérisée par quatre niveaux : (1) *objet de la décision et esprit de recommandation*, (2) *analyse des conséquences et élaboration des critères*, (3) *modélisation des préférences et approches opérationnelles pour l'agrégation des performances*, (4) *procédures d'investigation et élaboration de la recommandation*. Les méthodes d'aide à la décision trouvent leur intérêt dans de nombreuses applications recensées notamment dans [ZOPOUNIDIS et DIMITRAS, 1998]. Pour la suite du mémoire, nous considérerons que le *décideur* est l'utilisateur du système d'aide à la décision. Il peut être un homme ou un groupe de responsables qui s'appuie sur la méthode pour prendre une décision face à un problème et doit apporter les informations nécessaires au fonctionnement de la méthode. Dans notre cas, rappelons que l'utilisateur peut être un chef de projet qui s'interroge sur les méthodes et critères d'évaluation disponibles au regard de ses propres critères, un évaluateur, voire un apprenant en interaction Homme(s) – Machine(s).

Dans ce chapitre, la structure que doit adopter un système d'aide à la décision pour pouvoir répondre à ses objectifs sera d'abord étudiée. Dans la seconde partie, nous ferons une synthèse des différentes approches d'aide à la décision en justifiant l'approche choisie pour nos travaux. Dans la troisième, nous expliquerons la démarche suivie par le système développé dans le cadre de cette thèse, ainsi que la manière dont les données et les connaissances ont été modélisées.

I. Principe et structure d'un système d'aide à la décision

L'aide à la décision est à la fois simple et complexe, dans la mesure où elle met en scène à la fois de nombreux acteurs, méthodes et procédures. Sa structure est caractérisée par la problématique qui reflète le type de problème à traiter, par l'identification des critères de décision, par la définition des alternatives ou actions, par la modélisation des préférences et par le choix d'une méthode d'aide adaptée au type de problème à résoudre (voir figure III.1).

I.1. Problématique

La problématique est la façon dont le problème de décision est posé. La figure III.1 présente les quatre problématiques de référence décrites par ROY [1985]. Une fois qu'une problématique

relative au type de problème à traiter est choisie, cela permet de déduire l'approche d'aide à la décision correspondante.

Quatre problématiques représentatives ont été identifiées : (1) la problématique $P.\alpha$ caractérise le choix d'un sous-ensemble contenant les « meilleures » actions (optimums) ou, à défaut, des actions « satisfaisantes » ; (2) la problématique $P.\beta$ éclaire la décision par un tri résultant d'une affectation de chaque action à une catégorie, les catégories étant définies *a priori* en fonction de normes ayant trait à la suite à donner aux actions qu'elles sont destinées à recevoir ; (3) la problématique $P.\gamma$ s'intéresse au problème de rangement de classes d'équivalence composées d'actions, ces classes étant ordonnées de façon complète ou partielle et, enfin, (4) la problématique $P.\delta$ qui concerne une description, dans un langage approprié, des actions et de leurs conséquences.

Comme nous le verrons dans la suite du texte, nos travaux correspondent à la problématique de type $P.\alpha$. Une fois la problématique identifiée, on déterminera et on évaluera des critères de décision. Par là, le décideur aura à introduire ses choix de stratégies et ses préférences. Ces informations seront triées, traitées et agrégées par la méthode d'aide à la décision pour en déduire le choix de la meilleure alternative. Notons que les autres éléments présents dans la figure III.1 seront décrits par la suite.

Dans la figure III.1, les exemples présentés ont été choisis parmi d'autres pour illustrer chaque approche car ils sont les plus usuels et pour certains (ELECTRE par exemple), ils ont fait l'objet de nombreux travaux tant sur le plan de la recherche que sur le plan industriel.

I.2. Alternatives

Après avoir identifié le type de problématique, *l'homme d'étude* va dans un premier temps définir des actions potentielles que le décideur peut envisager. Cet ensemble est désigné par $A = \{a_1, a_2, \dots, a_n\}$ où les a_i sont les actions potentielles. Les actions potentielles sont celles qui sont provisoirement jugées possibles par l'un des intervenants au moins ou présumées telles par l'homme d'étude en vue de l'aide à la décision [ROY, 1985]. En pratique, il est opportun de procéder déjà, à ce stade, à un certain tri, en ne constituant l'ensemble A que par des *actions efficaces*. Néanmoins, il n'est pas toujours facile de dégager *a priori* ces actions. De plus, il faut définir la liste des actions de la façon la plus précise et la plus exhaustive possible afin de faire évoluer le système suivant tout son potentiel.

Rappelons qu'une *action efficace* est une action potentielle qui n'est dominée par aucune autre action de A . Selon MAYSTRE *et al.* [1994], une action a_i domine une action a_k si et seulement si $g_l(a_i) \geq g_l(a_k)$ pour un critère l appartenant à l'ensemble des critères $\{1, 2, \dots, n\}$; l'inégalité est stricte pour un critère l au moins, où $g_l(a_i)$ est l'évaluation de l'action a_i selon le critère l . Nous n'allons pas nous attarder sur le concept d'actions car seules celles concernant la problématique $P.\alpha$ nous intéressent ici.

Rappelons toutefois que pour les autres problématiques, comme la problématique $P.\beta$, on parlera d'*action de référence* qui est une action servant de référence par rapport à laquelle les *actions potentielles* sont examinées. Ces types d'actions servent de limites à des catégories auxquelles les actions potentielles sont affectées.

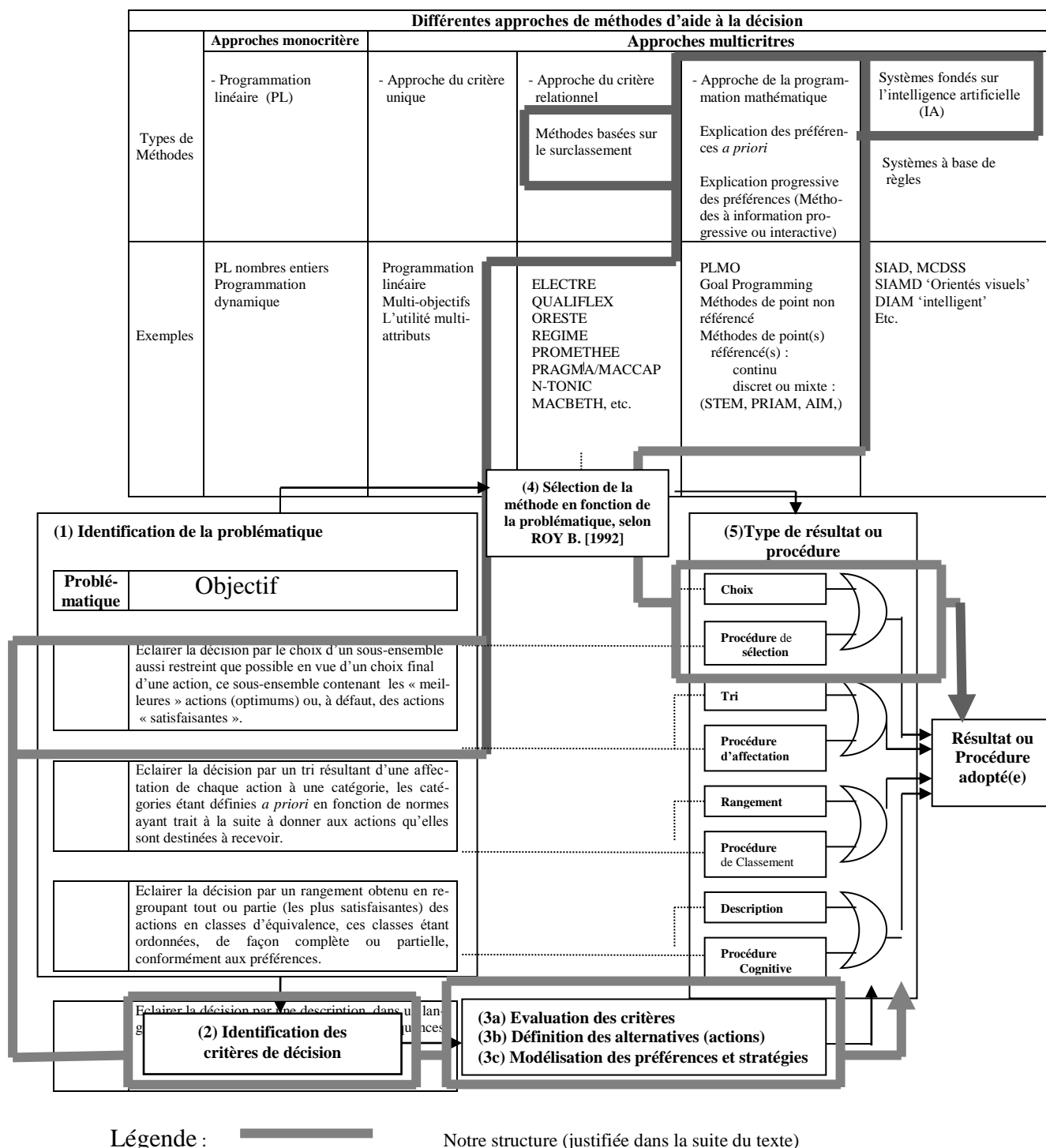


Figure III.1 : Principe et structure de l'aide à la décision (inspiré de [ROY, 1992])

Après avoir déterminé l'ensemble des actions, le choix d'une approche d'aide à la décision peut être envisagé. Si l'ensemble des actions ou alternatives est clairement déterminé, nous avons affaire à une aide à la décision multi-attributs. Par contre, si l'ensemble des alternatives peut évoluer au cours du temps, l'aide à la décision est qualifiée de multi-objectifs.

I.3. Les critères et leur évaluation

Un critère est une expression qualitative ou quantitative de points de vue, objectifs, aptitudes ou contraintes relative au contexte réel, permettant de juger des personnes, des objets ou des événements [MAYSTRE *et al.*, 1994]. Pour qu'une expression soit considérée comme un critère, elle doit être utile pour le problème considéré, fiable et dotée d'une structure de préférence. A cet effet, MAYSTRE *et al.* [1994] donnent un exemple en disant que : « si le problème est le choix parmi plusieurs candidats pour le poste d'ingénieur de recherche, la taille des candidats ne constitue pas un critère (absence d'utilité) ». Par contre, si le problème est le choix de méthodes ou techniques d'évaluation des systèmes interactifs (objet de notre étude), comme nous le verrons par la suite, le « temps alloué à l'évaluation » va être un critère de décision valide (forte utilité).

L'établissement des critères n'est pas chose facile. La difficulté réside donc dans un compromis à trouver entre la prise en compte de tous les aspects du problème de décision et la dépendance entre les critères sélectionnés. Après avoir déterminé l'ensemble des actions, l'homme d'étude va analyser, en étroite collaboration avec le décideur, les conséquences que telle ou telle décision (réalisation de telle ou telle action potentielle) pourrait avoir. Chaque action est jugée selon chaque critère. Ce jugement est appelé évaluation de l'action a_i et noté $g_j(a_i)$ (se lit évaluation de l'action a_i selon le critère j). L'ensemble des évaluations est appelé matrice des évaluations (ou tableau des performances). L'évaluation des critères se réalise sous deux angles dont le premier fournit la valeur réelle des critères et le second exprime les préférences du décideur.

I.4. Stratégies et préférences

Aider à la décision suppose de prendre en compte les valeurs et les préférences d'un ou de plusieurs acteurs dans un processus de décision. Dans cette optique, on admet généralement quatre situations fondamentales incompatibles d'indifférence, de préférence stricte, de préférence faible et d'incompatibilité qui conviennent pour définir une représentation réaliste des préférences d'un décideur quelles que soient les actions potentielles en cause, le point de vue adopté pour les comparer et l'information disponible, le décideur étant chargé de les juger.

En ce qui concerne la préférence faible, l'affirmation « a_i faiblement préférée à a_k » traduit le fait que l'homme d'étude qui juge en son nom s'estime en mesure d'affirmer que « a_k n'est pas préférée à a_i » tout en hésitant entre « a_i est strictement préférée à a_k » et « a_i est indifférente à a_k », les éléments dont il dispose pour trancher entre ces deux dernières possibilités ne lui paraissant pas suffisamment probants. Selon ZOPOUNIDIS et DIMITRAS [1998], l'originalité des méthodes d'aide à la décision transparait particulièrement à ce niveau avec l'introduction dans la méthode des aspects entièrement subjectifs que sont les choix stratégiques et les préférences. Ces considérations doivent faire évoluer le système dans le sens voulu par le décideur. L'une des difficultés lorsqu'il s'agit de résoudre un problème réside dans le fait que le nombre élevé des décideurs ne facilite pas l'agrégation de leurs préférences.

Dans certains cas, le tableau d'évaluation des critères est fourni par un seul décideur qui exprime alors ses préférences correspondant à la situation donnée de façon implicite [ROY,

1985] ; c'est la démarche que nous avons adoptée pour des raisons de simplicité. En effet, comme le mentionne ROY [1985], trop de décideurs compliquent le problème puisqu'il va falloir agréger des préférences parfois très différentes. Rappelons que la valeur d'un critère n'est pas la valeur réelle de ce critère mais l'évaluation ressentie par le décideur pour la situation actuelle.

Dans cette optique, le décideur n'a pas besoin d'avoir une appréciation globale de ses préférences, mais une bonne connaissance de ses préférences vis-à-vis de la situation actuelle. Dans certaines mesures, le décideur peut exprimer ses préférences de manière globale et doit, par conséquent, avant toute utilisation de la méthode d'aide à la décision, fournir ses préférences pour toute situation pouvant être rencontrée. Dans ce cas, une analyse préliminaire permet de comparer les préférences et les valeurs réelles des critères afin d'obtenir un tableau d'évaluation des critères. Le choix final est opéré à l'aide des opérateurs mathématiques présents dans les méthodes selon les différentes stratégies adoptées.

I.5. Les types de résultats

Le résultat est défini ici comme étant le type de résolution effectué par la méthode et correspond à une problématique donnée. En effet, la problématique $P.\alpha$ traite les problèmes de *choix* et de *procédure de sélection*, la problématique $P.\beta$ traite des problèmes de *tri* et de *procédure d'affectation*, alors que la problématique $P.\gamma$ s'intéresse au problème de *rangement* et de *classement* et, enfin, la problématique $P.\delta$ traite des problèmes de description et de procédure cognitive. Par exemple, si nous considérons la problématique $P.\alpha$, nous constatons que, suivant la méthode d'aide considérée, le *choix* sur les actions varie de l'unique action sélectionnée à un classement plus ou moins complet dans l'ordre de préférence. Dans ce cas, les actions doivent être présentées au décideur de manière ergonomique, par l'intermédiaire de l'interface Homme(s) – Machine(s).

I.6. Conclusion

Après avoir mis en évidence la structure de l'aide à la décision, nous pouvons dire que notre problématique est de type $P.\alpha$. En effet, notre système est fondé sur le fait qu'il éclaire la décision par le choix d'un ensemble d'actions optimales en vue du choix final d'une action. De plus, notre problème est caractérisé par le fait qu'il intègre beaucoup de critères de décision. A cet effet, la section suivante nous permettra de donner une justification de l'approche choisie.

II. Approches possibles en aide à la décision et justification de l'approche suivie

Pour tous les modèles d'évaluation des actions, les préférences d'un décideur impliqué dans un processus de décision ne sont pas toujours rigoureusement définies *a priori* et stables. L'appréciation de certaines des conséquences peut être hypothétique et décrite en termes de probabilités. Dans cette section, nous verrons les différentes approches d'aide à la décision monocritère et multicritère selon un classement qui respecte les différentes écoles. Concernant le jugement de préférence, on peut dire qu'il varie selon l'importance plus ou moins grande que le décideur accorde à tel ou tel aspect des conséquences. Dans cette optique, ce jugement apparaît

donc fréquemment comme résultant d'aspects conflictuels qui s'affrontent, tant dans l'esprit d'un décideur donné qu'entre acteurs différents au sein du processus de décision. Selon JACQUET-LAGREZE [1981], aider à décider, c'est, en tout premier lieu, aider à clarifier la formation, la transformation et l'argumentation des préférences. A ce niveau, le concept clé est celui du *critère*.

Les méthodes d'aide à la décision existantes sont nombreuses et variées. Nous allons brosser un portrait de ces principales méthodes classées selon les différentes « écoles » en nous inspirant de GANDIBLEUX [1995] et ROY [1985]

II.1. Analyse monocritère

Lorsque le spectre des conséquences n'est pas trop complexe, on peut vouloir expliciter d'emblée un critère unique appréhendant toutes les conséquences pertinentes. Une telle analyse *monocritère* suppose que l'on puisse valoriser chaque action potentielle sur un axe de signification unique choisi *a priori*. Ce dernier peut avoir un sens plus ou moins concret : bénéfique, taux de rentabilité, gain pour la collectivité, utilité pour le décideur. Ce premier mode d'analyse impose en particulier de chiffrer chacune des conséquences dans une unité commune, propre à l'axe de signification choisi. Trois écueils sont alors à éviter [MAYSTRE *et al.*, 1994] :

- exclure du chiffrage (plus ou moins consciemment) les aspects des conséquences qui se laissent mal appréhender dans un tel système de représentation ;
- recourir à des prix de référence, des taux d'équivalence, des barèmes de convention, etc., à la fois nécessaires pour ramener des conséquences hétérogènes à une conséquence commune mais difficiles à évaluer de façon objective ;
- aboutir à une formule complexe, fondée sur une logique obscure et peu propice à la communication.

Comme le rappelle GANDIBLEUX [1995], l'option normative met en évidence un problème d'optimisation qui se définit de la manière suivante :

- un univers Ap d'actions potentielles,
- une fonction objectif $g = f(a)$, $a \in Ap$,
- une direction Opt à optimiser (maximisation ou minimisation).

Le problème se représente habituellement par :

$$\begin{array}{l} \text{Optimiser} \quad g = f(a) \\ \text{sous contraintes} \quad a \in Ap \end{array}$$

Dans l'optimisation en programmation linéaire (PL), la fonction objectif et les contraintes sont linéaires en x :

$$\begin{array}{l} \text{Opt} \quad g = kx, \text{ avec } k \in \mathbb{R}^n \\ \text{Sous contraintes } Ax \leq m, \text{ où } x \in \mathbb{R}^n, m \in \mathbb{R}^m \text{ et } A \in \mathbb{R}^{m \times n} \end{array}$$

Tous les modèles développés dans le cadre des méthodes quantitatives de gestion, sous l'angle de la recherche opérationnelle, étaient souvent issus du calcul économique. Ces modèles

prenaient toujours en compte un *critère unique*, censé représenter la direction optimale dans laquelle le système étudié doit évoluer pour déterminer les meilleures *décisions* [ROY, 1992]. Un ensemble de techniques d'optimisation a conclu ces travaux et s'est distingué pour des problèmes qui peuvent être isolés du processus de gestion (comme des problèmes techniques, tels la recherche d'un mélange ou d'une découpe optimale, la recherche du plus court chemin...).

Opter pour une optimisation, c'est se situer implicitement dans une démarche à *critère unique* qui consiste ainsi à isoler l'économie (si c'est cet aspect qui est étudié par exemple) de tout son contexte. Et s'il est parfois commode d'un point de vue théorique d'adopter cette attitude, il ne faut pas oublier qu'il s'agit d'une simplification permettant d'élaborer des théories pour traiter les problèmes relativement limités. Les techniques d'optimisation sont dès lors confinées à une classe bien définie de problèmes.

Le motif de l'échec de l'optimisation sur les décisions de gestion qui ne peuvent pas être isolées de leur contexte (notamment tous les problèmes qui mettent en jeu des relations entre hommes, ou entre des hommes et leur environnement) réside dans le fait que la réalité est *multicritère* [SCHÄRLIG, 1985]. C'est pourquoi, dans le souci de mieux représenter la réalité complexe de la plupart des problèmes de décision, une autre façon de penser et de concevoir la modélisation est apparue : *l'approche multicritère des problèmes de décision*.

Ce nouveau paradigme consiste à accepter le fait que plusieurs critères, généralement non commensurables et bien souvent conflictuels, peuvent coexister au sein d'une modélisation. Adopter la vision multicritère en matière de décision, c'est avant tout prendre ses distances vis-à-vis de l'optimisation : c'est quitter la recherche opérationnelle classique pour rejoindre *l'aide à la décision* [SCHÄRLIG, 1985]. Ce courant de pensée est apparu il y a une vingtaine d'années. Il se base sur une **conception prescriptive** de l'aide à la décision.

II.2. L'aide multicritère

Lorsque l'hétérogénéité des conséquences est telle que les trois écueils précédents ne peuvent être évités, il est préférable de procéder à une analyse *multicritère* [ROY, 1985]. Celle-ci consiste à expliciter une famille de plusieurs critères, chacun d'eux n'appréhendant qu'une catégorie relativement homogène de conséquences. Chacun de ces critères doit évidemment éviter les écueils cités ci-dessus. La famille qu'ils constituent doit d'autre part satisfaire à un certain nombre d'exigences techniques. Elle doit enfin être intelligible pour les différents acteurs que l'aide à la décision concerne et acceptée par eux comme base de travail.

Dans une étape ultérieure de la modélisation, les n critères d'une famille peuvent être agrégés en un seul plus complexe. Il est important de ne pas confondre cette dernière manière de procéder avec l'analyse *monocritère* qui mélange et amalgame très tôt dans l'analyse, ce qui est de l'ordre des conséquences reconnues par tous avec ce qui relève de leur traitement en vue de les réduire à une unité commune par une procédure mathématique, inévitablement emprunte d'arbitraire et généralement influencée par un système de valeurs particulier.

II.2.1. Approches opérationnelles pour l'agrégation des performances : exemples représentatifs d'approches multicritères conduisant à un critère unique de synthèse

On note un accroissement des travaux de l'aide multicritère depuis le début des années 1980, notamment avec ACHILLES *et al.* [1979], ZELENY [1982], SISKOS et HUBERT [1983] et VINCKE [1989]. On a souvent trouvé l'explication de ce développement en faisant remarquer que la réalité elle-même était *multicritère* et que toute décision impliquait de peser 'le pour et le contre' [ZOPOUNIDIS et DIMITRAS, 1998]. SHÄRLIG [1985] estime que le *monocritère* n'apparaît que comme un cas limite et dégénéré du *multicritère*. Cette section est consacrée aux approches opérationnelles d'agrégation des performances dont les plus représentatives sont les Procédures d'Agrégation MultiCritère (PAMC).

Les PAMC visent à prendre appui sur une règle (ou procédé) apportant une réponse synthétique exhaustive au problème de l'agrégation des performances. Ainsi, un PAMC permet d'établir, sur la base d'un tableau des performances et d'informations inter-critères, un ou plusieurs systèmes relationnels de préférences sur l'ensemble A des actions. Une règle prend la forme d'un critère unique de synthèse agrégeant les n critères de la famille par le biais d'une fonction d'agrégation V en posant :

$$g(a) = V(g_1(a), g_2(a), \dots, g_n(a)).$$

Selon ROY [1992], les PAMC se différencient par les propriétés et le pouvoir discriminant qu'elles confèrent à ce critère de synthèse. Il est naturel et courant de fonder l'aide à la décision sur un modèle de préférences issu d'un critère g défini sur l'ensemble des actions A. Lorsqu'une analyse préalable a montré l'intérêt que pouvait présenter la prise en compte d'une famille de critères, ceux-ci peuvent être synthétisés en un seul par ce que ROY et BOUYSSOU [1993] appellent une *procédure d'agrégation multicritère*. Les PAMC que nous allons voir dans la suite de ces travaux visent à bâtir sur A une fonction g à valeurs réelles de telle sorte que g puisse être considéré comme un critère, et donc en particulier que :

$$g(a) \geq g(b) \Rightarrow a S b, \text{ avec } a, b \in A,$$

où S est une relation de surclassement globale prenant en compte les n critères de la famille.

II.2.1.1. Critère unique de synthèse : exemple représentatif de cas où l'on veut bâtir un vrai-critère¹

La plupart des PAMC visant à bâtir un critère de synthèse cherchent à en faire un vrai-critère. Cette section est réservée à ces PAMC. Elles conduisent à une modélisation des préférences globales telle que :

¹Un vrai-critère est un critère de type fonction-critère (on parle également de pseudo-critère) : c'est une fonction g à laquelle se trouvent associées deux fonctions seuils $q_g(g(a))$ et $p_g(g(a))$ vérifiant : $\forall a, b \in A, [q_g(g(a)) - p_g(g(a))] / [g(a) - g(b)] \geq -1$ [ROY et SKALKA, 1984].

$$\begin{aligned}
g(a) &= V(g_1(a), g_2(a), \dots, g_n(a)). \\
a P b &\Leftrightarrow g(a) > g(b) \\
a I b &\Leftrightarrow g(a) = g(b) \\
&\text{avec } a, b \in A,
\end{aligned}$$

c'est-à-dire que, dans le cas où l'évaluation des actions de A sur chacun des critères conduit à un petit nombre de valeurs, on peut vouloir chercher à définir la fonction V en se contentant d'associer une et une seule des relations P ou I à la comparaison de tous les couples de vecteurs $(g(a), g(b))$.

La PAMC la plus connue dans cette classe est dénommée **Agrégation additive** [ROY, 1985] [MARESCHAL, 1988]. Elle part du principe suivant : un des modèles les plus simples qu'on puisse envisager, en dehors de la somme pondérée, revient à supposer que l'on peut donner une valeur à $g(a)$ en additionnant de façon adéquate les diverses composantes de $g(a)$. Ceci revient à considérer un modèle tel que $a, b \in A$:

$$\begin{aligned}
g(a) &= \sum_{i=1}^n v_i(g_i(a)) \\
a P b &\Leftrightarrow g(a) > g(b), \\
a I b &\Leftrightarrow g(a) = g(b)
\end{aligned}$$

où les v_i sont des fonctions monotones strictement croissantes. En fait, dans l'agrégation $g(a)$, les v_i vont être les coefficients d'importance (plus la valeur de l'action a est importante, d'un critère i à un critère $i+1$, plus v_i est croissant, c'est-à-dire que les coefficients v_i s'accroissent en valeurs). Prenons un exemple relevant de nos travaux : soit l'action a (*augmenter le temps alloué à l'évaluation*), prenons deux critères i (***budget d'évaluation, niveau d'expérience des évaluateurs***).

Soient les préférences de l'expert : pour *augmenter le temps d'évaluation*, l'expert préfère que le ***budget d'évaluation*** soit très important et que le ***niveau d'expérience des évaluateurs*** soit indifférent. Cela revient à dire que $v_1 > v_2$, c'est-à-dire que le premier coefficient v_1 de l'action a par rapport au critère 1 est plus important que le deuxième coefficient v_2 du même action par rapport au critère 2. On peut imaginer une suite de critères qui donne finalement une fonction croissante de v_i .

Dans les cas réels, on peut mettre $g(a)$ sous une forme qui permet de mieux en apprécier l'intérêt. Dans ces problèmes, on peut en effet supposer que dans $\forall i \in F$ (une famille de critères), l'échelle E_i du critère g_i est bornée inférieurement par une valeur e_{i^*} , ces deux valeurs étant distinctes. On notera $e^* = (e_{1^*}, e_{2^*}, \dots, e_{n^*})$ et $e^* = (e_{1^*}^*, e_{2^*}^*, \dots, e_{n^*}^*)$.

Posons $\forall i \in F$:

$$v_{i^*}^* = v_i(e_{i^*}^*),$$

$$v_{1*} = v_1(e_{i*}),$$

$$v^* = \sum_{i=1}^n v_{i*},$$

$$v_* = \sum_{i=1}^n v_{i*},$$

$$w_i(g_i(a)) = (v_i(g_i(a)) - v_{i*}) / (v_{i*} - v_{i*}),$$

$$k_i = \frac{(v_{i*} - v_{i*})}{(v^* - v_*)}$$

Considérons alors le critère g' défini par :

$$g'(a) = (g(a) - v_*) / (v^* - v_*).$$

Il est facile de montrer que g' est un codage de g . Compte tenu de ce qui précède, on peut, dans les problèmes réels, écrire l'expression de $g(a)$ en posant :

$$w_i(e_{i*}) = 0, w_i(e_{i*}) = 0$$

$$g'(a) = \sum_{i=1}^n k_i * w_i(g_i(a))$$

$$\sum_{i=1}^n k_i = 1$$

$g'(a)$ étant compris entre 0 et 1.

Sous la dernière expression de $g'(a)$, on voit que la PAMC additive revient :

- ◆ à trouver un codage de chaque critère et
- ◆ à appliquer un modèle de *somme pondérée* sur les codages des critères.

L'intérêt de la PAMC additive en aide à la décision tient à sa grande simplicité. Selon ROY [1985], cette PAMC est certainement la plus utilisée en aide à la décision, soit en tant que telle, soit sous forme de somme pondérée ou encore en Goal Programming, etc. Toutefois, selon ROY [1985], cette PAMC présente quelques difficultés pratiques et théoriques pour des problèmes complexes.

II.2.1.2. Agrégation conduisant à un Quasi-critère

La PAMC qui fait l'objet de cette section cherche à construire un critère unique de synthèse dont le pouvoir discriminant soit plus faible que celui de vrai-critère. Elle présente un grand intérêt conceptuel, en particulier lorsque l'on travaille sur une famille cohérente de critères ne contenant pas de vrai-critères [KRANTZ, 1967] [ROBERTS, 1979].

Le principe consiste à construire un critère de synthèse g muni d'un seuil q tel que :

$$\forall a, b \in A \quad g(a) = V(g_1(a), g_2(a), \dots, g_n(a)) :$$

$$g(a) \geq g(b) \Rightarrow \begin{cases} aPb & \text{si } g(a) > g(b) + q(g(b)) \\ aIb & \text{ailleurs} \end{cases}$$

avec

$$\frac{q(g(a)) - q(g(b))}{g(a) - g(b)} \geq \text{Sinon}$$

sachant que le seuil $q(g(a))$ peut toujours être rendu constant moyennant un codage approprié de g . Les conditions nécessaires et suffisantes pour obtenir un quasi-critère de façon additive sont données par FISHBURN [1970]. De même, on trouvera des systèmes axiomes suffisants pour obtenir un quasi-critère g à seuil constant dans LUCE [1973] et SUPPES *et al.* [1989].

Malgré son réel intérêt, cette approche n'a fait l'objet que de très peu de travaux, ce désintérêt s'expliquant peut-être par de réelles difficultés de mise en œuvre [ROBERTS, 1979]. Selon KRANTZ [1967], elle présente toutefois l'avantage de permettre de concilier la simplicité de l'agrégation additive avec l'obtention d'un critère de synthèse au pouvoir discriminant nuancé.

II.2.1.3. Procédure d'Agrégation MultiCritère cherchant à construire une mesure

Dans cette partie, nous nous intéressons à des PAMC cherchant à bâtir un critère g tel que :

$$\begin{aligned} g(a) &= V(g_1(a), g_2(a), \dots, g_n(a)). \\ g(a) \geq g(b) &\Rightarrow a S b \\ g(a) - g(b) \geq g(c) - g(d) &\Rightarrow (a\theta b) S^*(c\theta d) \end{aligned}$$

où S^* représente une relation de surclassement sur $A \times A$,
et $a\theta b$ représente un écart entre deux actions qui peut s'interpréter comme une mutation $b \rightarrow a$

Nous nous limiterons ici à des approches faisant de S^* un pré-ordre complet. C'est dire que les critères de synthèse g dont il sera question dans cette partie seront tels que :

$$\begin{aligned} g(a) &= V(g_1(a), g_2(a), \dots, g_n(a)). \\ a P b &\Leftrightarrow g(a) > g(b) \\ a I b &\Leftrightarrow g(a) = g(b) \\ (a\theta b) P^*(c\theta d) &\Leftrightarrow g(a) - g(b) > g(c) - g(d) \\ (a\theta b) I^*(c\theta d) &\Leftrightarrow g(a) - g(b) = g(c) - g(d), \end{aligned}$$

ce qui correspond au cas où g est une « mesure ».

Dans cette classe, la PAMC la plus représentative est certainement la **Procédure d'agrégation multiplicative** [DYER et SARIN, 1979].

Son principe est le suivant : dans le cadre d'une modélisation des écarts de préférences, on dira qu'une sous-famille de critères J est indépendante au sens des différences de préférences dans une famille de critères F si les jugements en termes d'écart de préférence sur des actions ne se différencient que sur des critères de J ne dépendant pas des valeurs fixées sur les critères F / J.

En considérant quatre actions a_1, a_2, a_3, a_4 ayant toutes des évaluations identiques sur des critères de F / J, on dira que la sous-famille J est indépendante au sens des différences de préférence dans F si

$$a_1 \theta a_2 \text{ P}^* a_3 \theta a_4 \text{ (respectivement I}^*) \Rightarrow b_1 \theta b_2 \text{ P}^* b_3 \theta b_4 \text{ (respectivement I}^*) \quad (1)$$

où b_1, b_2, b_3, b_4 sont des actions ayant des évaluations identiques sur les critères de F / J et ayant respectivement les mêmes évaluations que a_1, a_2, a_3 et a_4 sur les critères de J. Considérons une fonction g vérifiant la relation (1). Si les critères sont mutuellement indépendants au sens des préférences, si l'un des critères est indépendant au sens des préférences dans une famille de

$$\text{soit } g(a) = \sum_{j=1}^n k_j * v_j(a) \text{ si } \sum_{j=1}^n k_j = 1 \quad (\text{forme additive}),$$

critères F et si $n \geq 3$, alors :

$$\text{soit } g(a) = \frac{1}{k} \left[\prod_{j=1}^n (1 + k * k_j * v_j(g_j(a))) - 1 \right] \text{ si } \sum_{i=1}^n k_i \neq 1 \quad (\text{forme multiplicative}),$$

où $\forall i \in F$:

$$v_i(e_i^*) = 1, v_i(e_i) = 0,$$

$$g(\tilde{a}) = 1, g(\text{non } \tilde{a}) = 0,$$

$$k_i \in [0 ; 1],$$

$k > -1$ et $k \neq 0$, k étant solution de l'équation :

$$1 + k = \prod_{j=1}^n (1 + k * k_j).$$

Selon KRANTZ *et al.* [1971], le résultat ci-dessus fonde le choix entre la forme additive et la forme multiplicative sur la somme des coefficients k_j . En pratique, cependant, ce critère de choix impliquerait de disposer de valeurs extrêmement précises des k_j , ce qui est rarement le cas. Pour contourner cette difficulté, DYER et SARIN [1979] montrent que l'on peut trouver quatre actions a, b, c et d ayant des performances identiques sur $F / \{i, j\}$, dont les performances sur le couple de critères (i, j) valent respectivement $(x_i, y_j), (z_i, w_j), (x_i, w_j), (z_i, y_j)$ et telles que $a \theta b$ $I^* c \theta d$; la forme additive s'impose alors.

Un des avantages de raisonner en termes de mesures est qu'il devient possible de bâtir des codages v_j indépendamment les uns des autres. Il faut toutefois souligner que dans le cas de la forme multiplicative, le problème est plus complexe dans la mesure où l'on n'a plus affaire à des inégalités linéaires [ROY et SKALKA, 1984].

II.2.1.4. Théorie de l'Utilité Espérée

La Théorie de l'Utilité Espérée [ALLAIS et HAGEN, 1979] [YAARI, 1987] [MUNIER, 1989] fait également partie des Procédures d'Agrégation MultiCritère, à la différence qu'elle ne prend pas appui sur un tableau de performances, c'est-à-dire sur les évaluations des actions potentielles sur différents critères constituant une famille cohérente. La théorie de l'Utilité Espérée s'attache au cas où la distribution de vraisemblance dont il vient d'être question est une distribution de probabilités. Elle vise à construire un vrai-critère de synthèse directement sur la base de ces évaluations sans nécessairement recourir à une phase préalable de construction d'un tableau de performance. Au niveau de la théorie de l'Utilité Espérée, le vrai-critère de synthèse est du type :

$$\left. \begin{array}{l} a P b \Leftrightarrow g(a) > g(b) \\ a I b \Leftrightarrow g(a) = g(b) \end{array} \right\} \quad (2)$$

avec :

$$g(a) = \sum_{e_1 \in E_1} \sum_{e_2 \in E_2} \dots \sum_{e_n \in E_n} u(e_1, e_2, \dots, e_n) \cdot \delta a(e_1, e_2, \dots, e_n)$$

ce qui revient à associer, à chaque vecteur (e_1, e_2, \dots, e_n) , un coefficient $u(e_1, e_2, \dots, e_n)$ généralement appelé « Utilité » et à faire ensuite la somme de ces coefficients, pondérée par les probabilités $\delta a(e_1, e_2, \dots, e_n)$, d'où l'appellation « Utilité Espérée ». La fonction d'Utilité u intègre, de façon non dissociable, deux aspects des préférences, relatifs respectivement : (1) aux comparaisons d'écart de préférences résultant de variations certaines des conséquences, (2) aux jugements de préférences résultant de variations dans l'appréciation probabiliste des conséquences.

Le critère d'Utilité Espérée relatif à n dimensions se présente comme une PAMC agréant des critères définis eux-mêmes selon un principe identique sur une dimension.

Les travaux de FISHBURN [1970] donnent les conditions permettant de représenter le système relationnel de préférences par un vrai-critère construit selon la formulation de l'Utilité Espérée (2). Pour bâtir un tel critère, il faut donc pouvoir évaluer la fonction $u(e_1, e_2, \dots, e_n)$. Selon KEENEY et RAIFFA [1976], lorsque la fonction u ne fait intervenir qu'une seule dimension, il existe un certain nombre de techniques de questionnements permettant de lui donner une forme précise. Selon ALLAIS et HAGEN [1979], on pourrait songer à appliquer ces mêmes techniques dans le cas où u fait intervenir plusieurs dimensions.

Du point de vue de l'aide à la décision, YAARI [1987] estime que si l'on souhaite bâtir un vrai-critère, la simplicité de la forme fonctionnelle utilisée en théorie de l'Utilité Espérée peut

être un avantage déterminant. Notons cependant que depuis quelques années, beaucoup de travaux synthétisant des expériences en laboratoire remettent profondément en cause cette théorie, notamment en ce qui concerne ses possibilités de décrire le comportement observé en laboratoire d'un grand nombre d'individus [ALLAIS et HAGEN, 1979][MUNIER, 1989].

II.2.2. Surclassement de synthèse

Ces approches d'aide à la décision consistent à prendre appui sur une règle apportant une réponse synthétique, exhaustive et définitive au problème de l'agrégation des performances. Elles prennent la forme d'un ensemble de conditions conduisant à accepter ou à rejeter un surclassement au niveau global. Ces approches visent à caractériser les surclassements qu'il est possible d'établir de façon suffisamment solide. Les approches dont il est question ici se répartissent en deux groupes, successivement décrits :

- ◆ celles qui aboutissent à un unique système relationnel de préférences : les principales procédures étudiées sont ELECTRE I, ELECTRE IS et TACTIC ;
- ◆ celles qui aboutissent à plusieurs systèmes relationnels de préférences, les principales procédures étudiées étant ELECTRE II, ELECTRE III, ELECTRE IV, PROMETHEE et les procédures par compensation.

II.2.2.1. Approches aboutissant à un unique système de préférences

Définir une procédure de ce type équivaut à formuler des conditions faisant intervenir un couple quelconque de vecteurs de performances ($g(a)$, $g(b)$), dès lors qu'elles sont satisfaites, et validant l'affirmation 'a S b'. Etant donné une paire d'actions (a, b), il se peut que ces conditions ne soient vérifiées ni par le couple (a, b), ni par le couple (b, a) : les deux actions sont alors incomparables (a R b).

A titre de premier exemple, **ELECTRE I** [ROY, 1968] [MAYSTRE *et al.*, 1994] part du principe que les formules d'agrégation s'appliquent au cas où une famille de critères F ne comporte que des vrai-critères. Les formules font intervenir, pour chacun d'eux : (1) un coefficient d'importance $k_j > 0$; (2) un seuil de veto $v_j(g_j) > 0$.

Ce sont là des données « inter-critères » que l'on retrouvera dans la plupart des PAMC dont il va être question dans la suite de cette section. La façon de leur attribuer en pratique une valeur numérique est la suivante :

$$C(b S a) = \{j \in F ; g_j(b) \geq g_j(a)\}$$

$$\forall C \subset F, k[C] = \sum_{j \in C} k_j *$$

Dans ELECTRE I, la proposition b S a est regardée comme valide si et seulement si les deux conditions suivantes sont satisfaites :

- Condition de concordance :

$$\frac{k[C(bSa)]}{k[F]} \geq s,$$

$$\frac{1}{2} \leq s \leq 1 - \frac{\min_{j \in F} k_j}{k[F]}. \quad (3)$$

- Condition de non veto :

$$\forall j \in F, g_j(b) + v_j[g_j(b)] \geq g_j(a)$$

La valeur de s est ce que l'on appelle le niveau exigé de concordance. Selon ROY [1968], dans ELECTRE I, il est facile de constater qu'en multipliant par une constante les formules utilisées, tous les coefficients d'importance ne changent pas le système de préférences produit par ELECTRE I. Il n'est pas restrictif, en pratique, de supposer que les coefficients d'importance sont rationnels. Après multiplication par une constante, on peut donc les supposer tous entiers. Le coefficient k_j peut donc s'interpréter comme le nombre de voix attribuées au critère j dans le cadre d'une procédure de vote. Le niveau exigé de concordance s s'interprète alors comme un niveau de majorité exigé (majorité simple pour $s=1/2$, majorité des deux tiers pour $s=2/3$, etc.).

Une fois que s dépasse la valeur s^* , avec :

$$s^* = 1 - \frac{\min_{j \in F} k_j}{k[F]},$$

La condition (3) ci-dessus équivaut à requérir l'unanimité, c'est-à-dire que $C(b S a) = F$ est automatiquement vérifiée.

Cette approche a été conçue par référence à une famille F de critères contenant au moins trois critères, et ce n'est même qu'à partir de cinq ou six qu'elle acquiert sa pleine signification. C'est la une contrainte et une difficulté pour cette approche. Elle a toutefois l'avantage d'avoir ouvert la voie aux autres approches PAMC selon ROY [1968]. Il faut de même souligner sa simplicité de mise en œuvre et insister sur le fait qu'elle ne repose pas sur des *a priori* comme la plupart des méthodes utilisant une fonction d'utilité. La contrepartie des avantages est que le résultat fourni par ELECTRE I peut être beaucoup moins net, ce qui d'ailleurs est normal puisque la méthode respecte les nuances. Elle ne conduit pas, en général, à la mise en évidence directe de la « meilleure » action potentielle. Son objectif est de proposer un cadre rationnel fondé sur des hypothèses relativement simples. Une extension d'ELECTRE I est ELECTRE IS [ROY et SKALKA, 1984]. La Procédure d'Agrégation MultiCritère sur laquelle s'appuie ELECTRE IS est une extension de la précédente visant à prendre en compte :

- d'éventuels seuils d'indifférence et de préférence non nuls pour certains critères de la famille F et corrélativement,

- un renforcement de l'effet de veto lorsque l'importance de la coalition concordante décroît.

Comme c'était déjà le cas avec ELECTRE I, soulignons que pour un critère, quel qu'il soit, on peut toujours supprimer l'effet de veto : il suffit pour cela de donner une valeur assez grande au seuil de veto v_j .

Dans son principe, cette approche s'applique au cas où les n critères de F sont des pseudo-critères, chacun étant muni d'un coefficient d'importance et d'un seuil de veto. Tous les seuils (indifférence, préférence et veto) peuvent varier le long de l'échelle du critère. Notons que :

$$C(b S a) = \{j \in F : g_j(b) + q_j[g_j(b)] \geq g_j(a)\}$$

$$C(b S a) = \{j \in F : g_j(b) + q_j[g_j(b)] < g_j(a) \leq g_j(b) + p_j[g_j(b)]\}$$

$$k[C(bSa), C(aQb)] = \sum_{j \in C(bSa)} k_j + \sum_{j \in C(aQb)} \varphi_j * k_j$$

$$\text{avec : } \varphi_j = \frac{g_j(b) + p_j[g_j(b)] - g_j(a)}{p_j[g_j(b)] - q_j[g_j(b)]}.$$

ELECTRE IS constitue une amélioration par rapport à ELECTRE I (problème α), dans la mesure où elle permet d'utiliser des pseudo-critères, que l'on retrouve en l'occurrence dans la logique floue [MAYSTRE *et al.*, 1994].

ELECTRE IS est d'un apport important dans la compréhension des relations entre actions, aussi bien à l'intérieur des circuits qu'entre les circuits et les autres éléments d'un graphe. Les indicateurs de stabilité qu'elle utilise permettent de connaître la stabilité des circuits et donc d'anticiper l'effet éventuel d'une modification des relations entre actions. ELECTRE IS offre des outils d'analyse permettant de connaître, pour chaque circuit de graphe, son taux de cohésion interne (relations entre les actions) et son taux de liaison externe (relations avec les autres éléments du graphe, action ou circuit).

II.2.2.2. Autres approches aboutissant à plusieurs systèmes de préférences

Les systèmes de préférences dont il est ici question sont le fruit de PAMC qui cherchent avant tout à séparer les situations de préférence stricte des autres. Dans ces procédures, les autres situations en question sont l'*indifférence* et l'*incomparabilité*. Selon ROY et SKALKA [1984], ces procédures, qui sont conçues avant tout pour les vrai-critères ou les quasi-critères, ignorent les situations de préférence faible. Définir une procédure de ce type équivaut donc à formuler les conditions faisant un couple quelconque de vecteurs de performances ($g(a)$, $g(b)$), conditions qui, dès lors qu'elles sont satisfaites, valident l'affirmation $b P A$.

Dans cette catégorie d'approche, l'outil d'aide TACTIC [VANSNICK, 1986] est très représentatif. Dans son principe, il s'applique au cas où une famille de critères F ne comporte que des vrai-critères ou des quasi-critères. Tout comme les méthodes ELECTRE, elle fait intervenir, pour chaque critère, un coefficient d'*importance* $k_j > 0$. Un aspect négatif dans

TACTIC est que la validation porte sur la préférence stricte et non sur le surclassement, ce qui constitue une différence essentielle avec ELECTRE I. En particulier, cela rend difficile d'isoler, dans TACTIC, les situations d'incomparabilité des situations d'indifférence. L'avantage que présente cette méthode réside dans sa simplicité lors de la mise en œuvre.

II.2.2.3 Systèmes de type aboutissant à plusieurs systèmes de préférences

Dans cette section, on s'intéresse aux méthodes qui font intervenir des conditions à satisfaire pour valider des propositions telles que bSa ou bPa , mais cette validation ne se fera plus en termes : vrai ou faux. Par contre, elle va être guidée par référence à une hiérarchie de crédibilité. Les divers niveaux qui caractérisent cette hiérarchie peuvent être imposés *a priori* : à chacun d'eux est associé un système de préférences d'autant plus pauvre en incomparabilité que le niveau de crédibilité baisse. On peut au contraire chercher à apprécier directement, sur une échelle allant de 0 à 1, le niveau de crédibilité avec lequel l'affirmation en question est validée. Dans cette catégorie d'approches, on peut citer : ELECTRE II, ELECTRE III, ELECTRE IV et PROMETHEE.

A titre d'exemple représentatif, prenons **ELECTRE II** [ROY et BERTIER, 1973], qui fut la première méthode d'aide à la décision prenant appui sur une séquence de systèmes de préférences emboîtés. Dans son principe, elle s'applique dans le cas où une famille de critères ne compte que des vrai-critères. Elle fait intervenir pour chacun d'eux :

- un coefficient d'importance $k_j > 0$;
- deux seuils de veto v_j^1 et v_j^2 ($v_j^1 \leq v_j^2$), dont le premier est fixé de telle sorte que, lorsqu'il autorise le surclassement, celui-ci apparaisse comme absolument indiscutable, tandis que le deuxième vise à exploiter la marche d'indétermination de cette valeur du veto pour atténuer l'effet restrictif.

La proposition $bS^i a$ est considérée valide si et seulement si les deux conditions suivantes sont satisfaites :

- Condition de concordance :

$$\frac{k[C(bSa)]}{k[F]} \geq s^i \quad \text{et} \quad k[C(bSa)] \geq k[C(aSb)].$$

- Condition de non veto :

$$\forall j \in F, \quad g_j(b) + v_j^i \geq g_j(a).$$

On remarque que si la condition de non veto est identique à celle d'ELECTRE I, la condition de concordance diffère puisqu'elle requiert en plus que : $k[C(bSa)] \geq k[C(aSb)]$. Cette

inégalité additionnelle a été introduite afin de n'aboutir à la configuration bS^a et aS^b (cas d'indifférence) que si $k[C(bPa)] = k[C(aPb)]$.

