



HAL
open science

Conception d'un système de pilotage distribué, supervisé et multicritère pour les systèmes automatisés de production

Damien Trentesaux

► **To cite this version:**

Damien Trentesaux. Conception d'un système de pilotage distribué, supervisé et multicritère pour les systèmes automatisés de production. Automatique / Robotique. Institut National Polytechnique de Grenoble - INPG, 1996. Français. NNT: . tel-00536218

HAL Id: tel-00536218

<https://theses.hal.science/tel-00536218>

Submitted on 15 Nov 2010

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.

THESE

présentée par

TRENTESAUX Damien

Ingénieur diplômé de

L'Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble

pour obtenir le grade de DOCTEUR

de l'INSTITUT NATIONAL POLYTECHNIQUE DE
GRENOBLE

(Arrêté ministériel du 30 mars 1992)

Spécialité : Automatique / Productique

Conception d'un système de pilotage distribué, supervisé et
multicritère pour les systèmes automatisés de production

Soutenue le 24 Janvier 1996 devant la commission d'examen composée de:

Messieurs	P. LADET	Président du jury, co-directeur de thèse
	J.C. POMEROL	Rapporteur
	J. RICHARD	Rapporteur
	C. TAHON	co-directeur de thèse
	R. SOENEN	Examineur

Thèse préparée au sein du Laboratoire d'Automatique et de Mécanique Industrielles et Humaines de
l'université de Valenciennes et du Hainaut-Cambrésis

Pour Annabelle,

pour son amour et son soutien tout au long de ces trois années,
que cette période marque le début de notre vie commune.

Pour mes parents,
ma soeur et mon frère,

pour leurs encouragements.

Remerciements

Le travail présenté dans ce mémoire a été réalisé au Laboratoire d'Automatique Industrielles et Humaines de l'université de Valenciennes et du Hainaut-Cambrésis, au sein de l'équipe Génie Industriel et Logiciel dirigée par le professeur René Soënen. Je tiens à lui exprimer toute ma gratitude pour m'avoir accueilli au sein du laboratoire.

Ce travail est issu d'une collaboration entre le Laboratoire d'Automatique de Grenoble et le Laboratoire d'Automatique et de Mécanique Industrielles et Humaines de Valenciennes. Je tiens à remercier Mme Cheruy, M. Ladet, M. Soënen et M. Tahon qui ont permis cette co-tutelle, en espérant que celle-ci marque le début d'une collaboration des plus fructueuses entre les deux laboratoires.

Cette thèse a été co-encadrée par M. Tahon, professeur à l'université de Valenciennes et responsable de l'équipe Gestion de Production, et par M. Ladet, professeur à l'I.N.P.G, directeur du G.S.I.P. Je tiens tout particulièrement à les remercier pour leur aide et leurs précieux conseils tout au long de ces années.

Mes remerciements vont également à M. Pomerol, professeur à l'université de Paris VI, et à M. Richard, professeur à l'université de Nancy I pour avoir accepté d'être rapporteurs de cette thèse.

Je remercie vivement M. Roubens, professeur de l'université de Liège, Belgique pour m'avoir consacré plusieurs heures de son temps.

Je tiens enfin à exprimer ma gratitude à M. Delstanche et à M. Détrait, de la société Précimétal, Belgique pour avoir contribué à la validation des modèles présentés dans ce mémoire.

Il me tient à coeur de remercier l'équipe Gestion de Production, qui a réussi tant bien que mal à s'accommoder d'un grenoblois et de ses réseaux de Petri. Qu'ils soient remerciés pour leur accueil ainsi que pour les longues discussions constructives au coin du Macintosh.

Sommaire

Chapitre I: Vers une structure de pilotage distribuée et supervisée à ordonnancement prévisionnel partiel

1. Introduction	1
2. Le pilotage des systèmes automatisés de production	2
2.1. Les systèmes automatisés de production et la gestion de production	2
2.2. Les fonctions de pilotage.....	4
2.3. Problématique générale du pilotage temps réel.....	5
3. Typologie des structures de pilotage.....	6
3.1. Caractérisation suivant l'existence ou non d'un ordonnancement prévisionnel.....	6
3.2. Caractérisation selon la structure de pilotage	8
4. Pilotage utilisant un ordonnancement prévisionnel total.....	11
4.1. Niveau 1.....	11
4.2. Niveau 2.....	11
4.3. Conclusion.....	12
5. Structures de pilotage sans ordonnancement prévisionnel	13
5.1. Niveau 1.....	13
5.2. Niveau 2.....	13
6. Conclusion	19
7. Structures de pilotage utilisant un ordonnancement prévisionnel partiel.....	20
7.1. Description fonctionnelle	20
7.2. Problématique de la structure à ordonnancement partiel	23
7.3. Conclusion.....	27
8. Synthèse	27
8.1. Point de vue structurel	27
8.2. Point de vue ordonnancement.....	28
8.3. Proposition.....	28
9. Conclusion	29