II.3. Approches utilisant les techniques de l'intelligence artificielle (IA)

Plusieurs travaux ont fait l'objet de l'apport de l'intelligence artificielle [McBRIDE, 1993][HOLZMAN, 1989][GOTTINGER et WEIMANN, 1992]. En effet, ces travaux soulignent l'importance du rapprochement entre l'IA et l'aide à la décision multicritère. L'objectif étant dans cette optique d'assister complètement le processus de décision. Beaucoup de systèmes industriels ont mis à contribution l'IA et l'aide à la décision multicritère. On peut citer par exemple un système de choix de machine de tri des paquets postaux [LEVINE *et al.*, 1990]. L'avantage des techniques fondées sur l'IA est qu'on ne soit pas obligé de quantifier les échelles de préférences. En effet, dans cette optique, les opérations portent sur les données qualitatives et les descripteurs utilisés sont du type : mauvais, moyen, bon, etc. Ces descripteurs peuvent ainsi grâce à l'IA passer des données quantitatives à qualitatives.

Un exemple de règle utilisée en IA est du type :

Si (taille de Jeanne > taille de Pierre)
et (taille de Jeanne < taille de François et taille François < taille André)
Alors (le plus grand = André).

II.4. Approches basées sur les ensembles flous

Nous venons de voir des approches dans lesquelles on peut quantifier de manière précise les paramètres nécessaires au fonctionnement des méthodes à l'exception d'ELECTRE III qui utilise le flou dans sa démarche. Ce paragraphe s'intéresse au problème de propositions floues dans leur contenu ou d'assertions floues dans leur expression, celui du raisonnement en présence des types de connaissances suivants :

- des concepts ou prédicats flous comme *grand, chaud, aimable, etc.* ;
- des faits flous : *la Terre est à peu près ronde, légèrement aplatie aux pôles, etc.* ;
- des appréciations de probabilité floues : *vraisemblable, fort, possible, etc.* ;
- des quantificateurs flous comme *certain, beaucoup de, la plupart des, peu de* ou encore *généralement, souvent, parfois, etc.* ;
- des opérateurs de comparaison flous : *aussi... que, moins... que, etc.* ;
- des adverbes de quantité que nous appellerons modificateurs flous : *peu, très, assez, plutôt, modérément, presque, etc.*

ZADEH [1975] est le créateur de la *logique floue* qui offre un cadre formel pour le raisonnement à partir de connaissances floues ou vagues. Il a d'abord proposé la théorie des ensembles flous puis s'est appuyé sur cette théorie pour développer la théorie des possibilités en reliant la notion d'appartenance à un ensemble flou à celle de distribution de possibilité.

Un *ensemble flou* A sur le domaine T est défini par la donnée d'une fonction μ_A à valeur dans $[0, 1]$. $\mu_A(t)$ est le *degré d'appartenance* de $t \in T$ à A. Prenons l'exemple du domaine T des

temps alloués possibles pour une évaluation de systèmes interactifs. La proposition le *temps alloué* est *important* contient le prédicat imprécis *important*. L'ensemble des tailles qui peuvent être qualifiées de grandes est un intervalle flou. Il est défini par une fonction du domaine des tailles dans l'intervalle [0, 1].

Sur la figure III.2, chaque point de la courbe représente le degré d'appartenance à l'ensemble flou *important* d'un *temps alloué* particulier. Remarquons le caractère subjectif d'une telle fonction, laissée à l'appréciation de la personne qui la définit. Il est des situations où ce caractère subjectif peut être évité en procédant, par exemple, par enquête sur un nombre d'évaluations représentatif.

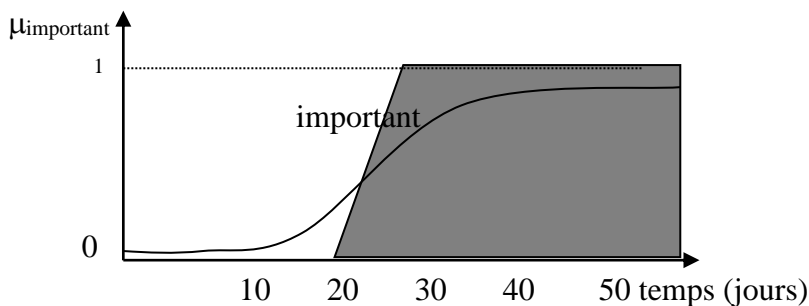


Figure III.2 : Fonction d'appartenance d'un temps alloué à l'ensemble des temps importants

Opérations élémentaires sur les ensembles flous

Concernant les opérations élémentaires sur les ensembles flous, nous dirons que diverses familles d'axiomes de définition ont été proposées. Les définitions les plus couramment adoptées sont les suivantes :

- l'inclusion : $\mu_A(s) \leq \mu_B(s)$ si $B \supseteq A$;
- l'égalité : $\mu_A(s) = \mu_B(s)$ si $A = B$;
- l'intersection : $\mu_{A \cap B}(s) = \min(\mu_A(s), \mu_B(s))$;
- l'union : $\mu_{A \cup B}(s) = \max(\mu_A(s), \mu_B(s))$;
- la complémentation : $\mu_{\bar{A}}(s) = 1 - \mu_A(s)$;
- le produit cartésien : $\mu_{A \times B}(s, t) = \min(\mu_A(s), \mu_B(t))$;
- la cardinalité : $\text{card}(A) = \sum_{s_j \in A} \mu_A(s_j)$ pour tous les éléments s_j de A si A est dénombrable ou passage au continu dans le cas contraire.

Avec $s \in S$ (ensemble quelconque) et $t \in T$ (ensemble des tailles possibles).

D'autres choix pourraient être effectués. C'est ainsi que l'intersection peut s'exprimer également par le produit $\mu_A * \mu_B$ ou la fonction $\max(0, \mu_A + \mu_B - 1)$.

Les distributions de possibilité peuvent être utilisées à plusieurs fins :

- Pour représenter des données précises ou imprécises (figure III.3) :

- Le temps alloué à une évaluation donnée vaut exactement 2 jours,
- Le temps alloué de l'évaluation X est de l'ordre de 3 jours. Il existe un degré d'appartenance flou du temps alloué de l'évaluation X sur l'échelle des temps possibles. Une telle courbe est souvent approchée par un trapèze, comme l'indique la figure III.4.b,
- L'évaluation de mon système dure entre 4 et 6 jours. Le temps est donné sous forme d'un intervalle non flou ;

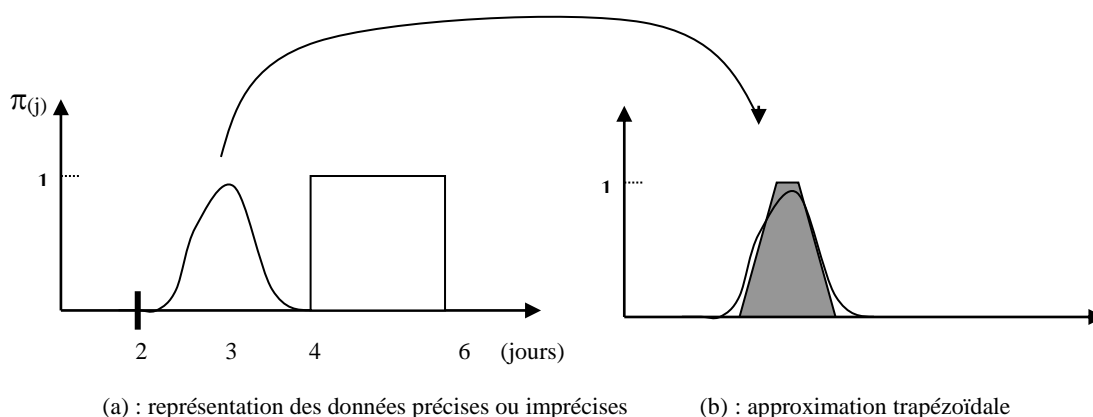


Figure III.3 : Distributions de possibilité

- Pour définir des prédicats flous, par exemple *important* (figure III.2) ou *Faible* (figure III.4). Dans ce dernier exemple, on note que la forme continue peut, dans la pratique, être approchée par une ligne brisée ;

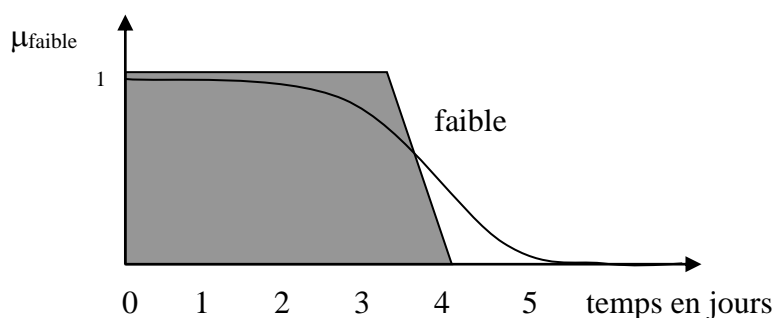


Figure III.4 : Fonction d'appartenance d'un temps alloué à l'ensemble des temps faibles

- Pour définir des ensembles flous de proportions traduisant des modificateurs (*pas*, *très*) ou des quantificateurs flous comme *assez* ou *généralement*, ce qui, par composition, permet de définir des ensembles flous comme *Assez_Faible*. Les fonctions d'appartenance associées sont définies sur $[0, 1]$ et prennent leur valeur sur $[0, 1]$. La figure III.5 montre le degré d'appartenance à l'ensemble flou *Généralement* d'une proportion comprise entre 0 et 100% [ZADEH, 1975]. Elle traduit le fait que

plus le terme *généralement* s'applique, plus grand est le degré d'appartenance. Notons qu'une telle fonction peut elle-même avoir des valeurs floues, ce qui permet de parler d'ultraflou (cf. figure III.5).

Notons la différence de statut entre *très* et *assez* et la différence de leurs effets sur μ_{faible} . Des *distributions de possibilité trapézoïdales* sont fréquemment utilisées pour approcher les fonctions vues plus haut. Les notions d'*ensembles flous* et de *possibilités* peuvent être considérées comme une généralisation de la probabilité [DUBOIS et PRADE, 1985]. L'objectif de ces modèles est de casser la logique booléenne contenue dans le *modus ponens*. On définit P l'ensemble, supposé fini, des propositions de la logique des propositions.

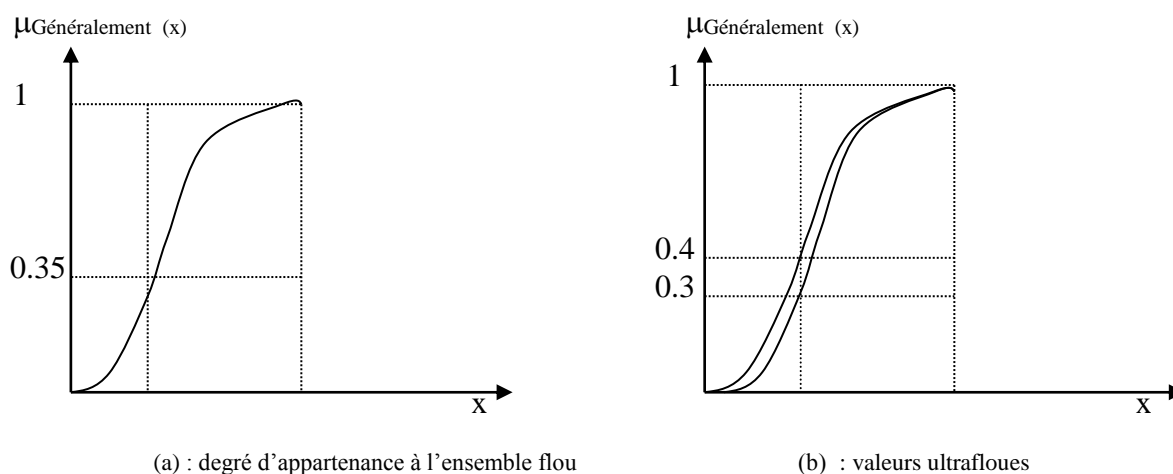


Figure III.5 : Mise en correspondance d'une proportion v avec généralement

La figure III.6 illustre l'effet de modificateurs sur l'appartenance à l'ensemble flou *Faible* :

- *Très_faible* peut se définir par :
si $t \in T$, alors $\mu_{\text{Très_faible}}(t) = [\mu_{\text{faible}}(t)]^2$;
- *Assez_faible* résulte de la composition de la fonction traduisant *assez* avec μ_{Faible} .

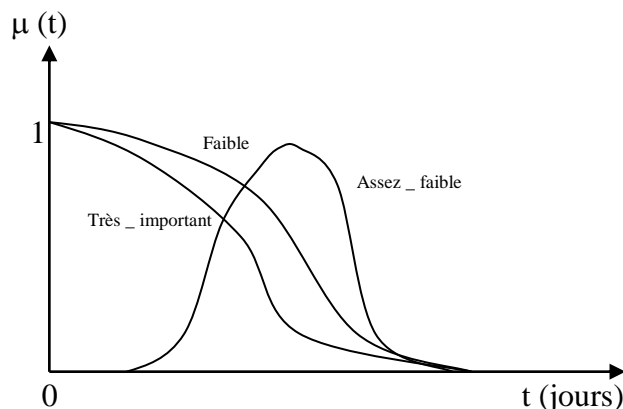


Figure III.6 : Composition de fonctions d'appartenance

Soit FA la proposition toujours fausse et VR la proposition toujours vraie. On considère une application p de P dans $\{0, 1\}$ telle que $p(FA) = 0$, $p(VR) = 1$ et $p(A \vee B) = \text{Max}(p(A), p(B))$. Une telle fonction p s'appelle une *mesure de possibilité* sur P . La mesure duale définie par $n(A) = 1 - p(\neg A)$ est appelée *mesure de nécessité*; elle vérifie $n(A \vee B) = \text{min}(n(A), n(B))$. Une généralisation possible du *modus ponens* donnant la *possibilité* et la *nécessité* de B connaissant celles de A et de $(A \Rightarrow B)$ est la suivante :

$$\begin{aligned} & \text{Si } p(A \Rightarrow B) \geq a \text{ et } n(A \Rightarrow B) \geq b \\ & \text{et } p(A) \geq c \text{ et } n(A) \geq d \\ & \text{Alors } p(B) \geq \max(a \cdot v(a+d), c \cdot v(b+c)) \\ & \text{et } n(B) \geq \text{min}(d, b) \end{aligned}$$

$$\text{où } v(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

Remarquons que l'on a, pour tout A , $\text{Max}(p(A), p(\neg A)) = 1$. par conséquent, l'égalité $\text{Max}(p(A), 1 - n(A)) = 1$ doit, dans le schéma ci-dessus, être vérifiée pour les propositions A ($A \Rightarrow B$) et B . Lorsqu'il existe une incertitude sur les critères, il devient pertinent de prendre en compte les concepts de l'incertitude et de l'imprécision dans l'algorithme d'aide à la décision. Or, ces concepts n'apparaissent pas dans les méthodes classiques d'aide à la décision. Les ensembles flous s'imposent alors. Dans la section III de ce chapitre, nous entrerons dans le détail de la démarche suivie, basée sur ceux-ci.

II.5. Synthèse et positionnement des différentes méthodes d'aide à la décision

Avant de présenter les caractéristiques de notre outil d'aide à la décision, nous avons dans un premier temps classé les différentes méthodes d'aide à la décision en familles de méthodes afin de nous positionner (voir figure III.7). Ce classement est fait à partir d'une synthèse de GANDIBLEUX [1995].

Dans la même logique, en passant de la classe 1 à la classe 5, on se dirige vers les méthodes plus interactives. Par rapport à l'axe des échelles de données, on peut lire que les méthodes de programmation linéaire (CL1 et CL2) traitent des informations plus quantitatives que qualitatives alors que les méthodes des classes CL4 et CL5 peuvent traiter des données qualitatives plus que celles des classes 1 et 3.

De même lorsqu'on s'oriente vers les classes les plus élevées, on se dirige plus vers l'aide à la décision, alors que vers les classes 1 et 2, c'est beaucoup plus de la prise de décision (descriptive) que de l'aide à décision (constructive) [GANDIBLEUX, 1995] dont il s'agit. Enfin, cette figure nous informe que les classes les plus élevées sont dirigées dans le sens de l'axe « approches multicritères - prescriptives », alors que les classes les plus faibles sont plus orientées dans le sens « approches monocritères - normatives ». Le tableau III.1 est une autre représentation de la figure III.7. Elle présente un autre type de synthèse des outils d'aide à la décision. Cette représentation plus lisible nous permettra de justifier l'approche d'aide à la décision pour laquelle nous avons opté pour nos travaux.

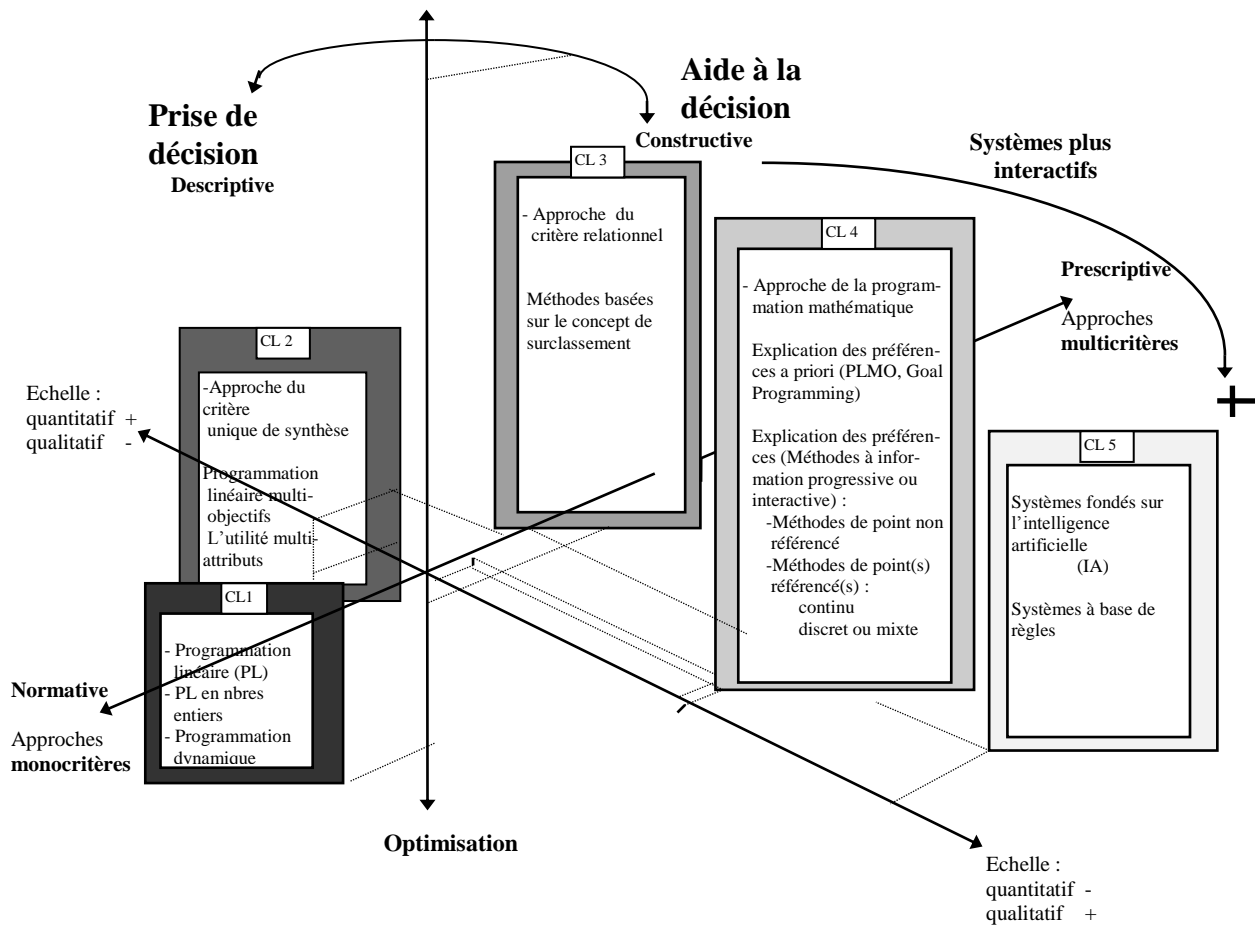


Figure III.7 : Synthèse et classement des différentes méthodes d'aide à la décision

Sur la figure III.7, nous constatons qu'on peut regrouper les différentes méthodes d'aide à la décision en cinq classes de CL1 à CL5. On constate que la première classe (CL1) est au niveau le plus bas par rapport aux axes horizontaux ; cela revient à dire par exemple que les méthodes de programmation linéaire (PL) traitent plus de l'optimisation que de l'aide à la décision [GANDIBLEUX, 95], l'optimisation étant dirigée selon l'axe vertical-bas.

Grâce aux différents paramètres qui caractérisent les différentes approches, nous pourrions choisir une méthode d'aide à la décision dont les caractéristiques permettent de résoudre le type de problème qui fait l'objet de nos travaux.

Justification de l'approche choisie

Pour choisir l'approche correspondant à nos travaux, nous avons utilisé une démarche qui consiste à maintenir dans le tableau III.1 uniquement des paramètres qui concernent le type de

problème que nous avons à traiter (cf. ci-après). Nous obtenons un nouveau tableau (tableau III.2).

Paramètres de choix ou de distinction des différentes approches	CL 5	CL 4	CL 3	CL 2	CL 1
	Systèmes fondés sur l'intelligence artificielle (IA) Systèmes à base de règles :	Approche de la programmation mathématique Explication des préférences a priori Explication des préférences (Méthodes à information progressive ou interactive)	Approches du critère relationnel Approches basées sur le concept de surclassement	Approche du critère unique de synthèse Programmation linéaire multi-objectifs	Programmation linéaire-nombres entiers Programmation dynamique
Type de résolution					
Aide à la décision	■	■	■	■	□
Prise de décision	□	□	□	□	□
Optimisation	□	■	■	■	■
Nature des données					
Qualitatives	■	■	■	■	□
Quantitatives	■	■	■	■	■
Nature des critères					
Multicritère	■	■	■	■	□
Monocritère	□	□	□	□	■
Buts					
Prescriptive	■	■	■	■	□
Normative	□	□	□	□	■
Contexte					
Constructive	■	■	■	■	□
Descriptive	□	□	■	□	■
Connaissance					
Règles de production	■	■	□	□	□

Légende : ■ approche orientée à 100% vers l'aide à la décision

(ex ; pour le type de résolution)

□ approche orientée à 0% vers l'aide à la décision ou 100% vers la prise de décision

■ (aide à la décision) approche plus orientée vers l'aide à la décision que vers la prise de décision ou l'optimisation

Tableau III.1 : Synthèse des approches d'aide à la décision

Nos travaux sont caractérisés par :

- Type de résolution :
 - *aide à la décision* (et non d'optimisation ou de prise de décision comme l'indique le tableau III.1 qui fait apparaître d'autres types de résolutions).
- Nature des données :

Notre système traite à la fois des données *qualitatives* et *quantitatives* :

- critères d'acceptabilité socio-culturelle
- des critères de décision tels que le niveau d'expérience des évaluateurs, les possibilités de financer les experts dont dispose ou non l'équipe d'évaluation, les moyens financiers disponibles pour l'évaluation, les moyens matériels, etc.
- Nature des critères :
 - *aide multicritère* (nous avons adopté une démarche multicritère car, comme l'indique le tableau III.2, les approches basées sur les sous-ensembles flous sont plus orientées vers le multicritère que vers le monocritère).
- But
 - *prescriptif* (et non normatif car nos travaux utilisent les modèles à critères multiple).
- Contexte
 - *constructive* (et non normative car les méthodes utilisées s'éloignent de la programmation linéaire).
- Connaissance
 - pratiquement pas de règles de production (très peu), caractéristique des problèmes de type nouveau où il n'existe pas encore de connaissances sous forme de règles de production [WAWAK, 1996].

Du tableau III.2, nous voyons que deux types d'approches se dégagent :

- Groupe 1 :
Systèmes fondés sur l'intelligence artificielle (IA), systèmes à base de règles
- Groupe 2 :
Approches du critère relationnel, systèmes basés sur les sous-ensembles flous, approches basées sur le concept de surclassement

II.6. Conclusion

En définitive, au vu des caractéristiques du problème à traiter, nous avons opté pour un système combiné basé sur les *sous-ensembles flous* et le *surclassement* à l'image d'Electre I (groupe 2), et en lien avec un système à base de règles (groupe 1). En effet, le flou est adapté au type de problèmes nouveaux où il n'existe pas encore de règles de production bien définies et où de surcroît l'imprécision et l'incertitude se dégagent ; c'est notre situation. La résolution de problème sera réalisée à l'aide d'un langage de *l'intelligence artificielle* (Prolog III), dont

l'intérêt majeur est la résolution de contraintes (données d'entrée de notre système), tout en offrant des outils de réalisation de systèmes d'aide à la décision (groupe 1). Les deux groupes (1 et 2) sont orientés multicritère, ce qui nous convient naturellement. Nous allons maintenant aborder les différentes étapes de la démarche globale du système d'aide à la décision en évaluation conçu et réalisé dans le cadre de cette thèse.

Paramètres de choix ou de distinction des différentes approches	CL 5	CL 4	CL 3	CL 2	CL 1
	Systèmes fondés sur l'intelligence artificielle (IA) Systèmes à base de règles:	Approche de la programmation mathématique Explication des préférences <i>a priori</i> Explication des préférences (Méthodes à information progressive ou interactive)	Approche du critère relationnel Approches basées sur le concept de surclassement	Approche du critère unique de synthèse Programmation linéaire Multi-objectifs	Programmation linéaire en nombres entiers Programmation Dynamique
Type de résolution					
Aide à la décision	■				
Nature des données					
A la fois Qualitatives et Quantitatives			■		
Nature des critères					
Multicritères	■		■		
Buts					
Prescriptive	■	■	■	■	
Contexte					
Constructive	■	■		■	
Connaissance			■		
Règles de production très faible quantité			■		

Tableau III.2 : Synthèse réduite correspondant aux caractéristiques de notre problématique

III. Démarche globale basée sur les ensembles flous

Après le choix d'une approche d'aide à la décision, nous allons dans cette partie, exposer les différentes étapes de la démarche suivie (figure III.8).

C'est une combinaison d'une faible quantité de règles de production « floues » et une agrégation des degrés de compatibilité flous (DCF) issus de l'adéquation entre l'évaluation des critères et les préférences de l'expert. Les différentes étapes de notre algorithme sont visibles en figure III.8.

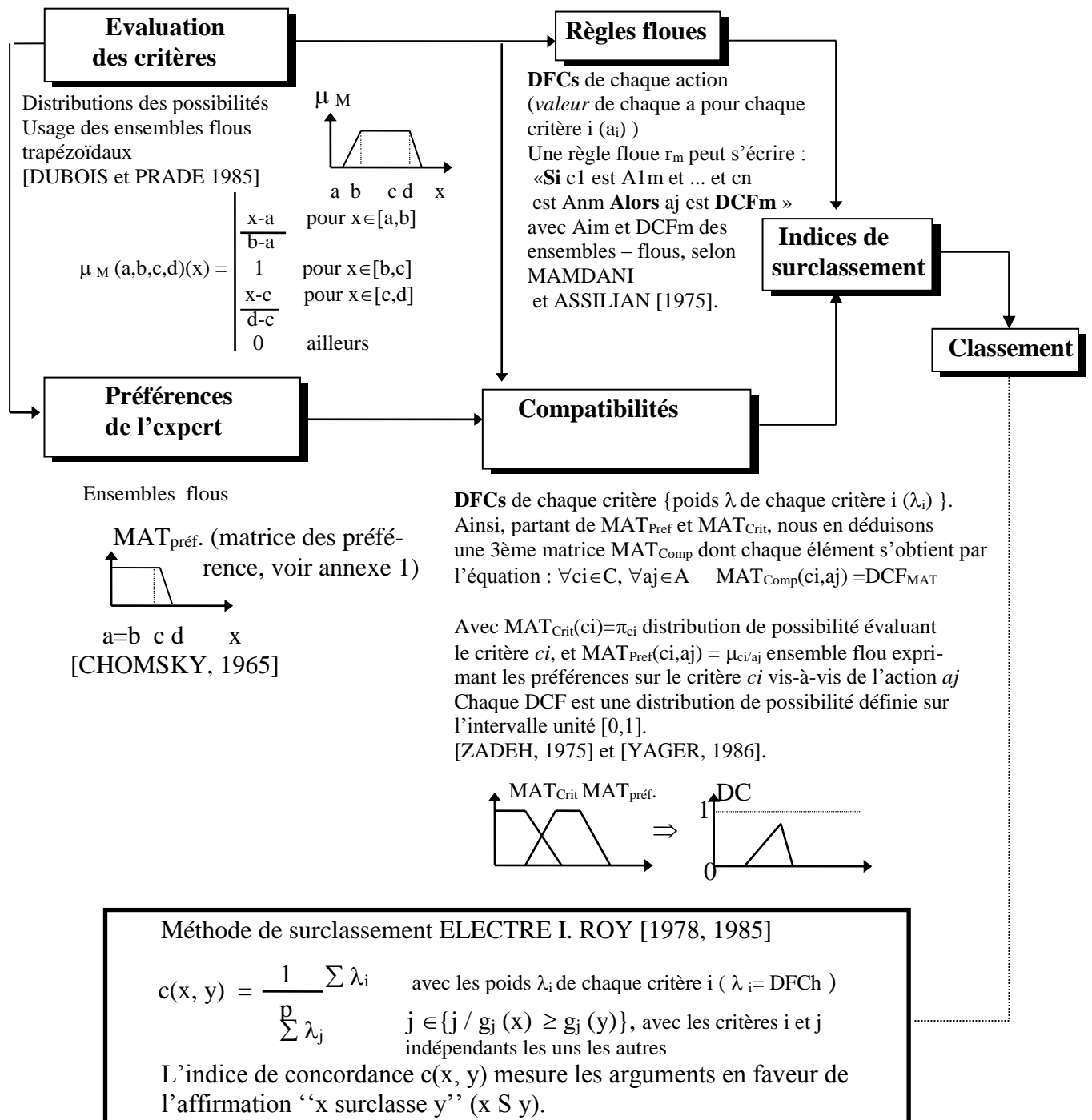


Figure III.8 : les différentes étapes de notre démarche (inspirée de WAWAK [1996])

III.1. Evaluation des critères et préférences de l'expert

Pour l'expression linguistique des préférences, il s'agit de définir un sous-langage du langage naturel (un langage artificiel) permettant d'exprimer une expertise sous forme de préférences vis-à-vis d'un critère ci concernant une action aj. Ce sous-langage nécessite un

vocabulaire et une syntaxe clairement établis [CHOMSKY, 1965]. Comme le rappelle WAWAK [1996], il faut y associer une sémantique permettant de quantifier les expressions découlant du sous-langage. ZADEH propose à ce sujet le concept de *variable linguistique* dans le cadre particulier de l'algèbre floue [ZADEH, 1975]. Une telle variable génère un sous-langage dont les expressions sont quantifiées par des ensembles flous. Elle se définit par un quintuplé (L, V_T , X, Gr, R_S) avec :

- L le nom de la variable linguistique,
- V_T l'ensemble des termes liés à la variable,
- X l'univers du discours,
- Gr la règle syntaxique,
- R_S la règle sémantique.

Une *variable linguistique* [ZADEH, 1975] est une variable dont les valeurs sont des mots ou des phrases du langage (naturel ou artificiel). Par exemple, Age ayant pour valeurs *jeune, très jeune, vieux, etc.*, est une variable linguistique. Ces valeurs peuvent en général être engendrées à partir d'un terme primaire (comme *jeune*), son antonyme (*vieux*), des modificateurs (*pas, très, plus ou moins, tout à fait, pas très,...*) et les connecteurs *et* et *ou*. Ainsi, une valeur de Age peut être *pas très jeune* et *pas très vieux*. A chaque valeur d'une variable linguistique peut être associée une distribution de possibilité. Selon ZADEH [1975], dans la construction d'une base de connaissances, un cognitifien est parfois confronté au problème de la formalisation d'expressions linguistiques floues, notamment en commande de procédés industriels (*la température doit être assez élevée,...*). Une méthode dite d'approximation linguistique a été proposée pour exprimer de telles connaissances en termes de flou [ESHRAUGH et MAMDANI, 1979].

Voici un exemple de préférence de l'expert dans notre cas pour le *critère temps alloué à l'évaluation* :

« Pour augmenter le *critère temps alloué à l'évaluation*, l'expert préfère que *Les moyens financiers disponibles soient plutôt bons* ». Nous savons que l'action concernée est « *ajuster les moyens humains* » (ajuster les moyens humains signifie : augmenter ou diminuer l'effectif de l'équipe des évaluateurs) vis-à-vis du critère *Moyens financiers disponibles*. La variable linguistique L à utiliser s'intitule « *Préférences de l'expert concernant l'action ajuster les moyens humains vis-à-vis du critère Moyens financiers disponibles* » et l'expression à évaluer est *plutôt bonne*. Aussi *bonne* doit-il être un terme descripteur et *plutôt* un terme modificateur.

Prenons à ce sujet l'exemple de la figure III.9, tiré de la figure III.3

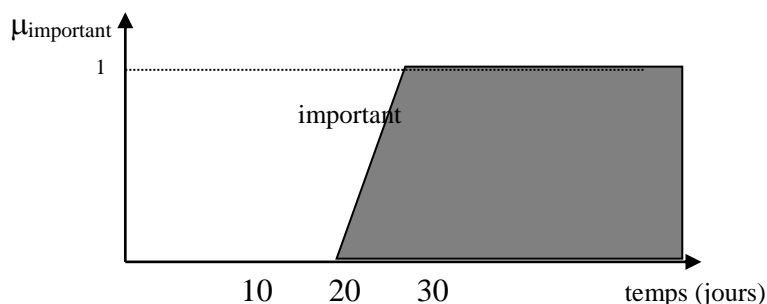


Figure III.9 : Fonction d'appartenance d'un temps alloué à l'ensemble des temps importants

En utilisant l'expression des degrés d'appartenance ci-dessous, on a :

$$\mu_M(a,b,c,d)(x) = \begin{cases} \frac{x-a}{b-a} & \text{pour } x \in [a,b] \\ 1 & \text{pour } x \in [b,c] \\ \frac{x-c}{d-c} & \text{pour } x \in [c,d] \\ 0 & \text{ailleurs} \end{cases}$$

On obtient avec par exemple ($a = 20$, $b = 30$, $c = 100$, $d = 100$), ces valeurs devant être définies en liaison avec un expert du domaine, la courbe (figure III.10) :

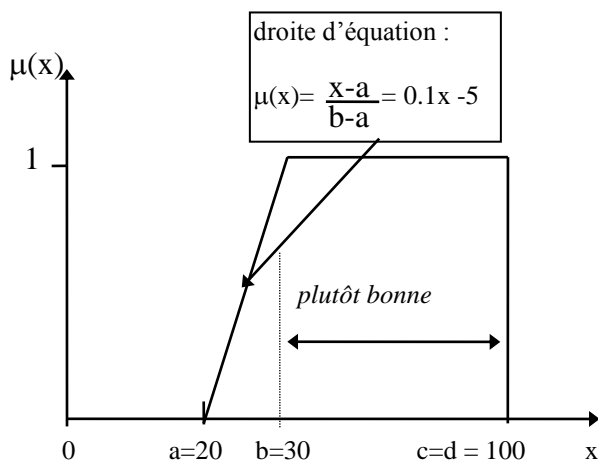


Figure III.10 : Préférence de l'expert pour les critères *Moyens humains et financiers*

En reprenant l'exemple ci-dessous, cette courbe s'interprète de la manière suivante :

Entre 20 et 30 jours d'évaluation, l'expert estime que les *Moyens financiers disponibles* sont *plutôt bons* (sur la figure, lire *plutôt_bonne*) pour augmenter le *temps alloué à l'évaluation*.

III.2. Comparaison préférences-critères (Degré de Compatibilité Flou)

Le degré de compatibilité ($DCF_{A/B}$) évalue la compatibilité d'un ensemble flou A vis-à-vis d'un autre ensemble flou B. Dans notre application, il s'agit de comparer les valeurs des critères à leurs préférences. Chaque critère permet d'obtenir un DCFh (hauteur du sous-ensemble DCF et ceci pour chaque action, ce qui conduit à une matrice MAT_{Comp} de dimension $n \times p$ (avec p le nombre de critères et n le nombre d'actions). Ainsi, partant des matrices MAT_{Pref} et MAT_{Crit} , nous en déduisons une troisième matrice MAT_{Comp} dont chaque élément s'obtient par l'équation [WAWAK, 1996] :

$$\forall ci \in C, \forall aj \in A \quad MAT_{Comp}(ci, aj) = DCF_{MAT}$$

avec $MAT_{Crit}(ci) = \pi_{ci}$ distribution de possibilité évaluant le critère ci , et $MAT_{Pref}(ci, aj) = \mu_{ci/aj}$ ensemble flou exprimant les préférences sur le critère ci vis-à-vis de l'action aj .

L'expression générale de la compatibilité entre deux ensembles A et B ($\tau_{A/B}$) [ZADEH 78] est :

$$\mu_{\tau_{A/B}} = \begin{cases} \sup (\mu_B(x) / x \in X, a = \mu_A(x)), & \mu_{\tau_{A/B}} = \text{DCF}_{\text{MAT}} \\ 0 & \text{si } \mu_A^{-1}(a) = \emptyset \end{cases}$$

Chaque DCFh est dans notre cas une distribution de possibilité définie sur l'intervalle unité [0,1]. La prochaine étape va consister à agréger les critères de chaque action pour en déduire une évaluation unique de chaque action. A cet effet, Notre choix s'est orienté vers la moyenne arithmétique comme moyen d'agrégation des degrés de compatibilité :

$$\text{arith}(\text{DCF1}, \text{DCF2}) = [1/2(\text{DCF1}^e \oplus \text{DCF2}^e)]_r.$$

Ce choix se justifie par le fait que les opérateurs ne conservent pas tous les mêmes propriétés. Notamment, pour l'agrégation d'ensembles flous trapézoïdaux, il n'est pas évident de retrouver un ensemble flou trapézoïdal après agrégation. Par exemple, la moyenne arithmétique (arith) conserve les segments de droite alors que la moyenne géométrique (geom) les aurait déformés.

III.3. Règles floues

Une règle floue r_m peut s'écrire :

« **SI** c1 est A1m et ... et cn est Anm **ALORS** aj est DCFm » [MAMDANI et ASSILIAN, 1975].

avec Aim et DCFm ensembles flous de référence.

En prenant un exemple de notre démarche, on peut s'orienter sur la règle suivante :

« **SI** critère(c1) *Moyens financiers disponibles* est 0.8 et **SI** critère (c2) *présence Ergonome* est 0 et **SI** critère (3) *temps alloué* à l'évaluation est 0.3 **ALORS** action (aj) *ajuster moyens humains* est 0.9 »

Cela peut se lire : Si le critère *Moyens financiers disponibles* représente 80% du budget global et que l'évaluateur ergonomiste est absent (0) du groupe des évaluateurs et si le critère *temps alloué* à l'évaluation représente 30 % du temps moyen de l'évaluation d'une IHM du même type, alors l'action *ajuster les moyens humains* est recommandée à 90%. C'est-à-dire, concrètement, que le système va proposer un ajustement des moyens humains ; en d'autres termes il sera conseillé de revenir sur les ressources de départ et de corriger la donnée en mettant « ergonomiste présent » lorsque le système pose une question du type : Le groupe d'évaluation dispose-t-il d'un ergonomiste ? En corrigeant cette donnée, cela revient à dire que le groupe fera appel à un ergonomiste (extérieur ou non) au moment de l'évaluation : les méthodes et techniques que proposera le système d'aide nécessiteront la présence d'un ergonomiste dans le groupe des évaluateurs. Rappelons que lorsqu'il s'agit d'évaluer une IHM en terme d'utilisabilité (critères ergonomiques), l'apport d'un spécialiste en ergonomie prime sur d'autres spécialités, même si en principe certaines méthodes sont à tort ou à raison utilisées par des non spécialistes en ergonomie (cf. par exemple [RAVDEN et JOHNSON, 1989]).

III.4. Méthode de surclassement répondant à notre problématique (provenant de ELECTRE I)

Il s'agit de construire la relation de surclassement selon ROY [1978]. Dans un premier temps, on suppose que chaque valeur de critère g_j est une valeur dans un ensemble ordonné. A chaque critère j est attribué un poids λ_j qui représente l'importance du critère ($\lambda_j \geq 0$, $j = 1 \dots p$) ; λ_j est d'autant plus grand que le critère considéré est important. On définit alors, pour chaque couple d'éléments $(x, y) \in A^2$, l'indice de concordance $c(x, y)$ pour le surclassement de a sur b , est défini par :

$$c(x, y) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum_{i=1}^m \lambda_i, \quad m = j \text{ pour } j \in \{j / g_j(y) \geq g_j(x)\}$$

L'indice de concordance $c(x, y)$ mesure les arguments en faveur de l'affirmation "x surclasse y" ($x S y$). Cependant, les performances de x risquent d'être tellement éloignées des performances de b sur un critère qu'il ne serait pas raisonnable d'admettre $x S y$. A cette fin, on définit $d(a, b)$ l'indice de discordance qui refuse le surclassement de y par x si une discordance trop grande est observée au moins sur un critère [ROY, 1985] pour les différentes formes de $d(x, y)$. A partir de deux seuils, l'un pour la concordance \hat{c} et l'autre pour la discordance \hat{d} , on convient que :

$x S y$ si les deux conditions suivantes sont vérifiées : $c(x, y) \geq \hat{c}$ et $d(x, y) \ll \hat{d}$, avec

$$d(x, y) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum_{i=1}^m \lambda_i, \quad m = j \text{ pour } j \in \{j / g_j(y) \geq g_j(x)\}$$

Pour notre cas, à chaque couple d'actions $(a, b) \in A^2$, l'indice de concordance $c(a, b)$ pour le surclassement de a sur b , est défini par :

$$c(a, b) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum_{i=1}^m \lambda_i$$

Les critères j sont des critères définis tels que $g_j(a) \geq g_j(b)$, c'est-à-dire que seuls sont pris en compte ici des critères j tels que la valeur de l'action a ($g_j(a)$) soit supérieure à la valeur de l'action b ($g_j(b)$).

Pour mieux comprendre, prenons un exemple issu de nos travaux. Soient trois critères i_1, i_2, i_3 tels que :

- i_1 : *moyens humains*
- i_2 : *moyens financiers*
- i_3 : *moyens matériels*.

Soit à classer deux actions a et b :

a : *ajuster les moyens financiers*

b : *ajuster les moyens humains,*

alors le/les critère(s) à prendre en considération ici, pour $g_j(a) \geq g_j(b)$ (c'est-à-dire que l'action a soit prioritaire sur l'action b) est le critère $j = i_2$: *moyens financiers*.

Ainsi dans l'expression de $c(a,b)$, l'expression de $\sum \lambda_i$ concernera les degrés de compatibilité (DFCh) des critères i (i_1, i_2, i_3) ; alors l'expression $\sum \lambda_j$ concerne un critère j ($j=i_2$).

De même le classement des méthodes d'évaluation se fait avec la même logique. Nous dirons :

à chaque couple de méthodes $(M1, M2) \in M^2$, l'indice de concordance $c(M1, M2)$ pour le surclassement de a sur b est définie par :

$$c(M1, M2) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum_{i=1}^m \lambda_i$$

Les critères j sont des critères définis tels que $g_j(M1) \geq g_j(M2)$, c'est-à-dire que seuls sont pris en compte ici des critères j , tels que le rapport $g_j(M1)$ entre la méthode M1 et le critère j , soit plus important que le rapport $g_j(M2)$ entre la méthode M2 et le critère j .

Pour mieux comprendre, prenons un exemple issu de nos travaux. Soient trois critères i_1, i_2, i_3 tels que :

i_1 : *Moyens humains* (absence d'ergonome dans l'équipe d'évaluation)

i_2 : *Moyens financiers* (insuffisance des moyens financiers)

i_3 : *Moyens matériels* (existence de matériel spécialisé en évaluation) .

Soit à classer deux méthodes M1 et M2 telles que :

M1 : *Enregistrement vidéo*

M2 : *Traces écrites du Travail,*

alors le/les critère(s) à prendre en considération ici, pour $g_j(M1) \geq g_j(M2)$, est le critère :

$j = i_3$: *moyens matériels* (existence de matériel spécialisé en évaluation) .

c'est le critère qui est le plus en rapport avec la méthode M1.

En fonction des informations fournies par l'utilisateur de système d'aide à la décision, le système doit classer progressivement deux méthodes à la fois. Il prend d'abord la première méthode de la base (M1) et la classe respectivement par rapport à M2, M3, M4, etc., il recommence avec le classement par la méthode M2 par rapport à M3, M4, M5, etc. et ainsi de suite.

Lorsque M_k est mieux classée que M_1 , alors le système abandonne M_1 et classe M_k par rapport à $M(k+1)$, $M(k+2)$, $M(k+3)$, etc.

III.5 Conditions d'arrêt

Chaque méthode d'évaluation est capable de vérifier et de valider une IHM, pour un nombre de critères d'évaluation donné (attention : ne pas confondre les critères d'évaluation, comme par exemple la *facilité d'utilisation* qui est un critère ergonomique, et les critères d'aide à la décision comme les *moyens humains* qui représentent une contrainte).

Le système s'arrête lorsque l'ensemble des critères liés à des méthodes déjà sélectionnées peuvent couvrir tous les aspects d'une IHM à évaluer. En d'autres termes, une méthode M_1 est liée à r_1 critères d'évaluation et le système contient n critères d'évaluation pour un type d'IHM que l'évaluateur ou l'utilisateur aura à définir en répondant aux questions du système. Le système teste l'égalité entre (1) l'ensemble r des critères $\{r_1, r_2, r_3, \dots\}$ que les méthodes M_x peuvent vérifier et (2) l'ensemble m $\{m_1, m_2, m_3, \dots\}$ des critères existants dans la base du système.

En d'autres termes, chaque méthode M_1 donnée est capable de vérifier r_1 critères d'évaluation, et après que l'utilisateur d'ADHESION ait saisi les contraintes décrivant le contexte d'évaluation, cela permet au système de définir l'ensemble des critères d'évaluation possibles pouvant être pris en compte ; ainsi, lorsqu'il commence à sélectionner les méthodes d'évaluation les unes après les autres par le processus de classement que nous venons de voir, il arrête de sélectionner les méthodes lorsque l'ensemble des critères correspondant aux méthodes déjà sélectionnées (r) est égal à l'ensemble des critères (m). Ceci se traduit par :

Critère d'arrêt/ avec M_d la dernière proposition de méthode du système :

SI ($m_1 \in r$) et ($m_2 \in r$) et ($m_3 \in r$)... ($m_i \in r$) **ALORS** ($M_k = M_d$)

Conclusion

Dans ce chapitre, nous avons vu les différentes approches d'aide à la décision. Nous avons justifié la raison pour laquelle nous avons opté pour une approche basée sur le flou (car elle répond mieux à notre problème). Lorsque nous étudions une méthode d'aide à la décision floue, nous devons en premier lieu nous interroger sur le rôle de la modélisation floue dans cette méthode. En effet, le « flou » peut intervenir dans la modélisation des entrées du système d'aide à la décision (les valeurs des critères), mais aussi sur les préférences du décideur. De plus, nous pouvons envisager différentes problématiques concernant les sorties (décisions prises par le système). Nous pouvons vouloir déterminer la meilleure action ou un classement de ces actions. Toutes ces possibilités vont induire autant de méthodes différentes.

Le chapitre suivant décrit le système d'aide à la décision en évaluation basé sur les principes expliqués dans ce chapitre.