Chapitre II: Spécification d'un système de pilotage hybride

1. Introduction	30
2. Résolution distribuée de problèmes.....	30
2.1. Agent	31
2.2. Tâche et objet.....	31
2.3. Communication.....	31
2.4. Contrôle.....	32
3. Pilotage de système de production à structure distribuée	32
3.1. La station intégrée de pilotage: SIP	33
3.2. Protocole de communication	34
3.3. Processus d'allocation	35
3.4. Décisions portant sur la gestion des files d'attente	36
3.5. Décisions portant sur la gestion de la reconfiguration	36
3.6. Décisions portant sur la gestion des aléas.....	36
4. Spécification de la structure de pilotage distribuée et supervisée	37
5. Spécification de l'ordonnancement prévisionnel partiel des tâches sur les ressources goulets	38
5.1. Modèle de gestion dynamique de contraintes	38
5.2. Suivi du processus de fabrication par le superviseur.....	39
5.3. Caractérisation du sous-ensemble des opérations planifiables.....	39
5.4. Pilotage par la gestion prioritaire et dynamique des ressources goulets.....	40
5.5. Détection et ordonnancement du goulet.....	44
5.6. Généralisation de l'approche à une famille de SIP	45
6. Capacité d'auto-adaptation du pilotage	46
6.1. Réactivité	46
6.2. Flexibilité	48
7. Conclusion	50

Chapitre III: Spécification d'un système interactif d'aide multicritère à la décision distribué

1. Introduction	51
2. Processus de décision	51
3. Processus décisionnel au niveau pilotage	52
3.1. Phase d'information.....	53
3.2. Phase de conception	56
3.3. Phase de choix	56
3.4. Phase d'évaluation.....	56

3.5. Intégration d'une coopération homme-machine	56
3.6. Vers un système interactif d'aide multicritère à la décision.....	57
3.7. Conclusion	60
4. Système interactif d'aide à la décision et pilotage	60
4.1. Concepts SIAD	60
4.2. SIAD et pilotage.....	65
4.3. Synthèse	72
5. L'aide multicritère à la décision.....	72
5.1. Approche multicritère pour la résolution d'un problème	72
5.2. L'approche multicritère en milieu industriel.....	76
5.3. Conclusion	76
6. Spécification d'un SIAMD pour le pilotage hybride	77
6.1. Structure du SIAMD.....	77
6.2. Niveau de connaissance.....	78
6.3. Niveau de coopération	78
6.4. Processus décisionnels du pilotage hybride.....	80
6.5. Modèle de décision multicritère: spécification de la matrice de décision	81
7. Conclusion	84

Chapitre IV: Modèle de données et modèle de traitement multicritère

1. Introduction	86
2. Modélisation des données	86
2.1. Approche mathématique choisie	86
2.2. Modèle des données incertaines	88
2.3. Enrichissement d'une donnée incertaine.....	91
2.4. Avantages du modèle.....	92
2.5. Application au cas du pilotage hybride.....	92
3. Structure et représentation des matrices génériques de décision multicritère	94
3.1. Matrice générique des critères	94
3.2. Matrice générique des alternatives	95
3.3. Matrice générique des évaluations	95
4. Choix de la méthode multicritère	96
4.1. La méthode Prométhée	97
4.2. Inconvénients.....	99
4.3. Avantages	99
4.4. Proposition.....	99
5. Première méthode M1	101
5.1. Etape 1: création des matrices génériques	102

5.2. Etape 2: instanciation des matrices génériques	102
5.3. Etape 3: Utilisation de la méthode multicritère Prométhée	104
5.4. Etape 4: mémorisation des résultats	105
5.5. Itération et intégration de l'analyse de sensibilité	108
5.6. Comparaison avec une méthode d'analyse de sensibilité	108
5.7. Etape 5: exploitation des résultats	109
6. Deuxième méthode M2	110
6.1. Etape 3: méthode multicritère basée sur les degrés de non-dominance.....	110
6.2. Etape 4: mémorisation des résultats	112
6.3. Etape 5: exploitation des résultats	113
7. Conclusion	113

Chapitre V: Proposition d'une structure de communication et intégration des processus décisionnels

1. Introduction	114
2. Structure de communication	114
2.1. Modélisation des niveaux T et A	116
2.2. Intégration de la robustesse au sein des niveaux T et A.....	125
3. Déroulement des processus décisionnels.....	132
3.1. Processus non structurés	133
3.2. Processus peu structurés	133
4. Conclusion	136