Chapitre IV : Première maquette de faisabilité d'un système d'Aide à la Décision en évaluation : ADHESION

Introduction

- I. Problème à traiter**
- II. Critères de décision, actions potentielles et choix stratégiques envisagés par la démarche du système**
- III. Description globale du système ADHESION**
- IV. Module de raisonnement flou implanté en Prolog III**

Conclusion

Introduction

Dans le chapitre précédent, nous avons présenté l'approche d'aide à la décision choisie ainsi que les différentes étapes de notre démarche basée sur les sous-ensembles flous. C'est une combinaison d'une très faible quantité des règles de production « floues » et d'une agrégation des degrés de compatibilité flous (DCF) issus de l'adéquation entre l'évaluation des critères et les préférences de l'expert. Dans ce chapitre sont exposés les aspects pratiques et fonctionnels de la première maquette d'un outil d'aide à la décision, baptisé **ADHESION** (Aide à la **D**écision **H**umaine pour l'**E**valuation des **S**ystèmes **I**nteractifs et leur validati**ON**) [NENDJO ELLA *et al.*, 1999a]. Avant d'exposer les aspects logiciels, nous présenterons d'abord le problème à traiter, celui du choix des méthodes et critères d'évaluation, suivi d'une définition des actions potentielles et critères d'aide à la décision. A l'issue de cette étape, il sera question des choix stratégiques et de l'architecture du système. Ce chapitre se poursuit par une description globale du système et des explications sur un ensemble de pages-écrans. Il se termine par une présentation détaillée du module de raisonnement flou.

I. Problème à traiter

Notre objectif est de mettre au point un système d'aide à la décision (ADHESION) pour le choix des méthodes et techniques d'évaluation, dont le principe est visible à la figure IV.1. L'approche choisie est basée sur les sous-ensembles flous et les possibilités. Comme nous l'avons expliqué dans le chapitre précédent, cette approche a été choisie car elle épouse bien la nature des données à manipuler. En effet, en plus des critères habituels d'acceptabilité pratique (utilité, utilisabilité), nous avons des critères d'acceptabilité socio-culturelle où la « subjectivité » est considérable. Or, les ensembles flous se prêtent autant aux données qualitatives que quantitatives, et c'est pour cela que cette approche a été privilégiée. Le support envisagé pour le développement est Prolog III qui manipule des données aussi bien qualitatives que quantitatives et intègre bien la notion de contraintes [DELAHAYE, 1988].

Pour atteindre nos objectifs, nous avons défini un ensemble de contraintes (exemple : moyens humains, moyens financiers disponibles, type d'application à évaluer, etc.) susceptibles d'influencer le choix des méthodes et techniques. Une partie de ces contraintes a été définie dans le cadre des travaux de BALBO et COUTAZ [1992], GRISLIN [1995], STANTON et NEVILLE [1996]¹, et a été complétée par des contraintes socio-culturelles [NENDJO ELLA et JACQUES, 1999]. Ces facteurs situationnels constituent les données d'entrée (les faits au sens de la conception des systèmes IA) du système. Ces données d'entrée sont introduites dans le système par le biais de réponses des utilisateurs (voir figure IV.1) répondant aux questions posées par le système. A la suite de cette série de questions – réponses, le système génère des premières propositions de méthodes et critères d'évaluation. Les utilisateurs de notre système d'aide à la décision peuvent être un chef de projet qui s'interroge sur les méthodes et critères d'évaluation

¹ Ces trois sources bibliographiques ont été choisies dans la mesure où ces auteurs recensent pédagogiquement des contraintes que l'on retrouve habituellement dans les situations d'évaluation. Bien entendu, ces contraintes ne sont pas présentées de la même manière dans ces trois sources, certaines ne sont pas présentes d'ailleurs dans une ou deux de ces sources. Il a donc été nécessaire d'en faire une synthèse, sans visée d'exhaustivité, dans la mesure où l'objectif de cette recherche est plutôt de prouver la faisabilité de la démarche d'aide à la décision en évaluation. Notons que de telles contraintes peuvent se retrouver dans d'autres documents. Elles sont par exemple disséminées dans [WILSON et CORLETT, 1990].

disponibles au regard de ses propres critères, un évaluateur inexpérimenté, un expert intéressé par des aspects spécifiques tels que les critères socio-culturels, voire un apprenant en interaction Homme(s)–Machine(s). Dans certains cas, les concepteurs peuvent également s'en servir, notamment pour prendre connaissance des critères d'évaluation que l'évaluateur est tenu de vérifier. Le système d'aide à la décision pour le choix des méthodes, techniques et critères peut donc être utilisé pour des objectifs divers, autres que l'évaluation proprement dite.

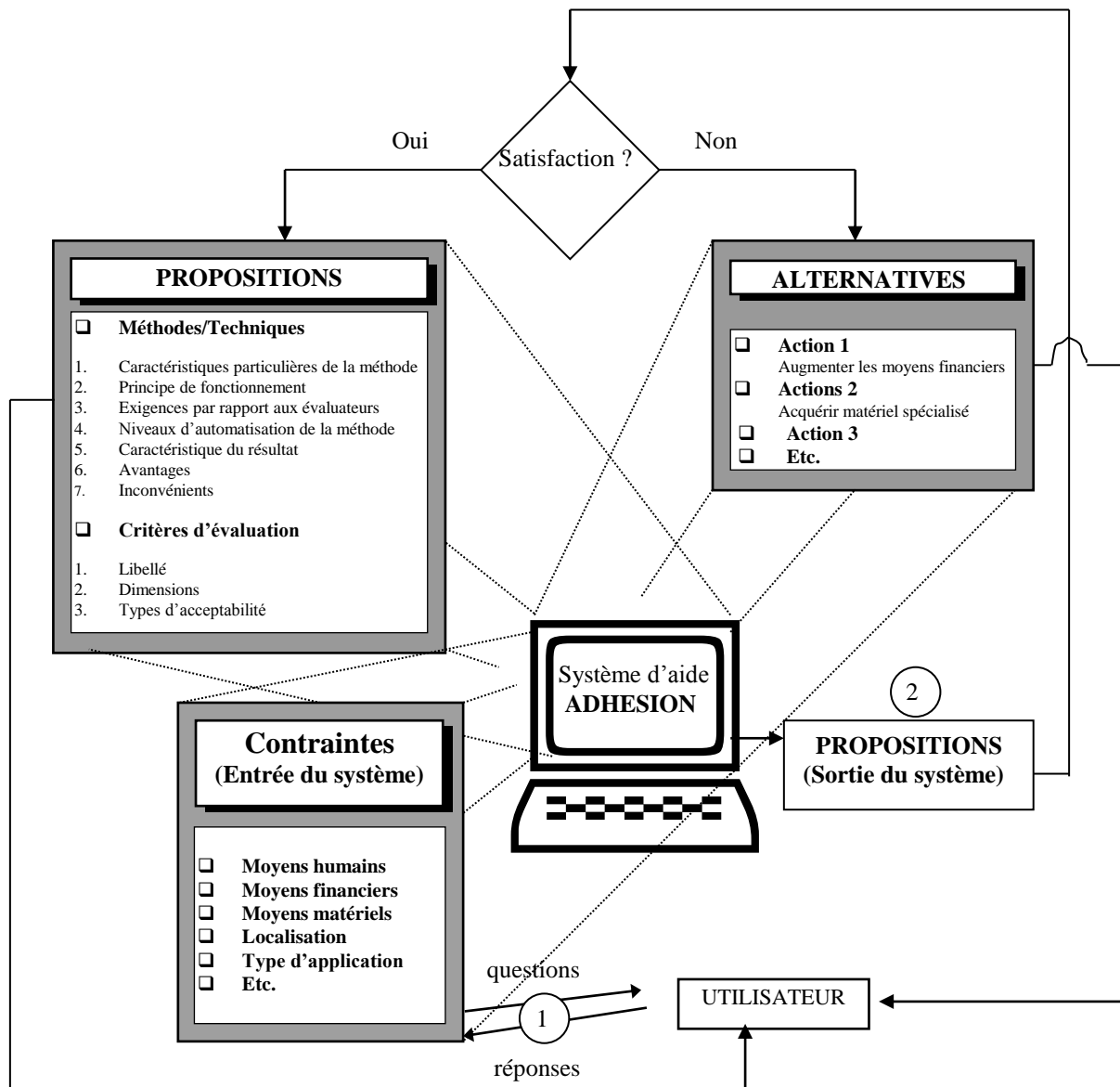


Figure IV.1 : Principe général d'utilisation du système d'aide à la décision ADHESION

En cas de non satisfaction, une aide à la décision intervient en proposant aux utilisateurs des alternatives sous forme d'actions (exemple : *Augmenter les moyens financiers disponibles*) à accomplir pour obtenir d'autres propositions de méthodes, techniques et critères. Rappelons que la non satisfaction peut être due soit à des propositions de méthodes qui ne sont pas convenables

aux yeux des utilisateurs, soit, tout simplement, à une absence de propositions. Pour ce dernier cas, il existe en effet des contraintes qui ne permettent pas de proposer aux utilisateurs du système d'aide des propositions de méthodes appropriées à leur contexte d'évaluation.

II. Critères de décision, actions potentielles et choix stratégiques envisagés par la démarche du système

Après avoir présenté le problème à traiter, les étapes de notre démarche d'aide à la décision pour le choix des méthodes, techniques et critères d'évaluation présentées au chapitre précédent, nous conduisent à la détermination des critères de décision et des actions potentielles.

II.1. Critères et actions potentielles

MAYSTRE *et al.* [1994] définissent le critère de décision comme une expression qualitative ou quantitative de points de vue, objectifs, aptitudes ou contraintes relatives au contexte réel et permettant de juger des personnes, des objets ou des événements. Selon les auteurs, pour qu'une telle expression puisse devenir un critère, elle doit être utile pour le problème considéré et fiable. Ils donnent ainsi l'exemple d'un projet de localisation d'une usine de compostage, exemple dans lequel les 5 critères de décision vont être : le prix du terrain, le prix de transport des déchets par rail, l'état de l'environnement dans la zone, le nombre de résidents touchés par les nuisances et la compétition avec d'autres projets de la zone considérée.

L'établissement d'une famille de critères de décision doit respecter une condition essentielle : la famille doit être comprise et acceptée par tous les intervenants du processus de décision. On parle alors de *famille cohérente* de critères. Cette dernière doit être conforme à trois exigences :

- ❖ exigence d'*exhaustivité* (pour cela MAYSTRE *et al.* [1994] rappellent : *il ne faut pas oublier de critères*) ;
- ❖ exigence de *cohérence* : il s'agit de trouver une cohérence entre les préférences locales de chaque critère et les préférences globales ; si l'évaluation de a_i est égale à celle de a_k sur tous les critères sauf un et qu'elle est meilleure sur ce seul critère, alors l'affirmation « a_i est préférée à a_k » peut-être émise ; ceci se traduit par :

$$g_j(a_i) = g_j(a_k) \text{ pour tout } j \neq l \text{ et } g_l(a_i) > g_l(a_k) \Rightarrow g_l(a_i) S g_l(a_k)$$
- ❖ exigence de *non-redondance* : il ne faut pas qu'il y ait des critères qui se dupliquent, donc qu'ils s'avèrent plus nombreux que nécessaire.

Dans cette optique, nous avons déterminé l'ensemble des critères de décision qui respectent ces exigences et le tableau IV.1 présente les critères de décision nécessaires pour notre système. Le sens attribué à chaque critère a fait l'objet de la fin du chapitre I. De même, l'homme d'étude va définir des *actions potentielles* que le décideur peut envisager. Cet ensemble est désigné par $A = (a_1, a_2, \dots, a_n)$ où les a_i sont des *actions potentielles*. Une action potentielle est une action jugée possible par un des intervenants au moins ou présumée telle par l'homme d'étude en vue de l'aide à la décision. Selon ROY [1985], en pratique, il est opportun de constituer déjà l'ensemble A par les *actions efficaces*. Rappelons qu'une *action efficace* est une action potentielle qui n'est

dominée par aucune autre action de l'ensemble A. Par exemple, dans notre cas du tableau IV.1, on peut voir que l'action : « faire appel à un ergonomiste extérieur » n'est pas une action efficace car elle est dominée par l'action « augmenter le budget d'évaluation ». En d'autres termes, la deuxième action est dominée par la première. En effet, si les moyens ne permettent pas d'augmenter le budget d'évaluation, on ne peut bénéficier d'une expertise extérieure en ergonomie.

	Critères principaux	Sous-critères	
	Critères de décision	Performance effective	<i>Accessibilité aux utilisateurs représentatifs (sujets d'évaluation)</i>
<i>Temps alloué à l'évaluation</i>			
<i>Localisation (site, laboratoire, pays)</i>			
<i>Type d'application</i>			
<i>Environnement où évoluera l'IHM</i>			
<i>Budget de l'évaluation</i>			
<i>Type d'interaction</i>			
<i>Etape du cycle de développement</i>			
<i>Nature des données</i>			
<i>Objet d'évaluation</i>			
<i>Support pour la capture des données</i>			
<i>Niveau d'abstraction</i>			
<i>Nature des informations (qualitative ou quantitative)</i>			
Performance potentielle des utilisateurs		<i>Origine des experts</i>	
	<i>Expérience des utilisateurs</i>		
	<i>Spécialités requises</i>		
Actions (Alternatives)	Actions potentielles		
	Actions efficaces	<i>Augmenter le budget d'évaluation</i>	
		<i>Rechercher les sujets finaux ou représentatifs</i>	
		<i>Faire l'évaluation sur site et non au laboratoire</i>	
		<i>Faire l'évaluation en laboratoire et non sur site</i>	
		<i>Attendre l'étape du cycle de développement suivant</i>	
	<i>Faire appel à un cogniticien</i>		
	<i>Augmenter le temps d'évaluation prévu</i>		
	<i>Faire appel à un ergonomiste extérieur</i>		
	<i>Faire appel à un sociologue extérieur</i>		
	<i>Faire appel à un spécialiste en assurance qualité</i>		
	<i>Faire appel à un spécialiste en droit et informatique</i>		
	<i>Acquérir un système oculométrique ou faire appel à une expertise extérieure</i>		
	<i>Acquérir un système d'enregistrement vidéo</i>		
	<i>Acquérir une expertise extérieure pour la description (modèle de l'utilisateur, de tâche et des spécifications externes)</i>		
	<i>Acquérir un ordinateur pour l'évaluation automatique</i>		

Tableau IV.1 : Critères de décision et actions

Dans le tableau IV.1, on peut distinguer 15 actions potentielles contre 5 actions efficaces. Dans le même tableau, on dénombre 16 critères de décision définis en tenant compte des travaux existants dans ce domaine et cités précédemment.

II.2. Choix stratégiques et schéma fonctionnel du système

La fonction de notre outil d'aide à la décision en évaluation est basée sur des choix stratégiques qui sont fonction de certains paramètres tels que le niveau de satisfaction des utilisateurs du système (voir figure IV.2).

Pour mieux comprendre les différentes stratégies et aspects fonctionnels du système, nous décrivons maintenant la figure IV.2 qui résume les différentes phases du système d'aide ADHESION.

- La première phase, dénommée (1) *contraintes situationnelles*, et représentant la première colonne de la figure IV.2, est la phase initiale de notre système d'aide à la décision. Cette phase est caractérisée par l'introduction des contraintes dans le système par l'utilisateur. Ces contraintes, qui correspondent aux faits du système au sens de la conception de systèmes IA, représentent les critères de décision définis précédemment.
- La deuxième phase, dénommée (2) *niveau de satisfaction*, est en fait le premier des niveaux de satisfaction de la stratégie globale. Cette *satisfaction* ou *non satisfaction* est exprimée par l'utilisateur une fois qu'il a introduit les contraintes dans le système. En fait, lorsque l'utilisateur a fini de saisir les contraintes, deux situations sont possibles : (1) soit il n'exprime aucun avis parce que les contraintes (telle le *budget de l'évaluation*) sont d'après lui largement suffisantes pour conduire l'évaluation à terme et dans ce cas il est satisfait et fait exécuter le système d'aide, (2) soit il n'est pas satisfait des contraintes qu'il vient d'introduire et dans ce cas le système lui laisse la possibilité de modifier celles qu'il estime ajustables. Dans ce dernier cas, il validera les nouvelles modifications et fera exécuter le système.
- La troisième phase, qualifiée de phase de (3) *proposition de classes de méthodes* d'évaluation, correspond au premier niveau de proposition des méthodes après l'étape (2). Les propositions fournies par le système d'aide dans cette phase sont regroupées en classes de méthodes et techniques. Chaque classe est justifiée et influencée par un groupe de contraintes de l'étape (1). Par exemple, la classe 1 est constituée par des méthodes et techniques d'évaluation qui sont très influencées par le critère de décision *temps alloué*. Ce sont généralement des méthodes qui nécessitent un temps très important pour la mise en œuvre ; lorsque les évaluateurs disposent d'un temps très court pour l'évaluation sans aucune possibilité d'allonger le temps initialement prévu, il est clair que toutes les méthodes ne peuvent appartenir à cette classe.
- La quatrième phase, dénommée (4) *niveau de satisfaction*, donne deux options à l'utilisateur : soit celui-ci est satisfait de la classe de méthodes et il la valide pour aller à la

phase (5), soit les méthodes des premières classes ne sont pas à son goût et le système vient à son aide en lui proposant les actions à mener (exemple d'action : *augmenter le budget d'évaluation*), ce qui lui permet d'obtenir de nouvelles classes de méthodes et ainsi de suite jusqu'à satisfaction. Après satisfaction, il valide les modifications et passe à la phase (5).

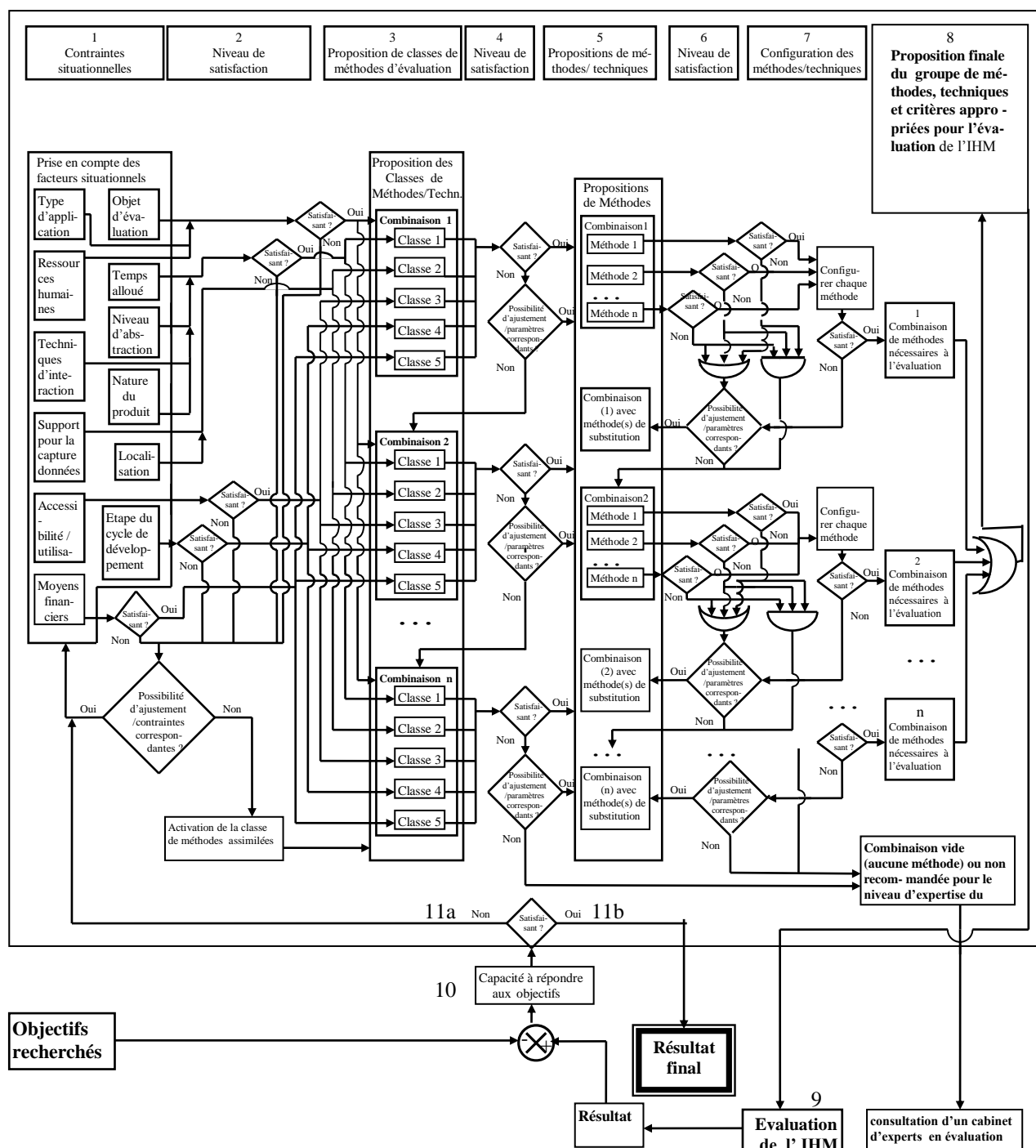


Figure IV.2 : Choix stratégiques et schéma fonctionnel

- ❑ La phase (5) est intitulée *propositions des méthodes et techniques d'évaluation*. Les méthodes ne sont pas ici fournies en classes, mais sont la combinaison des méthodes issues des différentes classes de l'étape (3). Cette fois, les méthodes et techniques d'évaluation sont les propositions devant permettre à l'utilisateur d'effectuer son évaluation.
- ❑ La phase (6) est celle du *niveau de satisfaction de l'utilisateur*, cette fois méthode par méthode. En d'autres termes, le système peut proposer un groupe de méthodes dont une ou plusieurs ne sont pas satisfaisantes à l'utilisateur. Dans ce dernier cas, le système propose à l'utilisateur des actions à mener (exemple d'action : *faire appel à un ergonome extérieur*). Concrètement, une telle action est exécutée en répondant **Très - oui** (Modificateur – Descripteur) à la question du système : « Le groupe d'évaluation dispose-t-il d'un ergonome ?, sachant qu'initialement la réponse était par exemple **Plutôt – Non**. Nous reviendrons sur ces réponses basées sur la logique « floue ». En cas de satisfaction, l'utilisateur peut configurer ou mettre en œuvre chaque méthode ; c'est l'objet de la phase (7) suivante.
- ❑ La phase suivante, dénommée (7) *configuration des méthodes*, permet à l'utilisateur de faire un essai de prise en compte et de mise en œuvre des méthodes qu'il juge satisfaisantes *a priori*. Cette phase peut être courte ou très longue selon le niveau d'expérience de l'utilisateur. Deux situations peuvent donc se présenter : la première est que l'utilisateur se rend compte que, contrairement à ce qu'il pensait, l'utilisation de la méthode engendrera des problèmes imprévus et dans ce cas le système vient en aide à l'utilisateur et lui propose des actions à mener. Dans le cas contraire, il est satisfait des méthodes et il passe à l'étape suivante (8).
- ❑ Enfin, la phase (8) s'intitule *proposition finale des méthodes*. L'utilisateur dispose à cette étape des méthodes, techniques et critères nécessaires pour l'évaluation. A l'issue de la phase (8), l'utilisateur va donc sortir du système d'aide schématisé par le grand cadre de la figure IV.2 et passe à la phase (9), celle de l'évaluation (rappelons que les phases à partir de la (9) ne sont pas traitées dans cette thèse). Après l'évaluation du système interactif, l'évaluateur compare (phase 10) les résultats de l'évaluation avec les objectifs initialement fixés par le groupe d'évaluation. Cette comparaison permet d'étudier si les méthodes ont permis de vérifier et de valider l'IHM suivant tous les aspects et exigences du cahier des charges. En cas de satisfaction (phase 11b), le résultat est considéré comme final ; dans le cas contraire (phase 11a) on revient au point de départ (phase 1) de la démarche (dans l'état actuel de cette recherche).

III. Description globale du système ADHESION

Dans cette section sont présentés les aspects logiciels de la première maquette de l'outil d'aide à la décision en évaluation. Nous justifierons le langage choisi et présenterons l'architecture et les fonctions disponibles.

III.1. Justification du choix du support d'implémentation

Pour réaliser notre programme, nous avons utilisé le langage Prolog III. Il s'agit dans un premier temps d'en faire une présentation globale et d'en rappeler les principes.

III.1.1. Présentation et principes

Prolog III est un langage qui s'appuie sur le concept de programmation logique par contraintes. Si l'on y retrouve, bien sûr, les principes généraux qui ont fait de Prolog un langage de l'Intelligence Artificielle, Prolog III représente surtout une refonte totale du cœur même du langage, l'unification DELAHAYE [1988]. Ceci se traduit par une manipulation plus fine des arbres, finis et infinis, qui demeurent les objets principaux utilisés par le langage, admis également par la génération de la notion fondamentale de *résolution des contraintes*. C'est grâce à cette notion que l'on peut, en Prolog III, dégager une approche du traitement numérique et du traitement de l'algèbre de Boole fondamentalement différente des essais qui ont été tentés dans ce domaine par l'adjonction de prédicats évaluables.

Le cœur de l'interprète est cinquante fois plus gros que celui de Prolog II, et le pari de s'attaquer à des problèmes exponentiels, comme le traitement de l'algèbre de Boole ou de la simplification de systèmes d'équations linéaires, amène parfois des déconvenues du point de vue des temps d'exécution [TREILLET, 1987]. Il faut donc rester conscient du fait que la puissance d'expression du langage voile parfois la complexité de la résolution qui en découle.

Nous pouvons, à partir de cette présentation de Prolog III, dire que notre choix est justifié par le fait que le système ADHESION doit résoudre des problèmes de contraintes. Le choix de ce langage se justifie aussi par sa faculté à manipuler des données qualitatives et quantitatives et par ses capacités de modélisation des contraintes.

III.1.2. Aspects de la programmation et primitives dans Prolog III

III.1.2.1. Les contraintes

Dans un programme écrit en Prolog III, les contraintes sont généralement exprimées au sein de systèmes de contraintes. Un système de contraintes est une suite finie de contraintes séparées par des virgules et encadrées par des accolades. Il est permis de réunir plusieurs contraintes $<$, \leq , $>$ ou \geq , à condition qu'elles soient toutes du type $<$, \leq , ou toutes du type $>$, \geq . On pourra donc écrire :

$\{2 < t < r \leq 4\}$ à la place de $\{2 < t, t < r, r \leq 4\}$

Quelques contraintes syntaxiquement correctes sont décrites ci-dessous :

- $A \# B(X, X+1)$ (1)
- $\langle \text{Methode1}, \text{"Technique d'enregistrement video"} \rangle = \langle \text{"Questionnaires"}, \text{Methode2} \rangle$ (2)
- $X \Rightarrow Z \ \& \ T$ (3)
- $0 < F < 3Y/7 + 5H$ (4)

La première expression exprime l'inégalité entre A et l'arbre B(X, X+1). La deuxième exprime l'égalité entre deux tuples de type $\langle a_1, a_2 \rangle$, donc de longueur ; pour le premier tuple, le deuxième élément est une chaîne de caractère : "Technique d'enregistrement video". La troisième expression montre une relation d'implication de X en Z & T. La quatrième exprime des contraintes à l'aide de relations d'infériorité. Il ne faut pas oublier que, contrairement aux opérateurs & et \Rightarrow , les symboles =, #, < ne sont pas des opérateurs. Ces symboles servent à exprimer des relations, lesquelles servent à écrire des contraintes.

III.1.2.2. Combinaisons entre les règles et les contraintes

Dans un programme Prolog III, une règle est de la forme :

$$A_0 \rightarrow A_1 A_2 A_3 \dots A_n, S, \quad (1)$$

où $A_0, A_1 A_2 A_3 \dots A_n$ sont des termes et **S** est un ensemble de **contraintes**. Nous savons que chaque affectation B de l'ensemble des variables de cette règle transforme les termes A_i en des arbres $X_i = A_i/B$ et l'ensemble S de contraintes en un ensemble $C=S/B$ de conditions sur des arbres. Par conséquent, une règle telle que (1) engendre tout un ensemble de règles portant sur des arbres, une pour chaque affectation B qui rend vraies les conditions $C=S/B$, de la forme :

$$X_0 \rightarrow X_1 X_2 X_3 \dots X_n ; \quad (2)$$

Une telle règle sans variable exprime une propriété logique, qui s'énonce :

« **Si** $X_1 X_2 X_3 \dots X_n$ sont des faits, **Alors** X_0 est un fait »

Dans le cas où $n=0$, on a une **règle** sans queue :

$$X_0 \rightarrow ;$$

La propriété logique correspondante est tout simplement « X_0 est un fait ».

Par exemple soit un programme de prise en compte ou de sélection des **actions** ayant un **DCF** (Degré de Compatibilité Flou) inférieur ou égal à 0,5. Le **programme** s'écrira :

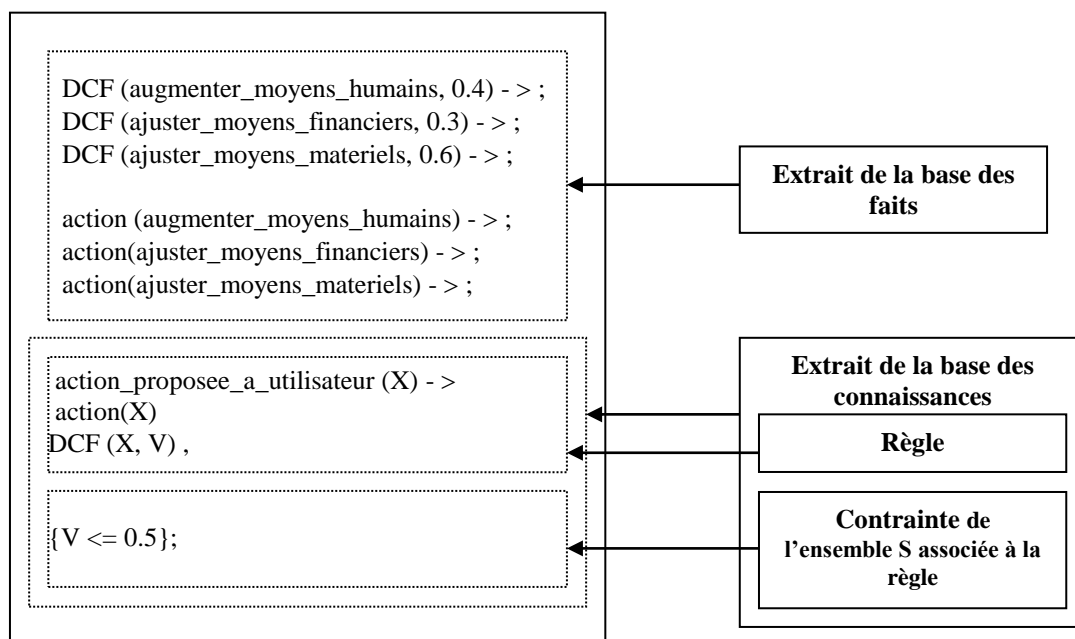


Figure IV.3 : Extrait du programme de prise en compte ou de sélection des actions ayant un $DCF \leq 0,5$

III.1.2.3. Exécution d'un programme visant à résoudre un problème

Examinons en quoi consiste l'exécution par la machine du programme. L'exécution vise à résoudre le problème suivant : « étant donné une suite $X_1X_2X_3...X_n$ de termes et un système **S** de contraintes, trouver les valeurs des variables qui transforment les termes X_i en des faits définis par le programme et les contraintes de **S** en des conditions vérifiées ». Ce problème sera soumis à la machine sous la forme d'une requête :

$$A_1A_2A_3... A_n, S.$$

A chaque fois qu'il parviendra à un état de la forme :

$$(W, , S),$$

c'est-à-dire que la suite de buts est vide, Prolog III fournira comme réponse la solution du système **S** sur l'ensemble de variables **W**.

En reprenant les bases de faits et de connaissances fournies en III.1.2., à partir de la **requête** très simple :

$$\text{action_proposee_a_utilisateur}(X),$$

des états de la machine seront (on ne montre pas les étapes promises à échouer) :

$$\{X\}, \text{action_proposee_a_utilisateur}(X), \{ \};$$

utilisation de la règle **action_proposee_a_utilisateur(X) - > etc...**

$$\{X\}, \text{action}(X) \text{ DCF}(X', V'),$$

$$\{V' \leq 0.5, \text{action_proposee_a_utilisateur}(X) = \text{action_proposee_a_utilisateur}(X')\};$$

qui se simplifie en :

$$\{X\}, \text{action}(X') \text{ DCF}(X', V'), \{V' \leq 0.5, X = X'\};$$

utilisation de la règle **action(ajuster_moyens_humains) - > ;**

$$\{X\}, \text{DCF}(X', V'),$$

$$\{V'' \leq 0.5, X=X', \text{action}(X') = \text{action}(\text{ajuster_moyens_humains})\};$$

qui se simplifie à son tour en :

$$\{X\}, \text{DCF}(X', V'),$$

$$\{V'' \leq 0.5, X=X', X' = \text{ajuster_moyens_humains}\};$$

utilisation de la règle **DCF(ajuster_moyens_humains, 0.4) - > ;**

$$\{X\}, , \{V'' \leq 0.5, X=X', X' = \text{ajuster_moyens_humains},$$

$DCF(X', V') = DCF(\text{ajuster_moyens_humains}, 0.4)$;

lequel se simplifie en :

$\{X\}, \{V' \leq 0.5, X=X', X' = \text{ajuster_moyens_humains}, V' = 0.4\}$;

La suite de buts étant devenue vide, Prolog III affiche alors la solution du système courant :

$S = \{V' \leq 0.5, X=X', X' = \text{ajuster_moyens_humains}, V' = 0.4\}$, ou plutôt la partie intéressante de cette **solution**, soit : $\{X = \text{ajuster_moyens_humains}\}$

Prolog III explore ensuite les autres exécutions possibles de la machine et on aura comme autre solution :

$\{X = \text{ajuster_moyens_financiers}\}$ (dont le DCF est de 0.3).

III.1.2.4. Primitives graphiques

En plus de sa capacité à résoudre les problèmes de contraintes, une bibliothèque des primitives est mise à la disposition du programmeur Prolog III pour la réalisation d'interfaces graphiques. Ce dernier suppose une certaine pratique des environnements graphiques multi-fenêtres et de leurs boîtes à outils. On trouve également dans Prolog III des primitives des objets structurés. Ces primitives nous ont permis de créer des sources d'écrans graphiques, qui ont ensuite été intégrées dans notre application ADHESION. Prenons un exemple très simple consistant en la réalisation d'une interface de questionnement « voulez-vous recommencer le test ? » :

Le programme sera :

```
>gr_dialog( : ccol([' Voulez-vous recommencer le test ?',
                  : buttonD,
                  : button('Annuler ', fail)]), [ ]
, L);
```

La forme générale de la primitive `gr_dialog` est définie par :

gr_dialog(D, I1, I2) où,

D est un arbre décrivant un dialogue, soit l'identificateur d'une règle,

I1 est une liste (éventuellement vide) de paires [nom_de_zone | valeur initiale] ; cette liste est utilisée pour définir ou redéfinir le contenu des zones,

I2 est la sortie ; si elle est une variable, elle est unifiée avec la liste des couples [nom_de_zone | valeur finale] du dialogue,

ccol (l) décrit une colonne centrée verticalement, dont le contenu est décrit par la séquence l.

Le résultat obtenu est celui de la figure IV.4 ci-dessous.

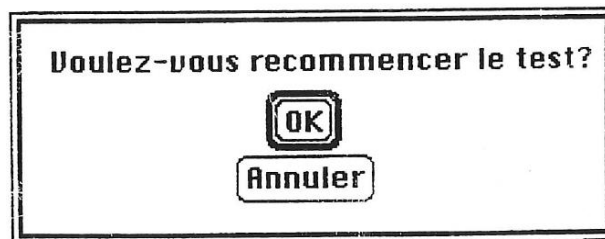


Figure IV.4 : Exemple très simple d'utilisation des primitives graphiques

III.2. Architecture du système d'aide à la décision ADHESION et fonctionnalités

Le système d'aide ADHESION se compose d'un module de génération de méthodes et techniques et d'un module de contrôle (MC) qu'exploite l'évaluateur à partir d'une IHM spécialisée pour établir un ensemble de modifications permettant de compenser les propositions non satisfaisantes (figure IV.4). Rappelons que cette première maquette a été développée à l'aide de Prolog III (sur MacIntosh d'Apple).

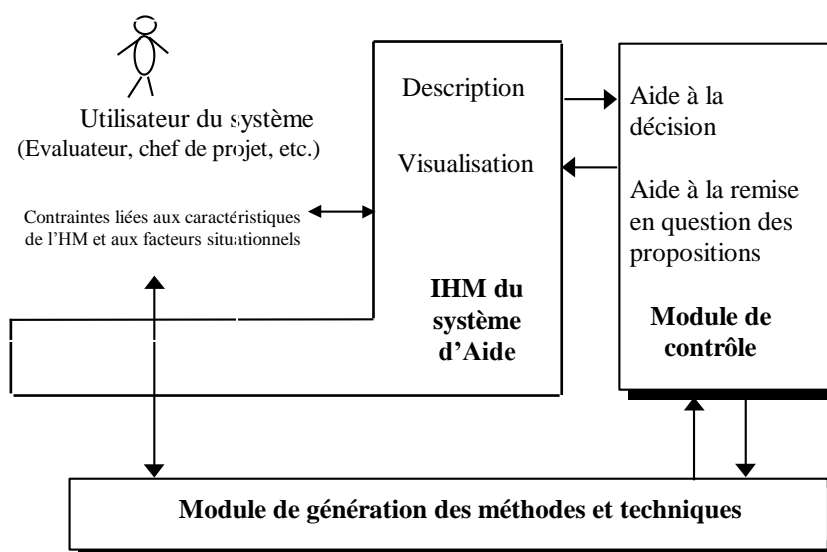


Figure IV.4 : Architecture du système d'aide à la décision [NENDJO ELLA *et al.*, 1999b]

* **Le module de contrôle (MC)** intègre les différents éléments destinés à assister le processus décisionnel de l'évaluateur. A partir de l'IHM, l'évaluateur exploite le MC pour établir un ensemble de modifications permettant de compenser des propositions de méthodes et techniques non satisfaisantes.

* **L'interface (IHM)** permet deux formes d'interaction : la première permet de dialoguer avec le module de génération des méthodes et techniques ; la seconde permet de solliciter le MC.

* **Le module de génération des méthodes et techniques** propose des méthodes et techniques même en cas de non satisfaction par l'évaluateur ; il reçoit ainsi des consignes du MC pour effectuer d'autres propositions de méthodes et techniques.

III.3. Explications liées à la maquette développée

Après avoir présenté l'architecture du système ADHESION, il est utile de continuer cette description sur des pages-écrans représentatives. Nous commençons par la page-écran d'accueil du système ADHESION (voir figure IV.5). Cette page ne fera bien entendu pas l'objet d'explications particulières, au contraire des suivantes. Après la page-écran d'accueil, le système présente immédiatement à l'utilisateur des explications sur certains termes qui pourraient compliquer la compréhension du système d'aide à la décision (voir figure IV.7).

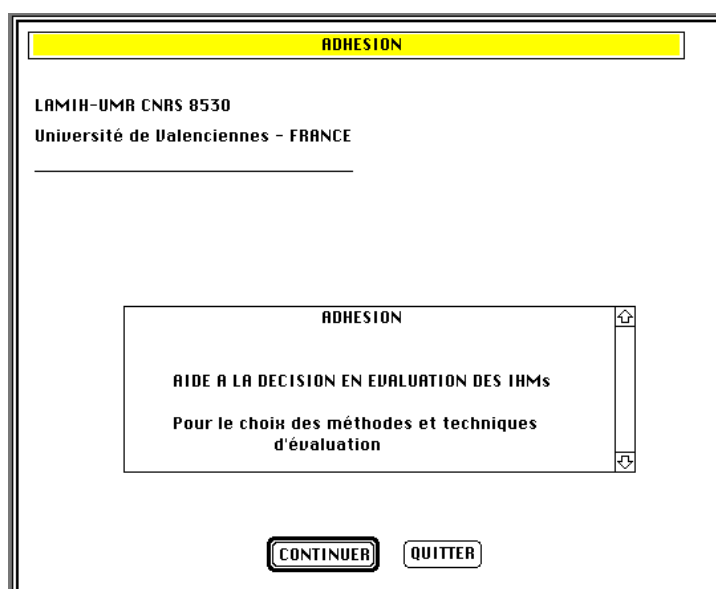


Figure IV.5 : Page-écran d'accueil du système d'aide ADHESION

La page-écran visible figure IV.6 fournit des explications sur des termes que l'utilisateur rencontrera au fur et à mesure de sa progression dans le système ADHESION. On distingue sur la partie gauche de la page-écran 14 termes qui sont expliqués. Il lui est conseillé d'en prendre connaissance avant de poursuivre, en agissant sur la barre de défilement de droite pour accéder aux explications correspondantes. Par exemple, il comprendra que le premier terme « *objet d'évaluation* » est le produit du processus de développement, tels le modèle de tâche, le modèle de l'utilisateur, les spécifications externes, la maquette, le produit final, etc. [BALBO, 1994]. Après cette page-écran d'explication des termes, l'utilisateur accède à la page-écran principale permettant de saisir les contraintes et constituer ainsi la base des faits (voir figure IV.7). Sur cette

figure, on peut distinguer 3 grandes zones : une zone 1 (en haut à gauche) de sélection du pays de destination de l'IHM à évaluer, une zone 2 (en haut à droite) d'explication du mode de réponse de la zone 3 et, enfin, une zone 3 essentielle dite de « questions-réponses ».

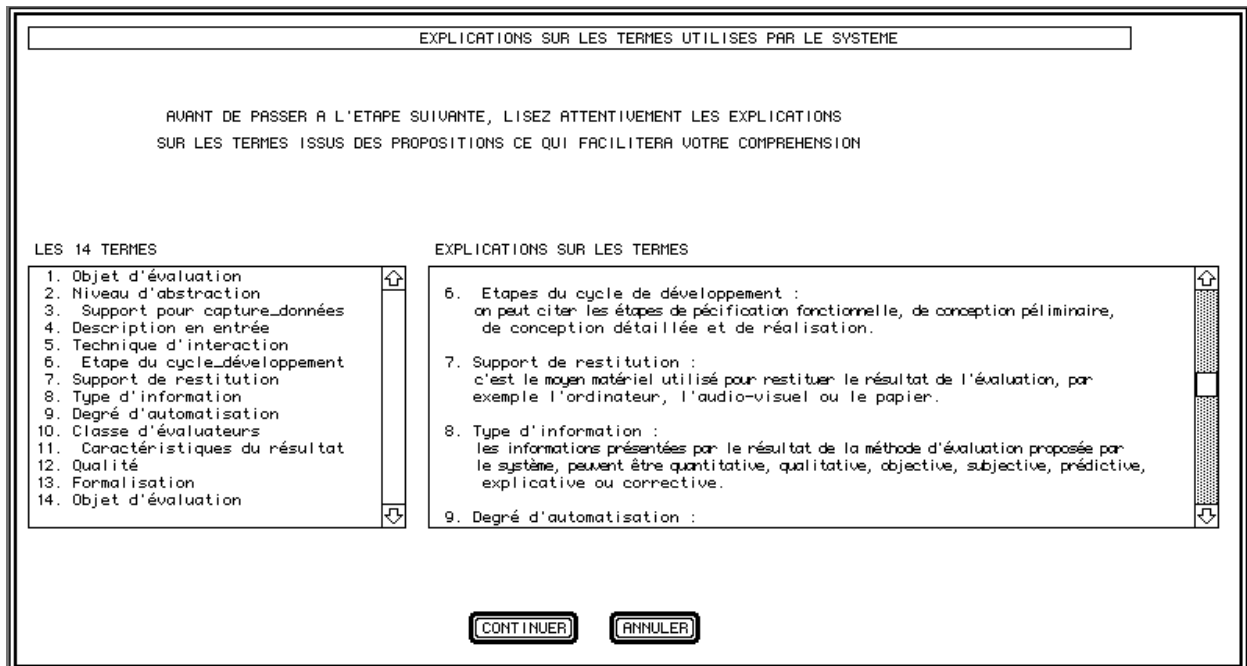


Figure IV.6 : Page-écran des explications des termes dans le système ADHESION

- La première zone¹ permet à l'utilisateur de sélectionner le pays de destination de l'IHM qu'il désire évaluer, dans un but final de prise en compte des aspects socio-culturels. Par exemple, une entreprise basée en France et qui doit évaluer un système interactif destiné à être utilisé en Albanie sélectionnera « Albanie ». Cette base contient plus de 212 pays et îles dont les aspects socio-culturels ont été étudiés avec notre collaboration par une spécialiste de la socio-technique du CLERSE (Centre Lillois d'Etudes et de Recherches Sociologiques et Economiques) [NENDJO ELLA et JACQUES, 1999], comme nous l'avons expliqué dans le chapitre I, section III.3.
- La deuxième zone représente l'aide à la compréhension des questions (visible en zone 3). Cette partie permet à l'utilisateur de comprendre le sens des questions et le mode de réponse qui utilise le raisonnement « flou ». Cette deuxième partie reprend chacune des questions de la zone 3 et explique le sens des réponses.

¹Dans cette partie, on peut sélectionner le pays de destination ou assimilé : ce terme assimilé a été prévu pour des territoires devenus pays indépendants. Par exemple, si le Kosovo devenait un pays indépendant, le système considérerait que la séparation du Kosovo de la Yougoslavie n'engendrerait pas de différence lors de l'évaluation, sous l'angle des critères socio-culturels. Et si le Kosovo, qui ne figure pas parmi les 212 pays, devenait donc indépendant, on sélectionnerait comme pays assimilé la Yougoslavie.

- ❑ La principale zone (la 3) est celle liée aux questions-réponses que l'on retrouve également dans la figure IV.1 intitulée « Principe général d'utilisation du système d'aide à la décision ADHESION ».
- ❑ Cette partie permet à l'utilisateur d'introduire des contraintes et de contribuer ainsi à la réalisation de la base de faits nécessaire au fonctionnement du système. La saisie de ces contraintes correspond de surcroît à la prise en compte des critères de décision par le système. Ayant adopté un système d'aide à la décision basé sur les sous-ensembles « flous », nous avons opté, pour des raisons de cohérence de programmation, pour un système de réponses selon l'esprit de la logique floue. L'utilisateur doit ainsi manipuler des termes issus des sous-ensembles flous, tels le *descripteur* (exemple : **oui/bon, moyen, mauvais/non**) et le *modificateur* (exemple : **plutôt, modérément, très**). Le modificateur a pour rôle de nuancer la réponse caractérisée par le *descripteur*.

DESCRIPTION DES CONTRAINTES

CETTE TECHNIQUE DE RESOLUTION VOUS OBLIGE A REPENDRE
A TOUTES LES QUESTIONS CI-DESSOUS

Sélectionnez le pays de destination (ou assimilé) de l'IHM à évaluer

Afghanistant

- Afrique du Sud
- Albanie
- Algérie
- Allemagne
- Andorre
- Angola
- Antigua-et-Barbuda
- Antilles Neerlandaises
- Arabie Saoudite

AIDE A LA COMPREHENSION DES QUESTIONS

Avez-vous accès aux utilisateurs finaux ?
Répondre Plutôt-Moyen, signifie que vous avez près de 50% de chance d'accéder aux futurs utilisateurs de l'IHM que vous voulez évaluer, alors que Très - Oui signifie 100% de chance.