Chapitre VI: Validation

1. Introduction	137
2. Objectifs de l'évaluation	137
2.1. Contraintes expérimentales	137
2.2. Evaluation mono-poste.....	138
2.3. Evaluation multi-poste	139
3. Structuration logicielle du pilotage hybride	139
3.1. Objets de base.....	139
3.2. Structure logicielle SIP niveau distribué.....	139
3.3. Structure logicielle superviseur.....	148
4. Mise en oeuvre, cas mono-poste.....	149
5. Evaluation.....	150
5.1. Cas d'étude industriel.....	150
5.2. Objectifs.....	151

6. Support matériel pour la mise en oeuvre de la maquette	151
6.1. Efficacité de l'ordonnancement a posteriori	152
6.2. Intégration de l'auto-adaptation et la robustesse	154
6.3. Conclusion	161
7. Evaluation multi-postes.....	162
8. Conclusion	163

Annexe A: Gestion de production et fonctions prévisionnelles

Annexe B: Compléments sur la notion d'enrichissement

Annexe C: Exemple d'une liste de critère appliquée à un processus de fabrication de type Job-shop. Application au pilotage hybride

Annexe D: Modèle fonctionnel et structure logicielle du gestionnaire de contraintes

Annexe E: Parallèle entre les niveaux A et T et la norme OSI de l'ISO

Annexe F: Liste des processus décisionnels et fonctions d'interface

Annexe G: Modélisation générique d'un ordonnancement. Caractérisation d'un ordonnancement juste à temps

Annexe H: Calcul des charges potentielles. Ordonnancement juste à temps de la ressource goulet

Introduction générale

Les responsables d'entreprise sont actuellement confrontés à un fort accroissement de la complexité de la production en terme:

- de contraintes externes: concurrence, prix du marché, diversité des produits, cycle de vie produit réduit, etc.
- de contraintes internes: inertie des processus de fabrication, aléas de fonctionnement, etc.

Les contraintes externes imposent aux industries de réduire les cycles de conception- fabrication des produits. Cette réduction nécessite la flexibilité du système de production afin de faire face au développement d'une diversité de produits dont la durée de vie est sans cesse réduite. La problématique de la production est globalisée et concerne désormais non seulement le cycle de vie produit mais aussi l'évolution de la structure du système de production.

Par ailleurs, les contraintes internes réduisent cette capacité à suivre les évolutions du marché. L'inertie du processus de production détermine fortement les potentiels de vente d'un produit. En outre, la présence de perturbations de natures diverses altère considérablement le fonctionnement et l'efficacité du système de production. La réactivité devient une caractéristique clef de la gestion de production.

Les méthodes actuelles de gestion répondent difficilement à cette dynamique de l'évolution croissante. Elles sont très souvent développées pour un besoin spécifique et ce manque de genericité réduit fortement la capacité des entreprises à évoluer dans des conditions favorables. En particulier, la gestion temps-réel (pilote) des systèmes de production constitue un niveau critique et complexe.

Le cadre de travail dans lequel s'inscrit cette thèse porte sur le développement d'une structure de pilotage prenant en compte simultanément les contraintes

prévisionnelles et les contraintes réelles du système de production. Ces premières s'expriment en terme d'objectifs globaux et les secondes en terme de réactivité et de flexibilité.

Un premier élément de solution a été apporté par J.F.N. Tchako qui a proposé une structure de pilotage distribuée. Tchako a montré le haut niveau de flexibilité de cette structure [Tchako 94] et indiqué sa réactivité potentielle (résistance aux aléas, etc.). Plusieurs études ont montré les possibilités offertes par la structure distribuée pour absorber les fonctionnements dégradés.

L'objectif de ce mémoire est d'étendre et de compléter la structure de pilotage de Tchako afin:

- de permettre la prise en compte de la notion d'objectif global,
- d'accroître la capacité intrinsèque de réactivité de l'approche distribuée.

Nous montrons d'une part que l'intégration de la notion d'objectif global peut être obtenue par l'adjonction d'un niveau de supervision au modèle de Tchako. D'autre part, afin de conserver un niveau de réactivité satisfaisant, nous montrons qu'un compromis efficace entre les besoins en termes de réactivité et de gestion prévisionnelle (objectifs globaux) consiste à ordonnancer une partie des opérations à réaliser (ordonnancement partiel). Nous élaborons une liste des différentes approches possibles concernant les opérations (ou les ressources de production) à ordonnancer, et justifions une approche par ordonnancement partiel d'une ressource critique. Nous désignons la structure de pilotage correspondante sous le terme de pilotage hybride.

L'accroissement de la capacité intrinsèque de réactivité résulte de la conception et du développement d'une structure de communication décomposée en deux niveaux, intégrant chacun une gestion de la robustesse. Ce point constitue une conceptualisation et une extension de la structure de communication de Tchako. Un autre point porte sur le développement d'un module de gestion de contraintes dynamiques portant sur le processus de fabrication.

Par ailleurs, l'étude des processus de décision dans un cadre de pilotage temps-réel et distribué montre que les décisions sont de nature multicritère. L'étude de faisabilité que nous présentons a porté sur l'intégration d'une approche multicritère au niveau des processus décisionnels, prenant en compte les contraintes imposées au niveau du pilotage hybride. Elle a mis en évidence le faible

niveau de structuration des processus décisionnels et conduit à analyser les possibilités d'intégration de l'opérateur humain via un système interactif d'aide multicritère à la décision.

Le plan de ce mémoire est ainsi le suivant:

Nous étudions dans le chapitre I les structures de pilotage temps réel des systèmes de production. Nous établissons une typologie permettant la caractérisation:

- de la structure de pilotage: centralisée, hiérarchisée, coordonnée, etc.
- de la nature de l'ordonnancement prévisionnel: total, partiel ou inexistant.

Cette étude nous conduit à proposer une structure de pilotage hybride s'appuyant sur:

- une structure distribuée et supervisée selon les concepts de l'intelligence artificielle distribuée,
- l'ordonnancement prévisionnel partiel d'une ressource goulet.

Les spécifications de cette structure sont proposées dans le chapitre II. Nous mettons en évidence l'importance de la gestion particulière de la ressource goulet et nous développons le concept de goulet dynamique en spécifiant la notion de charge potentielle maximale. Le développement de ce concept permet en particulier d'accroître l'intégration de la réactivité. Nous détaillons alors la notion d'auto-adaptativité qui représente la caractéristique pour un système de pilotage à intégrer la gestion des événements (aléatoires ou non) en fonction de l'état du système de production.

Le chapitre III détaille en premier lieu la notion de processus de décision. Nous montrons que les processus de décision au niveau du pilotage hybride sont nécessairement multicritères et doivent s'appuyer sur la modélisation de données certaines et/ ou incertaines. Nous discutons de l'intérêt de l'interaction entre l'(les) opérateur(s) et le système de pilotage en terme d'accroissement de la réactivité et de la flexibilité. Nous donnons alors les spécifications générale à adopter lors du développement d'un système interactif d'aide multicritère à la décision distribué, et ce, en fonction des besoins et des contraintes liées au pilotage hybride.

Le chapitre IV détaille un modèle de données incertaines et le modèle d'enrichissement puis détaille une mise en oeuvre multicritère des processus décisionnels pour le pilotage hybride à partir de ces modèles.

Le déroulement des processus décisionnels nécessite d'organiser la structure de communication entre les différents centres de décisions. Le chapitre V modélise cette structure en proposant deux niveaux de communication et de robustesse. L'intégration des processus décisionnels selon cette structure est alors détaillée.

Le chapitre VI propose une mise en oeuvre informatique des concepts présentés dans ce mémoire. La validation est proposée sur un cas industriel. Nous mettons en évidence la difficulté avec laquelle le modèle de Tchako intègre la notion d'objectif global et l'amélioration sensible induite par la structure de pilotage hybride. La faisabilité de l'intégration de l'approche multicritère est en outre mise en évidence.

Enfin, nous indiquons les perspectives de ce travail qui portent principalement sur le développement d'un système interactif d'aide multicritère à la décision distribué.

Chapitre I

Vers une structure de pilotage distribuée et supervisée à ordonnancement prévisionnel partiel

1. Introduction

Les évolutions actuelles imposent aux entreprises de pouvoir faire face simultanément à la concurrence et aux besoins du marché. La logique de production à moindre coût est efficace pour supporter la concurrence mais ne l'est plus lorsqu'il faut intégrer la dynamique du marché. Les systèmes de production doivent donc être capables d'établir un compromis entre les coûts de production et le besoin de flexibilité en terme de produits. Ce compromis peut être atteint notamment par une amélioration des performances du système de production en terme de réactivité et de flexibilité.

Une solution consiste alors à intégrer cette flexibilité et cette réactivité au niveau de la gestion temps-réel ou "pilotage" du système de production [Dindeleux 92].

Le but de ce chapitre est:

- d'établir un état de l'art des différentes approches pour la gestion temps réel du système de production. Ces approches sont caractérisées par la structure de pilotage adoptée et la présence ou non d'une fonction d'ordonnancement en amont du pilotage,
- de montrer l'insuffisance de ces systèmes à assurer à la fois une intégration réactive et flexible tout en prenant en compte les objectifs globaux de production.