Le Groupe d'évaluation, dispose t-il d'un ergonomes ?
Répondre Très - Non signifie que vous ne disposez pas du tout d'ergonomes, alors que Plutôt-Moyen signifie qu'il existe une chance sur deux pour que vous obtenez un ergonomes, cela peut être le cas d'une attente

***** Descripteurs ***** ***** Modificateurs *****

	Bon/Oui	Moyen	Mauvais/Non	Plutôt	Modérément	Très
Avez-vous accès aux utilisateurs finaux?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un ergonomes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un sociologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un psychologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en droit-informatique?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en assurance qualité?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le budget-évaluation vous semble t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le temps alloué à l'évaluation vous semble t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous financer le dépassement du temps prévu?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous financer les spécialistes d'évaluation manquant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le système à évaluer évoluera t-il dans un environ.nuisible?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Disposez-vous du matériel spécialisé (système oculométrique.,enreg.vidéo,)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avez-vous une expérience en évaluation ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous assimiler le système à évaluer à un type grand public?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
L'IHM à évaluer a t-elle franchit l'étape de réalisation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure IV.7 : Page-écran de saisie des contraintes

La philosophie du « flou » est l'antithèse du oui ou du non, du 1 ou 0, du « tout ou rien ». C'est ainsi qu'au lieu de répondre tout simplement par oui, non ou mauvais, on fera intervenir un modificateur qui va modifier, atténuer, nuancer et introduire l'incertitude dans la réponse par l'addition des modificateurs. Par exemple, à la première question : « Avez-vous accès aux utilisateurs ? » (ce qui correspond au critère de décision *Accessibilité aux utilisateurs représentatifs*, tableau IV.1), l'utilisateur répondra en fonction de son contexte, par exemple par « **oui** (*descripteur*) – **modérément** (*modificateur*) ». Si on avait répondu tout simplement par « **oui** » (réponse classique), cela pourrait s'interpréter par un accès facile et constant aux utilisateurs représentatifs, ce qui n'est certainement pas le cas ici².

Ainsi, le raisonnement flou introduit le *modificateur modérément* qui permet en fait de ne pas prendre « pour argent comptant » les prévisions de l'utilisateur qui pourrait être confronté au dernier moment à un problème lors de la recherche des utilisateurs représentatifs (sujets d'évaluation). Il se pourrait, par exemple, que ces sujets ne soient pas présents ou représentatifs, etc. Les réponses basées sur le raisonnement flou, certes plus incertain, se veulent donc plus réalistes en laissant en quelque sorte une marge de sécurité pour les imprévus que l'évaluateur pourrait rencontrer au moment de l'évaluation. Ainsi « **oui – modérément** » signifierait que l'on est sûr avec une certitude située autour de 50%.

La figure IV.8 est un exemple de page-écran de proposition de méthode. Sur cette figure apparaissent six grandes zones : la première donne des informations supplémentaires sur la méthode proposée ainsi que sur les caractéristiques de l'IHM à évaluer ; la deuxième zone en haut à droite explique le principe de fonctionnement de la méthode, ce qui donne déjà à l'utilisateur une première idée des difficultés qu'il pourra rencontrer lors de l'utilisation effective de la méthode ; une troisième zone présente les exigences physiques et mentales que peut globalement nécessiter cette méthode ; trois autres zones fournissent enfin des informations sur les niveaux d'automatisation de la méthode, les critères d'évaluation, les avantages et inconvénients et les caractéristiques du résultat d'évaluation obtenu avec la méthode proposée par le système.

□ La première zone est constituée par :

- le *libellé de la méthode*, qui indique le nom de cette dernière,
- la *méthode ou technique complémentaire*, qui mentionne la méthode que l'utilisateur peut également prendre en considération pour renforcer son évaluation,
- la *description en entrée*, qui précise si la mise en œuvre de cette méthode nécessite une activité de description d'un modèle de tâche, d'un modèle utilisateur, etc.,
- la *disponibilité de la méthode*, qui permet de savoir si la méthode a dépassé l'étape de prototype par exemple et est prête à être convenablement utilisée,
- la *formalisation*, qui indique si c'est une méthode formelle,

²Nous nous sommes contentés de ce premier niveau de description dans cette première maquette du système ADHESION. Une perspective de recherche consisterait bien entendu à détailler beaucoup plus cette description.

- la *technique d'interaction*, qui mentionne le mode d'interaction (exemple : système multi-utilisateur) que la méthode proposée est capable d'évaluer ; même remarque pour l'objet d'évaluation,
- la *classe d'évaluateurs exigée* rappelle enfin la catégorie d'évaluateurs pouvant en principe utiliser cette méthode.

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°3				
DESCRIPTION DE LA METHODE ET DES CRITERES				
CARACTERISTIQUES DE L'HM A EVALUER		PRINCIPE DE FONCTIONNEMENT DE LA METHODE		
Libellé de la méthode	: Enregistrement vidéo	Le principe consiste à enregistrer objectivement l'activité de l'utilisateur pour estimer sa performance et celle du système homme-machine (vidéo, interviews ou questionnaires). Cette approche permet d'identifier les moments où les difficultés sont rencontrées en recherchant les explications a posteriori en fonction des faiblesses de l'interface.		
Méthode ou Technique complémentaire	: Questionnaires			
Description en entrée (ex: description de tâche)	: Pas nécessaire			
Disponibilité de la méthode	: Oui			
Formalisation	: Méthode non formelle			
Technique d'interaction	: Manipulation directe			
Objet d'évaluation	: Produit final ou Prototype			
Classe d'évaluateurs exigée	: Experte			
Support pour la Capture de données	: Vidéo			
Etapes du cycle de développement concernées	: Spécification/conception/après réalisation			
Usage de métriques	: Pas nécessaire			
POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS		CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION		
moins important important Très import.		Qualité : Inspectabilité		
Requêtes des évaluateurs	: <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Support de restitution : Vidéo		
Conflits entre évaluateurs	: <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>	Support d'analyse des résultats : Vidéo		
Exigence mentale	: <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	Niveau d'abstraction : Tous		
Exigence physique	: <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>	Type d'information : Explicative		
NIVEAUX D'AUTOMATISATION DE LA METHODE		Performance de la méthode ou Technique : <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne		
Capture automatique des données	: <input checked="" type="checkbox"/>			
Restitution automatique des données après traitement	: <input type="checkbox"/>			
Analyse automatique	: <input type="checkbox"/>			
Critique automatique	: <input type="checkbox"/>			
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION		AVANTAGES		
N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.	INCONVENIENTS
1.	Contrôle explicite	Pratique	utilisabilité	Obligatoire Non Non 2
2.	Groupements	Pratique	utilisabilité	
3.	Guidage	Pratique	utilisabilité	
4.	Qualité des messages	Pratique	utilisabilité	
5.	Signification des codes	Pratique	utilisabilité	
6.				
7.				
REFERENCE SUR LA METHODE		: Youmans, 1987		
<input type="button" value="CONTINUER"/> <input type="button" value="ANNULER"/>				

Figure IV.8 : Exemple de proposition de méthode

- La zone « Positionnement par rapport aux évaluateurs » apporte des informations complémentaires :
 - les *requêtes des évaluateurs* indiquent si cette méthode oblige les évaluateurs à recourir à beaucoup d'informations pour la mise en œuvre,
 - les *conflits entre évaluateurs* mentionnent l'importance des désaccords possibles entre les évaluateurs lors de l'évaluation à l'aide de cette méthode ; il en est de même pour les exigences physiques et mentales,

- l'*exigence mentale* indique si la méthode ou technique proposée par le système ADHESION exige des évaluateurs un effort mental important (même remarque pour l'*exigence physique*).
- La zone «Niveaux d'automatisation de la méthode» apporte les informations suivantes :
- la *capture automatique des données* précise si la méthode ou technique permet de capturer automatiquement les données qui seront traitées et analysées par la suite,
 - la *restitution automatique des données après traitement* indique si la méthode ou technique restitue automatiquement les données déjà traitées, l'*analyse automatique* précisera si la proposition de méthode ou technique est capable de faire une analyse automatique dans le cadre de l'évaluation ; même remarque pour la *critique automatique*.
- La zone «Caractéristiques du résultat» apporte des informations sur certains aspects du résultat :
- la *qualité* regroupe des critères généraux qui dénotent la robustesse des résultats (l'utilisateur du système ADHESION prendra connaissance des explications supplémentaires dans la page-écran *explications des termes* de la figure IV.7),
 - les *supports de restitution ou d'analyse des résultats* mentionnent les moyens utilisés (ordinateur, crayon papier, etc.) pour restituer ou analyser les résultats,
 - le *niveau d'abstraction* des informations peut correspondre à : lexical, physique et syntaxique ou à niveaux des tâches,
 - le *type d'information* exprime si les informations issues du résultat sont subjectives, objectives, etc.
- La zone «Avantages et inconvénients» apporte des informations sur les limites et les apports de la proposition de méthode ou technique :
- les *coûts* représentent globalement les frais engendrés par les évaluateurs (salaires, frais divers, etc.), le matériel, les sujets d'évaluation, etc.,
 - le *temps moyen* est le temps approximatif qu'il faut pour effectuer l'évaluation,
 - l'*accessibilité aux utilisateurs* indique si l'usage de cette méthode ou technique oblige à recourir aux utilisateurs représentatifs ou non, ce qui représente un inconvénient en terme de coût de revient de mise en œuvre de la méthode, mais un avantage en terme de résultat,
 - cette partie donne également des informations telle que la *facilité d'utilisation* ou de mise en œuvre de la méthode, le *nombre moyen des sujets* d'évaluation et des *évaluateurs*,
- La zone «Critères d'évaluation correspondant à la proposition» permet à l'évaluateur de prendre connaissance des critères d'évaluation que cette méthode ou technique est capable de vérifier :

- le *libellé* donne le nom du critère,
- l'*acceptabilité* précise si le critère est d'ordre pratique ou socio-culturel,
- la *dimension* indique celle dont dépend le critère (utilité, utilisabilité, culturelle, éthique, etc.).

III.4. Conclusion sur la première version d'ADHESION

Après cette description globale du système, illustrée de pages-écrans, nous pouvons dire qu'ADHESION présente une qualité ergonomique et des aspects fonctionnels suffisants pour prouver la faisabilité de la démarche visée. Les apports et limites d'ADHESION seront d'ailleurs discutés dans la suite du mémoire. La section suivante est consacrée au module de raisonnement flou implanté en Prolog III.

IV. Module de raisonnement flou implanté en Prolog III

Dans cette section, il sera question de la constitution de la base des faits, des algorithmes de détermination des DCFs (Degrés de Compatibilité Flou) [DUBOIS et PRADE, 1985], des propositions des méthodes et actions à mener (alternatives) à l'aide de la méthode ELECTRE I [ROY, 1978 ; 1985] fondée sur la notion de surclassement. Ces phases correspondent aux étapes de notre démarche (voir figure IV.9, ci-après). La manière dont elles sont implantées en Prolog III sont successivement décrites.

IV.1. Constitution de la base des faits et algorithme

Nous avons vu en III.3 comment l'utilisateur devait introduire les contraintes, qui sont à la fois des critères de décision, selon le raisonnement de la logique floue. En terme de programmation, nous avons utilisé le modèle des ensembles flous trapézoïdaux (voir figure IV.9) [DUBOIS et PRADE, 1985], qui permettront une évaluation des critères de décision. Veillons ici à ne pas confondre l'évaluation des IHMs et l'évaluation de critères de décision que l'on retrouve dans la démarche d'aide à la décision, de même qu'il ne faut pas confondre les critères de décision et les critères d'évaluation.

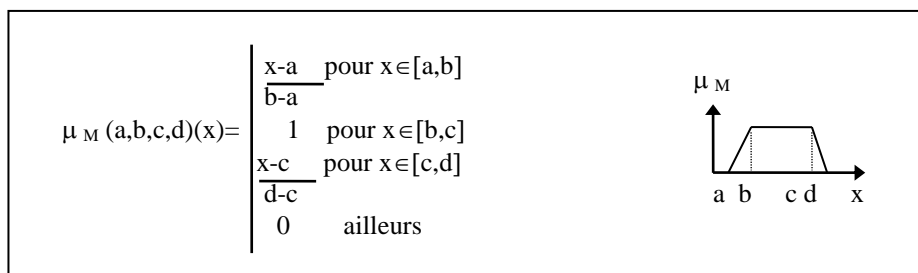


Figure IV.9 : Le modèle des sous-ensembles flous trapézoïdaux [ZADEH, 1975], [DUBOIS et PRADE, 1985]

Rappelons que dans le chapitre III, nous avons donné la signification des variables a, b, c, d et précisé que les valeurs de la fonction μ_M sont comprises entre 0 et 1.

Les utilisateurs répondent aux questions à l'aide des descripteurs et modificateurs comme nous l'avons vu en III.3. Pour pouvoir utiliser la méthode de ZADEH [1975], les réponses seront modélisées par des descripteurs ($x_1 = \text{Bon/Oui}$, $x_2 = \text{Moyen}$, $x_3 = \text{Mauvais/Non}$), modifiés par des modificateurs ($x_1' = \text{Modérément}$, $x_2' = \text{Plutôt}$, $x_3' = \text{Très}$), que l'on modélisera respectivement par deux vecteurs D et M de la manière suivante :

- $D(x_1, x_2, x_3)$ pour les descripteurs, et
- $M(x_1', x_2', x_3')$ pour les modificateurs.

Ainsi, pour tout couple de réponses (x_i, x_i') issu des deux vecteurs $D(x_1, x_2, x_3)$ et $M(x_1', x_2', x_3')$, le système construit une fonction trapézoïdale $\mu_M(a,b,c,d)$. En fait, pour le système, ceci revient à trouver tout simplement les valeurs de a, b, c, d en fonction de chaque couple de réponses (x_i, x_i'), comme le couple (Oui, modérément) par exemple. Le système, dans sa conception, contient déjà toutes les combinaisons des couples (x_i, x_i') et chaque fois que l'utilisateur les introduit, le système les reconnaît et affecte immédiatement la fonction $\mu_M(a,b,c,d)$ correspondante.

Par exemple, pour (Bon, Très), le système ira chercher dans sa base de données la fonction correspondante : ($a = 0.3, b = 0.4, c = 0.6, d = 0.8$). Les différentes valeurs (a, b, c, d) existantes dans le système ont été introduites dans la base lors de la programmation et proviennent d'une étude préliminaire qui met en jeu les différents rapports d'importance entre les critères de décision.

Prenons l'exemple suivant pour mieux comprendre le travail préliminaire qui a permis d'introduire dans le système les 15 combinaisons (a, b, c, d) correspondant aux différents couples (x_i, x_i') (rappelons qu'il y a 15 combinaisons car il existe comme nous le verrons ci-après 15 couples de réponses (x_i, x_i') possibles) :

Soit le critère de décision *budget de l'évaluation*. A la question : le budget d'évaluation vous semble-t-il suffisant ?, supposons que l'utilisateur réponde par *oui* (comme descripteur) et choisisse un modificateur entre *très, modérément et plutôt*, ce qui donne un pourcentage qui va de 100% à près de 50%. Ainsi, si l'utilisateur choisit le couple (*oui, modérément*), dans l'étude préliminaire on affectera un pourcentage à ce couple qui reste une estimation avec une marge non négligeable.

Ainsi dans cet exemple, on affectera au couple de réponse (*oui, modérément*) un pourcentage entre 60 et 100% et pour ce critère (*budget d'évaluation*) le pourcentage (%) est relatif au budget que l'on estime largement suffisant pour ce type d'évaluation (ce qui nécessite pour l'utilisateur la connaissance des budgets affectés à d'autres évaluations effectuées dans les mêmes conditions pour positionner son budget par rapport aux autres). Le système ADHESION prend une base de 50 à 100% d'un budget représentatif d'une évaluation donnée et estime donc qu'entre 50 et 100% de ce budget d'évaluation « idéal », le critère budget d'évaluation devrait être suffisant. Ceci correspond à la fonction suivante :

$$\mu_{\text{plutôt bonne}} = M(50,60,100,100 \text{ en } \%) = M(0.5,0.6,1,1) \text{ (voir figure IV.10).}$$

En utilisant l'expression des degrés d'appartenance ci-dessous, on a :

$$\mu_M(a,b,c,d)(x) = \begin{cases} \frac{x-a}{b-a} & \text{pour } x \in [a,b] \\ 1 & \text{pour } x \in [b,c] \\ \frac{x-c}{d-c} & \text{pour } x \in [c,d] \\ 0 & \text{ailleurs} \end{cases}$$

On obtient avec (a =50, b = 60, c=100, d = 100) la courbe ci-dessous :

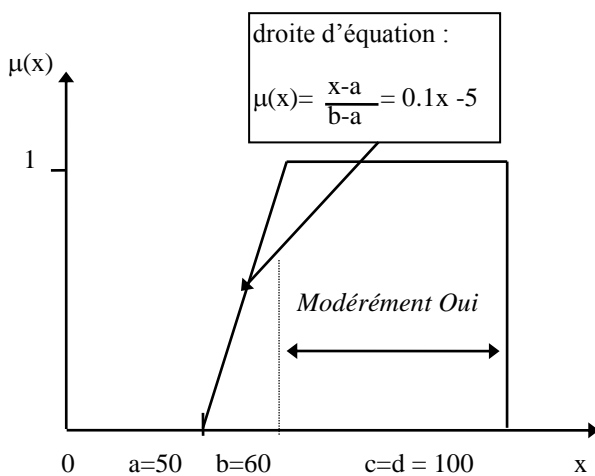


Figure IV.10 : Evaluation du critère de décision *budget d'évaluation*

Les différents couples de réponses sont stockés dans une base $f\{(y_1, y_1'), (y_2, y_2'), (y_3, y_3'), \dots, (y_n, y_n')\}$ permettant d'évaluer les critères de décision, sachant que dans le système, à chaque couple (y_i, y_i') correspond un vecteur $\mu_M(a,b,c,d)$ prédéterminé. La **partie du programme** (figure IV.11) et la solution qui donne la base $f\{(y_1, y_1'), (y_2, y_2'), (y_3, y_3'), \dots, (y_n, y_n')\}$ est présentée à la figure 12.

Ceci correspond aux combinaisons existantes dans le système $\{(y_1, y_1'), (y_2, y_2'), (y_3, y_3'), \dots, (y_n, y_n')\}$

L'utilisateur va donc introduire le couple (x_i, x_i') et le système va chercher à le reconnaître en le comparant à chaque couple (y_i, y_i') existant dans le système.

Tout ceci revient à tester la contrainte $(x_i, x_i') = (y_i, y_i')$, c'est-à-dire que l'on *ajoutera* au programme ci-dessus la règle suivante :

Choix2 ([A, B]) -> Descripteur (A) Modificateur (B) {A=X, B=Y} ;

(extrait de la base de connaissances)

avec $\{A=X, B=Y\}$ comme contrainte.

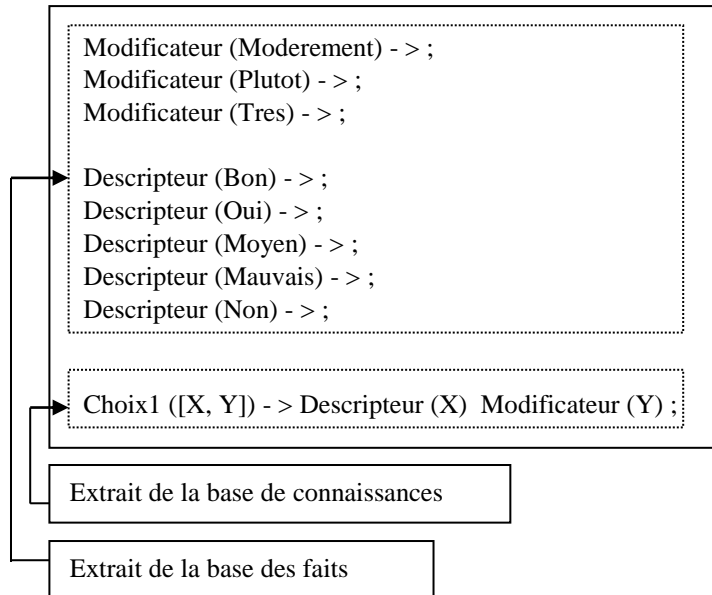


Figure IV.11 : Extrait du programme

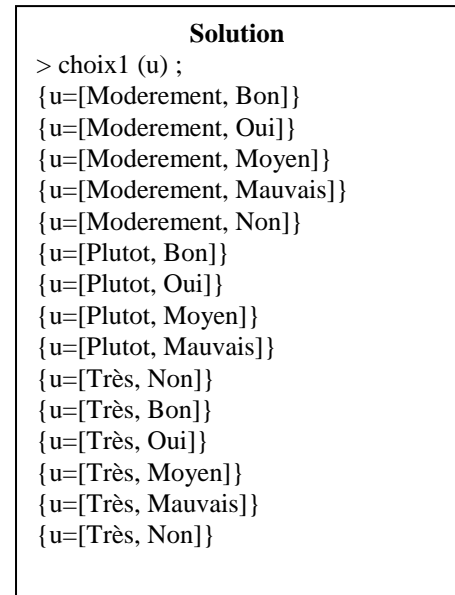


Figure IV.12 : Extrait de solution
(obtenue ici dans l'environnement de l'interpréteur de Prolog III)

Donc, à chaque couple (x_i, x_i') , ici (A, B) , fourni par l'utilisateur, la résolution consistera à trouver le couple correspondant (y_i, y_i') , sachant que dans le système il existe un programme qui fait correspondre à chaque couple (y_i, y_i') un vecteur $\mu_M(a,b,c,d)$ (sous-ensemble trapézoïdal représentant la mesure ou l'évaluation du critère). Cette évaluation sera comparée par la suite aux préférences de l'expert afin de déterminer le DCF (Degré de Compatibilité Flou).

$$\mu_{\tau A/B} = \sup (\mu_B(x) / x \in X, a = \mu_A(x)) , \quad \mu_{\tau A/B} = \text{DCF}_{\text{MAT}} \\ = 0 \text{ si } \mu_A^{-1}(a) = \emptyset$$

Nous obtenons à l'issue de cette étape de l'algorithme une matrice MAT_{Comp} de dimension $n \times p$ contenant les DCFs de chaque critère vis-à-vis des préférences pour chaque action. Chaque DCF est dans notre cas une distribution de possibilité définie sur l'intervalle unité $[0,1]$. La prochaine étape va consister à agréger les critères de chaque action pour en déduire une évaluation unique de chaque action.

IV.2. Algorithme de calcul des DCFs (Degrés de Compatibilité Flous)

Nous avons vu dans le chapitre 3 que l'expression permettant de déterminer les degrés de compatibilité flous s'exprime sous la forme :

$$\forall c_i \in C, \forall a_j \in A \quad \text{MAT}_{\text{Comp}}(c_i, a_j) = \text{DCF}_{\text{MAT}} \text{ avec,}$$

c_i un critère de décision,

a_j une action,

MAT_{Comp} est égale au DCF et est déterminée à partir des matrices MAT_{Pref} et MAT_{Crit} . $MAT_{Crit}(ci)=\pi_{ci}$ correspond à une distribution de possibilités évaluant le critère ci , et $MAT_{Pref}(ci,a_j)=\mu_{ci/a_j}$ un ensemble flou exprimant les préférences sur le critère ci vis-à-vis de l'action a_j .

L'expression générale de la compatibilité entre deux ensembles A et B ($\tau_{A/B}$) [ZADEH, 1978] est donnée par :

$$\begin{aligned}\mu_{\tau_{A/B}} &= \sup (\mu_B(x) / x \in X, a = \mu_A(x)) \\ &= 0 \text{ si } \mu_A^{-1}(a) = \emptyset\end{aligned}$$

$$\mu_{\tau_{A/B}} = DCF_{MAT}$$

Cette expression donne une évaluation floue sur [0,1] du degré de compatibilité de la mesure d'un critère par rapport à un modèle de préférence de l'expert. Pour mieux comprendre, prenons un exemple dans nos travaux, celui du critère de décision *temps alloué*. La figure IV.14, montre la détermination par le système du DCF connaissant l'évaluation du critère de décision et la préférence de l'expert.

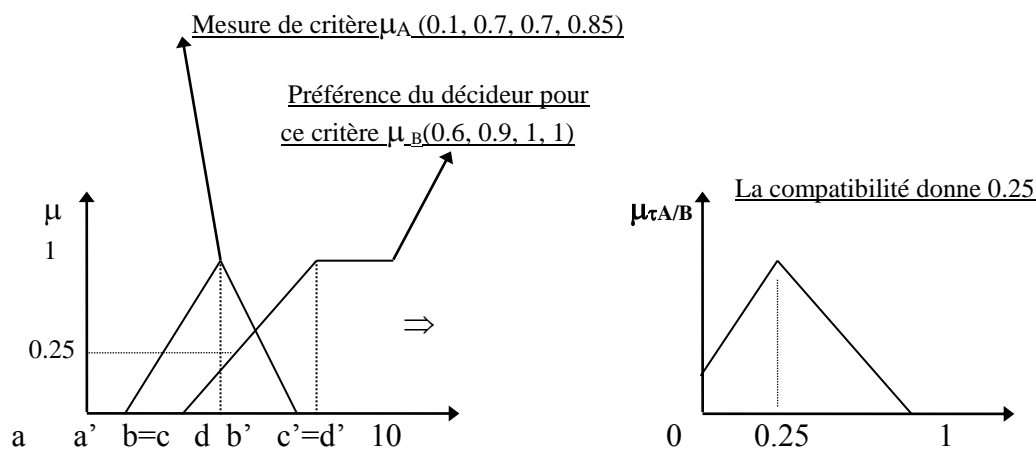


Figure IV.13 : Détermination des DCFs

C'est la valeur du DCF qui nous intéresse ici pour nos travaux. Etudions comment procède le système ADHESION pour déterminer les DCF : connaissant déjà l'évaluation de chaque critère $\mu_M(a,b,c,d)$ (exemple ci-dessous $\mu_M(0.1, 0.7, 0.7, 0.85)$ pour le critère *temps alloué*) et la préférence de l'expert $\mu_M(a',b',c',d')$ (exemple $\mu_M(0.6, 0.9, 1, 1)$), le système procède de manière itérative, de la façon suivante :

connaissant déjà $\mu_M(a,b,c,d)$ et $\mu_M(a',b',c',d')$, le système calcule les différences $a'-b'$ et $c'-d'$ de même que $a-b$ et $c-d$, ce qui nous permet d'obtenir directement les pentes suivantes (figure IV.14) : $1/(a'-b')$, $1/(c'-d')$, $1/(a-b)$, $1/(c-d)$.

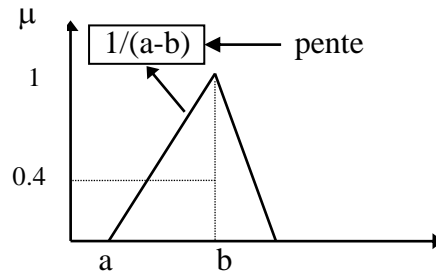


Figure IV.14 : Vue d'une pente

On calcule ensuite les rapports des pentes entre $1/(a-b)$ et $1/(a'-b')$, soit $(a'-b')/(a-b)$, de même qu'entre $1/(c-d)$ et $1/(c'-d')$, soit $(c'-d')/(c-d)$, ainsi que a/a' , b'/b et b/a' . Par exemple, lorsque le système trouve comme valeurs :

$(a'-b')/(a-b) = 1$ et $(c'-d')/(c-d) = 0$, ce qui correspond au cas de la figure IV.15 suivant :

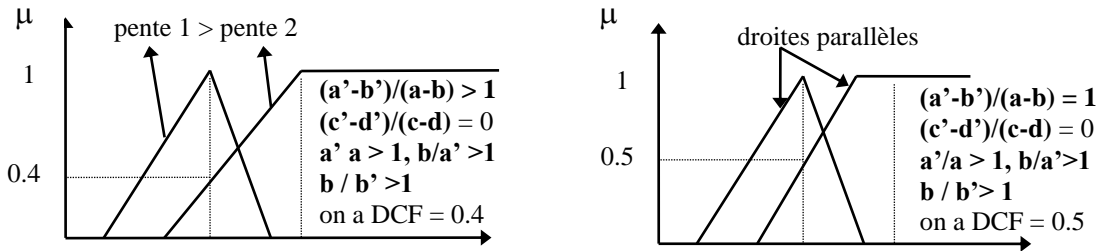


Figure IV.15 : Evolution des pentes connaissant $a, b, c, d, a', b', c', d'$

Le système travaille finalement avec cinq paramètres pour estimer avec une bonne précision la valeur du DCF.

Les cinq paramètres appartiennent à la fonction $T(t1, t2, t3, t4, t5)$, avec :

$t1 = (a'-b')/(a-b)$

$t2 = (c'-d')/(c-d)$

$t3 = a' / a$

$t4 = b / b'$

$t5 = b/a'$

Une fois que le système a calculé ces cinq paramètres, ces valeurs seront comparées à des vecteurs standards $T'(t1', t2', t3', t4', t5')$ contenus dans le système et qui correspondent chacun à des DCFs prédéterminés.

Par exemple, si le vecteur $T(t1, t2, t3, t4, t5) = T(1, 0, 1.4, 1.6, 1.08)$, la valeur DCF correspondant et système est 0.2 et correspond à la figure IV.16 ci-dessous :

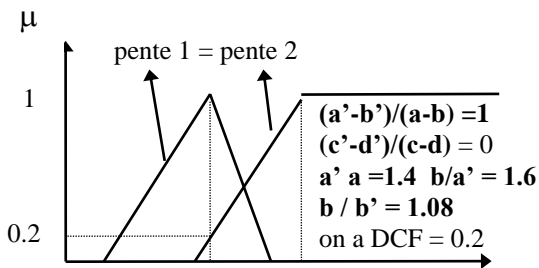


Figure IV.16 : Exemple de vecteur T

La partie du **programme** correspondant (figure IV.17) sera du type :

```

Param1 (0.1) - > ;
...
Param1 (5) - > ;
Param2 (0.1) - > ;
...
Param2(5) - > ;
Param4 (0.1) - > ;
...

Choix1 ([X, Y, Z, U, P]) - > Param1 (X) Param2 (X) Param3 (X) Param4 (X) Param5 (X) ;
Choix2 ([A, B, C, D, E]) - >
Param1 (A) Param2 (B) Param3 (C) Param4 (D) Param5 (D),
{A=X, B=Y, C=Z, D=U, E=P};

{A=X, B=Y, C=Z, D=U, E=P} étant la contrainte.

```

Figure IV.17 : Extrait de la programmation

Notons que les valeurs 5 et 0.1 sont déterminés comme étant les valeurs maxi et mini que l'on n'atteint pas dans la pratique sinon on arrive à des aberrations, sauf deux cas particuliers de paramètres possibles : 0/0 *indétermination* ou *infinie*. Dans ce cas, le système ne traite que la partie gauche respectivement et la partie droite selon les éléments a,b,c,d,a',b',c',d' responsables de ces valeurs particulières. Après avoir trouvé les DCFs, il faut passer à l'étape finale qui permettra de proposer les méthodes et actions à mener pour des alternatives, notre démarche étant celle de la figure IV.10. Pour ce faire, notre programme s'appuie sur la méthode de suclassement Electre I [ROY, 1985], déjà présentée dans le chapitre précédent :

$$c(x, y) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum \lambda_i \quad j \in \{j / g_j(x) \geq g_j(y)\}$$

avec les poids λ_i de chaque critère i pris égal au DCF du critère i ($\lambda_i = \text{DFC}$).

L'indice de concordance $c(x, y)$ mesure les arguments en faveur de l'affirmation "x surclasse y" ($x S y$).

Pour notre cas, nous avons expliqué au chapitre précédent que, à chaque couple d'actions $(a, b) \in A^2$, l'indice de concordance $c(a, b)$ pour le surclassement de a sur b est égal à :

$$c(a, b) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum \lambda_i$$

Les critères j sont des critères définis tels que $g_j(a) \geq g_j(b)$, c'est-à-dire que seuls sont pris en compte ici des critères j tels que la valeur de l'action a ($g_j(a)$) soit supérieure à la valeur de l'action b ($g_j(b)$). Pour la sélection des méthodes, on dira qu'à chaque couple de méthodes $(M1, M2) \in M^2$, l'indice de concordance $c(M1, M2)$ pour le surclassement de a sur b prend la forme :

$$c(M1, M2) = \frac{1}{\sum_{j=1}^p \lambda_j} \sum \lambda_i$$

Les critères j sont des critères définis de telle sorte que $g_j(M1) \geq g_j(M2)$, c'est-à-dire que seuls sont pris en compte ici des critères j de telle sorte que le rapport $g_j(M1)$ entre la méthode $M1$ et le critère j soit plus important que le rapport $g_j(M2)$ entre la méthode $M2$ et le critère j . Prenons l'exemple ci-dessous :

Soit trois critères i_1, i_2, i_3 tels que :

i_1 : *Moyens humains* (absence d'ergonome dans l'équipe d'évaluation)

i_2 : *Budget d'évaluation* (insuffisance des moyens financiers)

i_3 : *Moyens matériels* (existence du matériel spécialisé en évaluation).

Soit à classer deux méthodes $M1$ et $M2$ telles que :

$M1$: *Enregistrement vidéo*

$M2$: *Traces écrites du Travail,*

alors les critères j pris en considération ici sont ceux qui vérifient la relation $g_j(M1) \geq g_j(M2)$.
Le critère

$j = i_3$: *moyens matériels* (existence du matériel spécialisé en évaluation)

est le critère qui est le plus en rapport avec la méthode $M1$; c'est aussi le seul critère qui remplira la condition $g_j(M1) \geq g_j(M2)$.

Conclusion

Ce système d'aide à la décision en évaluation pour le choix des méthodes, techniques et critères d'évaluation a été mis en œuvre sous la forme d'une première maquette de faisabilité qui ouvre la voie à de nouvelles recherches intégrant non seulement des aspects relevant de l'acceptabilité pratique, mais aussi de l'acceptabilité sociale. A ce stade, notre maquette répond aux fonctionnalités qui lui sont assignées.

Les travaux présentés devraient à terme apporter une contribution originale à l'aide à la décision pour l'évaluation des systèmes interactifs, domaine largement négligé et méconnu des équipes de développement. Rappelons que l'approche visée vise à mettre à la disposition des évaluateurs une description des méthodes et techniques appropriées en fonction du système interactif à évaluer.

Les premiers résultats du système, appelé ADHESION (**A**ide à la **D**écision **H**umaine pour l'**E**valuation des **S**ystèmes **I**nteractifs et leur validati**ON**), réalisé à l'aide de Prolog III, sont d'ores et déjà prometteurs, aussi bien en terme de modélisation du processus de décision, que de résultats qu'il est possible de fournir aux évaluateurs. Le chapitre 5 en propose une première validation technique sur deux scénarios inspirés de cas réels.

Chapitre V : Mise en œuvre et validation technique de la démarche et de la maquette sur deux cas d'étude

Introduction

- I. Mise en œuvre et validation sur un premier cas d'étude : le projet ASTREE**
- II. Mise en œuvre et validation sur un second cas d'étude : le projet MOGADOR**

Conclusion

Introduction

Ce chapitre est consacré à la mise en œuvre de la démarche et de la maquette du système d'aide à la décision dans un but de validation sur deux cas concrets datant de plusieurs années. IL s'agit en quelque sorte ici de faire de la RETRO-INGENIERIE. Le premier cas concerne la première version d'un système conçu et réalisé au LAMIH, entrant dans le cadre du projet ASTREE (SNCF - FRANCE) ; le deuxième cas concerne le projet MOGADOR (SEEG - Société d'Energie et d'Eau du Gabon), par l'intermédiaire d'un système conçu et réalisé par EDF (Electricité De France). Le projet ASTREE est un poste de régulation du trafic du réseau ferroviaire de la Société Nationale des Chemins de Fer français ; MOGADOR est quant à lui un système de gestion des bassins hydrauliques pour les centrales hydro-électriques de TCHIMBELE et KINGUELE au GABON.

En tenant compte des *contraintes réelles*, ces études de cas visent à comparer d'une part les méthodes et techniques utilisées pour évaluer ces systèmes, et d'autre part celles proposées par notre système d'aide à la décision en évaluation. L'objectif est de démontrer en quoi le système ADHESION aurait été aussi, voire plus efficace et approprié pour l'évaluation du système donné, dans ces deux cas de figure.

Notons à ce sujet que le système interactif conçu dans le cadre du projet ASTREE a été évalué à l'époque à l'aide de trois techniques d'évaluation : les *Questionnaires*, les *Observations des tactiles visuelles* et le *Mouchard électronique* ; MOGADOR a quant à lui été évalué à l'aide d'une méthode dénommée *Mesure de performance*.

I. Mise en oeuvre et validation sur un premier cas d'étude : le projet ASTREE

Dans cette partie, nous allons mettre en œuvre et valider le système ADHESION à l'aide du premier cas (ASTREE). Pour y parvenir, nous allons d'abord présenter le contexte du projet ASTREE sur le plan architectural et fonctionnel suivi d'une présentation des résultats d'évaluation obtenu à l'aide des méthodes et techniques réellement utilisées pour évaluer ASTREE. Après cette présentation du projet ASTREE, nous expliquerons la mise en œuvre proprement dite du système ADHESION suivie de la validation qui consistera à comparer les propositions de méthodes et techniques fournies par ADHESION et les méthodes et techniques réellement utilisées.

I.1. Contexte du projet ASTREE

Lors du projet ASTREE, le LAMIH a conçu, réalisé et évalué la première version d'un système interactifs destinés à des régulateurs du réseau ferroviaire. Cette recherche a été menée en collaboration avec la direction de la recherche de la SNCF département ASTREE. Cette recherche a été exposée dans plusieurs travaux [BENAISSA, 1993][EZZEDINE et al., 1995]. L'évaluation a été réalisée avec des utilisateurs finaux qui sont des opérateurs humains travaillant en salle de contrôle. Les opérateurs étaient appelés en particulier à régler des problèmes de conflits dans le réseau ce qui impliquait des stratégies. Le poste de régulation ASTREE (figure V.1) était équipé à l'époque de quatre écrans physiques. Le régulateur pouvait accéder à plusieurs familles d'informations : des informations relatives au suivi et présentées sur

les écrans 1 et 2, des informations relatives au dialogue représentées sur l'écran 3, et des informations alternant suivi et dialogue représentées sur l'écran 4.

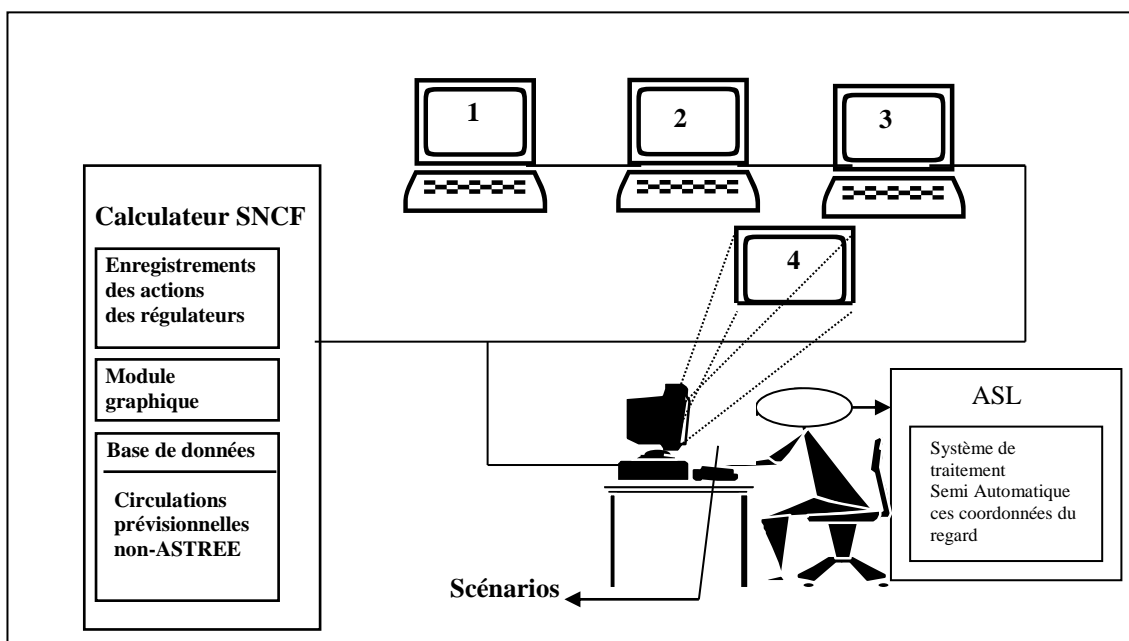


Figure V.1 : Poste de régulation ASTREE (avec intégration d'un oculomètre par les évaluateurs)

I.1.1. Les méthodes et techniques d'évaluation utilisées initialement

Dans cette partie seront présentées les différentes méthodes et techniques qui avaient été sélectionnées à l'époque de ce projet, et utilisées par des spécialistes en évaluation du LAMIH pour évaluer ASTREE. C'est ainsi qu'on présentera successivement : (1) *observations des tactiles visuelles*, (2) *mouchard électronique* et (3) *interviews et questionnaires*. Rappelons que ces méthodes ont été expliquées dans le chapitre II.

Observation des tactiles visuelles :

L'Observation des tactiles visuelles est basée sur une observation de l'activité oculaire au moyen d'un oculomètre pour estimer le comportement cognitif de l'utilisateur. L'un des avantages a résidé dans le fait que les mouvements oculaires enregistrés lors de sessions de travail pouvaient être rejoués en différé sur écran et commentés avec (et par) les régulateurs.

Le mouchard électronique :

Un mouchard électronique a permis de recueillir automatiquement des données objectives à l'aide d'un logiciel d'acquisition implémenté sur le calculateur SNCF. Ce logiciel a aussi permis de fournir des indicateurs sur les activités manuelles de l'opérateur humain

(régulateur du trafic ferroviaire), sur les performances de l'IHM (réactions aux situations dégradées, quantification des débits de circulation, quantification des charges de traitement du système, dénombrement de balises, discordance avec l'exploitation classique) et sur les performances de l'opérateur humain, en situation réelle (durée d'exécution d'une tâche, nombre de changements de vues, fréquence d'utilisation de menus et d'aide en ligne, nombre d'objets manipulés et non ou sous utilisés,...). L'avantage de cette technique a résidé dans la non perturbation de l'opérateur humain lors de son activité de régulation / contrôle du trafic ferroviaire : les données recueillies n'ont donc pas été biaisées.

Interviews et questionnaires :

Les questionnaires ainsi que des verbalisations lors d'interviews avec des régulateurs ont permis de recueillir des informations subjectives reflétant l'opinion et les critiques des régulateurs vis-à-vis de l'IHM. Notons que le recueil d'appréciations subjectives contribue au contrôle et à la validation des résultats expérimentaux obtenus par ailleurs. Le questionnaire qui a été choisi était un système à grille permettant d'avoir une notation simple. Ce questionnaire était orienté vers les problèmes ergonomiques liés à l'utilisation de l'IHM. Il a permis particulièrement de savoir, grâce à son classement par critère, si l'IHM était claire, fonctionnelle, etc., vis-à-vis des sujets.

I.1.2. Le protocole expérimental mis en œuvre initialement

La population de sujets pour cette expérimentation a été composée de véritables régulateurs, au nombre de sept, répartis par groupe de deux (un le matin, l'autre l'après-midi). Notons que les sujets ont manifesté beaucoup d'intérêt pour les différentes fonctionnalités du système ASTREE (sécurité, suivi) ainsi que pour le métier de régulation dans un tel système.

La durée totale de l'expérimentation a été de quatre heures et de 20 à 25 minutes avec le capteur oculométrique. Le protocole expérimental s'est déroulé en quatre phases :

1. Une phase de présentation du système ASTREE (durée = vingt minutes) ;
2. Une phase de formation sur le poste de régulation ASTREE (durée = deux heures) ;
3. Une phase d'exécution d'un jeu de scénarios relatifs au contrôle du réseau ferroviaire (durée = deux heures)
4. Une phase d'interview et de soumission du questionnaire aux différents sujets (une heure et vingt minutes).

I.1.3. Dispositif expérimental initialement mis en place

L'évaluation était menée dans la salle de régulation. L'une des caractéristiques majeures des sources d'information qu'offrait le poste de régulation ASTREE était la diversité ainsi que la complétude des informations par rapport au métier de régulation, par l'intermédiaire de divers outils dont disposait le régulateur. Ces outils prenaient en compte non seulement la gestion des circulations, travaux et roulements des engins, mais aussi la gestion des conflits qui restait le souci majeur du régulateur. La qualité de ces informations restait cependant tributaire du nombre attendu d'objectifs à atteindre par l'opérateur. Pour pouvoir évaluer des indices comme la durée

de fixation d'une zone ou déterminer des zones d'intérêt (représentant des informations pertinentes) et reproduire les séquences successives d'observation, le capteur utilisé était un oculomètre ASL (Applied Science Laboratories) modèle 4000, couplé à un système dédié à la prise en compte des mouvements de tête. Ce dispositif avait été décrit dans le chapitre II. Pour acquérir les différentes actions de l'opérateur humain et les différentes réactions de l'IHM, un enregistrement de tout changement de contexte était nécessaire, ceci étant réalisé grâce à un ordinateur industriel géré par l'équipe de la SNCF.

I.1.4. Présentation des résultats d'évaluation initialement obtenus

Dans cette section, les résultats d'évaluation obtenus à l'aide des différentes méthodes et techniques sont résumés¹ [EZZEDINE et ABED, 1997] [KADDOURI *et al.*, 1995].

I.1.4.1. Résultats obtenus à l'aide des questionnaires

La figure V.2 ci-dessous présente quelques résultats obtenus à l'aide de la méthode des *questionnaires*. Ces résultats obtenus par l'équipe d'évaluation présentent le pourcentage des sujets satisfaits et insatisfaits en fonction des critères d'évaluation pris en compte.

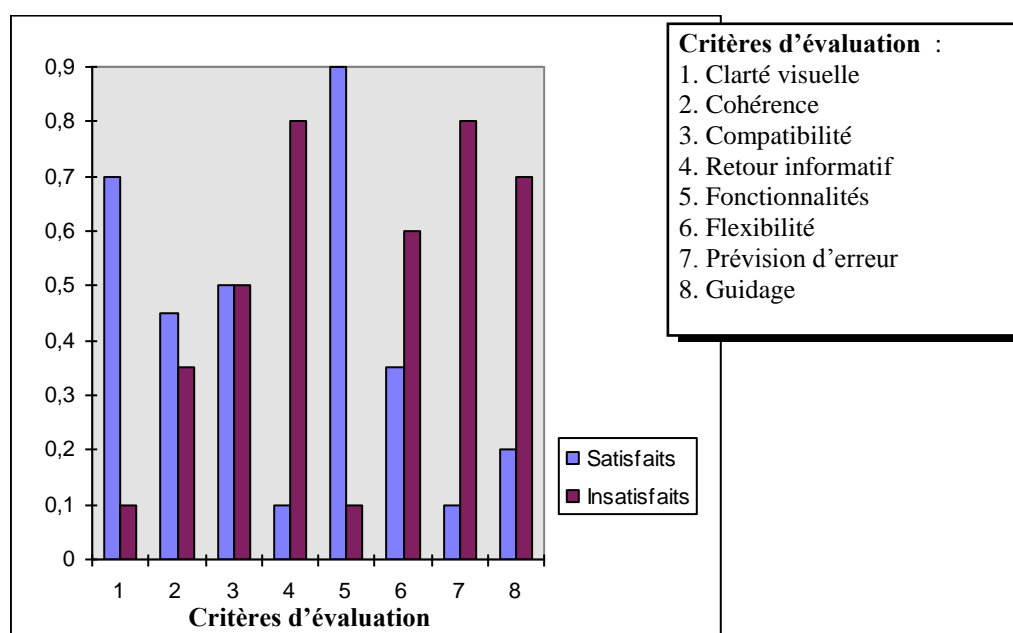


Figure V.2 : Taux de satisfaction et d'insatisfaction (pour tous les sujets)

Il est clair que les régulateurs, malgré la nouveauté du système, ont été plutôt satisfaits. A part quelques lacunes concernant les guides, l'aide en ligne et la complexité du dialogue, l'IHM a plutôt été vue comme un outil d'aide pour la régulation.

¹ Le lecteur intéressé par les résultats détaillés est invité à lire les documents cités.

I.1.4.2. Résultats obtenus à l'aide des Observations des tactiles visuelles

Même si la fonction de chaque écran a été globalement comprise par les régulateurs, quelques ambiguïtés subsistaient. On constatait en effet que les régulateurs tendaient à confondre l'écran 1 et l'écran 4. Les régulateurs travaillaient comme si le poste de régulation se réduisait aux deux seuls écrans 3 et 4. La figure IV.3 donne un aperçu de l'activité visuelle du 7^{ème} sujet durant les expérimentations :

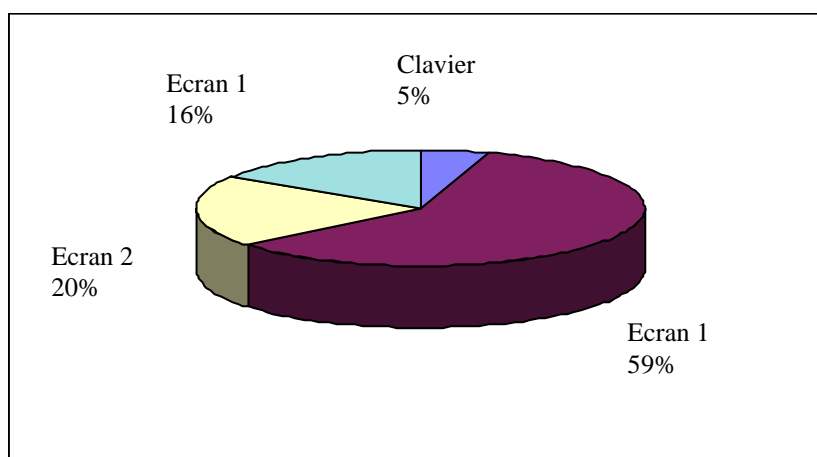


Figure V.3 : Taux de consultation de chaque écran du poste de régulation ASTREE (sujet 7)

On remarque que les entrées clavier ne représentaient que 5% de l'activité du régulateur. Ceci était en grande partie dû à l'aspect interactif du système. Cette interactivité a été plutôt bien accueillie par les régulateurs, même ceux venant des postes de conduite non informatisés. La cause de cette répartition était liée d'une part aux contraintes fonctionnelles établies par les concepteurs et, d'autre part, à un sentiment de confort des régulateurs à utiliser deux écrans au lieu de quatre.

I.1.4.3. Résultats obtenus à l'aide du mouchard électronique

Le mouchard électronique avait permis d'acquérir, d'une part toutes les actions de l'opérateur humain lors de la réalisation des tâches manuelles en vue de répondre aux objectifs de chaque scénario et de la durée d'exécution de chaque tâche et, d'autre part, tout changement de vue de l'Interface Homme-Machine suite aux actions de l'opérateur. Par ailleurs, ce mouchard a renseigné les évaluateurs sur les performances du système Homme(s)-Machine(s) en mesurant les temps d'accès à la base de données.

II.2 Mise en œuvre du système d'aide ADHESION

Après avoir pris connaissance du contexte, des méthodes et techniques d'évaluation initialement utilisés ainsi que des résultats de l'évaluation d'ASTREE dans la section précédente,

nous allons voir dans cette partie les propositions de méthodes et techniques fournies par notre système d'aide à la décision en évaluation (ADHESION).

Celles-ci seront comparées avec les méthodes et techniques réellement utilisées. Cette comparaison tient compte de certains facteurs tels que les objectifs d'évaluation, les tâches principales effectuées par les opérateurs humains, etc. Pour y parvenir, nous allons d'abord présenter les différentes étapes suivies pour la mise en œuvre du système d'aide ADHESION. Ces phases commencent par une étape de préparation des données qui précède la phase de saisie des contraintes proprement dite. Rappelons que la phase de saisie des contraintes représente la constitution de la base de faits du système ADHESION.

II.2.1. Préparation des données (contraintes)

Avant de passer à la phase de saisie des contraintes, une préparation est nécessaire afin de fournir au système ADHESION des informations fiables. Pour ce premier cas d'étude, les données ont été élaborées avec un membre expérimenté de l'équipe d'évaluation d'ASTREE. Cette préparation consiste à exprimer les facteurs situationnels, le contexte d'évaluation et les caractéristiques de l'IHM à évaluer. Une fois ces contraintes exprimées selon un raisonnement naturel, elles sont destinées à être traduites selon un raisonnement « flou ». Cette préparation se fait en tenant compte des critères de décision du tableau V.1.

Critères de décision	Critères principaux	Sous-critères	
	Performance du système et du contexte		<i>Type d'application</i>
			<i>Objet d'évaluation</i>
			<i>Type d'interaction</i>
			<i>Support pour la capture des données</i>
			<i>Accessibilité aux utilisateurs représentatifs (sujets d'évaluation)</i>
			<i>Etape du cycle de développement</i>
			<i>Budget de l'évaluation</i>
			<i>Localisation (site, laboratoire, pays)</i>
			<i>Environnement où évoluera l'IHM</i>
			<i>Temps alloué à l'évaluation</i>
		<i>Moyens humains</i>	
	Performance potentielle des utilisateurs		<i>Origine des experts</i>
			<i>Expérience des utilisateurs</i>
		<i>Spécialités requises</i>	

Tableau V.1 : Critères de décision

Après avoir proposé une synthèse des critères de décision (voir tableau V.1), cette section consiste à préparer les données en justifiant les réponses aux questions posées par le système. Cette phase est très importante car une description qui ne reflète pas le contexte réel de l'évaluation a des conséquences néfastes pour les propositions de méthodes et techniques fournies par ADHESION. Les propositions fournies par le système risquent en effet de ne pas

permettre d'atteindre les objectifs d'évaluation. C'est l'une des limitations du système ASDHESION.

Le système propose des méthodes en fonction des informations qu'il reçoit. Par exemple, à la question : «le groupe d'évaluation dispose-t-il d'un spécialiste en assurance qualité ?», certains évaluateurs n'ayant reçu aucune formation initiale en assurance qualité, impliqués dans des projets d'évaluation industriels, auraient peut-être tendance à répondre par (**Très, Oui**), soit par crainte de ne pas être capable de justifier leur présence dans une équipe d'évaluation, soit pour tout autre raison ; alors qu'une réponse de type (**Modérément, Oui**), (**Plutôt, Oui**) ou (**Très, Moyen**) serait plus réaliste. De fausses informations ou des informations non conformes sont de nature à rendre le système ADHESION moins efficace, d'où l'importance de cette phase de préparation.

□ Type d'application (premier critère)

Description du contexte par rapport au critère :

Selon les évaluateurs, ASTREE est de très loin une application qui n'est pas de type grand public ; c'est un système qui est destiné à des utilisateurs spécialement formés. Ils étaient sûrs et certains de ces propos.

Question du système :

Pouvez-vous assimiler le système à un type grand public ?

Choix de réponse selon un raisonnement « flou » :

Descripteur			Modificateur		
Bon/Oui	Moyen	Mauvais/Non	Plutôt	Modérément	Très

Dans cette étape de préparation des données, l'évaluateur de l'équipe d'évaluation d'ASTREE et nous avons choisi le couple de réponses (Modificateur = Très , Descripteur = Non). C'était à notre avis le choix de réponse qui correspondait le mieux à la description du contexte.

Le Descripteur **Non** a été choisi au détriment du moyen à la mention : « de très loin » (dans la description du contexte).

Le Modificateur **Très** est choisi car les évaluateurs sont sûrs et certains de leur affirmation. Il faut une formation préalable pour utiliser et comprendre le système. Nous sommes donc très éloignés du type grand public, d'où ces choix de réponses extrêmes.

□ Qualité du support pour la capture des données

Description du contexte par rapport au critère :

L'équipe disposait de matériel spécialisé pour l'évaluation. Les évaluateurs étaient sûrs et certains que le matériel n'était pas défaillant et que vu l'état du matériel, le risque d'un dysfonctionnement était presque nul.

Question de l'outil d'aide :

Disposez-vous de matériel spécialisé (système oculométrique, enregistreur vidéo,...) ?

Réponse des évaluateurs :

(**Très, Oui**)

□ Accessibilité aux utilisateurs

Description du contexte par rapport au critère :

L'équipe d'évaluation n'a connu aucune difficulté à trouver des sujets pour l'évaluation d'ASTREE. Les sujets étaient les futurs utilisateurs du système. Ils étaient donc très représentatifs.

Question de l'outil d'aide :

Avez-vous accès aux utilisateurs finaux ?

Réponse des évaluateurs :

(Très , Oui)

Etape du cycle de développement

Description du contexte par rapport au critère :

Le système à évaluer était déjà réalisé, c'est-à-dire que le contexte se situait au-delà des étapes de spécification fonctionnelle, de conception préliminaire et de conception détaillée.

Question de l'outil d'aide :

L'IHM a-t-elle franchi l'étape de réalisation ?

Réponse des évaluateurs :

(Très , Oui)

Moyens financiers

Description du contexte par rapport au critère :

L'équipe d'évaluation a confirmé que le budget de l'évaluation était suffisant, qu'elle était capable de financer des jours au-delà du temps prévu.

Question 1 de l'outil d'aide :

1. Le budget d'évaluation vous semble-t-il suffisant ?

Réponse des évaluateurs :

(Très , Oui)

Question 2 de l'outil d'aide :

2. Pouvez-vous financer le dépassement du temps prévu ?

Réponse des évaluateurs :

(Très , Oui)

Question 3 de l'outil d'aide :

3. Pouvez-vous financer les spécialistes d'évaluation manquants ?

Réponse des évaluateurs :

(Très , Oui)

Localisation (site, laboratoire, pays)

Description du contexte par rapport au critère :

L'évaluation se passait sur site et non en laboratoire et les pays devant adopter le système à évaluer étaient la France et l'Allemagne.

Question 1 de l'outil d'aide :

Le système à évaluer évoluera-t-il dans un environnement nuisible ?

Réponse des évaluateurs :

(Très , Oui)

Question 2 de l'outil d'aide :

Choisissez le pays de destination de l'IHM à évaluer.

Réponse 1 des évaluateurs : **France**

Réponse 2 des évaluateurs : **Allemagne**

Le Temps alloué

Description du contexte par rapport au critère :

Les évaluateurs d'ASTREE ont estimé qu'ils avaient largement le temps pour l'évaluation.

Question de l'outil d'aide :

Le temps alloué à l'évaluation vous semble-t-il suffisant ?

Réponse des évaluateurs :

(Très , Oui)

Moyens humains

Description du contexte par rapport au critère :

L'équipe ne disposait pas d'une personne faisant office de spécialiste en assurance qualité. Elle ne disposait pas de psychologue, ni de sociologue ou d'un spécialiste en droit et informatique. Par contre, l'équipe disposait d'un ergonome et les membres, au nombre de 3, avaient une bonne expérience en évaluation.

Question 1 de l'outil d'aide :

1. Le groupe d'évaluation dispose-t-il d'un ergonome ?

Réponse des évaluateurs :

(Plutôt , Oui)

Question 2 de l'outil d'aide :

2. Le groupe d'évaluation dispose-t-il d'un spécialiste en assurance qualité ?

Réponse des évaluateurs :

(Modérément , Oui)

(pas de formation initiale en assurance qualité, auto-formation)

Question 3 de l'outil d'aide :

3. Le groupe d'évaluation dispose-t-il d'un psychologue ?

Réponse des évaluateurs :

(Très , Non)

Question 4 de l'outil d'aide :

4. Le groupe d'évaluation dispose-t-il d'un sociologue ?

Réponse des évaluateurs :

(Très , Non)

Question 5 de l'outil d'aide :

5. Le groupe d'évaluation dispose-t-il d'un spécialiste en droit et informatique ?

Réponse des évaluateurs :

(Très , Non)

□ Expérience des évaluateurs - Habileté

Description du contexte par rapport au critère :

Les membres de l'équipe d'évaluation avaient tous une bonne expérience en évaluation.

Question de l'outil d'aide :

Avez-vous une expérience en évaluation ?

Réponse des évaluateurs : (Très , Oui).

II.2.2. Tableau de synthèse de la phase de préparation des données

Après avoir préparé les données nécessaires au fonctionnement d'ADHESION, une synthèse s'impose pour des raisons d'organisation. Cette synthèse sous-forme de tableau est l'ultime étape avant d'utiliser le système d'aide à la décision ADHESION. Pour ce premier cas étudié (ASTREE), le tableau IV.2 présente la synthèse de la phase de préparation.

Questions	Réponses	
	Modificateur	Descripteur
Pouvez-vous assimiler le système à un type grand public ?	Très	Non
Disposez-vous du matériel spécialisé (système oculométrique,...) ?	Très	Oui
Avez-vous accès aux utilisateurs finaux ?	Très	Oui
Le budget d'évaluation vous semble-t-il suffisant ?	Très	Oui
Pouvez-vous financer le dépassement du temps prévu ?	Très	Oui
Pouvez-vous financer les spécialistes d'évaluation manquants ?	Très	Non
Le système à évaluer évoluera-t-il dans un environnement nuisible ?	Très	Oui
Le temps alloué à l'évaluation vous semble-t-il suffisant ?	Très	Oui
Le groupe d'évaluation dispose-t-il d'un ergonome ?	Plutôt	Oui
Le groupe d'évaluation dispose-t-il d'un spécialiste en assurance qualité?	Modérément	Oui
Le groupe d'évaluation dispose-t-il d'un psychologue ?	Très	Non
Le groupe d'évaluation dispose-t-il d'un sociologue ?	Très	Non
Le groupe d'évaluation dispose-t-il d'un spécialiste en droit informatif ?	Très	Non
Avez-vous une expérience en évaluation ?	Très	Oui
L'IHM a-t-elle franchi l'étape de réalisation ?	Très	Oui
Pays de destination du système à évaluer ASTREE	France Allemagne	

Tableau V.2 : Tableau de synthèse de la phase de préparation des données

I.3. Saisie des contraintes

Après la phase de préparation des données, la saisie des contraintes est l'étape principale lors de l'utilisation du système d'aide à la décision. L'introduction des données dans le système se fait en cochant les cases correspondant aux réponses aux questions posées (voir figure V.4 pour le cas français), et ceci en utilisant le tableau de synthèse (tableau V.2).

Pour le cas allemand, il a nous fallu sélectionner simplement l'Allemagne car les autres informations sont identiques.

La figure V.4 présente la saisie des contraintes supposées refléter les facteurs situationnels pour l'évaluation d'ASTREE. On peut remarquer que le pays sélectionné est la France car le

système est destiné à la France ; de même, pour l'Allemagne, on sélectionnera l'Allemagne, les évaluateurs d'ASTREE ayant confirmé que le système était destiné à l'Allemagne et la France. On constate en fait que les réponses aux questions sont identiques selon les informations fournies par l'un des membres de l'équipe d'évaluation avec qui nous avons préparé les contraintes pour ce premier cas d'étude. Rappelons que malgré le fait que ce soit le même système à évaluer, les réponses pouvaient être différentes entre la France et l'Allemagne si par exemple le système devait être utilisé dans un environnement nuisible en Allemagne et non en France.

DESCRIPTION DES CONTRAINTES

CETTE TECHNIQUE DE RESOLUTION VOUS OBLIGE A REPONDRE
A TOUTES LES QUESTIONS CI-DESSOUS

Sélectionnez le pays de destination (ou assimilé) de l'IHM à évaluer

Etats-Unis

Ethiopie

Finlande

France

Gabon

Gambie

Gaza et Cisjordanie

Géorgie

Ghana

Grèce

Avez-vous accès aux utilisateurs finaux ?
Répondre Plutôt-Moyen, signifie que vous avez près de 50% de chance d'accéder aux futurs utilisateurs de l'IHM que vous voulez évaluer, alors que Très - Oui signifie 100% de chance.

Le Groupe d'évaluation dispose t-il d'un ergonome ?
Répondre Très - Non signifie que vous ne disposez pas du tout d'ergonome, alors que Plutôt-Moyen signifie qu'il existe une chance sur deux pour que vous obtenez un ergonome, cela peut être le cas d'une attente

***** Descripteurs ***** ***** Modificateurs *****

	Bon/Oui	Moyen	Mauvais/Non	Plutôt	Modérément	Très
Avez-vous accès aux utilisateurs finaux?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un ergonome?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un sociologue?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un psychologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en droit-informatique?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en assurance qualité?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le budget-évaluation vous semble t-il suffisant?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le temps alloué à l'évaluation vous semble t-il suffisant?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pouvez-vous financer le dépassement du temps prévu?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pouvez-vous financer les spécialistes d'évaluation manquant?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le système à évaluer évoluera t-il dans un environ.nuisible?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Disposez-vous du matériel spécialisé (système oculométrique, enreg.vidéo,)?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Avez-vous une expérience en évaluation ?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pouvez-vous assimiler le système à évaluer à un type grand public?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
L'IHM à évaluer a t-elle franchit l'étape de réalisation?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figure V.4 : Saisie des contraintes pour le premier cas ASTREE (cas où le pays sélectionné est la France)

Comme nous l'avons souligné dans le chapitre IV au niveau de la section 'Explications sur les pages-écrans', l'utilisateur peut à chaque fois consulter le sens donné aux questions en cas de difficulté dans la compréhension. Cette aide se justifie par le souci de répondre aux exigences ergonomiques. La section suivante est consacrée aux résultats obtenus pour ce premier cas d'étude.

I.4. Présentation des propositions de solution fournies par le système d'aide ADHESION pour le cas étudié (Projet ASTREE)

Après avoir saisi les contraintes, l'utilisateur du système d'aide ADHESION doit valider les réponses en cliquant sur « continuer » (figure V.5). Le système propose alors des méthodes, techniques et critères d'évaluation. Il peut être dans l'impossibilité d'en proposer et dans ce cas il propose des actions (voir listes des actions possibles dans le tableau V.1) de modifications des contraintes. Les figures V.5a à V.5c présentent les différentes propositions obtenues par la mise en œuvre d'ADHESION pour le premier cas étudié (ASTREE). On constate ici que les propositions de méthodes sont identiques pour la France et l'Allemagne car toutes les autres contraintes sont identiques, en dehors des aspects socio-culturels. Mais le système minimisant les différences culturelles entre la France et l'Allemagne [NENDJO ELLA et JACQUES, 1999], les réponses ne sauraient être différentes. La figure V.5a montre la première méthode obtenue (l'analyse de protocoles).

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°1

DESCRIPTION DE LA METHODE ET DES CRITERES

CARACTERISTIQUES DE L'IHM A EVALUER

Libellé de la méthode : Analyse des protocoles

Méthode ou Technique complémentaire : Questionnaires

Description en entrée (ex: description de tâche) : Pas nécessaire

Disponibilité de la méthode : Oui

Formalisation : Méthode non formelle

Technique d'interaction : Toutes

Objet d'évaluation : Produit final ou Prototype

Classe d'évaluateurs exigée : Experte/Première expérience

Support pour la Capture de données : Papier

Etapes du cycle de développement concernées : Après réalisation

Usage de métriques : Pas nécessaire

PRINCIPE DE FONCTIONNEMENT DE LA METHODE

Le principe consiste à modéliser, au moyen d'un arbre, l'activité intellectuelle par des observations et des entretiens.

Dans cette approche, l'utilisateur explique ce qu'il fait lors de l'interaction avec la machine (possibilité d'enregistrement, pour analyse ultérieure) : problèmes types, verbalisations prononcées, informations à la demande, étude de la mémorisation, pensée à voix haute...

POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS

moins important important Très import.

Requêtes des évaluateurs :

Conflits entre évaluateurs :

Exigence mentale :

Exigence physique :

CHARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION

Qualité : Sensibilité aux hypothèses

Support de restitution : Papier

Support d'analyse des résultats : Papier

Niveau d'abstraction : Action physique

Type d'information : subjective/Objective

Performance de la méthode ou Technique : Moy. Bonne Très Bonne

NIVEAUX D'AUTOMATISATION DE LA METHODE

Capture automatique des données :

Restitution automatique des données après traitement :

Analyse automatique :

Critique automatique :

CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION

N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.
1.	Connaissances_système	Pratique utilisabilité	
2.	Enchaînement	Pratique utilisabilité	
3.	Activité	Pratique utilisabilité	
4.	Problèmes types	Pratique utilisabilité	
5.	Stratégies	Pratique utilisabilité	
7.			
8.			
9.			

AVANTAGES

INCONVENIENTS

Coûts : Faibles

Temps moyen : 1 jour

Accessibilité aux utilisateurs finaux : Non

Facilité d'utilisation : Oui

Version automatisée : Non

Nombre moyen de sujets : 0

Nombre moyen d'évaluateurs : 2

REFERENCE SUR LA METHODE : Ericsson et al., 1984

CONTINUER

ANNULER

Figure V.5a : La première proposition (Analyse des protocoles) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation d'ASTREE

Rappelons que l'analyse des protocoles est caractérisée par sa capacité à prendre en compte des critères d'évaluation tels que les *connaissances sur le système*, les *stratégies*, les *problèmes types*, l'*enchaînement*, l'*activité*. Les informations concernant cette méthode ont été présentées au chapitre II. La figure V.5b montre une deuxième proposition (le *questionnaire*) du système d'aide à la décision. Les informations concernant cette technique d'évaluation ont aussi été présentées au chapitre II. C'est une technique très utilisée, qui vient généralement en complément à d'autres méthodes et techniques d'évaluation.

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°2																																																									
DESCRIPTION DE LA METHODE ET DES CRITERES																																																									
<table border="1"> <thead> <tr> <th colspan="2">CARACTERISTIQUES DE L'IHM A EVALUER</th> </tr> </thead> <tbody> <tr> <td>Libellé de la méthode</td> <td>: Questionnaires</td> </tr> <tr> <td>Méthode ou Technique complémentaire</td> <td>: Interviews</td> </tr> <tr> <td>Description en entrée (ex: description de tâche)</td> <td>: Pas nécessaire</td> </tr> <tr> <td>Disponibilité de la méthode</td> <td>: Oui</td> </tr> <tr> <td>Formalisation</td> <td>: Méthode non formelle</td> </tr> <tr> <td>Technique d'interaction</td> <td>: Toutes</td> </tr> <tr> <td>Objet d'évaluation</td> <td>: Produit final ou Prototype</td> </tr> <tr> <td>Classe d'évaluateurs exigée</td> <td>: Experte/Novice</td> </tr> <tr> <td>Support pour la Capture de données</td> <td>: Ordinateur/papier</td> </tr> <tr> <td>Etapes du cycle de développement concernées</td> <td>: Spécification/conception/après réalisation</td> </tr> <tr> <td>Usage de métriques</td> <td>: Pas nécessaire</td> </tr> </tbody> </table>	CARACTERISTIQUES DE L'IHM A EVALUER		Libellé de la méthode	: Questionnaires	Méthode ou Technique complémentaire	: Interviews	Description en entrée (ex: description de tâche)	: Pas nécessaire	Disponibilité de la méthode	: Oui	Formalisation	: Méthode non formelle	Technique d'interaction	: Toutes	Objet d'évaluation	: Produit final ou Prototype	Classe d'évaluateurs exigée	: Experte/Novice	Support pour la Capture de données	: Ordinateur/papier	Etapes du cycle de développement concernées	: Spécification/conception/après réalisation	Usage de métriques	: Pas nécessaire	<table border="1"> <thead> <tr> <th colspan="2">PRINCIPE DE FONCTIONNEMENT DE LA METHODE</th> </tr> </thead> <tbody> <tr> <td colspan="2">Le principe consiste à apprécier de manière subjective les besoins des utilisateurs auprès d'un groupe sélectionné.</td> </tr> <tr> <td colspan="2">La technique des questionnaires existe également sous-forme, d'utilisation sur la qualité de l'interface. Cette Technique est finalisée par une analyse des données recueillies par des méthodes statistiques</td> </tr> </tbody> </table>	PRINCIPE DE FONCTIONNEMENT DE LA METHODE		Le principe consiste à apprécier de manière subjective les besoins des utilisateurs auprès d'un groupe sélectionné.		La technique des questionnaires existe également sous-forme, d'utilisation sur la qualité de l'interface. Cette Technique est finalisée par une analyse des données recueillies par des méthodes statistiques																											
CARACTERISTIQUES DE L'IHM A EVALUER																																																									
Libellé de la méthode	: Questionnaires																																																								
Méthode ou Technique complémentaire	: Interviews																																																								
Description en entrée (ex: description de tâche)	: Pas nécessaire																																																								
Disponibilité de la méthode	: Oui																																																								
Formalisation	: Méthode non formelle																																																								
Technique d'interaction	: Toutes																																																								
Objet d'évaluation	: Produit final ou Prototype																																																								
Classe d'évaluateurs exigée	: Experte/Novice																																																								
Support pour la Capture de données	: Ordinateur/papier																																																								
Etapes du cycle de développement concernées	: Spécification/conception/après réalisation																																																								
Usage de métriques	: Pas nécessaire																																																								
PRINCIPE DE FONCTIONNEMENT DE LA METHODE																																																									
Le principe consiste à apprécier de manière subjective les besoins des utilisateurs auprès d'un groupe sélectionné.																																																									
La technique des questionnaires existe également sous-forme, d'utilisation sur la qualité de l'interface. Cette Technique est finalisée par une analyse des données recueillies par des méthodes statistiques																																																									
<table border="1"> <thead> <tr> <th colspan="4">POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS</th> </tr> <tr> <td></td> <td>moins important</td> <td>important</td> <td>Très import.</td> </tr> </thead> <tbody> <tr> <td>Requêtes des évaluateurs</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>Conflits entre évaluateurs</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>Exigence mentale</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>Exigence physique</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table>	POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS					moins important	important	Très import.	Requêtes des évaluateurs	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Conflits entre évaluateurs	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Exigence mentale	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Exigence physique	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<table border="1"> <thead> <tr> <th colspan="2">CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION</th> </tr> </thead> <tbody> <tr> <td>Qualité</td> <td>: Inspectabilité</td> </tr> <tr> <td>Support de restitution</td> <td>: papier</td> </tr> <tr> <td>Support d'analyse des résultats</td> <td>: papier/Ordinateur</td> </tr> <tr> <td>Niveau d'abstraction</td> <td>: Tous</td> </tr> <tr> <td>Type d'information</td> <td>: subjective/Explicative</td> </tr> <tr> <td>Performance de la méthode ou Technique</td> <td>: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne</td> </tr> </tbody> </table>	CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION		Qualité	: Inspectabilité	Support de restitution	: papier	Support d'analyse des résultats	: papier/Ordinateur	Niveau d'abstraction	: Tous	Type d'information	: subjective/Explicative	Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne																		
POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS																																																									
	moins important	important	Très import.																																																						
Requêtes des évaluateurs	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																																						
Conflits entre évaluateurs	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																						
Exigence mentale	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																						
Exigence physique	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																						
CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION																																																									
Qualité	: Inspectabilité																																																								
Support de restitution	: papier																																																								
Support d'analyse des résultats	: papier/Ordinateur																																																								
Niveau d'abstraction	: Tous																																																								
Type d'information	: subjective/Explicative																																																								
Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne																																																								
<table border="1"> <thead> <tr> <th colspan="4">NIVEAUX D'AUTOMATISATION DE LA METHODE</th> </tr> </thead> <tbody> <tr> <td>Capture automatique des données</td> <td>: <input type="checkbox"/></td> </tr> <tr> <td>Restitution automatique des données après traitement</td> <td>: <input type="checkbox"/></td> </tr> <tr> <td>Analyse automatique</td> <td>: <input type="checkbox"/></td> </tr> <tr> <td>Critique automatique</td> <td>: <input type="checkbox"/></td> </tr> </tbody> </table>	NIVEAUX D'AUTOMATISATION DE LA METHODE				Capture automatique des données	: <input type="checkbox"/>	Restitution automatique des données après traitement	: <input type="checkbox"/>	Analyse automatique	: <input type="checkbox"/>	Critique automatique	: <input type="checkbox"/>																																													
NIVEAUX D'AUTOMATISATION DE LA METHODE																																																									
Capture automatique des données	: <input type="checkbox"/>																																																								
Restitution automatique des données après traitement	: <input type="checkbox"/>																																																								
Analyse automatique	: <input type="checkbox"/>																																																								
Critique automatique	: <input type="checkbox"/>																																																								
<table border="1"> <thead> <tr> <th colspan="4">CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION</th> </tr> <tr> <th>N°</th> <th>*Libellé</th> <th>*Acceptabilité</th> <th>*Dimension d'évaluat.</th> </tr> </thead> <tbody> <tr> <td>1.</td> <td>Actions minimales</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>2.</td> <td>Actions explicite</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>3.</td> <td>Capacité fonctionnelle</td> <td>Pratique utilité</td> <td></td> </tr> <tr> <td>4.</td> <td>Charge mentale</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>5.</td> <td>Compatibilité</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>6.</td> <td>Concision</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>7.</td> <td>Contrôle explicite</td> <td>Pratique utilisabilité</td> <td></td> </tr> <tr> <td>8.</td> <td>Flexibilité</td> <td>Pratique utilisabilité</td> <td></td> </tr> </tbody> </table>	CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION				N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.	1.	Actions minimales	Pratique utilisabilité		2.	Actions explicite	Pratique utilisabilité		3.	Capacité fonctionnelle	Pratique utilité		4.	Charge mentale	Pratique utilisabilité		5.	Compatibilité	Pratique utilisabilité		6.	Concision	Pratique utilisabilité		7.	Contrôle explicite	Pratique utilisabilité		8.	Flexibilité	Pratique utilisabilité		<table border="1"> <thead> <tr> <th>AVANTAGES</th> <th>INCONVENIENTS</th> </tr> </thead> <tbody> <tr> <td>Coûts</td> <td>: Faibles</td> </tr> <tr> <td>Temps moyen</td> <td>: 2 jours</td> </tr> <tr> <td>Accessibilité aux utilisateurs finaux</td> <td>: Obligatoire</td> </tr> <tr> <td>Facilité d'utilisation</td> <td>: Oui</td> </tr> <tr> <td>Version automatisée</td> <td>: Non</td> </tr> <tr> <td>Nombre moyen de sujets</td> <td>: Dizaine</td> </tr> <tr> <td>Nombre moyen d'évaluateurs</td> <td>: 2</td> </tr> </tbody> </table>	AVANTAGES	INCONVENIENTS	Coûts	: Faibles	Temps moyen	: 2 jours	Accessibilité aux utilisateurs finaux	: Obligatoire	Facilité d'utilisation	: Oui	Version automatisée	: Non	Nombre moyen de sujets	: Dizaine	Nombre moyen d'évaluateurs	: 2
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION																																																									
N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.																																																						
1.	Actions minimales	Pratique utilisabilité																																																							
2.	Actions explicite	Pratique utilisabilité																																																							
3.	Capacité fonctionnelle	Pratique utilité																																																							
4.	Charge mentale	Pratique utilisabilité																																																							
5.	Compatibilité	Pratique utilisabilité																																																							
6.	Concision	Pratique utilisabilité																																																							
7.	Contrôle explicite	Pratique utilisabilité																																																							
8.	Flexibilité	Pratique utilisabilité																																																							
AVANTAGES	INCONVENIENTS																																																								
Coûts	: Faibles																																																								
Temps moyen	: 2 jours																																																								
Accessibilité aux utilisateurs finaux	: Obligatoire																																																								
Facilité d'utilisation	: Oui																																																								
Version automatisée	: Non																																																								
Nombre moyen de sujets	: Dizaine																																																								
Nombre moyen d'évaluateurs	: 2																																																								
<table border="1"> <thead> <tr> <th colspan="2">CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION</th> </tr> </thead> <tbody> <tr> <td>REFERENCE SUR LA METHODE</td> <td>: Rushinek et al., 1986</td> </tr> </tbody> </table>	CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION		REFERENCE SUR LA METHODE	: Rushinek et al., 1986																																																					
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION																																																									
REFERENCE SUR LA METHODE	: Rushinek et al., 1986																																																								
<input type="button" value="CONTINUER"/> <input type="button" value="ANNULER"/>																																																									

Figure V.5b : La deuxième proposition (Questionnaire) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation d'ASTREE

La figure V.5c présente la troisième proposition de technique d'évaluation (l'*enregistrement vidéo*). Les principes techniques de cette méthode d'évaluation ont également été présentés dans le chapitre II. La méthode de l'enregistrement vidéo a pour but d'enregistrer objectivement l'activité de l'opérateur humain pour estimer sa performance et celle du système Homme(s)-Machine(s) (figure V.5c).

La méthode de l'enregistrement vidéo a pour but d'enregistrer objectivement l'activité de l'opérateur humain pour estimer sa performance et celle du système Homme(s)-Machine(s) (figure V.5c).

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°3																																																												
DESCRIPTION DE LA METHODE ET DES CRITERES																																																												
<table border="1"> <thead> <tr> <th colspan="2">CARACTERISTIQUES DE L'IHM A EVALUER</th> </tr> </thead> <tbody> <tr> <td>Libellé de la méthode</td> <td>: Enregistrement vidéo</td> </tr> <tr> <td>Méthode ou Technique complémentaire</td> <td>: Questionnaires</td> </tr> <tr> <td>Description en entrée (ex: description de tâche)</td> <td>: Pas nécessaire</td> </tr> <tr> <td>Disponibilité de la méthode</td> <td>: Oui</td> </tr> <tr> <td>Formalisation</td> <td>: Méthode non formelle</td> </tr> <tr> <td>Technique d'interaction</td> <td>: Manipulation directe</td> </tr> <tr> <td>Objet d'évaluation</td> <td>: Produit final ou Prototype</td> </tr> <tr> <td>Classe d'évaluateurs exigée</td> <td>: Experte</td> </tr> <tr> <td>Support pour la Capture de données</td> <td>: Vidéo</td> </tr> <tr> <td>Etapes du cycle de développement concernées</td> <td>: Spécification/conception/après réalisation</td> </tr> <tr> <td>Usage de métriques</td> <td>: Pas nécessaire</td> </tr> </tbody> </table>	CARACTERISTIQUES DE L'IHM A EVALUER		Libellé de la méthode	: Enregistrement vidéo	Méthode ou Technique complémentaire	: Questionnaires	Description en entrée (ex: description de tâche)	: Pas nécessaire	Disponibilité de la méthode	: Oui	Formalisation	: Méthode non formelle	Technique d'interaction	: Manipulation directe	Objet d'évaluation	: Produit final ou Prototype	Classe d'évaluateurs exigée	: Experte	Support pour la Capture de données	: Vidéo	Etapes du cycle de développement concernées	: Spécification/conception/après réalisation	Usage de métriques	: Pas nécessaire	<table border="1"> <thead> <tr> <th colspan="2">PRINCIPE DE FONCTIONNEMENT DE LA METHODE</th> </tr> </thead> <tbody> <tr> <td colspan="2">Le principe consiste à enregistrer objectivement l'activité de l'utilisateur pour estimer sa performance et celle du système homme-machine (vidéo, interviews ou questionnaires). Cette approche permet d'identifier les moments où les difficultés sont rencontrées en recherchant les explications a posteriori en fonction des faiblesses de l'interface.</td> </tr> </tbody> </table>	PRINCIPE DE FONCTIONNEMENT DE LA METHODE		Le principe consiste à enregistrer objectivement l'activité de l'utilisateur pour estimer sa performance et celle du système homme-machine (vidéo, interviews ou questionnaires). Cette approche permet d'identifier les moments où les difficultés sont rencontrées en recherchant les explications a posteriori en fonction des faiblesses de l'interface.																																
CARACTERISTIQUES DE L'IHM A EVALUER																																																												
Libellé de la méthode	: Enregistrement vidéo																																																											
Méthode ou Technique complémentaire	: Questionnaires																																																											
Description en entrée (ex: description de tâche)	: Pas nécessaire																																																											
Disponibilité de la méthode	: Oui																																																											
Formalisation	: Méthode non formelle																																																											
Technique d'interaction	: Manipulation directe																																																											
Objet d'évaluation	: Produit final ou Prototype																																																											
Classe d'évaluateurs exigée	: Experte																																																											
Support pour la Capture de données	: Vidéo																																																											
Etapes du cycle de développement concernées	: Spécification/conception/après réalisation																																																											
Usage de métriques	: Pas nécessaire																																																											
PRINCIPE DE FONCTIONNEMENT DE LA METHODE																																																												
Le principe consiste à enregistrer objectivement l'activité de l'utilisateur pour estimer sa performance et celle du système homme-machine (vidéo, interviews ou questionnaires). Cette approche permet d'identifier les moments où les difficultés sont rencontrées en recherchant les explications a posteriori en fonction des faiblesses de l'interface.																																																												
<table border="1"> <thead> <tr> <th colspan="4">POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS</th> </tr> <tr> <th></th> <th>moins important</th> <th>important</th> <th>Très import.</th> </tr> </thead> <tbody> <tr> <td>Requêtes des évaluateurs</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>Conflits entre évaluateurs</td> <td><input checked="" type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Exigence mentale</td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>Exigence physique</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> </tbody> </table>	POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS					moins important	important	Très import.	Requêtes des évaluateurs	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Conflits entre évaluateurs	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Exigence mentale	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Exigence physique	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<table border="1"> <thead> <tr> <th colspan="2">CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION</th> </tr> </thead> <tbody> <tr> <td>Qualité</td> <td>: Inspectabilité</td> </tr> <tr> <td>Support de restitution</td> <td>: Vidéo</td> </tr> <tr> <td>Support d'analyse des résultats</td> <td>: Vidéo</td> </tr> <tr> <td>Niveau d'abstraction</td> <td>: Tous</td> </tr> <tr> <td>Type d'information</td> <td>: Explicative</td> </tr> <tr> <td>Performance de la méthode ou Technique</td> <td>: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne</td> </tr> </tbody> </table>	CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION		Qualité	: Inspectabilité	Support de restitution	: Vidéo	Support d'analyse des résultats	: Vidéo	Niveau d'abstraction	: Tous	Type d'information	: Explicative	Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne																					
POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS																																																												
	moins important	important	Très import.																																																									
Requêtes des évaluateurs	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																									
Conflits entre évaluateurs	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>																																																									
Exigence mentale	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																																									
Exigence physique	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																																									
CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION																																																												
Qualité	: Inspectabilité																																																											
Support de restitution	: Vidéo																																																											
Support d'analyse des résultats	: Vidéo																																																											
Niveau d'abstraction	: Tous																																																											
Type d'information	: Explicative																																																											
Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne																																																											
<table border="1"> <thead> <tr> <th colspan="3">NIVEAUX D'AUTOMATISATION DE LA METHODE</th> </tr> </thead> <tbody> <tr> <td>Capture automatique des données</td> <td>:</td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>Restitution automatique des données après traitement</td> <td>:</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Analyse automatique</td> <td>:</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Critique automatique</td> <td>:</td> <td><input type="checkbox"/></td> </tr> </tbody> </table>	NIVEAUX D'AUTOMATISATION DE LA METHODE			Capture automatique des données	:	<input checked="" type="checkbox"/>	Restitution automatique des données après traitement	:	<input type="checkbox"/>	Analyse automatique	:	<input type="checkbox"/>	Critique automatique	:	<input type="checkbox"/>																																													
NIVEAUX D'AUTOMATISATION DE LA METHODE																																																												
Capture automatique des données	:	<input checked="" type="checkbox"/>																																																										
Restitution automatique des données après traitement	:	<input type="checkbox"/>																																																										
Analyse automatique	:	<input type="checkbox"/>																																																										
Critique automatique	:	<input type="checkbox"/>																																																										
<table border="1"> <thead> <tr> <th colspan="3">CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION</th> </tr> <tr> <th>N°</th> <th>*Libellé</th> <th>*Acceptabilité *Dimension d'évaluat.</th> </tr> </thead> <tbody> <tr> <td>1.</td> <td>Contrôle explicite</td> <td>Pratique utilisabilité</td> </tr> <tr> <td>2.</td> <td>Groupements</td> <td>Pratique utilisabilité</td> </tr> <tr> <td>3.</td> <td>Guidage</td> <td>Pratique utilisabilité</td> </tr> <tr> <td>4.</td> <td>Qualité des messages</td> <td>Pratique utilisabilité</td> </tr> <tr> <td>5.</td> <td>Signification des codes</td> <td>Pratique utilisabilité</td> </tr> <tr> <td>6.</td> <td></td> <td></td> </tr> <tr> <td>7.</td> <td></td> <td></td> </tr> </tbody> </table>	CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION			N°	*Libellé	*Acceptabilité *Dimension d'évaluat.	1.	Contrôle explicite	Pratique utilisabilité	2.	Groupements	Pratique utilisabilité	3.	Guidage	Pratique utilisabilité	4.	Qualité des messages	Pratique utilisabilité	5.	Signification des codes	Pratique utilisabilité	6.			7.			<table border="1"> <thead> <tr> <th colspan="2">AVANTAGES</th> <th colspan="2">INCONVENIENTS</th> </tr> </thead> <tbody> <tr> <td>Coûts</td> <td>: Moyens</td> <td></td> <td></td> </tr> <tr> <td>Temps moyen</td> <td>: 2 jours</td> <td></td> <td></td> </tr> <tr> <td>Accessibilité aux utilisateurs finaux</td> <td>:</td> <td>Obligatoire</td> <td></td> </tr> <tr> <td>Facilité d'utilisation</td> <td>:</td> <td>Non</td> <td></td> </tr> <tr> <td>Version automatisée</td> <td>:</td> <td>Non</td> <td></td> </tr> <tr> <td>Nombre moyen de sujets</td> <td>:</td> <td>2</td> <td></td> </tr> <tr> <td>Nombre moyen d'évaluateurs</td> <td>: 3</td> <td></td> <td></td> </tr> </tbody> </table>	AVANTAGES		INCONVENIENTS		Coûts	: Moyens			Temps moyen	: 2 jours			Accessibilité aux utilisateurs finaux	:	Obligatoire		Facilité d'utilisation	:	Non		Version automatisée	:	Non		Nombre moyen de sujets	:	2		Nombre moyen d'évaluateurs	: 3		
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION																																																												
N°	*Libellé	*Acceptabilité *Dimension d'évaluat.																																																										
1.	Contrôle explicite	Pratique utilisabilité																																																										
2.	Groupements	Pratique utilisabilité																																																										
3.	Guidage	Pratique utilisabilité																																																										
4.	Qualité des messages	Pratique utilisabilité																																																										
5.	Signification des codes	Pratique utilisabilité																																																										
6.																																																												
7.																																																												
AVANTAGES		INCONVENIENTS																																																										
Coûts	: Moyens																																																											
Temps moyen	: 2 jours																																																											
Accessibilité aux utilisateurs finaux	:	Obligatoire																																																										
Facilité d'utilisation	:	Non																																																										
Version automatisée	:	Non																																																										
Nombre moyen de sujets	:	2																																																										
Nombre moyen d'évaluateurs	: 3																																																											
<table border="1"> <thead> <tr> <th colspan="2">REFERENCE SUR LA METHODE</th> </tr> </thead> <tbody> <tr> <td>:</td> <td>Youmans, 1987</td> </tr> </tbody> </table>	REFERENCE SUR LA METHODE		:	Youmans, 1987																																																								
REFERENCE SUR LA METHODE																																																												
:	Youmans, 1987																																																											
<input type="button" value="CONTINUER"/> <input type="button" value="ANNULER"/>																																																												

Figure V.5c : La troisième proposition (Enregistrement vidéo) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation d'ASTREE

I.5. Comparaison entre les méthodes utilisées pour l'évaluation d'ASTREE et celles proposées par ADHESION

Dans cette section, nous allons comparer les méthodes d'évaluation utilisées (Questionnaires, Observations des tactiles visuelles, Mouchard électronique) et les propositions de méthodes du système d'aide ADHESION (Analyse des protocoles, Questionnaires, Enregistrements vidéo). Pour procéder à cette comparaison, nous allons d'abord définir des critères de comparaison. Ces derniers doivent tenir compte des objectifs d'évaluation recherchés, des exigences de la tâche principale que doivent réaliser les opérateurs humains et des conditions particulières dans lesquelles évolueront les futurs utilisateurs du système étudié.

I.5.1. Elaboration des critères de comparaison

Cette section concerne l'élaboration préliminaire des critères de comparaison entre les méthodes et techniques réellement utilisées et les propositions du système ADHESION. Pour y parvenir, il a notamment fallu mettre en évidence toutes les caractéristiques particulières du système à évaluer et les problèmes types qu'il engendre. Ces paramètres ont permis d'établir une synthèse de critères ou paramètres de comparaison (voir tableau V.3). Ces paramètres sont les suivants :

Caractéristiques particulières du système à évaluer

□ Activité principale

- Complexité de la tâche
- Spécialisation de la tâche
- Classe des opérateurs humains
 1. Spécialiste
 2. Novice
 3. Première expérience
- Forte densité des tâches
- Forte charge de travail
- Exigences physiques
- Exigences mentales
- Forte fréquence de présentation des informations à l'écran
- Nature des informations présentées à l'écran
 1. Graphique
 2. Numérique
 3. Alphanumérique
- Type d'utilisation
 1. Multi-utilisateur
 2. Mono-utilisateur

□ Par rapport à l'évaluation

- Exigibilité d'une formation préalable des sujets
- Classe des évaluateurs
 1. Spécialiste
 2. Novice
 3. Première expérience

□ Type de problèmes engendrés par l'activité principale

- Base de résolution
 1. Stratégies
 2. Plans de reconnaissance
 3. Expérience personnelle
- Exigences physiques
- Exigences mentales
- Mode de fonctionnement du système
 1. Enchaînement des séquences
 2. Mode « pas à pas »
 3. Mode continu

❑ Localisation (pays de destination et environnement de travail)

- Environnement perturbé et nuisible
- Exigences socio-culturelles
 1. Faibles exigences
 2. Exigences moyennes
 3. Fortes exigences

I.5.2. Tableau de synthèse des critères de comparaison

Le tableau V.3 représente une synthèse de ces paramètres de comparaison que nous allons utiliser par la suite pour comparer les méthodes et techniques d'évaluation réellement utilisées et les propositions d'ADHESION.

Caractéristiques particulières du Système et de son environnement		Critères ergonomiques spécifiques à ces caractéristiques à prendre en compte absolument (C)	
❑ Activité principale	• Tâche complexe	Contrôle explicite (C 0)	
	• Tâche spécialisée	Expérience de l'utilisateur (C1)	
	• Classe des opérateurs	1. Spécialiste	Expérience de l'utilisateur (C 2)
		2. Novice	Expérience de l'utilisateur, Clarté (C 3)
		3. Première expérience	Expérience de l'utilisateur (C 4)
	• Forte densité des tâches	Charge mentale (C 5)	
	• Exigences mentales	Charge mentale (C 6)	
	• Exigences physiques	Retour informatif immédiat (C 7)	
	• Forte fréquence de présentation des informations à l'écran	Qualité des messages (C 8)	
	• Nature des Informations Présentées à l'écran	1. Graphique	Couleurs (C 9)
2. Numérique		Clarté, Concision (C 10)	
3. Alphanumérique		Clarté (C 11)	
• Système multi-utilisateur	Actions explicites (C 12)		
❑ Activité principale par rapport à l'évaluation	• Exigences d'une formation préalable des sujets	Signification des codes, Clarté, Flexibilité (C 13)	
	• Classe Evaluateurs	1. Spécialisée	Tous (C 14)
		2. Novice	Incitation, Clarté, Protection contre les erreurs, Correction des erreurs (C 15)
		3. Première expérience	Incitation, Clarté, Protection contre les erreurs, Correction des erreurs (C 16)
❑ Types de problèmes engendrés par l'activité principale • Conflits • Contrôle de seuils • Recherche • Traitement • Analyse	• Base de Résolution	1. Stratégies	Stratégies, Enchaînement (C17)
		2. Plans de reconnaissances	Problèmes types, activité (C 18)
		3. Expérience personnelle	Connaissances sur le système (C19)
	• Exigences physiques	Charge de travail (C 20)	
	• Exigences mentales	Charge mentale (C 21)	
	• Mode de fonctionnement du système	1. Enchaînements des séquences	Enchaînement (C22)
		2. Mode « pas à pas »	Guidage (C 23)
		3. Mode continu	Homogénéité et consistance (C 24)
	❑ Localisation (pays de destination et environnement de travail)	• Environnement perturbé et nuisible	Gestion des erreurs (C 25)
		• Exigences socioculturelles	1. Faibles exigences
2. Exigences moyennes			Dimensions d'éthique et culturelle (C 27)
3. Fortes exigences			Tous les critères socio-culturels (C 28)

Tableau V.3 : Tableau des critères de comparaison

Après avoir défini les critères de comparaison, nous allons passer à la comparaison entre les méthodes et techniques réellement utilisées et les propositions d'ADHESION (tableau V.4).