Dans cette optique, nous proposons tout d'abord une définition du pilotage d'atelier, puis nous présentons la problématique associée. Nous définissons une typologie des différentes méthodes en fonction de l'approche selon laquelle elles réalisent l'allocation des tâches avec ou sans ordonnancement pré-établi. Trois catégories sont définies:

- les méthodes à ordonnancement prévisionnel total,
- les méthodes sans ordonnancement prévisionnel,
- les méthodes à ordonnancement prévisionnel partiel.

Pour chacune de ces méthodes, nous détaillons la structure de pilotage et les caractéristiques associées.

Cette étude nous amène à proposer une méthode de pilotage hybride qui se fonde sur:

- un ensemble de décisions locales et temps réel,
- un ensemble de décisions globales ou locales qui peuvent être de nature prévisionnelle ou temps réel en fonction de la typologie des perturbations potentielles.

2. Le pilotage des systèmes automatisés de production

2.1. Les systèmes automatisés de production et la gestion de production

Nous nous intéressons plus particulièrement aux “systèmes automatisés de production” et à la gestion de tels systèmes.

2.1.1. Système automatisé de production

Un système automatisé de production (SAP) se caractérise par l'automatisation de certaines fonctions du système de production [Tchako 94]:

- les fonctions de transformation: machines automatisées, etc.
- les fonctions de manutention: convoyeurs automatisés, etc.
- le stockage: magasins automatisés, etc.
- le traitement de l'information: calculateurs, etc.

L'automatisation présente les principaux avantages suivants:

- elle permet de gérer de manière efficace la production [Giard 88],
- elle permet la création d'un suivi performant [Noyes 87],
- elle induit potentiellement une meilleure productivité et une meilleure flexibilité [Noyes 87].

En résumé, l'automatisation de la production permet de diminuer les temps de réponse et d'augmenter la flexibilité de la structure de production. Elle constitue par ailleurs une condition nécessaire pour la mise en place de la structure de pilotage décrite dans ce mémoire. Le concept de système automatisé de production à été associé à celui plus étendu de CIM (Computer Integrated Manufacturing) depuis la fin des années 70.

2.1.2. Gestion d'un système automatisé de production

De façon générale, la gestion d'un SAP comporte deux niveaux (figure 1):

- la gestion prévisionnelle (niveau 1): elle assure l'anticipation et la programmation d'un ensemble d'actions ou de décisions.
- la gestion temps réel (niveau 2): elle élabore les décisions et les actions qui sont réalisées en temps-réel et qui sont donc déclenchées par un ensemble d'événements liés à l'état courant du système de production.

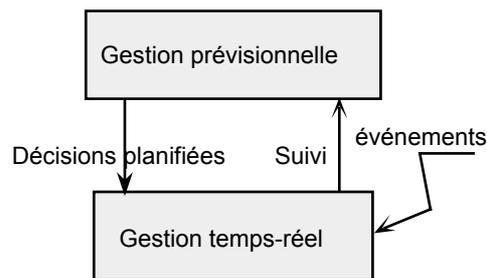


Figure 1: interactions entre les deux niveaux de gestion d'un SAP.

Nous allons décrire plus précisément ces deux niveaux.

2.1.2.1. Niveau 1: gestion prévisionnelle

La complexité et l'inertie du système de production conduisent à mettre en oeuvre une gestion prévisionnelle [Mercé 87].

2.1.2.1.1. La complexité du système de production

Plusieurs éléments sont à l'origine de cette complexité. Notons principalement:

- la diversité des types des ressources composant le système: ressources informatiques, mécaniques, électriques ou humaines. La gestion d'un ensemble de ressources hétérogènes est rendue difficile notamment par les différents niveaux nécessaires de connaissance, d'intelligence et les diverses possibilités de communication entre ces ressources.
- la multiplicité et parfois la contradiction entre les objectifs de production en terme de coût, délai ou qualité. Par exemple, la production organisée suivant un objectif global de qualité impliquera certainement un rendement global plus faible qu'une production organisée selon un objectif portant sur la minimisation des délais. Ce point introduit la notion d'ordonnancement multi-objectifs qui sera détaillée ultérieurement.
- la complexité des calculs prévisionnels qu'il est nécessaire la plupart du temps d'effectuer hors ligne.

2.1.2.1.2. L'inertie du système de production

L'inertie est due principalement aux composantes mécaniques et humaines du

système de production. Elle est définie en terme de temps de production, de changement d'outil, de calendriers (postes de travail), etc.