Caractéristiques particulières du Système et de son environnement		Cas étudié	Capacité des méthodes à vérifier les critères spécifiques C (inspirée des travaux de GRISLIN [1995])						Critères spécifiques en rapport avec les méthodes	
			Méthodes et techniques d'évaluation utilisées			Propositions du système ADHESION				
			ASTREE	M1	M2	M3	P1	P2		P3
☐ Activité principale	• Tâche complexe	Oui		■	■				C0	
	• Tâche spécialisée	Oui			■				C1	
	• Classe des opérateurs	4. Novice	Non						C2	
		5. Première expérience	Oui			■			C4	
	• Forte densité des tâches	Oui	■		■				C5	
	• Exigences mentales	Oui					■		C6	
	• Exigences physiques	Non							C7	
	• Forte fréquence de présentation des informations à l'écran	Oui	■		■		■	■	C8	
	• Nature des informations présentées à l'écran	4. Graphique	Oui			■		■		C9
		5. Numérique	Non							C10
		6. Alphanumérique	Oui	■		■		■		C11
	• Système multi-utilisateur	Oui		■	■		■	■	C12	
☐ Activité principale Par rapport à l'évaluation	• Exigence d'une formation préalable des sujets	Oui		■	■		■	■	C13	
	• Classe Evaluateurs	4. Spécialisée	Oui			■			C14	
		5. Novice	Non						C15	
		6. Première expérience	Non						C16	
☐ Types de problèmes engendrés par l'activité principale	• Base de Résolution	4. Stratégies	Oui				■		C17	
		5. Plans de reconnaissances	Non						C18	
		6. Expérience personnelle	Oui			■		■	C19	
	• Exigences physiques	Non							C20	
	• Exigences mentales	Oui	■		■		■	■	C21	
	• Mode de fonctionnement du système	4. Enchaînement des Séquences	Non							C22
		5. Mode « pas à pas »	Non							C23
		6. Mode continu	Oui	■		■		■	■	C24
☐ Localisation (pays de destination et environnement de travail)	• Environnement perturbé et nuisible	Oui	■	■	■		■	■	C25	
	• Exigences socio-culturelles	6. Faibles exigences	Oui						C26	
		7. Exigences moyennes	Non						C27	
		8. Fortes exigences	Non						C28	

Légende : ■ Capacité totale de la méthode à vérifier les critères ergonomiques Ci selon la classification de GRISLIN [1995]

■ Capacité partielle

- M1 : moucharé électronique
M2 : Observations des tactiques visuelles
M3 : Questionnaire
P1 : Analyse des protocoles
P2 : Questionnaire
P3 : Enregistrement vidéo

Tableau V.4 : Tableau de synthèse de la phase de préparation des données

Le tableau V.4 présente une synthèse de cette comparaison. La comparaison consiste à vérifier les méthodes et techniques qui prennent en compte les critères d'évaluation spécifiques (notés Ci) lorsque ces critères concernent le système étudié (ce qui correspond aux cases « oui »). Rappelons que dans ce tableau, le système étudié est le système ASTREE.

Dans ce tableau, on constate que les deux classes de méthodes (celles utilisées réellement pour évaluer ASTREE et celles proposées par le système ADHESION) dégagent deux grandes tendances. En effet, en faisant une analyse horizontale du tableau, on constate que la première classe de méthodes (celles réellement utilisées : M1, M2, M3) est orientée vers la charge mentale (C5, C6) alors que les propositions d'ADHESION sont orientées vers les stratégies (C17) et le contrôle (contrôle explicite) (C12).

Ces deux grandes tendances sont révélatrices des critères d'évaluation les plus importants qui caractérisent l'activité principale des opérateurs humains, celle de la régulation du trafic ferroviaire de la SNCF. Le problème le plus important de cette activité est la gestion des conflits qui implique des stratégies, d'où la nécessité de disposer de méthodes qui sont capables de vérifier des critères tels que les *stratégies* et *enchaînements*. En définitive, les méthodes appropriées pour l'évaluation du système ASTREE sont celles qui prennent nécessairement en compte les critères de *charge de travail* et de *stratégies* en plus des autres critères ergonomiques qui améliorent la qualité du système ASTREE en terme d'utilisabilité.

En résumé du tableau de comparaison V.4, on peut dire que :

- les propositions de méthodes et techniques fournies par ADHESION ne permettent pas de vérifier totalement les critères de charge mentale (sous-critère de *la charge de travail*) et de clarté (sous-critère de *guidage*) si la deuxième proposition (P2), qui est la technique du questionnaire, ne contient pas des questions exhaustives. En effet, selon GRISLIN [1995], le questionnaire peut vérifier la plupart des critères ergonomiques si les questions sont exhaustives, ce qui n'est pas toujours facile dans la pratique.
- les propositions de méthodes et techniques fournies par ADHESION permettent de vérifier tous les critères d'évaluation que les méthodes et techniques de la première classe (M1, M2, M3) sont capables de vérifier, en dehors de la *charge mentale* et de la *clarté* qui sont conditionnées par l'exhaustivité du questionnaire.

Les méthodes et techniques de la première classe (M1, M2, M3), qui sont des méthodes et techniques réellement utilisées, ne permettent pas de vérifier le respect des critères tels que les *stratégies*, l'*enchaînement*, les *problèmes type*, les *connaissances sur le système* et l'*activité* qui sont des critères que peuvent par contre vérifier la méthode de l'analyse des protocoles [GRISLIN, 1995]. Or, ces critères sont très importants dans ce type de système (ASTREE) où les opérateurs humains sont appelés à résoudre des problèmes types tels que les conflits dans le réseau SNCF. La résolution des conflits implique absolument l'adoption de stratégies.

En conclusion, les propositions de méthodes et techniques du système ADHESION sont susceptibles de répondre aussi bien à ce contexte et à ce système (ASTREE) que celles de la première classe (les méthodes réellement utilisées par les spécialistes du LAMIH). En effet, les

propositions de méthodes ont en effet bien pris en compte la complexité du système ASTREE en proposant des critères spécifiques au problème qu'il engendre (la gestion des conflits), et le respect de ces critères peut être vérifié à l'aide de l'analyse des protocoles. Cette proposition est renforcée et complétée par les techniques d'enregistrement vidéo et les questionnaires. La répartition presque identique des couleurs du tableau V.4, entre les deux parties (méthodes réellement utilisées et méthodes proposées) indique des capacités en principe comparables des deux groupes de méthodes et techniques en terme d'évaluation d'ASTREE.

Ceci signifie que le système ADHESION aurait pu apporter des réponses pertinentes à une équipe de projet se posant la question des critères à considérer et des méthodes d'évaluation à utiliser.

II. Mise en œuvre et validation sur un second cas d'étude : le projet (MOGADOR)

Dans cette seconde partie, l'étude d'un autre cas est présentée, celui-ci entrant dans le cadre du projet MOGADOR [NENDJO ELLA, 1992]. Dans cette partie seront comparées les méthodes réellement utilisées durant ce projet (couvrant les années 1991-1992) et les propositions du système d'aide à la décision en évaluation ADHESION.

II.1. Contexte du projet MOGADOR

C'est un cas où l'auteur de cette thèse était à la fois évaluateur et formateur des opérateurs humains. Le but des travaux était d'évaluer le système de gestion d'eau (MOGADOR) de deux centrales hydroélectriques (voir figure V.6) et de former sur place les opérateurs humains ensuite. La décision d'utiliser une IHM pour gérer l'eau fait suite à une année de sécheresse exceptionnelle que le Gabon n'avait pas connue depuis 1983.

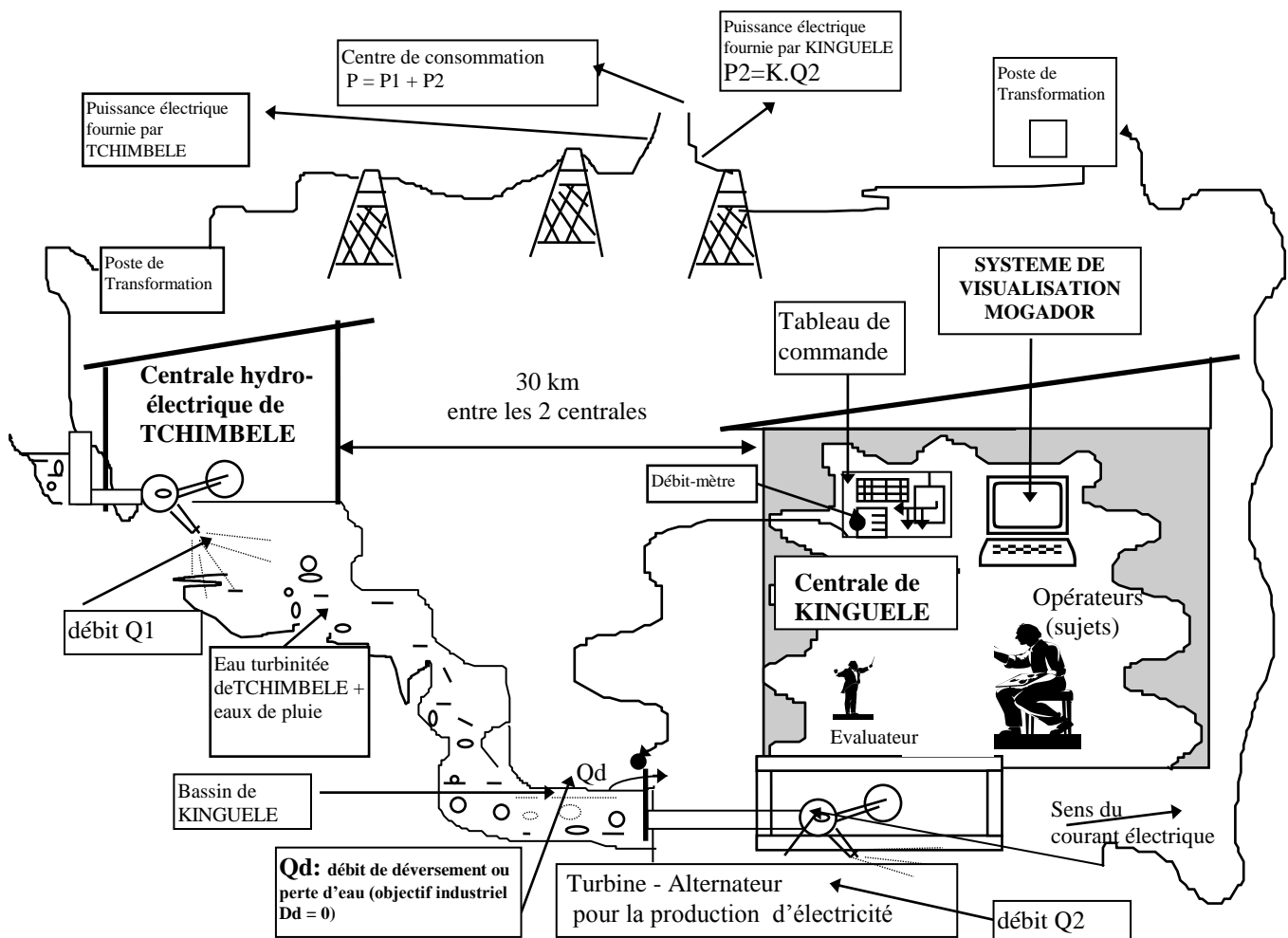


Figure V.6 : Gestion automatisée des bassins hydrauliques des centrales hydroélectriques de KINGUELE et TCHIMBELE

L'architecture de MOGADOR est simple. Il est équipé d'un écran de visualisation qui indique l'évolution de la charge (en MGW/h) en fonction du niveau d'eau dans le bassin hydraulique. Il ne comprend que des informations alphanumériques, aucune n'est graphique.

II.1.1. La technique d'évaluation utilisée initialement

La technique initialement utilisée au cours du projet MOGADOR consistait en une *Mesure de performance*. Rappelons dans un premier temps que l'estimation de la charge de travail s'effectue en mesurant une donnée représentative du niveau de difficulté de l'utilisateur lorsqu'il réalise sa tâche. Toute augmentation de la charge de travail engendre une dégradation de sa performance sur la tâche principale. Cette technique est limitée par le fait qu'il est difficile de définir la variable ou l'ensemble des variables dépendante(s) de la tâche principale, qui sera à même de refléter la charge de travail de l'utilisateur [ODGEN *et al.*, 1979].

Ainsi, au cours de l'évaluation, l'évaluateur a pris comme donnée représentative du niveau de difficulté le débit de déversement Qd. Ce débit varie en fonction de la charge de travail des opérateurs humains qui doivent gérer de manière simultanée deux centrales hydroélectriques en cascade. Ce débit varie également en fonction des erreurs faites par les opérateurs. L'objectif théorique de MOGADOR est d'obtenir un débit de déversement $Qd = 0 \text{ m}^3$, c'est-à-dire d'éviter des pertes en eau pendant les mois de faibles pluviométries (saison dite de sécheresse : juin, juillet, août, septembre).

II.1.2. Le protocole expérimental mis en œuvre initialement

La population concernée était celle des opérateurs qui devaient réellement piloter le système par la suite, ce qui représente une situation idéale. Il y avait 3 sujets dont 2 affectés à la centrale de KINGUELE, l'autre étant de la centrale de TCHIMBELE.

La phase d'évaluation a d'abord été précédée d'une phase de formation au cours de laquelle les sujets devaient apprendre et comprendre le fonctionnement du système. Compte tenu du niveau de formation des sujets, plusieurs jours de formation ont été nécessaires, courant plusieurs phases :

1. une phase de présentation du système MOGADOR pendant près de 5 heures ;
2. une phase de formation sur le principe de fonctionnement du système MOGADOR ;
3. une phase d'exécution d'un jeu de scénarios pour lesquels les courbes de charges (en Kwh) étaient fictives ;
4. une phase d'entretien pour d'éventuelles questions posées par les opérateurs humains.

II.1.3. Dispositif expérimental initialement mis en place

MOGADOR offre une source d'information importante pour l'activité de gestion des centrales hydroélectriques en « cascade », ce qui signifie que la centrale *i* réutilise l'eau débitée par la centrale *i-1* et ainsi de suite. L'utilisation de MOGADOR nécessite un relevé manuel de la courbe de charge grâce au débit-mètre installé en salle de contrôle (figure V.7) et qui permet de vérifier que MOGADOR a fait de bonnes prévisions de répartition de charges entre les deux centrales.

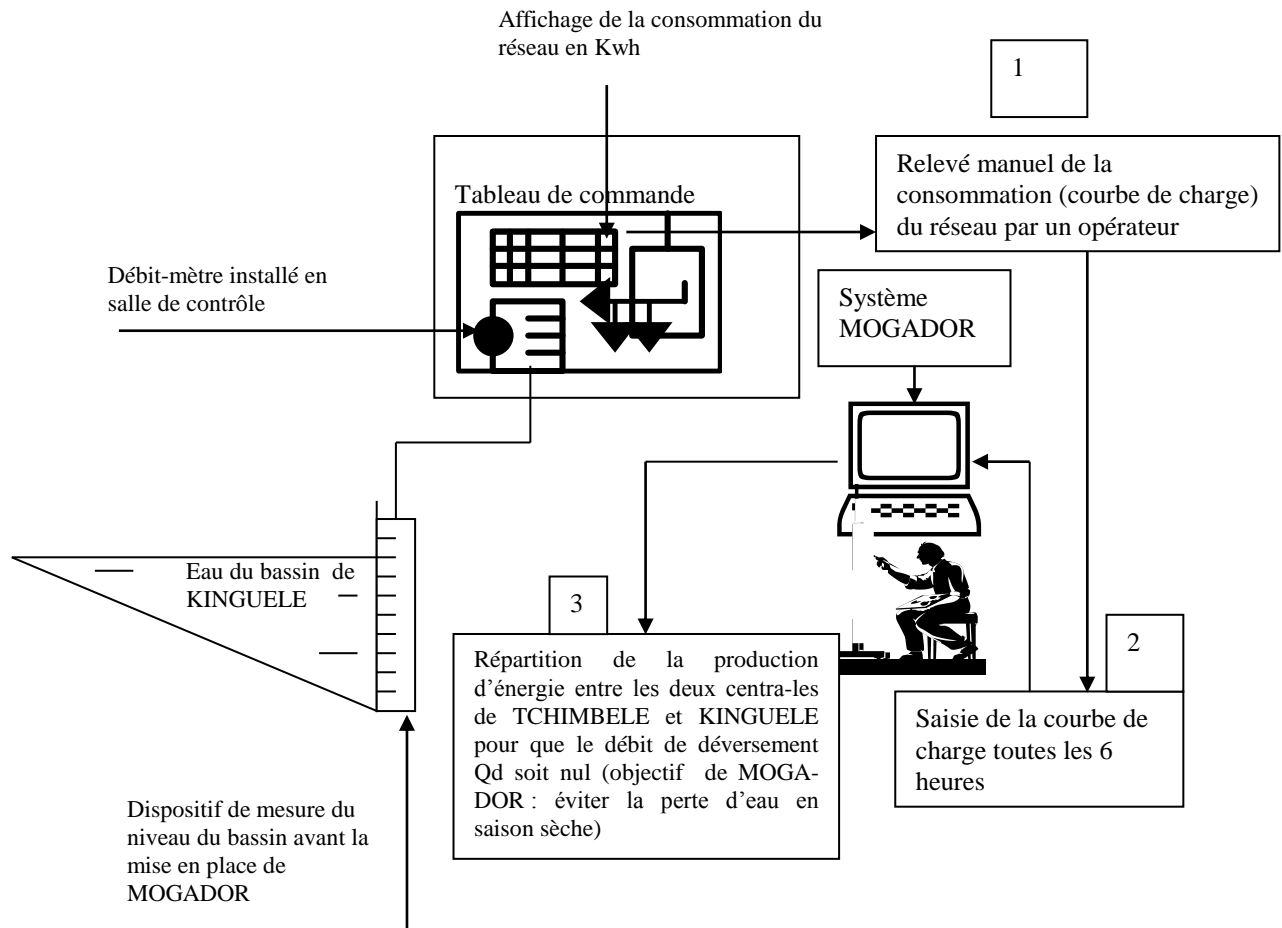


Figure V.7 : Architecture matérielle du système MOGADOR

On peut ainsi, grâce au débit-mètre, vérifier à chaque fois l'objectif ($Q_d = 0$, correspondant à un déversement nul). La qualité de ces informations reste cependant tributaire de la qualité du relevé de la courbe de charge qui demeure manuel et donc source d'erreurs. C'est l'une des faiblesses de MOGADOR, qui aurait dû être conçu pour enregistrer automatiquement la consommation du réseau (voir figure V.2).

II.1.4. Présentation des résultats initialement obtenus

Nous allons présenter dans cette section les résultats obtenus à l'aide de la méthode utilisée pour évaluer MOGADOR : la *mesure de performance*. Rappelons que cette méthode est basée sur une mesure de donnée représentative du niveau de difficulté de l'opérateur lorsqu'il réalise sa tâche (cette donnée peut correspondre à un temps de réponse, une erreur par rapport à la consigne, etc.). Dans notre cas, la consigne était $Q_d=0$ et il fallait mesurer les écarts par rapport à celle-ci. La figure V.8 présente la mesure de la performance de MOGADOR. Il exprime le débit de déversement par rapport au temps mis pour évaluer le système. Pour rendre cette courbe plus parlante, il avait été décidé d'utiliser un rapport $T = 3 Q_o / (Q_d + Q_o)$ en ordonnée et, en abscisse,

les dizaines de jours, avec Q_0 le débit maximal du bassin de KINGUELE. Nous sommes ici devant le cas d'une évaluation qui dure plusieurs dizaines de jours car il fallait évaluer le système selon tous les scénarios possibles (forte pluie, faible pluie et absence de pluie). Notre objectif était de valider le système dans un maximum de cas de pluviométrie possibles afin d'étudier si nous obtenions toujours $Q_d=0$.

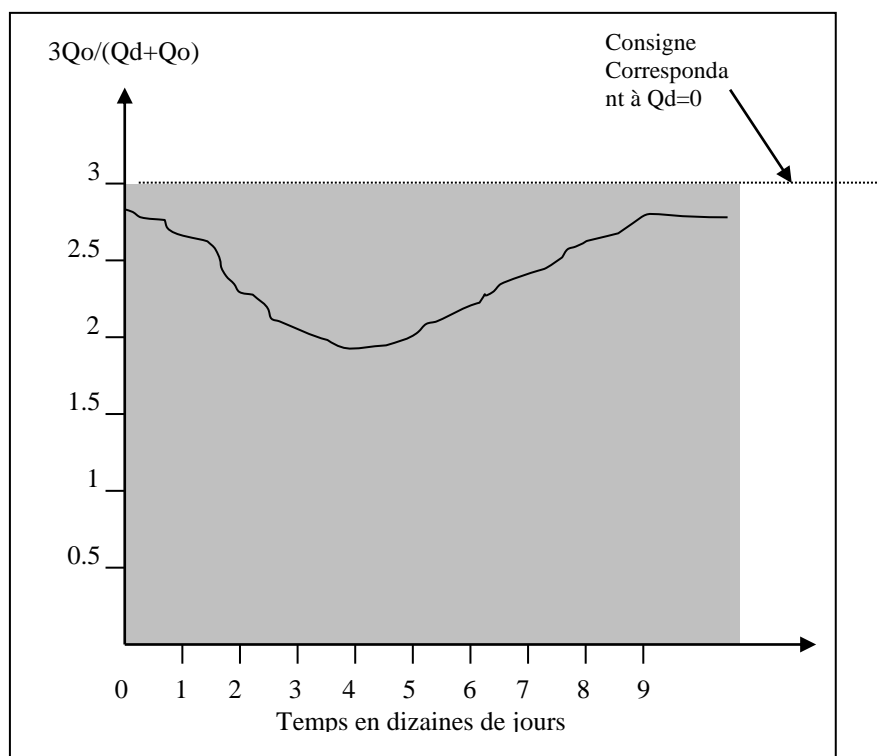


Figure V.8 : Taux d'erreurs obtenu

Sur cette courbe, le « creux » représente le volume d'eau perdu sur plusieurs dizaines de jours. Au vu de celui-ci, il est évident que l'utilisation de MOGADOR ne donne pas les résultats espérés et ceci pour plusieurs raisons. Tout d'abord, les relevés manuels sont sources d'erreurs et il faut noter, en second lieu, le refus constaté des opérateurs de passer d'un système de gestion totalement manuelle à un système automatique dont l'arrivée est signe de réduction des effectifs. Une enquête préalable et une sensibilisation à ce type de technologie aurait en fait été nécessaire avant d'effectuer ce transfert technologique.

En effet, l'arrivée de MOGADOR dans les centrales de TCHIMBELE et KINGUELE avait causé la frustration des opérateurs qui travaillaient manuellement depuis très longtemps. Le savoir-faire manuel est un signe de fierté dans cette région du monde, et constitue un acte de fierté. De plus, comme le rappelait un opérateur : « nous ne pouvons pas nous séparer de notre savoir-faire manuel ». Il y avait derrière cette parole des problèmes socio-culturels dans un environnement où la population a toujours effectué des travaux dont la connaissance se transmet de père en fils depuis des décennies.

Le système avait été abandonné après l'évaluation suite aux mauvaises prévisions de MOGADOR. Ces mauvaises prévisions étaient dues à la mauvaise volonté des opérateurs qui saisissaient des courbes de charges avec des erreurs grossières, ceci afin de démontrer à l'administration qu'un système comme MOGADOR ne pouvait remplacer leur savoir-faire. En fait, si on avait tenu compte des exigences des opérateurs humains, et ceci aurait été possible par exemple grâce à un questionnaire socio-technique, beaucoup de problèmes d'ordre social auraient été détectés. L'approche de cette application aurait pu être autre.

II.2. Mise en œuvre du système d'aide ADHESION

Après avoir constaté les résultats obtenus avec la méthode dite de mesure de performance, il est possible passer à l'étape de mise en œuvre d'ADHESION afin d'obtenir des propositions de méthodes qui auraient permis d'évaluer MOGADOR autrement. Pour y parvenir, nous allons procéder exactement comme pour le cas d'ASTREE, mais cette fois en présentant directement le tableau de synthèse de la phase de préparation des données.

II.2.1. Tableau de synthèse de la phase de préparation des données

Après avoir préparé les données nécessaires au fonctionnement d'ADHESION, nous avons résumé la situation du deuxième cas dans le tableau de synthèse ci-dessous (tableau V.5).

Questions	Réponses	
	Modificateur	Descripteur
Avez-vous accès aux utilisateurs finaux ?	Très	Oui
Le groupe d'évaluation dispose-t-il d'un ergonome ?	Plutôt	Moyen
Le groupe d'évaluation dispose-t-il d'un sociologue ?	Très	Non
Le groupe d'évaluation dispose-t-il d'un psychologue ?	Très	Non
Le groupe d'évaluation dispose-t-il d'un spécialiste en droit informatique?	Très	Non
Le groupe d'évaluation dispose-t-il d'un spécialiste en assurance qualité?	Modérément	Non
Le budget d'évaluation vous semble-t-il suffisant ?	Très	Oui
Le temps alloué à l'évaluation vous semble-t-il suffisant ?	Très	Oui
Pouvez-vous financer le dépassement du temps prévu ?	Très	Moyen
Pouvez-vous financer les spécialistes d'évaluation manquants ?	Plutôt	Non
Le système à évaluer évoluera-t-il dans un environnement nuisible ?	Très	Oui
Disposez-vous de matériel spécialisé (système oculométrique, enregistrement vidéo, ...) ?	Très	Non
Avez-vous une expérience en évaluation ?	Modérément	Oui
Pouvez-vous assimiler le système à un type grand public ?	Très	Non
L'IHM a-t-elle franchi l'étape de réalisation ?	Très	Oui
Pays de destination du système à évaluer MOGADOR	Gabon	

Tableau V.5 : Tableau de synthèse de la phase de préparation des données

Ayant déjà expliqué dans le premier cas en quoi consiste cette étape, nous allons immédiatement fournir les données issues de la phase de préparation et qui traduisent le contexte dans lequel évoluera l'évaluation.

II.2.2. Saisie des contraintes (données) par l'utilisateur de l'outil d'aide

Comme on l'a déjà dit, la phase de préparation des données est suivie de la saisie des contraintes. L'introduction des données dans le système consiste à cocher les cases correspondant aux réponses aux questions posées (voir figures V.9a à V.9c), en utilisant toujours le tableau de synthèse (tableau V.5).

DESCRIPTION DES CONTRAINTES

CETTE TECHNIQUE DE RESOLUTION VOUS OBLIGE A REPENDRE
A TOUTES LES QUESTIONS CI-DESSOUS

Sélectionnez le pays de destination (ou assimilé) de l'IHM à évaluer

Estonie	↑
Etats-Unis	
Ethiopie	
Finlande	
France	
Gabon	
Gambie	
Gaza et Cisjordanie	
Géorgie	
Ghana	↓

AIDE A LA COMPREHENSION DES QUESTIONS

Avez-vous accès aux utilisateurs finaux ?
Répondre Plutôt-Moyen, signifie que vous avez près de 50% de chance d'accéder aux futurs utilisateurs de l'IHM que vous voulez évaluer, alors que Très - Oui signifie 100% de chance.

Le Groupe d'évaluation, dispose t-il d'un ergonomes ?
Répondre Très - Non signifie que vous ne disposez pas du tout d'ergonomes, alors que Plutôt-Moyen signifie qu'il existe une chance sur deux pour que vous obtenez un ergonomes, cela peut être le cas d'une attente

***** Descripteurs ***** ***** Modificateurs *****

	Bon/Oui	Moyen	Mauvais/Non	Plutôt	Modérément	Très
Avez-vous accès aux utilisateurs finaux?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un ergonomes?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un sociologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un psychologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en droit-informatique?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en assurance qualité?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le budget-évaluation vous semble t-il suffisant?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le temps alloué à l'évaluation vous semble t-il suffisant?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pouvez-vous financer le dépassement du temps prévu?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pouvez-vous financer les spécialistes d'évaluation manquant?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le système à évaluer évoluera t-il dans un environ.nuisible?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Disposez-vous du matériel spécialisé (système oculométrique.,enreg.vidéo,)?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Avez-vous une expérience en évaluation ?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous assimiler le système à évaluer à un type grand public?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
L'IHM à évaluer a t-elle franchit l'étape de réalisation?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figure V.9 : Saisie des contraintes pour le deuxième cas (MOGADOR)

La figure V.9 présente la saisie des contraintes supposées refléter les facteurs situationnels pour l'évaluation de MOGADOR. On peut remarquer que le pays sélectionné est le GABON car le système est destiné à ce pays. La section suivante est consacrée aux résultats obtenus pour ce second cas d'étude.

II.2.3. Présentation des propositions de solutions fournies par le système d'aide ADHESION pour le cas étudié (projet MOGADOR)

Le système, en suivant la même procédure que celle décrite pour le premier cas d'étude, propose un certain nombre de méthodes, techniques et critères d'évaluation. Les figures V.10a à V.10c présentent ces différentes propositions obtenues par la mise en œuvre d'ADHESION. La figure V.6a montre la première méthode proposée (les *traces écrites du travail*).

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°1			
DESCRIPTION DE LA METHODE ET DES CRITERES			
CARACTERISTIQUES DE L'IHM A EVALUER		LE PRINCIPE DE FONCTIONNEMENT LA METHODE	
Libellé de la méthode	: Traces écrites du Travail	Le principe consiste à analyser les traces écrites des utilisateurs (notes écrites, documents de conduite, cahiers de doléances, enregistrements des entrées utilisateurs). Cette approche vise également à localiser des lacunes par les changements sollicités ainsi, que les problèmes empêchant d'exécuter une tâche efficacement.	
Méthode ou Technique complémentaire	: Interviews		
Description en entrée (ex: description de tâche)	: Pas nécessaire		
Disponibilité de la méthode	: Oui		
Formalisation	: Méthode non formelle		
Technique d'interaction	: Toutes		
Objet d'évaluation	: Produit final ou Prototype		
Classe d'évaluateurs exigée	: Experte/Novice		
Support pour la Capture de données	: Ordinateur/papier		
Etapes du cycle de développement concernées	: Réalisation		
Usage de métriques	: Pas nécessaire		
POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS		CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION	
moins important important Très import.		Qualité	: Inspectabilité
Requêtes des évaluateurs	: <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	Support de restitution	: Papier
Conflits entre évaluateurs	: <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>	Support d'analyse des résultats	: Papier/Ordinateur
Exigence mentale	: <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	Niveau d'abstraction	: Tous
Exigence physique	: <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	Type d'information	: Objective
NIVEAUX D'ATOMATISATION DE LA METHODE		Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input type="checkbox"/> Bonne <input checked="" type="checkbox"/> Très Bonne
Capture automatique des données	: <input type="checkbox"/>		
Restitution automatique des données après traitement	: <input type="checkbox"/>		
Analyse automatique	: <input type="checkbox"/>		
Critique automatique	: <input type="checkbox"/>		
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION		AVANTAGES	INCONVENIENTS
N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.
1.	Guidage	Pratique	utilisabilité
2.	Charge de travail	Pratique	utilisabilité
3.	Contrôle explicite	Pratique	utilisabilité
4.	Adaptabilité	Pratique	utilisabilité
5.	Gestion des erreurs	Pratique	utilisabilité
6.	Signification des codes	Pratique	utilisabilité
7.	Compatibilité	Pratique	utilisabilité
REFERENCE SUR LA METHODE		Coûts	: Faibles
		Temps moyen	: 1 jour
		Accessibilité aux utilisateurs finaux	: Obligatoire
		Facilité d'utilisation	: Oui
		Version automatisée	: Non
		Nombre moyen de sujets	: 3
		Nombre moyen d'évaluateurs	: 3
		<input type="button" value="CONTINUER"/> <input type="button" value="ANNULER"/>	

Figure V.10a : La première proposition (Traces écrites de travail) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation de MOGADOR

A travers l'analyse des traces écrites du travail des opérateurs humains, il aurait été possible de détecter leurs difficultés et leurs frustrations par rapport à ce système qui leur était étranger sur tous les points de vue puisque ces opérateurs utilisaient jusqu'à présent leur savoir-faire manuel pour résoudre les problèmes techniques de la centrale. La frustration avait pour conséquences un taux d'erreurs important dans le relevé de la courbe de charge par exemple. En

effet, en comparant les travaux d'un jour donné à ceux réalisés dans un passé immédiat, on aurait pu détecter la négligence et les erreurs intentionnelles. L'arrivée du nouveau système avait bouleversé leur organisation habituelle et leurs « habitudes » de travail, ce qui créait un choc culturel entre le savoir-faire manuel et l'automatisation dont ils ne voyaient pas bien l'intérêt. Une méthode comme les Traces écrites du travail aurait permis de ne pas les influencer et de récolter naturellement leurs difficultés et leur insatisfaction.

La figure V.10b montre une deuxième proposition du système d'aide à la décision, consistant en l'utilisation d'un *questionnaire socio-technique*.

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°2			
DESCRIPTION DE LA METHODE ET DES CRITERES			
CARACTERISTIQUES DE L'IHM A EVALUER		PRINCIPE DE FONCTIONNEMENT DE LA METHODE	
Libellé de la méthode	: Questionnaires socio-technique	Le principe consiste à mettre en évidence les problèmes d'ordre culturel, social et éthique. Pour y parvenir on sélectionne un groupe d'utilisateurs représentatifs du pays de destination de l'IHM à évaluer. Ainsi, une série de questions leur sont posées les uns après les autres. Cette étape est suivie par une analyse des données recueillies.	
Méthode ou Technique complémentaire	: Interviews		
Description en entrée (ex: description de tâche)	: Pas nécessaire		
Disponibilité de la méthode	: Oui		
Formalisation	: Méthode non formelle		
Technique d'interaction	: Toutes		
Objet d'évaluation	: Produit final ou Prototype		
Classe d'évaluateurs exigée	: Experte/Novice		
Support pour la Capture de données	: Ordinateur/papier		
Etapes du cycle de développement concernées	: Spécification/conception/après réalisation		
Usage de métriques	: Possible		
POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS		CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION	
Requêtes des évaluateurs	: <input type="checkbox"/> moins important <input type="checkbox"/> important <input checked="" type="checkbox"/> Très import.	Qualité	: Concision
Conflits entre évaluateurs	: <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	Support de restitution	: Papier
Exigence mentale	: <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Support d'analyse des résultats	: Papier/Ordinateur
Exigence physique	: <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Niveau d'abstraction	: Tous
NIVEAUX D'AUTOMATISATION DE LA METHODE		Type d'information	: Subjective/Explicative
Capture automatique des données	: <input type="checkbox"/>	Performance de la méthode ou Technique	: <input type="checkbox"/> Moy. <input checked="" type="checkbox"/> Bonne <input type="checkbox"/> Très Bonne
Restitution automatique des données après traitement	: <input type="checkbox"/>		
Analyse automatique	: <input type="checkbox"/>		
Critique automatique	: <input type="checkbox"/>		
CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION		AVANTAGES	INCONVENIENTS
N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.
1.	Confidentialité	Sociale	culturelle
2.	Langue	Sociale	culturelle
3.	Couleur	Sociale	culturelle
4.	Signe	Sociale	culturelle
5.	Moralité	Sociale	sociale
6.	Symbole	Sociale	culturelle
7.	Concision	Pratique	utilisabilité
8.	Coût	Pratique	utilisabilité
REFERENCE SUR LA METHODE		Coûts	: Faibles
		Temps moyen	: 2 jours
		Accessibilité aux utilisateurs finaux	: Obligatoire
		Facilité d'utilisation	: Oui
		Version automatisée	: Non
		Nombre moyen de sujets	: Dizaine
		Nombre moyen d'évaluateurs	: 2
CONTINUER		ANNULER	

Figure V.10b : La deuxième proposition (questionnaire socio-technique) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation de MOGADOR

La deuxième proposition de technique d'évaluation fournie par le système d'aide à la décision ADHESION est le questionnaire socio-technique (ANNEXE 3). Cette technique est intéressante puisqu'elle aurait pu permettre de détecter les incohérences culturelles ressenties par

les opérateurs humains. Le transfert technologique d'une culture à une autre nécessite effectivement la vérification du respect des critères socio-culturels. En proposant cette technique, le système ADHESION a en fait considéré que le Gabon fait partie des pays où les aspects socio-culturels sont importants.

La figure V.10c représente la troisième et dernière proposition de méthode ou technique d'évaluation (Mesure de la performance). On remarque que, contrairement au premier cas étudié, le système n'a ici proposé que trois méthodes et techniques d'évaluation. Cela n'est pas étonnant car il existe des contextes où le système ne propose pas de méthodes du tout ; dans ce cas, il doit proposer à l'utilisateur des modifications concernant les contraintes liées au projet.

PROPOSITION DE METHODES ET CRITERES D'EVALUATION : PROPOSITION N°3

DESCRIPTION DE LA METHODE ET DES CRITERES

CARACTERISTIQUES DE L'IHM A EVALUER

Libellé de la méthode : Mesure de performance
 Méthode ou Technique complémentaire : Interviews
 Description en entrée (ex: description de tâche) : Pas nécessaire
 Disponibilité de la méthode : Oui
 Formalisation : Méthode non formelle
 Technique d'interaction : Toutes
 Objet d'évaluation : Produit final ou Prototype
 Classe d'évaluateurs exigée : Experte
 Support pour la Capture de données : Ordinateur/Enregistreurs
 Etapes du cycle de développement concernées : Après réalisation
 Usage de métriques : Oui

PRINCIPE DE FONCTIONNEMENT DE LA METHODE

Le principe consiste à estimer la charge de travail en mesurant une donnée représentative du niveau de difficulté de l'utilisateur lorsqu'il réalise sa tâche (temps de réponse, erreurs par rapport à la consigne).
Toute augmentation de la charge de travail engendre une dégradation de sa performance sur la tâche principale. Voir par exemple la méthode de la double tâche qui consiste à estimer l'exigence de la tâche principale en ajoutant une tâche supplémentaire qui sature l'utilisateur.

POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS

moins important important Très import.

Requêtes des évaluateurs :

Conflits entre évaluateurs :

Exigence mentale :

Exigence physique :

CARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION

Qualité : Concision
 Support de restitution : Papier
 Support d'analyse des résultats : Papier/Ordinateur
 Niveau d'abstraction : Action physique
 Type d'information : Objective
 Performance de la méthode ou Technique : Moy. Bonne Très Bonne

NIVEAUX D'AUTOMATISATION DE LA METHODE

Capture automatique des données :

Restitution automatique des données après traitement:

Analyse automatique :

Critique automatique :

CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION

N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.
1.	Charge de Travail	Pratique	utilisabilité
2.			
3.			
4.			
5.			
6.			
7.			
8.			

AVANTAGES

Coûts :
 Temps moyen : 2 jours
 Accessibilité aux utilisateurs finaux :
 Facilité d'utilisation : Oui
 Version automatisée :
 Nombre moyen de sujets :
 Nombre moyen d'évaluateurs : 2

INCONVENIENTS

Assez importants
 Obligatoire
 Non
 3

REFERENCE SUR LA METHODE : Leplat et al., 1967

Figure V.10c : La troisième proposition (mesure de la performance) de méthode d'évaluation fournie par le système d'aide ADHESION pour l'évaluation de MOGADOR

Très appropriée pour les mesures industrielles, la détermination de la performance du système MOGADOR à l'aide de cette méthode d'évaluation est très convenable car elle est

simple et que l'élément représentatif choisi pour la mesure est laissé à l'appréciation des évaluateurs. Dans certaines situations industrielles, elle s'avère très efficace et révèle immédiatement les insuffisances du système que l'on désire évaluer.

II.3. Comparaison entre la méthode utilisée pour l'évaluation de MOGADOR et celles proposées par ADHESION

Dans cette section, nous allons comparer la méthode réellement utilisée pour évaluer MOGADOR et les propositions de méthodes et techniques fournies le système d'aide ADHESION. La démarche est ici la même que pour le premier cas, ce qui nous permet d'aller directement dans le tableau de comparaison (tableau V.6).

On constate ici une nette différence qui confirme l'affirmation selon laquelle une seule méthode ou technique d'évaluation ne peut à elle seule vérifier et valider un système. Rappelons que les propositions fournies par ADHESION peuvent être complétées par d'autres ; on peut en effet voir sur les figures V.10a à V.10c que les propositions peuvent être complétées par une autre technique (en l'occurrence ici des *interviews*).

En résumé du tableau de comparaison V.6, on peut dire que :



- La méthode de la mesure de la performance ne permet pas d'évaluer MOGADOR de façon suffisante. Cette méthode, à elle seule, ne permet pas de mettre en évidence la qualité ergonomique. Son usage exclusif avait en fait été motivé par des raisons économiques ; on n'a donc pas tenu compte des problèmes ergonomiques ou des aspects socio-culturels. La société a plutôt privilégié les aspects fonctionnels de MOGADOR ; l'orientation a été de se focaliser exclusivement sur la mesure de la performance pour répondre aux objectifs de l'entreprise, ce qui a été catastrophique car on a laissé de côté les aspects humains.
- Les propositions de méthodes et techniques fournies par ADHESION auraient permis d'évaluer MOGADOR aussi en terme d'acceptabilité sociale que pratique. L'acceptabilité sociale est très importante dans la plupart des pays en voie de développement, à un point tel que les aspects de fonctionnalité peuvent venir en second.

Conclusion

En définitive, comme dans ce dernier cas, on voit que les propositions d'ADHESION auraient apporté une aide appréciable à tout évaluateur, chef de projet ou concepteur, même si ce dernier ne connaît pas les réalités du pays destinataire du système à évaluer.

On peut ici noter l'apport considérable du raisonnement « flou » dans une aide à la décision où la subjectivité est très importante. Cette aide à la décision présente néanmoins l'inconvénient d'être très sensible aux contraintes (données d'entrée du système). En effet, lorsque les utilisateurs saisissent des informations, ces dernières doivent se rapprocher le plus possible du contexte réel d'évaluation. Si ce principe n'est pas respecté, le système d'aide ADHESION aura tendance à fournir des propositions ne permettant pas de vérifier et de valider convenablement un système.

Caractéristiques particulières du système et de son environnement		Cas étudié	Capacité des méthodes à vérifier les critères spécifiques C (inspirée des travaux de GRISLIN [1995])				Critère spécifique en rapport avec les méthodes	
			Méthodes et techniques d'évaluation utilisées	Propositions du système ADHESION				
				M1	P1	P2		P3
<input type="checkbox"/> Activité principale	• Tâche complexe	Non					C1	
	• Tâche spécialisée	Non					C2	
	• Classe des opérateurs	1. Spécialiste	Non					C4
		2. Novice	Oui					C3
		3. Première expérience	Oui					C4
	• Forte densité des tâches	Non					C5	
	• Exigences mentales	Non					C6	
	• Exigences physiques	Oui					C7	
	• Forte fréquence de présentation des informations à l'écran	Non					C8	
	• Nature des informations présentées à l'écran	1. Graphique	Non					C9
		2. Numérique	Oui					C10
		3. Alphanumérique	Oui					C11
• Système multi-utilisateur	Non					C12		
<input type="checkbox"/> Activité principale par rapport à l'évaluation	• Exigence d'une formation préalable des sujets	Oui					C13	
	• Classe des évaluateurs	1. Spécialiste	Non				C14	
		2. Novice	Non				C15	
		3. Première expérience	Oui				C16	
<input type="checkbox"/> Types de problèmes engendrés par l'activité principale	• Base de résolution	1. Stratégies	Non				C17	
		2. Plans de reconnaissance	Oui				C18	
		3. Expérience personnelle	Non				C19	
	• Exigences physiques	Oui					C20	
	• Exigences mentales	Oui					C21	
	• Mode de fonctionnement du système	1. Enchaînement des séquences	Non					C22
		2. Mode « pas à pas »	Non					C23
		3. Mode continu	Oui					C24
Localisation (pays de destination et environnement de travail)	• Environnement perturbé et nuisible	Oui					C25	
	• Exigences socioculturelles	4. Faibles exigences	Non				C26	
		5. Exigences moyennes	Non				C27	
		6. Fortes exigences	Oui				C28	

Légende :  Capacité totale de la méthode à vérifier les critères ergonomiques Ci selon la classification de GRISLIN [1995]
 Capacité partielle

M1 : Mesure de la performance

P1 : Traces écrites de travail

P2 : Questionnaire socio-technique

P3 : Mesure de performance

Tableau V.6 : Tableau de comparaison entre les propositions d'ADHESION et la technique d'évaluation réellement utilisée

Ces deux premières validations techniques montrent bien la potentialité et la faisabilité de cette approche d'aide à la décision.

Un long chemin reste encore à parcourir pour parvenir à des propositions pertinentes et précises du système AHESION dans tous les cas de figure. Ce chemin passe par un ensemble de perspectives de recherche décrites dans la dernière partie du mémoire.

CONCLUSION GENERALE
ET
PERSPECTIVES DE RECHERCHE

Conclusion générale

Entre le développement des systèmes Homme(s) – Machine (s) devenant de plus en plus interactifs et l'évolution des mentalités au sein de la société, les rapports entre l'Homme et la machine se sont complexifiés. En effet, face à la machine dont les fonctionnalités se sont accrues, l'Homme est devenu plus exigeant en terme d'acceptabilité pratique (utilité, utilisabilité, etc.) et d'acceptabilité sociale (culture, éthique, etc.). Face aux nouvelles exigences de l'Homme, certains auteurs à l'image de NIELSEN [1993] pensent que, pour parvenir à des systèmes interactifs dits acceptables, il faut les considérer selon deux angles : l'acceptabilité sociale et l'acceptabilité pratique. Dans l'acceptabilité pratique, on retrouve les notions d'utilité et d'utilisabilité largement étudiées dans la littérature vis-à-vis de la conception et de l'évaluation de systèmes interactifs. Par contre, la notion d'acceptabilité sociale revêt souvent un aspect que l'on pourrait presque qualifier de « mystérieux » et est largement négligée par les équipes de développement de systèmes interactifs, malgré le besoin croissant de systèmes utilisés dans un contexte international.

On peut dès lors imaginer que la prise en compte de l'acceptabilité pratique et sociale va faire appel à des connaissances pluridisciplinaires : les équipes d'évaluation ne seront plus constituées que par des informaticiens, des automaticiens, des ergonomes ou des spécialistes en assurance qualité ; il sera nécessaire d'y intégrer d'autres spécialistes tels que les socio-techniciens. L'évaluation des systèmes interactifs prend alors une nouvelle dimension et va amener à se poser un certain nombre de questions.

En plus du problème des nouvelles compétences à intégrer dans les équipes d'évaluation, il s'agira de résoudre les problèmes survenant lors du processus d'évaluation, en terme d'organisation et du choix des méthodes d'évaluation. En effet, la multiplicité (on compte aujourd'hui plus de 100 méthodes et techniques d'évaluation) et la diversité de plus en plus considérable des méthodes et techniques d'évaluation (avec l'arrivée de méthodes qui exigent de nouvelles connaissances, telles le questionnaire socio-technique décrit dans le chapitre II), il se pose naturellement le problème du choix de méthodes et techniques et des critères d'évaluation adéquats pour évaluer un système donné.

Ces problèmes de choix auxquels peuvent être confrontés des acteurs tels qu'un chef de projet qui s'interroge sur les méthodes et critères disponibles au regard de ses propres critères, un évaluateur, voire un apprenant en interaction Homme(s) – Machine(s), ont attiré notre attention et justifié l'objectif de nos travaux.

Ainsi, pour apporter un élément de réponse à ces différents acteurs, nous nous sommes fixés comme objectif de mettre à leur disposition un outil qui puisse leur permettre d'obtenir des propositions concernant des méthodes, techniques et critères adéquats et ceci en fonction des contraintes situationnelles.

Pour ce faire, nous avons étudié et mis au point la première version d'un système d'aide à la décision pour le choix des méthodes, techniques et critères d'évaluation. Pour y parvenir, il

fallu dans un premier temps déterminer les critères de décision pouvant être pris en compte, ceci ayant été rendu possible grâce aux travaux antérieurs sur le choix des méthodes et techniques d'évaluation [BALBO, 1992] [BALBO, 1992][GRISLIN et KOLSKI, 1996] [STANTON et NEVILLE, 1996] [NENDJO ELLA et JACQUES, 1999].

La complexité des paramètres influençant le choix des méthodes et leur nature (informations objectives et subjectives) nous a orienté vers une démarche d'aide à la décision en évaluation basée sur la théorie des sous-ensembles flous et des possibilités et implémentée à l'aide de Prolog III. Rappelons que les utilisateurs de notre système d'aide à l'évaluation peuvent être de plusieurs catégories : il peut s'agir de chefs de projets, d'évaluateurs novices (voire expérimentés), ou apprenants en interaction Homme(s)-Machine (s).