Ainsi, la gestion prévisionnelle permet l'intégration de cette inertie par la planification de certaines actions au niveau du système de production. Ce niveau peut être décomposé de manière générale en deux ou trois fonctions selon:

- la structure de pilotage adoptée,
- la présence ou non d'un ordonnancement prévisionnel, cf. Annexe A.

2.1.2.2. Niveau 2: gestion temps réel ou pilotage

Cette gestion se justifie par:

- l'altération des données sur lesquelles sont effectuées les prévisions du niveau 1: le système de production est soumis à des perturbations, à des modifications structurelles et/ou de son environnement,
- l'accessibilité, la fiabilité ou l'agrégation des données due au volume important d'informations à traiter et parfois contradictoires.

Ainsi, nous définissons le pilotage comme la structure de décision et d'information associée à la gestion temps réel.

2.2. Les fonctions de pilotage

La fonction pilotage temps réel comprend deux sous-fonctions: conduite et commande [Archimède 91], [Dindeleux 92], [Tchako 94], voir figure 2. La fonction conduite regroupe les activités décisionnelles qui sont traduites en ordres et transmises au niveau commande. Ce dernier niveau gère la réalisation du processus physique de fabrication et retourne au niveau conduite un ensemble de données de suivi. Nous détaillons ci-après chaque fonction.

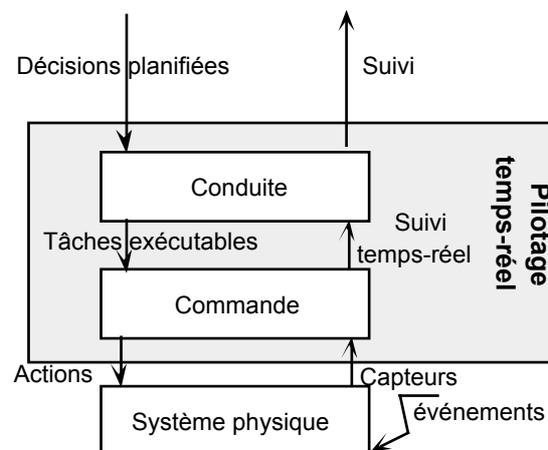


Figure 2: les fonctions de pilotage

2.2.1. La conduite

La conduite correspond au niveau décisionnel du pilotage dont le rôle principal est de mettre en oeuvre les décisions planifiées par le niveau prévisionnel. Si l'état du système de production ne permet pas d'appliquer correctement ces planifications, une décision peut être prise localement. Si aucune décision locale ne peut remédier au problème, il sera nécessaire de remettre en question les décisions du niveau supérieur.

La conduite peut être décomposée en sous-fonctions dont les principales sont:

- le contrôle: cette fonction assure la cohérence globale du système de pilotage, et gère les interactions entre les différents centres de décision, etc.
- la définition et la gestion de l'environnement: gestion de la logistique technique, des configurations, etc.

A ces deux fonctions est associée la fonction de suivi qui collecte les données nécessaires au suivi de fonctionnement du système, de ses performances, du respect des objectifs, etc.

2.2.2. La commande

Cette fonction correspond à la commande des équipements et automatismes. Elle regroupe notamment les fonctions de l'informatique industrielle.

2.3. Problématique générale du pilotage temps réel

Le pilotage doit établir constamment un compromis entre les ordres issus de la gestion prévisionnelle et les actions réellement effectuées au niveau du système de production. La problématique principale du pilotage est donc d'assurer la cohérence des décisions avec:

- les contraintes imposées par la gestion prévisionnelle en terme de décisions ou d'objectifs,
- les contraintes imposées par le système physique (perturbation, retard, etc.).

Ces deux aspects sont détaillés dans la suite du chapitre.

3. Typologie des structures de pilotage

3.1. Caractérisation suivant l'existence ou non d'un ordonnancement prévisionnel

A ce niveau, nous précisons quelques définitions afin de caractériser chaque type de pilotage des SAP. Nous utilisons la définition générale d'un ordonnancement donnée par Parunak [Parunak 91]:

“Un ordonnancement est un sous-ensemble du produit cartésien

Quoi*Où*Quand pour un sous-ensemble de l'ensemble des tâches à réaliser”.

Dans le domaine qui nous concerne, une tâche est une opération, c'est à dire un élément d'un ordre de fabrication. Ce dernier est établi selon les gammes opératoires du produit à réaliser. Nous différencions l'ordonnancement défini de manière prévisionnelle (a priori) de celui effectivement réalisé au niveau du système de production (a posteriori).

Nous définissons par ailleurs l'allocation comme un triplet Opération*Où*Quand. Un ordonnancement est alors défini comme une série d'allocations. De manière similaire, une allocation peut être prévisionnelle (a priori) ou réalisée en temps réel (allocation dynamique).