Cette réflexion pluridisciplinaire a conduit à la première version d'un système appelé **ADHESION** (Aide à la **D**écision **H**umaine pour l'**E**valuation des **S**ystèmes **I**nteractifs et leur validati**ON**), ayant pour entrée différentes contraintes (moyens humains, localisation, forme du produit, type d'application, etc.) pouvant influencer sur le choix d'une méthode donnée. Ces contraintes, exprimées d'abord par des réponses à des questions à l'utilisateur du système, permettent à l'outil de sélectionner automatiquement la/les méthode(s) ou technique(s) adaptée(s) au système à évaluer.

Avant cela, il nous a fallu dans un premier temps réfléchir sur les différents aspects et travaux en rapport avec notre problématique. C'est ainsi que nous avons commencé par identifier les problèmes relatifs à l'évaluation des systèmes interactifs après avoir précisé le principe de l'évaluation, les difficultés liées au choix des méthodes d'évaluation et souligné l'importance des critères socio-culturels [*chapitre I*]. Après cette étape, nous avons tracé le contour de la notion d'aide à l'évaluation en précisant différentes approches existantes, ce qui a conduit à une classification des différents outils d'aide à l'évaluation et à une synthèse des méthodes d'évaluation [*chapitre II*]. Ces deux premiers chapitres ont montré que ce sont les méthodes d'évaluation empiriques qui sont les plus utilisées.

Après avoir discuté de l'évaluation proprement dite et des différentes limites des outils d'aide à l'évaluation existants, le *troisième chapitre* était consacré aux problèmes liés à l'aide à la décision en évaluation. On a également présenté les différentes stratégies et les préférences qui ont été établies à partir des connaissances expertes et modélisées à l'aide des sous-ensembles flous (par des modèles trapézoïdaux). Nous avons par la suite présenté la démarche proposée pour l'aide à la décision en évaluation.

Dans le *quatrième chapitre*, nous avons présenté une description de l'algorithme utilisé et une présentation des pages-écrans de l'outil et le *dernier chapitre* a été consacré à la validation sur deux scénarios d'évaluation provenant de cas réels ; l'un de ces deux scénarios a mis en évidence la prise en compte des critères socio-culturels. Ces premiers résultats sont encourageants en terme de modélisation et de solutions fournies.

□ Contribution

Rappelons ici que l'objectif était de mettre à la disposition des évaluateurs ou autres acteurs intéressés de près ou de loin à des questions d'évaluation, une aide à la décision pour le choix des méthodes, techniques et critères d'évaluation. Cet objectif est renforcé par l'idée de NIELSEN [1993] qui pense que, pour parvenir à des systèmes interactifs dits acceptables, il faut les considérer selon deux angles : l'acceptabilité sociale et l'acceptabilité pratique.

Ce travail nous a permis d'intégrer des notions importantes dans le domaine de l'évaluation et de surcroît la conception des systèmes interactifs, en termes de :

- Apport d'une contribution originale d'aide à l'évaluation des systèmes interactifs destinés à être utilisés dans n'importe quel pays
- choix automatique des méthodes, techniques et critères d'évaluation adaptés à chaque situation : faisabilité prouvée,
- prise en compte du niveau d'expérience des évaluateurs,
- prise en compte des aspects des critères culturels et d'éthique,
- prise en compte des critères sociaux,
- modélisation du processus de décision,
- pluridisciplinarité (socio-technique, ergonomie, droit et informatique, etc.) dans les équipes d'évaluation.

□ Limites et perspectives

Malgré l'apport du système d'aide ADHESION dans le domaine de l'évaluation des systèmes interactifs, il faut souligner qu'il présente des limites et aborde un domaine nouveau où il n'existe pas de règles de production bien définies. C'est l'une des raisons qui a motivé le choix des sous-ensembles flous qui sont bien adaptés, selon WAWAK [1996], à des problèmes nouveaux où les règles de production sont inexistantes ou en très faible quantité, contrairement aux systèmes experts. La preuve est que, comme nous l'avons vu, s'il existe des travaux sur l'aide au choix de méthodes et techniques d'évaluation, dans l'orientation (c'est-à-dire que l'évaluateur est guidé dans le choix de méthodes, par exemple : FACE, Denley et LONG), il n'existe pas de démarches offrant automatiquement des méthodes et techniques d'évaluation.

Les limites du système d'aide sont diverses :

- Les résultats sont fortement tributaires de la véracité des contraintes introduites (en d'autres termes si l'utilisateur du système ADHESION fournit des contraintes qui sont très loin de la réalité de son contexte d'évaluation, le système fournira des propositions non adéquates).
- La convivialité de l'IHM d'ADHESION est à améliorer.

Les perspectives de recherche sont nombreuses, liées :

- (1) à l'enrichissement, au développement et aux tests des différents modules,

- (2) à l'enrichissement de la description de la situation de départ ; en effet, cette phase est la plus sensible lors de la mise en œuvre du système ADHESION dans la mesure où une description incomplète ou éloignée de la réalité, peut entraîner des biais ou lacunes lors de la résolution. Cette partie nécessite une réflexion qui tournerait autour de l'élaboration d'une démarche formelle de description de la situation de départ ; la possibilité de plusieurs niveaux de description serait également à étudier.
- (3) à l'expertise à intégrer dans la base de connaissance : l'expertise à intégrer dans la base de connaissance, peut faire l'objet d'un programme de recherche considérable, car les connaissances sont pluridisciplinaires et sont largement réparties parmi de nombreux experts au niveau international.
- (4) aux tests dans le cadre de projets réels ou simulés ; deux scénarios ont été utilisés pour valider la faisabilité de la méthode, il s'agira bien entendu dans les années à venir d'aller beaucoup plus loin en se basant sur un ensemble diversifié de situations d'évaluation à étudier.
- (5) à la prise en compte (1) de critères spécifiques, tel que le plaisir lié à l'interaction enfant-ordinateur et (2) de méthodes d'évaluation adaptés aux systèmes interactifs destinés aux enfants [MALONE, 1983][ROBERT *et al.*, 1999]. De nombreux domaines d'application pourraient être étudiés dans ce sens (Internet, CAO...).

En définitive, nous dirons que parmi les nouvelles approches d'aide à l'évaluation, telles que FACE et le plan d'aide de DENLEY et LONG qui guident l'évaluateur dans le choix de méthodes et techniques, le système ADHESION est actuellement la seule approche de ces nouvelles classes d'aide, qui fournit automatiquement des méthodes et techniques d'évaluation. C'est une contribution majeure dans le domaine d'évaluation des systèmes interactifs et dans une certaine mesure cela peut également apporter de l'aide dans la conception, dans la mesure où le système ADHESION fournit également des critères d'évaluation susceptibles d'être utilisés dans les phases amont de conception, lors d'un projet.

**REFERENCES
BIBLIOGRAPHIQUES**

REFERENCES BIBLIOGRAPHIQUES

- ABED M. (1990).** *Contribution à la méthodologie de la tâche de conception par des outils de spécification exploitant les mouvements oculaires : application à la conception et l'évaluation des Interfaces Homme-Machine.* Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Septembre.
- ABED M., ANGUE J.C. (1994).** A new method for conception, realisation and evaluation of man-machine. *Proceedings of the IEEE International conference on systems, man and cybernetics*, San Antonio, Texas, October 2-5, pp. 1420-1427.
- ACHILLES A., ELSTER K.H., NEHSE R. (1979).** Bibliographie zur Vektoroptimierung (Theorie and Anwendungen). *Mathematische Operationsforschung und Statistik*, 10, pp. 277-321.
- ACIER B. (1990).** *Contribution à la conception et à l'évaluation d'Interfaces Homme-Machine par l'analyse du comportement visuel d'un superviseur humain.* Thèse de Doctorat en Automatique industrielle et humaine, université de Valenciennes et du Hainaut-Cambrésis, Janvier.
- ADREIT F., OZKAN N., SENEGES V., VAILLANT P. (1992).** Ergonomie et méthodes de conception et d'évaluation des IHM. *Quatrièmes journées sur l'ingénierie des Interfaces Homme-Machine IHM'92*, Paris, 30 Novembre, pp. 1-2.
- AMALBERTI R., BATEJAT D., BATAILLE M., MENU J.C., SANTUCCI G. (1986).** Charge de travail mental, validation en France d'une technique d'évaluation subjective américaine : SWAT. *Revue de Médecine Aéronautique et Spatiale.*, Tome XXV, n° 98, pp. 158-163.
- BALBO S., COUTAZ J. (1992).** Un pas vers l'évaluation automatique des Interfaces Homme-Machine. *Conférence Internationale ERGO-IA'92*, Biarritz, France, 7-9 Octobre, pp. 367-382.
- BALBO S. (1994).** *Evaluation ergonomique des interfaces utilisateur : un pas vers l'automatisation.* Thèse de Doctorat en Informatique, Laboratoire de Génie informatique – IMAG, Université Joseph Fourier, Grenoble I, France, Septembre.
- BARNARD P.J. (1987).** Cognitive resources and the learning of human-computer dialogs. In *Interfacing thought*, J.M. Carroll (Ed.), The MIT Press, pp. 112-158.
- BARTHET M.F. (1988).** *Logiciels interactifs et ergonomie, modèles et méthodes de conception.* Editions Dunod Informatique, Paris.
- BASTIEN C. (1996).** *Les critères ergonomiques : un pas vers une aide méthodologique à l'évaluation des systèmes interactifs.* Thèse de Doctorat en Ergonomie cognitive, Université René Descartes, Paris V, Décembre.
- BASTIEN C., SCAPIN D.L. (1993).** Preliminary findings on the effectiveness of ergonomic criteria for the evaluation of human-computer interfaces. *Proceedings of the INTERCHI'93 conference*, S. Ashlund et al. (Eds.), Amsterdam, April.

- BELKIN N.J., BROOKS H.M., DANIELS P.J. (1987).** Knowledge elicitation using discourse analysis. *International Journal of Man-Machine Studies*, 27, pp. 127-144.
- BENAISSA M.L. (1993).** *Une démarche de conception, réalisation et évaluation d'IHM : application au projet ferroviaire ASTREE*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Décembre.
- BERGER T. (1991).** *Modélisation de la tâche d'un opérateur dans le poste propulsion (PCP) d'un SNA*. Rapport DRET, Contrat n° 89-1023 LAMIH Valenciennes / CERTSM Toulon, 64 p.
- BERNARD J.M. (1994).** *Exploitation des mesures oculométriques dans la modélisation de la tâche prescrite et de l'activité réelle des opérateurs par réseaux de Petri*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis.
- BIAS R.G., GILLAN D.J., TULLIS T.S. (1993).** Three usability enhancements to the human factors – design interface. *Proceedings of the Fifth International Conference HCI'93*, Salvendy and Smith (Eds.), Elsevier, Orlando, Florida, August 8-13, Vol. 2, pp. 169-174.
- BODKER S., GRONBAEK K. (1989).** Cooperative prototyping experiments – users and designers envision a dental case record system. *Proceedings of the first EC-CSCW'89*, J. Bowers and S. Benford (Eds.), UK, Computer sciences company.
- BONNARDEL N. (1992).** *Le rôle de l'évaluation dans les activités de conception*. Thèse de Doctorat en Psychologie cognitive, Université d'Aix-en-Provence, Juillet.
- BOUYSSOU D. (1993).** *Décision multicritère ou aide multicritère ?* Bulletin du groupe de travail européen 'Aide multicritère à la décision', Série 2, n° 2, p. 1.
- BROOKE J. (1996).** SUS : a 'quick and dirty' usability scale. In *Usability evaluation in industry*, P.W. Jordan *et al.* (Eds.), Taylor & Francis, London, pp. 189-194.
- BROWN C.M. (1988).** *Human-Computer Interface design guidelines*. NJ : Ablex.
- BYRNE M.D., WOOD S.D., SUKAVIRIYA P., FOLEY J.D., KIERAS D.E. (1994).** Automating interface evaluation. *Proceedings CHI'94 Conference*, pp. 232-237, ACM Press, Addison Wesley
- CACM (1993).** *Special issue about participatory design*. Communication of the ACM, 36(4), June.
- CALVARY G. (1998).** *Proactivité et réactivité : de l'assignation à la complémentarité en conception et évaluation d'Interfaces Homme-Machine*. Thèse de Doctorat en Informatique, Université Joseph Fourier, Grenoble I, Octobre.
- CARD S.K., MORAN T.P., NEWELL A. (1983).** *The psychology of Human-Computer Interaction*. Hillsdale, NJ : Lawrence Erlbaum Associates.
- CHOMSKY N. (1965).** *Aspects of the theory of syntax*. Mass : MIT Press, Cambridge, USA.
- COOPER G., HARPER B. (1969).** *The use of pilot rating in the evaluation of aircraft handling qualities*. Moffett Field, California : NASA Ames, Report TN-D-5153, April.

-
- COOPER M. (1999).** Evaluating accessibility and usability of web pages. In *Computer-Aided Design of User interfaces II, CADUI'99*, J. Vanderdonckt and A. Puerta (Eds.), Kluwer Academic Publishers, The Netherlands, pp. 33-42.
- COTE V., MARTIN J.F., COULETTE B. (1993).** Un modèle paramétrique pour l'évaluation des interfaces personne-système. *Proceedings on the Sixth International Conference on Software Engineering and its applications*, CNIT, Paris-La Défense, EC2, November 15-19.
- COUTAZ J. (1990).** *Interfaces homme-ordinateur : conception et réalisation*. Bordas, Paris.
- DEBERNARD S. (1993).** *Contribution à la répartition dynamique des tâches entre opérateur et système automatisé : application au contrôle du trafic aérien*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Janvier.
- DELAHAYE J.P. (1988).** *Cours de PROLOG avec TURBO PROLOG : éléments fondamentaux*. Editions Eyrolles, Paris.
- DENLEY I., LONG J. (1997).** A planning aid for human factors evaluation practice Behaviour and Information technology, vol. 16, n°415, pp. 203-219.
- DIENG R. (1990).** Méthodes et outils d'acquisition des connaissances. *Proceedings of the ERGO'IA'90 Conference on 'Ergonomie et informatique avancée'*, Biarritz.
- DUBOIS D., PRADE H. (1985).** A review fuzzy set agregation connectives. *Information Sciences*, n° 36, pp. 85-121.
- DUGUE D., GIRAULT M. (1969).** *Analyse de variance et plans d'expérience*. Dunod, Paris.
- DUNCAN K.D. (1981).** Training for fault diagnosis in industrial process plant. In *Human detection and diagnosis of system failures*, Rasmussen & Rouse (Eds.), Plenum Press, New York.
- DYER J.S., SARIN R.K (1979).** Group preference agregation rules based on strength of preference. *M. Science*, 25, pp. 822-831.
- EASON K.D. (1984).** Towards the experimental study of usability. *Behaviour and information technology*, 3(2), pp. 133-145.
- ERICSSON A.K., SIMON H.A. (1984).** *Protocol analysis verbal : reports as data*. Cambridge, MA : MIT Press.
- ESHRAGH F., MAMDANI E.H (1979).** A general approach to linguistic approximation. *International Journal of Man-Machine studies*, vol 11.
- EZZEDINE H., ANGUE J.C., LANCIEN D., KAMPMANN B., CAPDEVIELLE C. (1995).** Conception de l'interface homme-machine du procédé ferroviaire Astrée. *Revue Générale de l'Electricité*, n° 2, pp. 33-39.
- EZZEDINE H., ABED M. (1997).** Une méthode d'évaluation d'interface homme-machine de supervision d'un procédé industriel. *Journal Européen des Systèmes Automatisés*, 31(7), pp.1087-1110.

-
- FARENC C. (1997).** *ERGOVAL : une méthode de structuration des règles ergonomiques permettant l'évaluation automatique d'interfaces graphiques*. Thèse de Doctorat en Informatique, Université de Toulouse I, Janvier.
- FARENC C., PALANQUE P. (1999).** A generic framework based on ergonomics rules for computer aided design of user interface. In *Computer-Aided Design of User interfaces II, CADUI'99*, J. Vanderdonckt and A. Puerta (Eds.), Kluwer Academic Publishers, The Netherlands, pp. 281-292.
- FISHBURN P.C. (1970).** *Utility theory for decision making*. John Wiley and Sons, New York.
- FLANAGAN C. (1954).** The critical incident technique. *Psychology Bulletin* n° 51, pp. 327-386.
- GOTTINGER H.W., WEIMANN P. (1992).** Intelligent decision support systems. *Decision Support Systems*, n°8, pp. 317-332.
- GAMBOA-RODRIGUEZ F. (1998).** *Spécification et implémentation d'ALACIE : Atelier Logiciel d'Aide à la Conception d'Interfaces Ergonomiques*. Thèse de Doctorat en Sciences, Université de Paris Sud – Paris XI, Octobre.
- GANDIBLEUX X. (1995).** *Système d'aide à la décision pour la conduite de processus perturbés : une approche hybride fondée sur l'intelligence artificielle, la programmation linéaire et l'aide multicritère à la décision*. Thèse de Doctorat en Automatique Industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, janvier.
- GIBSON J.J. (1979).** *The ecological approach to visual perception*. Boston : Houghton-Mifflin.
- GILMORE W.E., GERTMAN D.I., BLACKMAN H.S. (1989).** *User-Computer Interface in process control : a human factors engineering handbook*. Academic Press.
- GRAY W.D., JOHN B.E., ATWOOD M.E. (1992).** The precis of project Ernestine or an overview of validation of GOMS. *Proceedings CHI'92, ACM Conference on human factors in computing systems*, Bauersfeld *et al.* (Eds.), Monterey, California, May 3-7, pp. 307-312.
- GRISLIN M. (1995).** *Définition d'un cadre pour l'évaluation a priori des Interfaces Homme-Machine dans les systèmes industriels de supervision*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Novembre.
- GRISLIN M., KOLSKI C. (1996).** Evaluation des Interfaces Homme-Machine lors du développement des systèmes interactifs. *Technique et Science Informatiques*, Vol. 15, n° 3, pp. 265-296.
- GRISLIN M., KOLSKI C., ANGUE J.C. (1995a).** Towards an organization of man-machine interface evaluation techniques using a usability criteria guid. *Proceedings XII European Annual Conference on Human decision making and manual control*, University of Kassel, Germany, June 22-24, pp. 4.6-4.13.
- GRISLIN M., KOLSKI C., ANGUE J.C. (1995b).** Human-Computer Interface evaluation in industriel complex systems : a review of usable techniques. *Proceedings of 6th IFAC/IFIP/IFORS/IEA Symposium on analysis, design and evaluation of MMS*, MIT Cambridge, USA, June 27-29.

-
- GUERRA T.M. (1991).** *Analyse de données objectifivo-subjectives : approche par la théorie des sous-ensembles flous.* Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Juillet.
- HAMMOND N., HINTON G., BARNARD P., MACLEAN A., LONG J., WHITEFIELD A. (1984).** Evaluating the interface of a document processor : a comparison of expert judgement and user observation. In *Human-Computer Interaction, INTERACT'84*, B. Schnackel (Ed.), Elsevier Science Publishers B.V. : North Holland, IFIP, pp. 725-729.
- HAMMONTREE M.L., HENDRICKSON J.J., HENSLEY B.W. (1992).** Integrated data capture and analysis tools for research and testing on graphical user interfaces. *Proceedings of the CHI'92 Conference*, ACM Press, Monterey, May 3-7, pp. 431-432.
- HANSON S.J., KRAUT R.E., FARBER J.M. (1984).** Interface design and multi-variate analysis of UNIX command use. *ACM Transaction of Office Information Systems*, 2, pp. 24-57.
- HART S.G. (1986).** Development of NASA-TLX (Task Load IndeX) : results of empirical and theoretical research. In *Human mental workload*, P.A. Hancock and N. Meshkati (Eds.), *Advances in psychology*, North Holland, pp. 139-184.
- HART S.G., STAVELAND L.E. (1988).** Theory and measurement of human workload. In J. Reidner (Ed.), *Human productivity enhancement training and human factors in systems design*, Praeger, New-York, pp. 396-456.
- HÛ O., TRIGANO P., CROZAT S. (1999).** Considering subjectivity in softwares evaluation. Application for teachware evaluations. In *Computer-Aided Design of User interfaces II, CADUI'99*, J. Vanderdonck and A. Puerta (Eds.), Kluwer Academic Publishers, The Netherlands, pp. 331-336.
- HULZEBOSCH R., JAMESON A. (1993).** A rapid method for tailored, multiperspective evaluation of user interfaces. In S. Ashlung, K. Mullet, A. Henderson, E. Hollnagel and T. White (Eds.), *Adjunct Proceedings of the ACM INTERCHI'93 Conference on Human factors in computing systems*, Amsterdam, The Netherlands : ACM, April 24-29, pp. 193-194.
- IRVING S., OLSON P., IRVING J.E. (1994).** A GOMS analysis of the advanced automated cockpit. *Proceedings of the CHI'94 Conference*, ACM Press, Addison-Wesley, Boston, Massachusetts, USA, April 24-28, pp. 344-350.
- JACQUES C., RABIER J.C. (1997).** *Dématérialisation des feuilles de soins : migration informatique, carrières et métiers dans les CMCAS EDF-GDF.* Rapport d'enquête, CLERSE-IFRESI-CNRS, Lille, Juin.
- JACQUET-LAGREZE E. (1981).** *Systèmes de décision et acteurs multiples. Contribution à une théorie de l'action pour les sciences des organisations.* Thèse de Doctorat d'Etat, Université de Paris-Dauphine.
- JAMES M.G. (1991).** PRODUSER : PROcess for Developing USER interfaces. In *Taking software design seriously*, J. Karat (Ed.), Academic Press.
- JOHN B.E. (1988).** *Contribution to engineering models of Human-Computer Interaction.* Doctoral dissertation, Carnegie Mellon University.

-
- JOHN B.E., VERA A.H. (1992).** A GOMS analysis of graphic, machine-paced, highly interactive task. *Proceedings of the ACM Conference on Human factors in computing systems*, Bauersfeld et al. (Ed.), Monterey, California, May 3-7, pp. 251-258.
- JOHNSEN G.I., CLEGG C.W., RAVDEN S.J. (1989).** Towards a practical method of user interface evaluation. *Applied Ergonomics*, 20(4), pp. 255-260.
- KADDOURI S.A., EZZEDINE H., SIMON P. (1995).** Evaluation du poste de régulation du trafic ferroviaire du centre opérationnel ASTREE de Pajol. Rapport d'étude, LAMIH, Octobre.
- KEENEY R., RAIFFA H. (1976).** *Decisions with multiple objectives : preferences and value trade-offs.* Wiley and sons, New York.
- KIERAS D., POLSON P.G. (1985).** An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22, pp. 365-394.
- KOLSKI C. (1989).** *Contribution à l'ergonomie de conception des interfaces graphiques homme-machine dans les procédés industriels : application au système expert SYNOP.* Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Janvier.
- KOLSKI C. (1997).** *Interfaces Homme-Machine : application aux systèmes industriels complexes.* Editions Hermès, Paris.
- KOLSKI C., MOUSSA F., MILLOT P. (1988).** *Formalisation de connaissances ergonomiques dans le domaine de l'ergonomie de conception d'interfaces graphiques Homme-Machine.* Congrès de la SELF, Montréal, 3-5 Octobre.
- KOLSKI C., MILLOT P. (1991).** A rule-based approach to the ergonomic « static » evaluation of man-machine graphic interface in industrial processes. *International Journal of Man-Machine Studies*, 35(5), pp. 657-674.
- KOLSKI C., RIERA B., BERGER T., MILLOT P. (1993).** *Analyse en salle de contrôle : lignes directrices pour la conception d'imagerie pour un procédé de même type.* Rapport établi dans le cadre de la convention SOLVAY/LAIH, Octobre.
- KOUBEK R.J., SALVENDY G. (1991).** Cognitive performance of super-experts on computer program modification tasks. *Ergonomics*, 8, pp. 1095-1112.
- KRANTZ D.H. (1967).** Extensive measurement in semiorders. *Philosophy of Science*, 34, pp. 348-362.
- LANNING T.R. (1991).** Let the users design ! In *Taking software design seriously*, J. Karat (Ed.), Academic Press, pp. 127-135.
- LEVINE P., POMEROL M. J., SANEH R. (1990).** Rules integrate data in a multicriteria decision support system. *IEEE Transactions on Systems*, vol. 20, n°3, May/June.
- LEWIS C., POLSON P., WHARTON C., RIEMAN J. (1990).** Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *Proceedings of the CHI'91 Conference on Human-Computer Interaction*, ACM New York, Seattle, April, pp. 235-242.

-
- LONG J. DOWELL J. (1989).** Conceptions of the discipline of the HCI : craft, applied science and engineering. *Proceedings of the Firth Conference of the BCS HCI SIG*, A. Sutcliffe and L. Macaulay (Eds.), Cambridge University Press.
- LOSLEVER P. (1988).** *Etude ergonomique du poste de bureautique : approche par les méthodes multidimensionnelles d'analyse des données.* Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis.
- LÖWGREN J., NORDQVIST T. (1992).** Knowledge-based evaluation as design support for graphical user interfaces. In *Proceedings of the ACM CHI'92 Conference on Human factors in computing systems*, Bauersfeld et al. (Eds.), Monterey, California, May 3-7, pp. 181-188.
- LUCE R.D. (1973).** Three axiom systems for additive semiorordered structures. *SIAM Journal on Applied Mathematics*, 25, pp. 41-53.
- MAGUIRE M., SWEENEY M. (1989).** System monitoring : garbage generator or basic for comprehensive evaluation system ? *Proceedings of the Fifth Conference of the British Computer Society*, HCI Specialist Group, University of Nottingham, September 5-8, pp. 375-394.
- MAHAJAN R., SCHNEIDERMAN B. (1995).** *A family of user interface consistency checking tools.* Report n° CS-TR-3472, College Park : University of Maryland.
- MAHFOUDI A. (1997).** TOOD : une méthodologie de description orientée objet des tâches utilisateur pour la spécification et la conception des interfaces homme-machine, application au contrôle du trafic aérien. Thèse de Doctorat en Automatique Industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, janvier.
- MAIRE F., BRUMENT J.M. (1988).** *La conduite de projet dans les industries de process.* Montrouge. ANACT, Collection Outils et méthodes.
- MAMDANI E.H., ASSILIAN S. (1975).** A linguistic synthesis with fuzzy logic comtroller. *International Journal of Man-Machine Studies*, n° 7, pp. 1-13.
- MARESCHAL B. (1988).** Weights stability intervals in multicriteria decision aid. *European Journal of Operational Research*, 33, pp. 69-77.
- MAYSTRE L.Y., PICTET J., SIMOS J. (1994).** *Méthodes multicritères ELECTRE. Description, conseils pratiques et cas d'application à la gestion environnementale.* Presses polytechniques et universitaires romandes, Lausanne.
- MCCLELAND I. (1990).** Product assessment and user trials. In *Evaluation of human work – a practical ergonomics methodology*, T. Wilson and F. Corlett (Eds.), pp. 218-247.
- MEINADIER J.P. (1991).** *L'interface utilisateur : pour une informatique plus conviviale.* Dunod Informatique, Paris.
- MILLOT P. (1988).** *Supervision des procédés automatisés et ergonomie.* Edition Hermès, Paris.
- MIZIO (1998).** Tuer un flic, c'est ce qui rapporte le plus. *Libération*, n° 579, janvier.

-
- MORAN T.P. (1981).** The command language grammar : a representation of the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, n° 15, pp. 3-50.
- MORAY N. (1986).** Monitoring behaviour and supervisory control. In *Handbook of perception and human performance*, K.R. Boff, L. Kaufman and J.P. Thomas (Eds.), Vol. II, Cognitive processes and performance, John Wiley and sons, Chapter 40, pp. 40.1-40.51.
- MORAY N., JONES B.G., SANDERSON P.M., REISING D.V., SHAHEEN S., RASMUSSEN J. (1995).** The 'Bird's Foot' integrated graphical interface for NPP operation. In *Symbiosis of Human and Artifact, Social aspects of Human-Computer Interaction*, Y. Anzai, K. Ogawa and H. Mori (Eds.), Elsevier Science B.V., pp. 995-998.
- MUNIER B. (1989).** New models of decisions under uncertainty. *European Journal of Operational Research*, 38, pp. 307-317.
- NEAL A.S., SIMON R.M. (1983).** Playback : a method for evaluating the usability of software and its documentation. In *Human factors in computing systems – I*, A. Janda (Ed.), ACM, North Holland, Amsterdam, pp. 78-82.
- NENDJO ELLA A. (1992).** *Projet MOGADOR : l'automatisation des centrales hydroélectriques de Tchimbele et Kinguele, provinces de l'estuaire du Gabon (Afrique centrale)*. Rapport technique, SEEG (Société d'Electricité et d'Eau du Gabon), Septembre.
- NENDJO ELLA A., GRISLIN M., KOLSKI C. (1996).** Towards a global approach to solve social aspects of usability and acceptability, especially in Developing Countries. *Proceedings of the XV European Annual Conference on Human decision making and manual control*, TNO Human Factors Research Institute, Soesterberg, The Netherlands, June 10-12.
- NENDJO ELLA A., KOLSKI C., WAWAK F., JACQUES C., YIM P. (1999a).** An approach of computer-aided choice of UI evaluation criteria and methods. In *Computer-Aided Design of User interfaces II, CADUI'99*, J. Vanderdonckt and A. Puerta (Eds.), Kluwer Academic Publishers, The Netherlands, pp. 319-329.
- NENDJO ELLA A., KOLSKI C., JACQUES C., YIM P. (1999b).** Vers un système d'aide à la décision pour l'évaluation des systèmes interactifs. *Onzième Conférence Francophone sur l'Interaction Homme-Machine, IHM'99*, Le Corum, Montpellier, 22-26 Novembre.
- NENDLO ELLA A., JACQUES C. (1999).** Vers une démarche d'évaluation adaptée aux PVD. *Colloque International sur l'Ethique et la qualité dans l'évaluation*, Reims, France, 25-26 Octobre.
- NEWELL A., SIMON H.A. (1972).** *Human problem solving*. Prentice-Hall, Englewood Cliffs.
- NIELSEN J. (1990).** Designing user interfaces for international use. Elsevier, Amsterdam, North Holland.
- NIELSEN J., MOLICH R. (1990).** Heuristic evaluation of user interfaces. *Proceedings of the CHI'90 Conference on Human-Computer Interaction*, Seattle, ACM Press New York, pp. 349-356.
- NIELSEN J. (1993).** *Usability engineering*. Academic Press, Boston, MA.

-
- NORMAN D.A. (1986).** Cognitive engineering. In *User centred system design : new perspectives on Human-Computer Interaction*, D.A. Norman and S.W. Draper (Eds.), Hillsdale, NJ : Erlbaum Associates, pp. 31-61.
- NOVARA F., BERTAGGIA N., ALLAMANO N. (1987).** Usability evaluation and feedback to designers : an experimental study. In *Human-Computer Interaction, INTERACT'87*, H.J. Bulinger and B. Shackel (Eds.), Elsevier Science Publishers B.V., North Holland, pp. 337-340.
- ODGEN E.C., LEVINE J.M., EISNER E.J. (1979).** Measurement of workload by secondary tasks. *Human Factor*, 21(5), pp. 529-548.
- OLSON J.R., RUETER H.H. (1987).** Extracting expertise from experts : methods for knowledge acquisition. *Experts systems*, Vol. 4, n° 3.
- OMBREDANE A., FAVERGE J.M. (1955).** *L'analyse du travail*. PUF, Paris.
- PARETO W. (1919).** *Traité de sociologie générale*. Editions Payot, Paris.
- PAYNE S.J., GREEN T.R.G. (1989).** Task-Action Grammar : the model and developments. In *Task analysis for Human-Computer Interaction*, D. Diaper (Ed.), John Wiley and Sons, pp. 75-107.
- POITRENAUD S. (1993).** *Le réseau sémantique : une alternative aux grammaires d'action*. Rapport de recherche, Laboratoire de Psychologie cognitive du traitement de l'information symbolique, CNRS – Université de Paris VIII.
- POLLIER A. (1991).** *Evaluation d'une interface par des ergonomes : diagnostics et stratégies*. Rapport de recherche, INRIA, Rocquencourt, France, n° 1391, Février.
- POLSON P.G., LEWIS C.H., RIEMAN J., WHARTON C. (1992).** Cognitive walkthroughs : a method for theory-based evaluation of use interfaces. *International Journal of Man-Machine Studies*, 36, pp. 741-773.
- POULAIN T. (1994).** *Contribution du génie logiciel à la conception et l'évaluation d'applications de supervision*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Février.
- PHILIPPE B. (1993).** *Système d'aide à la conception des protocoles expérimentaux en communication Homme-Machine, fondé sur une méthodologie d'analyse orientée par les objets*. Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Février.
- RASMUSSEN J., VICENTE K.J. (1989).** Coping with human errors through system design : implications for ecological interface design. *International Journal of Man-Machine Studies*, 31, pp. 517-534.
- RAVDEN S.J., JOHNSON G.I. (1989).** *Evaluating usability of Human-Computer Interface : a practical method*. Chichester : Ellis Horwood Books in Information Technology.
- REASON J. (1990).** *Human error*. Cambridge : Cambridge University Press.

- REID G.B., NYGREN T.E. (1988).** The subjective workload assessment technique : a scaling procedure form easuring mental workload. In *Human mental workload*, P.A. Hancock and N. Meshkati (Eds.), Advances in psychology, North Holland, pp. 139-184.
- REISNER P. (1984).** Formal grammars as a tool for analysing ease of use : some fundamental concepts. In *Human factors in computer systems*, J.C. Thomas and M.L. Schneider (Eds.), Norwood, N.J. : Ablex.
- RIERA B., GRISLIN M., MILLOT P. (1994).** Methodology to evaluate man-car interfaces. *IFAC/IFORS Transportation systems : theory and application of advanced technology*, Tianjin, China, August 24-26, preprints Vol. II, pp. 459-469.
- ROBERTS F.S. (1979).** *Measurement theory with applications to decision making, utility and the social sciences*. Addison-Wesley.
- ROOT R.W., DRAPER S. (1983).** Questionnaire as a software evaluation tool. In *Human factors in computing systems – I*, A. Janda (Ed.), ACM, Amsterdam, North Holland, pp. 83-87.
- ROWLEY D.E. (1994).** Usability testing in the field : bringing the laboratory to the user. *Proceedings of the Conference on Human factors in computing systems, CHI'94*, ACM Press, Addison-Wesley.
- ROY B. (1968).** Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *RIRO*, 8, pp. 57-75.
- ROY B. (1978).** Electre III : un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples. *Cahier du CERO*, Vol. 20, n° 1, Paris.
- ROY B. (1988).** Des critères multiples en recherche opérationnelle : pourquoi ?. In *Operational Research '87*, G.K. Rand (Ed.), Elsevier Science Publishers B.V., North Holland, pp. 829-842.
- ROY B. (1985).** *Méthodologie multicritère d'aide à la décision*. Editions Economica, Série Gestion, Paris.
- ROY B. (1992).** Science de la décision ou science de l'aide à la décision ?. *Revue Internationale de Systémique*, 6, pp. 497-529.
- ROY B., BERTIER P. (1973).** La méthode ELECTRE II : une application au média-planning. In *Operational Research '72*, M. Ross (Ed.), North Hollan Publishing Company, pp. 291-302.
- ROY B., SKALKA J.M. (1984).** *ELECTRE IS : aspects méthodologiques et guide d'utilisation*. Document du LAMSADE, Université de Paris-Dauphine.
- ROY B., BOUYSSOU D. (1993).** *Aide multicritère à la décision : méthodes et cas*. Editions Economica, Paris.
- RUSHINEK A., RUSHINEK S. (1986).** What makes users happy ? *Communication of ACM*, 29.

- RUSSO P., BOOR S. (1993).** How fluent is your interface ? Design for international users. *Proceedings of the INTERCHI'93 Conference on Human factors in computing systems*, ACM Press, Addison-Wesley, Morristown, pp. 342-347.
- SCAPIN D.L. (1987).** *Guide ergonomique de conception des Interfaces Homme-Machine*. Rapport technique n° 77, INRIA, 92 p.
- SCAPIN D.L. (1988).** *Vers des outils formels de description des tâches orientées conception d'interfaces*. Rapport de recherche n° 893, INRIA, 28 p.
- SCAPIN D.L., GOLBREICH C. (1989).** MAD : une méthode analytique de description des tâches. *Colloque sur l'ingénierie des Interfaces Homme-Machine*, INRIA, Sophia-Antipolis, Le Chesnay, 24-26 Mai, pp. 131-148.
- SCAPIN D.L. (1990a).** Des critères ergonomiques pour l'évaluation et la conception d'interfaces utilisateurs. *Actes du XXVIème Congrès de la SELF*, Montréal, Canada, 3-5 Octobre.
- SCHÄRLIG A. (1985).** *Décider sur plusieurs critères. Panorama de l'aide à la décision multicritère*. Presses Polytechniques et Universitaires Romandes, Lausanne, Suisse.
- SCHIELE F., GREEN T. (1990).** HCI formalisms and cognitive psychology : the case of Task-Action Grammar. In *Formal methods in computer interaction*, Harrison and Thimbleby (Eds.), Cambridge University Press, Cambridge, UK, pp. 9-62.
- SEARS A.L. (1993).** Layout appropriateness: a metric for evaluating user interface widget layout. *IEEE Transactions on Software Engineering*, 19 (7), pp. 707-719.
- SEARS A.L. (1994).** Automated metrics for user interface design and evaluation. *International Journal of Bio-Medical Computing*, 34, pp. 149-157.
- SEBILLOTTE S. (1994).** *Méthodologie pratique d'analyse de la tâche en vue de l'extraction de caractéristiques pertinentes pour la conception d'interface*. Rapport technique n° 163, INRIA, Mars.
- SENACH B., ALENGRY P. (1985).** Evaluation d'un dispositif d'assistance à la conduite et au dépannage de chaînes de fabrication automatisées. Rapport technique INRIA, février.
- SENACH B. (1990).** *Evaluation ergonomique des Interfaces Homme-Machine : une revue de la littérature*. Rapport n° 1180, INRIA, 75 p.
- SFEZ L. (1992).** *Critique de la décision*. Presses de la Fondation Nationale des Sciences Politiques, 4^{ème} édition, Paris.
- SHARATT B. (1987).** The incorporation of early interface evaluation into command language grammar specification. In *People and computers III*, D. Diaper and R. Winder (Eds.), Cambridge University Press : Cambridge, pp. 11-28.
- SHEPHERD A. (1993).** An approach to information requirements specifications for process control tasks. *Ergonomics*, Vol. 36, n° 11, pp. 1425-1437.

-
- SIMON P. (1993).** *Contribution de l'analyse des mouvements oculaires à l'évaluation de la charge de travail mental.* Thèse de Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis.
- SIMON P., BERGER J., BARNARD J.M, ANGUE J.C. (1991).** Towards an instantaneous workload assessment by means of the scanning movements. Tenth European Annual Conference on Human Decision Making and Manual control, Liège Belgique.
- SIOCHI A.C., HIX D. (1991).** A study of computer-supported user interface evaluation using maximal repeating pattern analysis. *Proceedings of the CHI'91 Conference on Human-Computer Interaction*, ACM Press, New Orleans, May, pp. 301-305.
- SISKOS J., HUBERT P. (1983).** Multicriteria analysis of the impacts of energy alternatives : a survey and a new comparative approach. *European Journal of Operational Research*, 13, pp. 278-299.
- SMITH S.L., MOSIER J.N. (1986a).** *A design evaluation checklist for user-system interface software.* Report MTR-9480 EDS-TR-84-358, The MITRE Corporation, Bedford, MA.
- SMITH S.L., MOSIER J.N. (1986b).** *Guidelines for designing user interface software.* Technical report MTR-10090, The MITRE Corporation, Bedford, MA 01730, Massachusetts, USA.
- SPERANDIO J.C. (1991).** Les méthodes d'analyse du travail en psychologie ergonomique. In *La recherche en psychologie (domaine et méthodes)*, J.P. Rossi (Ed.), Dunod, Paris.
- STAMMER R.B., MICHAEL S.C., ASTEY J.A. (1990).** Task analysis. In *Evaluation of human works : a practical ergonomics methodology*, J.R. Wilson and E.N. Corlett (Eds.), Taylor and Francis, pp. 134-160.
- STANTON N.A., BABER C. (1995).** Alarm-initiated activities : an analysis of alarm handling by operators using text-based alarm systems in supervisory control systems. *Ergonomics*, 38(11), pp. 2414-2431.
- STEPHANIDIS C., GRAMMENOS D., AKOUMIANAKIS C. (2000).** Integrated support for working with guidelines: the Sherlock guideline management system. *Interacting With Computers*, vol. 13, n°1.
- STREVELER D.J., WASSERMAN A.I. (1984).** Quantitative measures of the spatial properties of screen design. In *Human-Computer Interaction: INTERACT'84*, B. Shackel (Ed.), Amsterdam, North-Holland, pp. 81-89.
- SUPPES P., KRANTZ D.H., LUCE R.D., TVERSKY A. (1989).** *Foundations of measurement (Vol. 2)*. Academic Press, New York.
- SWEENEY M., MAGUIRE M., SHACKEL B. (1993).** Evaluation user-computer interaction : a framework. *International Journal of Man-Machine Studies*, Vol. 38, pp. 689-711.
- TABARY D., ABED M. (1998).** TOOD : an object oriented methodology for describing user task interface design and specification, an application for traffic control . *La lettre de l'intelligence artificielle*, vol. 134-135, pp.107-114.

-
- TABORIN V. (1989).** *Coopération entre opérateur et système d'aide à la décision pour la conduite de procédés continus : application à l'interface opérateur – système expert du projet ALLIANCE.* TH7SE DE Doctorat en Automatique industrielle et humaine, Université de Valenciennes et du Hainaut-Cambrésis, Mars.
- TARBY J.C. (1993).** *Gestion automatique du dialogue homme-machine à partir de spécifications conceptuelles.* Thèse de Doctorat en Informatique, Université de Toulouse I, Septembre.
- TAUBER M. (1990).** ETAG : Extended Task Action Grammar – A language for the description of the user's task language. In *Human-Computer Interaction, INTERACT'90*, D. Diaper *et al.* (Eds.), North Holland, pp. 163-168.
- THOMAS B. (1996).** 'Quick and dirty' usability tests. In *Usability evaluation in industry*, P.W. Jordan *et al.* (Eds.), Taylor & Francis, London, pp. 107-114.
- TREILLET M. (1987).** *Le livre de Turbo-Prolog.* Editions PSI, Paris.
- TULLIS T.S. (1988).** Screen design. In *Handbook of Human-Computer Interaction*, M. Helander (Ed.), Elsevier Science Publishers B.V., North Holland, pp. 377-411.
- VANDERDONCKT J. (1994).** *Guide ergonomique des Interfaces Homme-Machine.* Presses Universitaires de Namur, Namur, Belgique.
- VANSNICK J.C. (1986).** De Borda et Condorcet à l'agrégation multicritère. *Cahier du LAMSADE*, n°70, Université de Paris-Dauphine, 25p.
- VICENTE J. (1995).** Supporting operator problem solving through ecological interface design. *IEEE Transactions on Systems, Man and Cybernetics*, 25(4), pp. 529-545.
- VILLEMEUR A. (1988).** *Sûreté de fonctionnement des systèmes industriels : fiabilité, facteur humain, informatisation.* Eyrolles, Paris.
- VINCKE P. (1989).** *L'aide multicritère à la décision.* Collection Statistiques et mathématiques appliquées, Editions de l'Université de Bruxelles, Bruxelles.
- VOLLE (1981).** *Analyse des données.* Deuxième édition, Editions Economica, Paris.
- VRIES G. (DE), HARTEVELT M., OOSTERHOLT R. (1996).** Private camera conversation : a new method for eliciting user responses. In *Usability evaluation in industry*, P.W. Jordan *et al.* (Eds.), Taylor & Francis, London, pp. 147-155.
- WALLISER (1977).** *Introduction critique à l'analyse de systèmes.* Seuil, Paris.
- WAWAK F. (1996).** *Elaboration d'un système d'aide à la décision pour la supervision en téléopération : approche basée sur les théories des sous-ensembles flous et des possibilités.* Thèse de Doctorat en Automatique et informatique industrielle, Université des Sciences et Technologies de Lille, Mai.
- WHITEFIELD J., WILSON F., DOWELL J. (1991).** A framework for human factors evaluation. *Behaviour and Information Technology*, 10(1), January, pp. 65-79.

-
- WIERWILLE W.W. (1979).** Physiological measures of aircrew mental workload. *Human factors*, 21(5), pp. 575-593.
- WIERWILLE W.W., CASALI J.G. (1983).** Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Human Factors*, 25(1), pp. 1-16.
- WILSON J.R., CORLETT E.N. (1990).** *Evaluation of human works: a practical ergonomics methodology*. Taylor & Francis, London.
- WILSON G.F., O'DONNELL R.D. (1988).** Measurement of operator workload test battery. In *Human mental workload*, Hancock and Meshkati (Eds.), Advances in psychology, North Holland.
- WISNER A. (1984).** Ergonomics or anthropotechnology : a limited or wide approach to working conditions in technology transfert. In *Ergonomics of developing countries*, H. Shahnavaz (Ed.), C.E.D.C., Lulea University.
- WISNER A. (1985).** Ergonomics in developing countries. *Ergonomics*, 28(8), pp. 1213-1224.
- YAARI (1987).** The dual theory of choice under risk. *Econometrica*, 55, pp. 95-115.
- YAGER R.R. (1986).** A characterisation of the extension principle. *Fuzzy set and systems*, n° 18, pp. 205-217.
- YAGER R.R. (1991).** Connectives and quantifiers in fuzzy sets. *Fuzzy set and systems*, n° 40, pp. 39-75.
- YOUNG R.M., WHITTINGTON J. (1990).** A knowledge analysis of interactivity. *Proceedings of INTERACT'90*, D. Diaper, G. Cockton, B. Shackel (Eds.), Elsevier Scientific Publishers B.V., Bambridge, UK, August 27-31, pp. 207-121.
- ZADEH L.A. (1975).** The concept of a linguistic variable and its application to approximative reasoning. Part I, *Information Sciences*, n° 8, pp. 301-357 ; Part II, *Information Sciences*, n° 9, pp. 43-80.
- ZADEH L.A. (1978).** PRUF – A meaning representation language for natural language. *International Journal of Man-Machine Studies*, 10, n° 8, pp. 395-460.
- ZELENY M. (1982).** *Multiple criteria decision making*. McGraw-Hill.
- ZOPOUNIDIS C., DIMITRAS A.I. (1998).** *Multicriteria decision aid methods for the prediction of business failure*. Kluwer Academic Publishers, Netherlands.



ANNEXES

ANNEXE I : Ensembles et raisonnement flous

Nous allons aborder la question du raisonnement en présence d'imprécis. Nous allons pour cela nous intéresser à la théorie des ensembles flous, aux modes de représentation de l'imprécision à l'aide de fonctions d'appartenance à un ensemble, et à la façon d'associer à cette représentation un facteur de confiance traduisant l'incertain.

Nous ferons le lien avec la théorie des possibilités de façon à montrer qu'il est possible de traiter imprécis et incertain dans un même cadre. Des applications de ce dernier point de vue seront présentées dans le paragraphe suivant.

Propositions précises ou imprécises

a) Représentation du contenu d'une proposition floue

Par définition, une proposition floue est une proposition de la forme $X \text{ est } A$, où X est une variable prenant ses valeurs dans un domaine T et A un ensemble flou ; il en est ainsi de *Jean est de grande taille* ou plutôt *La taille de Jean est grande*. Cette proposition spécifie de façon imprécise la valeur de la taille de Jean. Notons que nous utilisons l'expression classique proposition floue pour parler plutôt d'une proposition dont on affirme la vérité mais contenant du flou, cette remarque valant pour la suite. Il en sera de même pour les propositions imprécises ou précises.

On appelle proposition élémentaire une proposition floue de la forme $X \text{ prend la valeur } t$ où $t \in T$. Une telle proposition est dite précise par rapport au référentiel T . C'est le cas de la proposition *La taille de Jean est 1m82*. Toute proposition non élémentaire (et différente de F) est imprécise. Il convient de faire la différence entre imprécis non flou (*Jean mesure entre 1m80 et 1m85*) et imprécis flou (*Jean mesure environ 1m82*).

La proposition toujours fautive F représente $X \text{ prend sa valeur sur } \emptyset$ et la proposition toujours vraie V correspond à $X \text{ prend sa valeur sur } T$.

Le contenu de la proposition imprécise $X \text{ est } A$ peut être représenté par l'ensemble flou des valeurs plus ou moins possibles que peut prendre la variable X . Ainsi, le contenu de la proposition *Jean est de grande taille* est représenté par l'ensemble flou des valeurs plus ou moins possibles que peut prendre la variable *taille de Jean*, sachant qu'elle est grande. Quel que soit $t \in T$, la possibilité pour que X prenne la valeur t vaut $\mu_A(t)$, ce qui s'écrit :

$$\pi_X(t) = \mu_A(t).$$

π_X est la distribution de possibilité attachée à X . Cela signifie que la possibilité que X soit égal à t est mesurée par le degré d'appartenance de t à A . Notons que cette égalité ne traduit pas une identité, les deux fonctions ne possédant pas les mêmes propriétés.

Notre exemple se lit : la possibilité que la taille de Jean vale t sachant que Jean est grand est égale au degré de vérité de la proposition *Jean est grand* sachant qu'il mesure t .

On appelle support de X le sous-ensemble non flou des valeurs de t telles que $\pi_X(t) > 0$ et on appelle noyau de X le sous-ensemble non flou des valeurs de t telles que $\pi_X(t) = 1$. Le support inclut toujours le noyau. Ces ensembles peuvent ne pas être finis.

Le rapport de proportionnalité de 1 entre π_X et μ_A conduit aux remarques suivantes :

- si X prend effectivement une valeur existant dans le domaine T , il existe au moins une valeur t telle que $\pi_X(t) = 1$. La distribution de possibilité est dite normalisée ;
- lorsque π_X ne vaut 1 que pour une valeur unique de t et 0 partout ailleurs, le fait est précis ;
- s'il existe des valeurs t telles que $\pi_X(t)$ est compris strictement entre 0 et 1, le fait est flou ;
- lorsque π_X vaut 1 pour plusieurs valeurs de t mais 0 partout ailleurs, le fait est imprécis mais non flou ;
- une valeur inconnue est représentée par une distribution de possibilité égale à 1 en tout point du domaine.

b) Distributions de possibilité

□ Description de l'imprécis

Comme nous l'avons vu plus haut, des distributions de possibilité peuvent être utilisées à plusieurs fins :

- pour représenter des données précises ou imprécises :
 - *La taille de l'enfant vaut exactement 1,40m,*
 - *Charles mesure (exactement) entre 1,80 et 1,90m.* La taille est donnée sous forme d'un intervalle non flou,
 - *La taille de Jeanne est de l'ordre de 1,60m ;*
- pour définir des prédicats flous, par exemple *Grand* ;
- pour définir des ensembles flous de proportions traduisant des modifieurs (*pas, très*) ou des quantificateurs flous comme *assez* ou *généralement*, ce qui, par composition, permet de définir des ensembles flous comme *Assez_jeune*. Les fonctions d'appartenance associées sont définies sur $[0, 1]$ et prennent leur valeur sur $[0, 1]$.

Notons la différence de statut entre *très* et *assez* (on ne peut dire d'un jeune enfant qu'il est assez jeune et, au delà d'un certain âge, une personne est de moins en moins assez jeune...) et la différence de leurs effets sur μ_{jeune} .

Des *distributions de possibilité trapézoïdales* sont fréquemment utilisées pour approcher les fonctions vues plus haut. Une telle distribution est complètement définie par la donnée du noyau et du support : un trapèze se définit par un quadruplet (α, a, b, β) où α et β désignent la largeur des intervalles du support autour du segment noyau $[a, b]$.

□ Association de l'incertain à l'imprécis

Parmi les techniques proposées pour traduire l'incertitude dans les distributions trapézoïdales, on peut retenir l'élargissement du trapèze [BOY, 1988]. Cette technique consiste à élargir la distribution de possibilité associée à une proposition lorsque l'incertitude croît. Lorsqu'une proposition est totalement incertaine (ignorance totale), la distribution de possibilité vaut 1 en tout point du domaine. Enfin, lorsque la certitude de la proposition contraire croît, il convient de considérer la complémentarité à 1 de la distribution de possibilité de la proposition en question. L'auteur a appliqué cette méthode dans un système de recherche de document en bibliothèque lorsque les informations sur la date de publication, l'auteur, etc., sont incertaines.

Adéquation entre deux propositions

Nous considérons dans la suite que la proposition $X \text{ est } \Delta$ est donnée et nous cherchons comment la mettre en relation avec la proposition $X \text{ est } \Phi$ qui joue le rôle d'un filtre : il est question de *filtrage flou*. Il peut s'agir de la mise en relation d'une donnée avec la partie prémisses d'une règle de production, de façon à rechercher l'adéquation entre les deux. Δ et Φ sont deux ensembles flous définis sur le même domaine S . La proposition $X \text{ est } \Delta$ peut se caractériser par deux quantités $\Pi(\Phi/\Delta)$ et $N(\Phi/\Delta)$ traduisant respectivement la possibilité et la nécessité que X soit Φ étant donné que $X \text{ est } \Delta$:

- $\Pi(\Phi/\Delta) = \text{degré d'intersection de } \Delta \text{ et } \Phi$
 $= \sup_{s \in S} \min(\mu\Phi(s), \mu\Delta(s)).$

Si les noyaux de Φ et de Δ ont une intersection non vide, cette quantité vaut 1 ; il est alors totalement possible que Δ s'apparie avec Φ . Si l'intersection des supports de Δ et de Φ est vide, la quantité est nulle ; il est impossible que Δ s'apparie avec Φ .

- $N(\Phi/\Delta) = \text{degré d'inclusion de } \Delta \text{ dans } \Phi$
 $= \inf_{s \in S} \max(\mu\Phi(s), 1 - \mu\Delta(s)).$

Si le support de Δ est inclus dans le noyau de Φ , la nécessité vaut 1 ; il est totalement certain que Δ s'apparie avec Φ . Si une partie du noyau de Δ se trouve en dehors du support de Φ , elle vaut 0 ; il n'est pas certain que Δ s'apparie avec Φ . Il est possible de vérifier que l'on a toujours $\Pi(\Phi, \Delta) \geq N(\Phi, \Delta)$.

Lorsque les ensembles flous sont définis par des distributions trapézoïdales, les quantités $\Pi(\Phi, \Delta)$ et $N(\Phi, \Delta)$ se calculent par des opérations simples à partir des caractéristiques des trapèzes (α, a, b, β) et $(\alpha', a', b', \beta')$ représentant respectivement Δ et Φ :

- $\Pi(\Phi, \Delta) = \max(0, \min(1, A, B))$;
 - $N(\Phi, \Delta) = \max(0, \min(1, C, D))$;
- expressions dans lesquelles :
- $A = 1 + (b - a')/(\alpha' + \beta)$; $B = 1 + (b' - a)/(\alpha + \beta')$;
 - $C = (a - a' + \alpha')/(\alpha + \alpha')$; $D = (b' - b + \beta')/(\beta + \beta')$.

Des techniques de calcul permettent de traiter les cas d'indétermination (cas d'étalements infinis). En guise d'exemple, prenons le cas de tailles t exprimées en cm, les notions de *très proche de t*, *proche de t* et *environ t* étant supposées représentées par les distributions données sur la figure A.1.

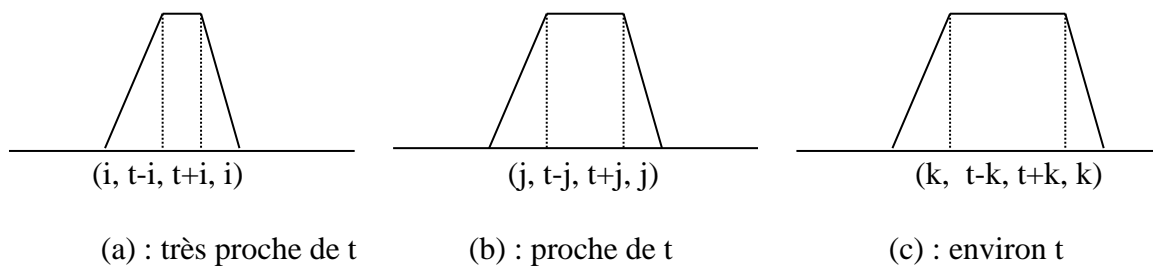


Figure A.1 : Trapèzes choisis pour les exemples proposés par [HATON *et al.*, 1990]

Pour chacune des situations possibles, 6 cas de combinaison des N et Π peuvent, s'il y a lieu, être distingués. Les commentaires correspondent aux cas suivants (μ et λ étant strictement compris entre 0 et 1 et μ étant inférieur ou égal à λ) :

$N = 1$	$\Pi = 1$	oui ;	$N = \mu$	$\Pi = 1$	plutôt oui ;
$N = \mu$	$\Pi = \lambda$	peut-être	$N = 0$	$\Pi = \lambda$	plutôt non ;
$N = 0$	$\Pi = 0$	non ;	$N = 0$	$\Pi = 1$	inconnu.

Règles d'inférences floues

Nous envisageons maintenant la propagation inférentielle de l'imprécision. Les règles peuvent se classer en règles catégoriques (sans quantificateurs flous) et règles dispositionnelles (un ou plusieurs éléments de prémisses contiennent, implicitement ou explicitement, le quantificateur *généralement*). Nous présentons un certain nombre d'entre elles, dans lesquelles X, Y, Z, \dots sont des variables prenant leurs valeurs sur des univers spécifiques du discours, et A, B, C, \dots sont des prédicats ou ensembles flous (pour plus de détails, voir [ZADEH, 1988]).

a) Principe de déduction (entailment)

X est A
$B \supseteq A$
X est B

où X est une variable prenant ses valeurs dans le domaine S , et A et B sont des ensembles flous de S .

Exemple : *Pierre est jeune* se traduit par $\text{Age}(\text{Pierre})$ appartient à *Jeune* où *Jeune* est un ensemble flou. Si *Jeune* est un sous-ensemble du complément de *Vieux*, il est possible d'inférer *Pierre n'est pas vieux* avec :

$$\mu_{\text{Jeune}}(s) < 1 - \mu_{\text{Vieux}}(s), s \in [0, 100].$$

b) *Règle de conjonction de conclusions*

Imaginons que les deux règles suivantes :

Si... alors la pression est haute

Si... alors la pression est très haute

ont leurs parties conditions vérifiées et sont appliquées. La conjonction des conclusions permet d'inférer la conclusion plus spécifique :

La pression est (haute et très haute),
représentée par (0.5, 76.5, 77, 0.5).

c) *Produit cartésien*

X est A

Y est B

(X, Y) est $A \times B$

où (X, Y) est une variable binaire et $A \times B$ est défini par :

$$\begin{aligned} \mu_{A \times B}(s, t) &= \mu_A(s) \wedge \mu_B(t), s \in S, t \in T ; \\ &= \min_{s, t}(\mu_A(s), \mu_B(t)). \end{aligned}$$

d) *Modus ponens généralisé*

X est A

Si X est B alors Y est C

Y est G

On est dans la situation d'une règle d'inférence de filtre *X est B* et en présence d'un fait connu *X est A* . Le *modus ponens* généralisé permet la dérivation d'une nouvelle connaissance sur Y : Y prend ses valeurs sur l'ensemble flou $G = A \circ (B \oplus C)$. \circ est une loi de composition et \oplus la somme floue. G est caractérisé par la fonction d'appartenance :

$$\begin{aligned} \mu_G(t) &= \sup_s (\mu_A(s) \wedge \mu_{B \oplus C}(s, t)) \\ \text{où } \mu_{B \oplus C}(s, t) &= 1 \wedge (1 - \mu_B(s) + \mu_C(t)), s \in S, t \in T. \end{aligned}$$

Nous pouvons à ce propos faire allusion à la forme de raisonnement par analogie qui conduit, dans une situation S , à prendre une décision par rapprochement avec une situation S' voisine. On suppose alors qu'une petite variation des variables de situation influe sous forme d'une petite variation sur la conclusion. La règle du *modus ponens* généralisé permet d'écrire :

$$\begin{array}{l} X \text{ est } B' \\ \text{Si } X \text{ est } B \text{ alors } Y \text{ est } C \end{array}$$

$Y \text{ est } C'$

Dans cette règle, les ensembles flous B et C se rapportent à la situation S . B' et C' sont des ensembles flous « proches » de B et C respectivement et caractérisant la situation S' . Il y a correspondance entre B et B' d'une part (définis sur le même domaine) et C et C' d'autre part. Dans ce type de raisonnement, la dépendance est représentée par l'inférence *Si... alors*.

e) Règles d'inférence graduelle

Enfin, le raisonnement expert qualitatif fait souvent apparaître des règles d'inférence graduelle (ou *topoi*) comme :

- *Plus X est A plus Y est B ;*
- *Plus les voitures roulent vite plus la notion de confort d'une chaussée est importante ;*
- *Plus une chaussée est dégradée moins elle est confortable.*

Leur application peut aussi recourir à des fonctions d'appartenance.

ANNEXE II : Evaluation du système ADHESION

Le système d'aide ADHESION a été évalué par un expert du LAMIH, dans le but de tester les fonctionnalités du système et le respect des critères ergonomiques (voir Chap. II, partie II, Expertise Humaine). Notons qu'il aurait été possible d'utiliser le système ADHESION pour savoir quelle méthode d'évaluation utiliser. Cette évaluation a permis de réaliser une seconde version du système ADHESION. Les pages-écrans A.2a et A.2b sont respectivement les nouvelle et ancienne versions du système obtenues après et avant l'évaluation du système ADHESION. Ces pages-écrans concernent les pages d'explications des termes utilisés par ADHESION et peuvent être difficilement compréhensibles par les utilisateurs d'ADHESION.

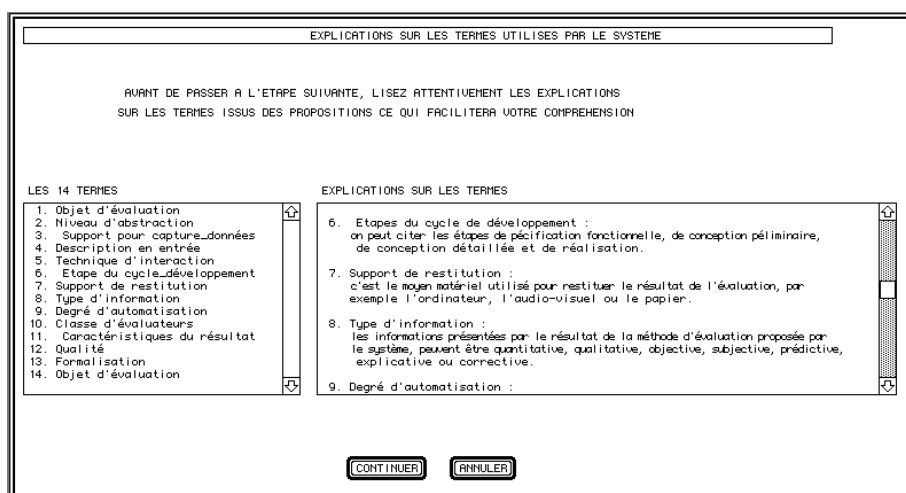


Figure A.2a : Page-écran d'explications des termes après évaluation

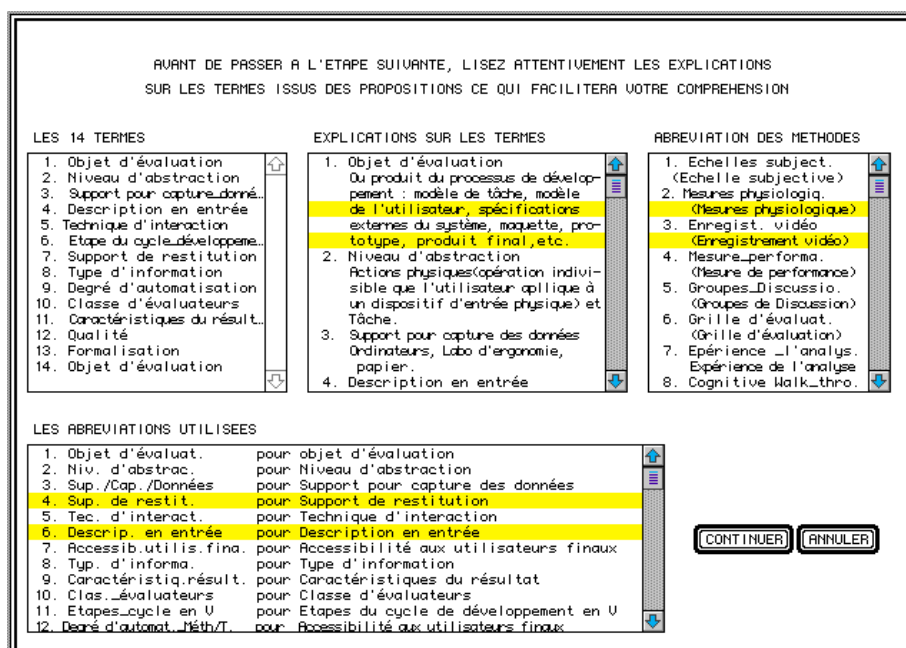


Figure A.2b : Page-écran d'explications des termes avant évaluation

En tenant compte du critère d'évaluation « facilité de compréhension », l'évaluateur, qui est ici un expert, estime que ce critère n'est pas respecté. Il propose de simplifier les choses en supprimant toutes les abréviations dans le développement. Par contre, il a estimé que les explications des termes sont nécessaires avant de passer à la saisie des contraintes proprement dite. De même, les figures A.3a et A.3b montrent respectivement les interfaces de saisie de contraintes après et avant l'évaluation par l'expert.

DESCRIPTION DES CONTRAINTES

CETTE TECHNIQUE DE RESOLUTION VOUS OBLIGE A REPONDRE
A TOUTES LES QUESTIONS CI-DESSOUS

Sélectionnez le pays de destination (ou assimilé) de l'IHM à évaluer

Afghanistan

- Afrique du Sud
- Albanie
- Algérie
- Allemagne
- Andorre
- Angola
- Antilles Néerlandaises
- Arabie Saoudite

AIDE A LA COMPREHENSION DES QUESTIONS

Avez-vous accès aux utilisateurs finaux ? Répondre Plutôt-Moyen, signifie que vous avez près de 50% de chance d'accéder aux futurs utilisateurs de l'IHM que vous voulez évaluer, alors que Très - Oui signifie 100% de chance.

Le groupe d'évaluation dispose t-il d'un ergonomes ?
Répondre Très - Non signifie que vous ne disposez pas du tout d'ergonomes, alors que Plutôt-Moyen signifie qu'il existe une chance sur deux pour que vous obtenez un ergonomes, cela peut être le cas d'une étudiante

	***** Descripteurs *****		***** Modificateurs *****		
	Bon/Oui	Moyen	Mauvais/Non	Très	Plutôt Modérément Très
Avez-vous accès aux utilisateurs finaux?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un ergonomes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un sociologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un psychologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en droit-informatique?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le groupe d'évaluation dispose t-il d'un spécialiste en assurance qualité?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le budget-évaluation vous semble t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le temps alloué à l'évaluation vous semble t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous financer le dépassement du temps prévu?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous financer les spécialistes d'évaluation manquant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le système à évaluer évoluera t-il dans un environ.nuisible?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Disposez-vous du matériel spécialisé (système oculométrique, enreg.vidéo,)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avez-vous une expérience en évaluation ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous assimiler le système à évaluer à un type grand public?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
L'IHM à évaluer a t-elle franchit l'étape de réalisation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure A.2b : Page-écran d'explications des termes avant évaluation

*****Cette technique de résolution vous oblige à répondre à toutes les questions ci-dessous*****

Inscrivez ici le nom du Pays destinataire(ou assimilé) de l'IHM à évaluer, pour la prise en compte des aspects socio-culturels; le système reconnaît 212 pays et îles.

AIDE à la compréhension des questions

Indonésie

Indonésie

Avez-vous accès aux utilisateurs finaux?
Répondre Plutôt-Moyen, signifie que vous avez près de 50% de chance d'accéder aux futurs utilisateurs de l'IHM que vous voulez évaluer.

	Descripteurs		Modificateurs		
	Bon/Oui	Moyen	Mauvais/Non	Très	Plutôt Modérément
Avez-vous accès aux utilisateurs finaux?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Groupe d'évalu. dispose t-il d'un ergonomes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Groupe d'évalu. dispose t-il d'un sociologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Groupe d'évalu. dispo. t-il d'un psychologue?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gr.d'év.dis.t-il d'un spéc.droit-informatique?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gr.d'év.disp.t-il d'un spéc.assurance qualité?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le budget-évaluati. vous semble t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le temps alloué à l'éval.v.sem.t-il suffisant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pouvez-vous financer le dépasse.du temps prévu?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P.vous financer les spécialistes d'év.manquant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
L'IHM évoluera t-elle dans un environ.nuisible?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accès mat.-spécial.(sys.oculomé.,enreg.vidéo,)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avez-vous une expérience en évaluation ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P.-vous assimiler l'IHM à un type grand public?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
L'IHM a t-elle franchis l'étape de réalisation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure A.3b : Page-écran de saisie de contraintes avant évaluation

Annexes

L'expert a ici mis l'accent sur les critères « facilité d'utilisation » et « lisibilité », ce qui nous a permis de supprimer les abréviations. De plus, il a estimé que les pays devaient être sélectionnés et non édités pour plusieurs raisons : méconnaissance possible des pays par l'utilisateur en termes d'orthographe, ce qui pourrait entraîner des erreurs difficilement détectables, ou en termes géographiques. Les figures A.4a et A.4b montrent un exemple de proposition obtenue pour le scénario choisi dans le cadre d'une évaluation.

PROPOSITION DE METHODES ET CRITERES D'EVALUATION - PROPOSITION N°2

DESCRIPTION DE LA METHODE ET DES CRITERES

CHARACTERISTIQUES DE L'INTI A EVALUER

Libellé de la méthode : Questionnaires
Méthode ou Technique complémentaire : Interviews
Description en entrée (ex: description de tâche) : Pas nécessaire
Disponibilité de la méthode : Oui
Formalisation : Méthode non formelle
Technique d'interaction : Toutes
Objet d'évaluation : Produit final ou Prototype
Classe d'évaluateurs exigée : Experte/Novice
Support pour la Capture de données : Ordinateur/papier
Etapes du cycle de développement concernées : Spécification/conception/après réalisation
Usage de métriques : Pas nécessaire

PRINCIPE DE FONCTIONNEMENT DE LA METHODE

Le principe consiste à apprécier de manière subjective les besoins des utilisateurs auprès d'un groupe sélectionné. La technique des questionnaires existe également sous-forme, d'utilisation sur la qualité de l'interface. Cette Technique est finalisée par une analyse des données recueillies par des méthodes statistiques

POSITIONNEMENT PAR RAPPORT AUX EVALUATEURS

soins important important Très import.

Requêtes des évaluateurs :
Conflits entre évaluateurs :
Exigence mentale :
Exigence physique :

NIVEAUX D'AUTOMATISATION DE LA METHODE

Capture automatique des données :
Restitution automatique des données après traitement :
Analyse automatique :
Critique automatique :

CRITERES D'EVALUATION CORRESPONDANTS A LA PROPOSITION

N°	*Libellé	*Acceptabilité	*Dimension d'évaluat.
1.	Actions minimales	Pratique	utilisabilité
2.	Actions explicites	Pratique	utilisabilité
3.	Capacité fonctionnelle	Pratique	utilité
4.	Charge mentale	Pratique	utilisabilité
5.	Compatibilité	Pratique	utilisabilité
6.	Conclusion	Pratique	utilisabilité
7.	Contrôle explicite	Pratique	utilisabilité

REFERENCE SUR LA METHODE : Rushinek et al., 1996

CONTINUER ANNULER

CHARACTERISTIQUES DU RESULTAT D'EVALUATION FAITE A L'AIDE DE CETTE PROPOSITION

Qualité : Inspectabilité
Support de restitution : papier
Support d'analyse des résultats : papier/Ordinateur
Niveau d'abstraction : Tous
Type d'information : subjective/Éplicative
Performance de la méthode ou Technique : Moy. Bonne Très Bonne

AVANTAGES INCONVENIENTS

Coûts : Faibles
Temps moyen : 2 jours
Accessibilité aux utilisateurs finaux : Obligatoire
Facilité d'utilisation : Oui
Version automatisée : Non
Nombre moyen de sujets : Dizaine
Nombre moyen d'évaluateurs : 2

Figure A.4a : Exemple de proposition obtenue après évaluation

Libellé : Questionnaires
Méthod./Tec. associ. : Interviews
Descrip. en entrée : Pas nécessaire
Disponibilité : Oui
Formalisation : Non formelle
Tec. d'interact. : Toutes
Objet d'évaluat. : Pr. final/Prot.
Clas.-évaluateurs : Experte/Novice
Sup./Capt./Données : Ordinateur/papier
Etapes-cycle en U : Spé./concep./réal.
Usag.-métriques : Pas nécessaire

Le principe de fonctio. de la Méthod./Tech.

Le principe consiste à apprécier de manière subjective les besoins des utilisateurs auprès d'un groupe sélectionné. La technique des questionnaires existe également sous-forme, d'utilisation sur la qualité de l'interface. Cette Technique est finalisée par une analyse des données recueillies

CONTINUER ANNULER

Degré d'automat. de Méthod./T. : Captu. aut. Restitu. au. Analy. aut. Critique au.

Caractéristiq. du résultat

Qualité : Inspectabi.
Niv.d'abstrac.: Tous
Sup. de restit.: papier

Type d'informat.: subj./Expl.
Performance_Méthode/Tec.: Moy. Bonne T.Bo.
Support d'analyse resul.: papier/Ordinateur

moins imp. import. Très important

Requ.-évaluateurs.:
Conflits-évaluat.:
Exigence mentale :
Exigence physique :

Référen. sur la Méthod./Tech.: Rushinek et al.86

Avantages Inconvénients

Critères d'évaluation conresp. à cette Résolu.

N°	*Libellé	*Acceptabilité	*Dimensl. d'éval.
1.	Actions minimales	Pratique	utilisabilité
2.	Actions explicites	Pratique	utilisabilité
3.	Capacité functio.	Pratique	utilité
4.	Charge mentale	Pratique	utilisabilité
5.	Confidentialité	Sociale	sociale
6.	Compatibilité	Pratique	utilisabilité
7.	Conclusion	Pratique	utilisabilité

Coûts : Faibles
Temps moyen : 2 jours
Accessib. utilis.fina.: Oui
Facilité d'utilisation: Oui
Version automatisée : Non
Nombre moyen de sujets : 15
Nomb.moy.d'évaluateurs: 2

Figure A.4b : Exemple de proposition obtenue avant évaluation

Ici, en plus des critères cités ci-dessus, l'expert a pris en considération des critères d'évaluation qui permettent de rendre plus compréhensibles les propositions du système.

ANNEXE III : Questionnaire socio-technique

QUESTIONNAIRE SOCIO-TECHNIQUE POUR L'EVALUATION DES INTERFACES H-M	NOTE SUR 5				
	1	2	3	4	5
DIMENSION SOCIALE					
Confidentialité					
- L'interface respecte-t-elle l'anonymat de l'opérateur ?.....					
- L'interface permet-elle d'obtenir des informations relatives à la vie privée de l'opérateur ?.....	4				
Sécurité-innocuité					
- Les dysfonctionnements de l'interface peuvent-ils être nuisibles à l'environnement et donc à la société ?.....					
- L'interface permet-elle l'intervention de l'homme en cas de dysfonctionnement ?.....					
Emploi/Travail					
- La mise en place de l'interface risque-t-elle de provoquer des licenciements ?.....					
- La mise en place de l'interface peut-elle entraîner une diminution du salaire des opérateurs ?.....					
- La mise en place de l'interface va-t-elle modifier l'organisation du travail ?.....					
- La mise en place de l'interface va-t-elle opérer une redéfinition des postes des travailleurs ?.....					
- L'utilisation de l'interface est-elle frustrante pour les opérateurs ?.....					
- L'interface laisse-t-elle une autonomie aux opérateurs dans l'exécution de leurs tâches ?.....					
Savoir-faire					
- Les opérateurs ont-ils déjà été en présence de ce type d'interface ?.....					
- Les opérateurs disposent-ils des « outils » intellectuels nécessaires à la mise en œuvre de l'interface ?.....					
- Les opérateurs ont-ils déjà travaillé avec ce type d'interface ?.....					
- L'interface remet-elle en cause le savoir-faire des opérateurs ?.....					
Compétence					
- Les compétences nécessaires à la compréhension de l'interface sont-elles présentes chez les opérateurs ?.....					
- L'interface est-elle adaptée au niveau de formation des opérateurs ?.....					
- Les opérateurs disposent-ils des compétences nécessaires à l'utilisation de l'interface ?.....					
DIMENSION D'ETHIQUE					
Moralité					
- L'interface met-elle en scène des actions immorales ?.....					
- L'interface est-elle compatible avec les principes éthiques de la population locale ?.....					
Violence					
- Les actions proposées sur l'interface peuvent-elles être considérées comme violentes par la population locale ?.....					
- L'interface met-elle en scène des objets susceptibles d'être utilisés dans le cadre d'actes violents ?.....					
- L'interface présente-t-elle des scènes conflictuelles ou laisse-t-elle supposer l'existence d'une certaine tension ?.....					
DIMENSION CULTURELLE					
Langue					
- La langue d'usage de l'interface est-elle la même que celle des utilisateurs ou de la population locale?.....					
- Les opérateurs connaissent-ils la langue d'usage de l'interface?.....					
- Les opérateurs comprennent-ils la langue d'usage de l'interface?.....					
Forme					
- Les formes présentes sur l'interface peuvent-elles heurter la sensibilité de la population locale ?.....					
- Les symboles représentés sur l'interface ont-ils des significations particulières ?.....					
Si oui, cette signification est-elle la même que celle que la population locale leur attribue?.....					
- La proximité de certains symboles peut-elle être choquante pour la population locale ?.....					
Couleur					
- Les couleurs présentes sur l'interface ont-elles une signification particulière ?.....					
Si oui, cette signification est-elle la même que celle que la population locale leur attribue?.....					
- L'assemblage des couleurs est-il de nature à heurter la population locale ?.....					
- Certaines des couleurs présentes sur l'interface sont-elles en désaccord avec son contexte ?.....					
Gestuelle					
- Les gestes représentés sur l'interface sont-ils des gestes quotidiens de la population locale ?.....					
- Certains des gestes représentés sur l'interface sont-ils interdits ?.....					
- Certains des gestes représentés sur l'interface ont-ils une signification particulière pour la population locale ?.....					
Costume					
- Les costumes proposés sur l'interface sont-ils admis par la population locale ?.....					
- Est-il interdit de porter certains des costumes présents sur l'interface ?.....					
- Les costumes proposés sur l'interface sont-ils ceux qui sont habituellement portés par la population locale ?.....					
Moeurs					
- Les actions représentées sur l'interface sont-elles effectuées couramment par la population locale ?.....					
- Les actions représentées sur l'interface sont-elles considérées comme choquantes par la population locale ?.....					
TOTAL : / 205					

ANNEXE IV : Éléments de modélisation des données et des connaissances pour l'aide à la décision en évaluation

Le concept de sous-ensemble flou est associé à une fonction d'appartenance. De façon classique, un sous-ensemble se définit comme contenant une partie des éléments de son ensemble de définition. Cela peut se modéliser en associant à chaque élément de l'ensemble de définition un 1 ou un 0 suivant que l'élément appartient ou non au sous-ensemble. L'idée du sous-ensemble flou consiste à étendre ce formalisme en associant aux éléments une appartenance graduelle représentée par un réel de l'intervalle unité qui sera appelé *dégré d'appartenance* au sous-ensemble flou.

Si \tilde{F} est un ensemble discontinu, les degrés d'appartenance d'un sous-ensemble flou peuvent être considérés individuellement (dans le cas discontinu), c'est-à-dire que l'ensemble ordinaire X est dénombrable, ou globalement au travers d'une fonction caractéristique appelée *fonction d'appartenance* (dans le cas continu), lorsque X est indénombrable. Soit X l'ensemble de définition du sous-ensemble flou \tilde{F} .

Considérons par exemple le sous-ensemble flou des <<contraintes que l'équipe d'évaluation doit satisfaire pour qu'un système d'aide à l'évaluation lui fournisse une méthode d'évaluation appropriée à l'IHM à évaluer ainsi qu'au contexte de l'évaluation>>.

Si ce sous-ensemble s'écrit :

$$\tilde{C}_{\text{Méthode}_i} = \tilde{F} = \{(\text{Moyens_financiers}, 0.8), (\text{Charge_de_travail}, 1), (\text{Clarté_visuelle}, 1), (\text{Psychologue}, 0), (\text{Expérience_en_évaluation}, 0.2), (\text{Matériel_d'enregistrement_vidéo}, 0), (\text{Financement_des_experts_extérieurs}, 0.9)\},$$

Cela signifie que l'équipe d'évaluation possède des moyens financiers (x_7) assez suffisants, qu'elle veut évaluer la Charge de Travail (x_5) comme attribut de l'utilisabilité ainsi que la Clarté visuelle (x_6), mais que l'équipe ne dispose pas de psychologue (x_1), qu'elle n'a pas d'expérience en évaluation (x_3), qu'elle ne dispose pas de matériel d'enregistrement vidéo (x_2) et que les moyens financiers nécessaires pour s'offrir les compétences d'experts extérieurs (x_4) sont suffisants (très bon) (ex : psychologue). La figure AIV.1 est une représentation de ces données.

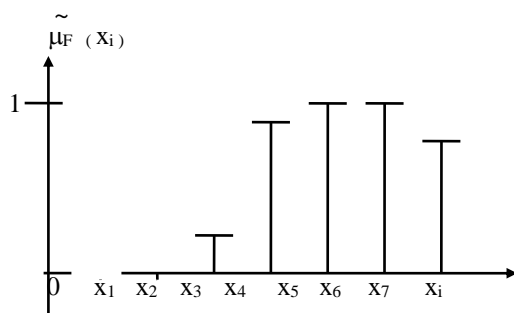


Figure AIV.1 : Un exemple d'ensemble flou discontinu

III.6.2 Ensembles flous continus

Si \tilde{F} est un **sous-ensemble flou continu**, la modélisation des connaissances change (figure AIV.2). Reprenons l'exemple de la figure III.3 dans le chapitre III pour modéliser à l'aide de la fonction trapézoïdale l'information suivante (sur le critère de décision *temps alloué* à l'évaluation) :

- *Le temps alloué* d'une évaluation est de l'ordre de 3 jours. Il existe un degré d'appartenance flou du *temps alloué* de l'évaluation sur l'échelle des temps possibles.

La figure ci-après représente la modélisation de cette contrainte. On voit bien qu'on est capable de transformer cette information en trapèze (μ_M).

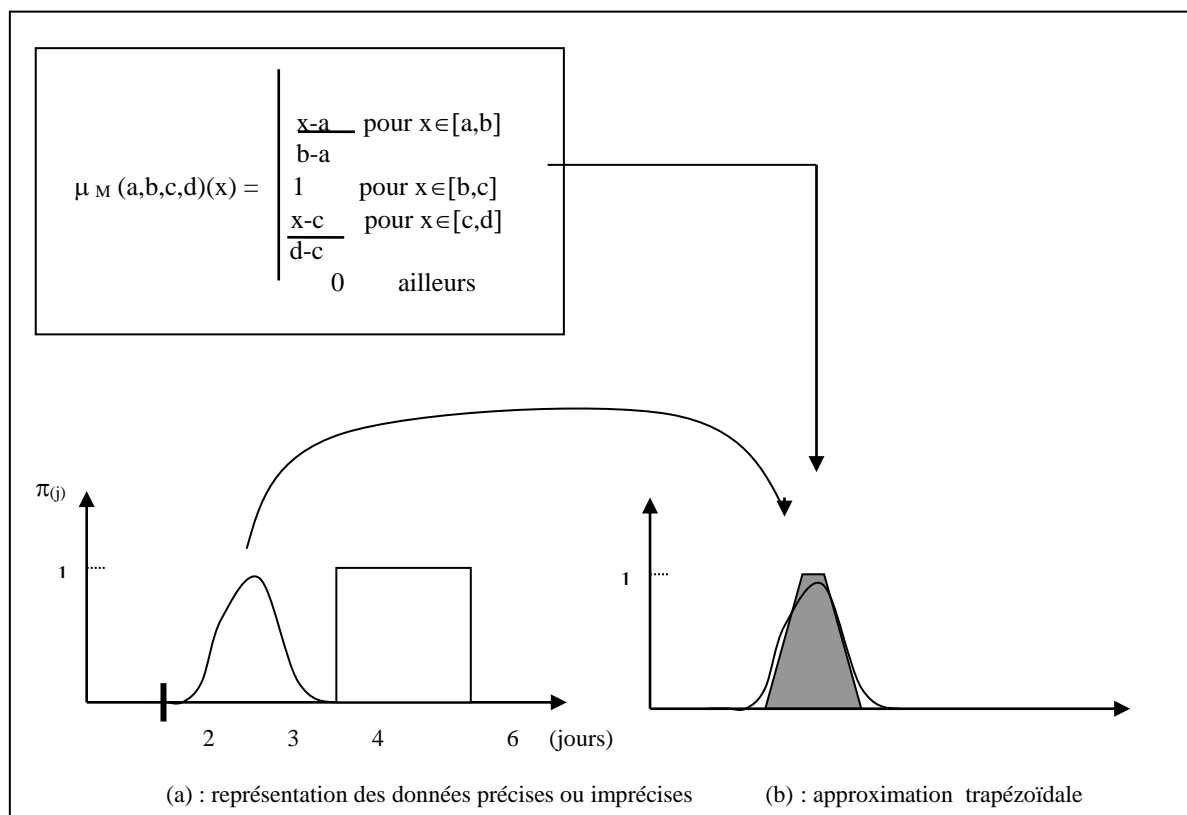


Figure AIV.2 : Distributions de possibilité (inspirée de [ROY, 1968])

Index des figures et tableaux

Index des figures

Figure 1	: Principales variables cibles pour l'évaluation.....	12
Figure I.1	: Principe de l'évaluation des IHMs en termes d'acceptabilité pratique et sociale complété, inspiré et précisé à partir des informations issues de [NIELSEN, 1993], [GRISLIN, 1995], [NENDJO ELLA et JACQUES, 1999].....	16
Figure I.2	: Positionnement de l'évaluation dans le processus de développement (inspiré de [GRISLIN, 1995]).....	17
Figure I.3	: Composant graphique de l'interface réalisée.....	26
Figure I.4	: Composant graphique modifié de l'interface.....	26
Figure I.5	: Composant graphique issu de Viewpoint Xerox.....	27
Figure I.6	: Exemple de boîte de dialogue disponible en plusieurs langues.....	28
Figure I.7	: Exemple de clarification d'usage de protections de disquette.....	29
Figure I.8	: Première position à valider des phases et critères d'évaluation associés, dans un cycle de développement en V.....	42
Figure II.1	: Interaction directe par interview entre utilisateur(s) et évaluateur(s).....	48
Figure II.2	: Salle d'évaluation à l'aide de la méthode Quick and Dirty chez PHILIPS.....	50
Figure II.3	: Illustration de la cabine vidéo.....	51
Figure II.4	: Analyse des traces écrites en tant que support de l'évaluation.....	52
Figure II.5	: La technique SUS.....	53
Figure II.6	: Technique de recueil de l'expertise.....	54
Figure II.7	: Exemple de note finale issue de l'évaluation.....	55
Figure II.8	: Principe du mouchard électronique.....	56
Figure II.9	: Extrait d'un questionnaire socio-technique.....	57
Figure II.10	: Analyse des protocoles en tant que méthode d'évaluation d'interface.....	58
Figure II.11	: Principe d'évaluation par analyse des mouvements oculaires : cas ASL-ASCENSION.....	59
Figure II.12	: Principe de la méthode SWAT.....	63
Figure II.13	: Principe de la méthode TLX.....	64
Figure II.14	: Echelle de COOPER-HARPER.....	65
Figure II.15	: Trois approches complémentaires pour les tests de conception.....	68
Figure II.16	: Démarche du prototypage.....	70
Figure II.17	: Principe du groupe de discussion.....	71
Figure II.18	: Extrait d'une grille d'évaluation.....	74
Figure II.19	: Schéma du fonctionnement cognitif de l'individu selon ICS.....	78
Figure II.20	: Description de la tâche par SADT et réseaux de Petri.....	83
Figure II.21	: Principe d'utilisation de KRI/AG.....	88
Figure II.22	: Fenêtre d'édition client.....	90
Figure II.23	: Décomposition de la boîte de dialogue.....	90
Figure II.24a	: La page initiale.....	91
Figure II.24b	: La page indiquant les erreurs d'accessibilité auxquelles on accède en cliquant sur « hat ».....	91
Figure II.25	: Synthèse et classification des méthodes et techniques d'évaluation des IHMs.....	93
Figure II.26	: Décomposition de la boîte de dialogue.....	96
Figure II.27	: Plan d'aide à l'évaluation des IHMs [DENLEY et LONG, 1997].....	98
Figure III.1	: Principe et structure de l'aide à la décision.....	104
Figure III.2	: Fonction d'appartenance d'un temps alloué à l'ensemble des temps importants.....	120
Figure III.3	: Distributions de possibilité.....	121
Figure III.4	: Fonction d'appartenance d'un temps alloué à l'ensemble des temps faibles.....	121
Figure III.5	: Mise en correspondance d'une proportion v avec générale.....	122

Figure III.6	: Composition de fonctions d'appartenance.....	122
Figure III.7	: Synthèse et classement des méthodes d'aide à la décision.....	124
Figure III.8	: Les différentes étapes de notre démarche.....	128
Figure III.9	: Fonction d'appartenance d'un temps alloué à l'ensemble des temps importants.....	129
Figure III.10	: Préférence de l'expert pour les critères <i>Moyens humains et financiers</i>	130
Figure IV.1	: Principe général d'utilisation du système d'aide à la décision ADHESION.....	139
Figure IV.2	: Choix stratégiques et schéma fonctionnel.....	143
Figure IV.3	: Extrait du programme de prise en compte ou de sélection des actions ayant un DCF ≤ 0.5	146
Figure IV.4	: Architecture du système d'aide à la décision.....	149
Figure IV.5	: Page-écran d'accueil du système d'aide à la décision.....	150
Figure IV.6	: Page-écran des explications des termes dans le système ADHESION.....	151
Figure IV.7	: Page-écran de saisie des contraintes.....	152
Figure IV.8	: Exemple de proposition de méthode.....	154
Figure IV.9	: Le modèle des sous-ensembles flous trapézoïdaux.....	156
Figure IV.10	: Evaluation du critère de décision <i>Budget d'évaluation</i>	158
Figure IV.11	: Extrait du programme.....	159
Figure IV.12	: Extrait de solution (obtenue ici dans l'environnement de l'interpréteur de PROLOG III).....	159
Figure IV.13	: Détermination des DCFs.....	160
Figure IV.14	: Vue d'une pente.....	161
Figure IV.15	: Evolution des pentes.....	161
Figure IV.16	: Exemple de vecteur T.....	161
Figure IV.17	: Extrait de la programmation.....	162
Figure V.1	: Poste de régulation ASTREE.....	167
Figure V.2	: Taux de satisfaction et d'insatisfaction.....	169
Figure V.3	: Taux de consultation de chaque écran du poste de régulation ASTREE.....	170
Figure V.4	: Saisie des contraintes pour le premier cas (ASTREE).....	176
Figure V.5a	: Une des proposition (analyse des protocoles) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation d'ASTREE.....	177
Figure V.5b	: Une des proposition (questionnaire) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation d'ASTREE.....	178
Figure V.5c	: Une des proposition (enregistrement vidéo) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation d'ASTREE.....	179
Figure V.6	: Gestion automatisée des bassins hydrauliques des centrales hydro-électriques de Tchimbele et Kinguele.....	185
Figure V.7	: Architecture matérielle du projet MOGADOR.....	187
Figure V.8	: Taux d'erreurs obtenu.....	188
Figure V.9	: Saisie des contraintes pour le deuxième cas (MOGADOR).....	190
Figure V.10a	: La première proposition (traces écrites de travail) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation de MOGADOR.....	191
Figure V.10b	: La deuxième proposition (questionnaire socio-technique) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation de MOGADOR.....	192
Figure V.10c	: La première proposition (mesure de la performance) des méthodes d'évaluation fournies par le système d'aide ADHESION pour l'évaluation de MOGADOR.....	193

Index des tableaux

Tableau I.1	: Critères ergonomiques selon Bastien et Scapin.....	23
Tableau I.2	: Exemples de caractéristiques culturelles de couleurs.....	30
Tableau I.3	: Analyse d'éthique pour le développement des IHMs.....	31
Tableau I.4	: Groupes de pays.....	38
Tableau I.5	: Importance des aspects socio-culturels selon le groupe d'appartenance.....	38
Tableau I.6	: Echantillon de l'étude.....	38
Tableau I.7	: Classement par ordre d'importance socio-culturelle.....	39
Tableau I.8	: Contraintes influençant le choix des méthodes et techniques d'évaluation.....	44
Tableau II.1	: Aide au choix des techniques d'évaluation dans FACE.....	95
Tableau III.1	: Synthèse des approches d'aide à la décision.....	125
Tableau III.2	: Synthèse réduite correspondant aux caractéristiques de notre problématique.....	127
Tableau IV.1	: Critères de décision et actions.....	141
Tableau V.1	: Critères de décision.....	171
Tableau V.2	: Tableau de synthèse de la phase de préparation des données.....	175
Tableau V.3	: Tableau des critères de comparaison.....	181
Tableau V.4	: Tableau de synthèse de la comparaison.....	182
Tableau V.5	: Tableau de synthèse de la phase de préparation des données.....	189
Tableau V.6	: Tableau de comparaison entre les propositions d'ADHESION et la technique d'évaluation réellement utilisée.....	185

Résumé

L'évaluation des systèmes interactifs est source de nombreuses difficultés, aussi bien pour les spécialistes de la communication Homme-Machine que pour les non-spécialistes. La qualité de l'évaluation dépend de paramètres variés et souvent contradictoires. En effet, en plus des problèmes de procédure, de méthodologie et d'organisation que peuvent rencontrer les évaluateurs des Interfaces Homme(s)-Machine(s) (IHM), le problème capital auquel sont confrontés les évaluateurs est celui du choix des méthodes et techniques appropriées pour l'évaluation d'un système interactif donné. Selon le contexte international lié à l'utilisation du système interactif, les aspects socio-culturels doivent dans certains cas être aussi considérés. Ainsi, pour aider les évaluateurs dans le processus d'évaluation, il existe plusieurs types d'outils d'aide à l'évaluation dont les niveaux d'aide sont très variables. On constate qu'aucun des outils existants ne propose aux évaluateurs les méthodes appropriées au contexte de l'évaluation et aucun ne permet aux évaluateurs de prendre en compte les aspects socio-culturels, en plus des aspects pratiques liés à l'utilité et l'utilisabilité du système visé. Pourtant, le problème du choix des méthodes et techniques d'évaluation des IHM constitue l'une des étapes les plus importantes dans le processus d'évaluation car les objectifs recherchés et la qualité de l'évaluation vont dépendre des méthodes et techniques choisies par les évaluateurs. C'est dans ce contexte que se situe notre recherche visant un système d'aide à la décision faisant appel aux sous-ensembles flous. La première maquette du système d'aide à la décision est implémentée avec Prolog III. La maquette a fait l'objet d'une validation technique sur deux scénarios réels.

Mots-clés

Système d'Aide à la Décision, Système Interactif, Interface Homme-Machine, Evaluation, Sous-ensembles Flous, Critères d'Acceptabilité Pratique et Socio-Culturelle.

Abstract

Interactive system evaluation is the source of many difficulties for human-machine communication specialists as well as for non specialists. The quality of the evaluation depends on many parameters which are often contradictory. A major problem which project managers (and even human evaluators) are confronted with is the choice of criteria and methods appropriate for the evaluation. Given the international context linked with the situation of use, social and cultural aspects must also be considered in certain cases. Our research is oriented towards the design of a Decision Support system (DSS) for computer-aided choice of UI evaluation criteria and methods. It is based on Fuzzy sets. A first prototype was implemented with Prolog III and was validated with two real cases.

Keywords

Decision Support System, Interactive System, Human-Machine Interface, Evaluation, Fuzzy Sets, Practical and Socio-Cultural Acceptability Criteria