On peut différencier les types de pilotage selon qu'un ordonnancement prévisionnel des tâches est réalisé ou non:

3.1.1. Pilotage sans ordonnancement prévisionnel

Le niveau prévisionnel n'élabore aucun ordonnancement. La gestion des allocations et des files d'attente se fait dynamiquement au fur et à mesure de l'évolution événementielle du système de production.

3.1.2. Pilotage à ordonnancement prévisionnel

Dans ce cas, l'ordonnancement est défini par le niveau prévisionnel.

Par définition, une opération, appartenant au sous-ensemble des opérations à réaliser, à laquelle est associée une ressource, une date de début et une date de fin de fabrication est une opération planifiée.

Nous distinguons l'ordonnancement prévisionnel partiel de l'ordonnancement prévisionnel total:

3.1.2.1. *Pilotage à ordonnancement prévisionnel partiel*

Un ordonnancement prévisionnel est dit partiel si le sous-ensemble des opérations planifiées est inclus (inclusion stricte) dans l'ensemble des opérations à réaliser. Les approches de pilotage à ordonnancement prévisionnel partiel se différencient selon:

- le choix de ce sous-ensemble,
- les méthodes d'ordonnancement partiel.

Le sous-ensemble complémentaire est donc composé des opérations non planifiées dont l'allocation est réalisée dynamiquement, cf. figure 3.

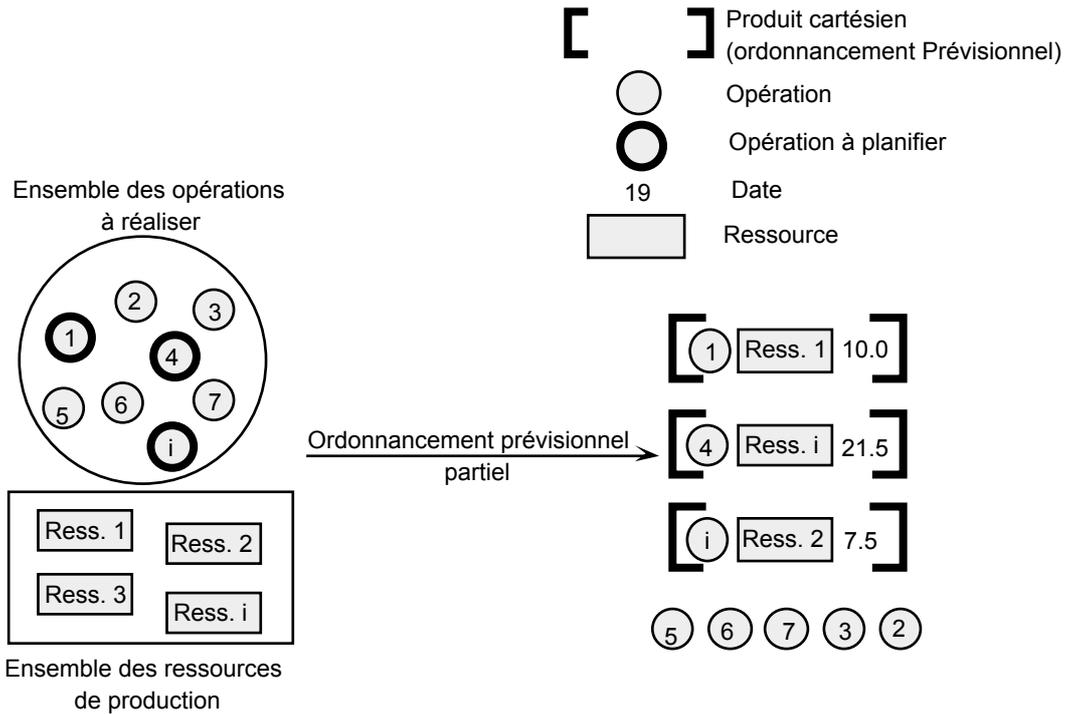


Figure 3: ordonnancement prévisionnel partiel

3.1.2.2. Pilotage à ordonnancement prévisionnel total

Un ordonnancement prévisionnel est dit total si le sous-ensemble des opérations planifiées est celui des opérations à réaliser, cf. figure 4.

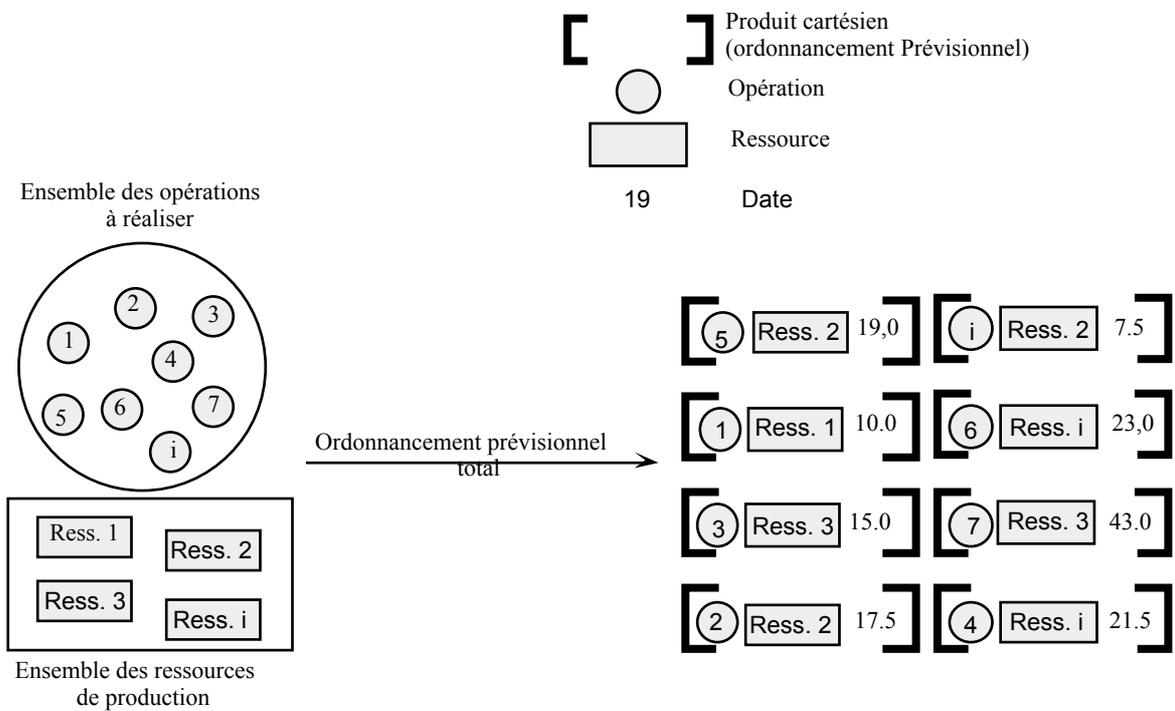


Figure 4: ordonnancement prévisionnel total

Le pilotage avec ordonnancement prévisionnel total constituait il y a quelques

années l'unique type de pilotage implémenté sur les sites industriels. Le Pilotage sans ordonnancement prévisionnel constitue une approche beaucoup plus récente et fait actuellement l'objet de nombreuses études. La théorie du pilotage à ordonnancement prévisionnel partiel est à ce jour relativement bien développée.

Pour chacun de ces trois types de pilotage, différentes structures peuvent être utilisées.

3.2. Caractérisation selon la structure de pilotage

Nous pouvons recenser plusieurs approches concernant la structuration du pilotage:

- l'approche pilotage centralisé,
- l'approche pilotage hiérarchisé,
- l'approche pilotage coordonné,
- l'approche pilotage distribué,
- l'approche pilotage distribué supervisé.

Nous en donnons tout d'abord une brève description. Dans la suite de chapitre, nous nous efforçons de préciser les caractéristiques des différentes approches en présentant un certain nombre de structures de pilotage s'appuyant sur ces différentes approches.

3.2.1. L'approche centralisée

L'approche centralisée est très classique (cf. figure 5). Elle se caractérise par un pilotage localisé au sein d'une ressource unique qui supervise la production et gère seule, en temps réel, les événements qui surviennent tout au long de la production.

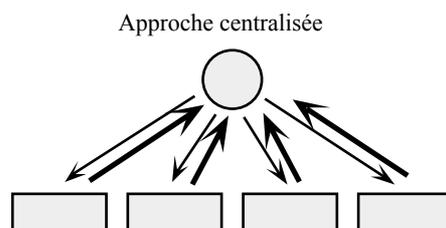
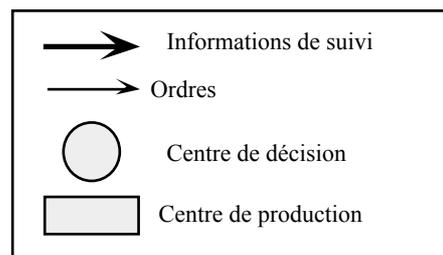


Figure 5: l'approche centralisée

3.2.2. L'approche hiérarchisée

Une présentation détaillée des nombreuses études réalisées est fournie dans [Gonthier 90], [Archimède 91], [Dindeleux 92], [Tchako 94]. La structure générale est représentée sur la figure 6.

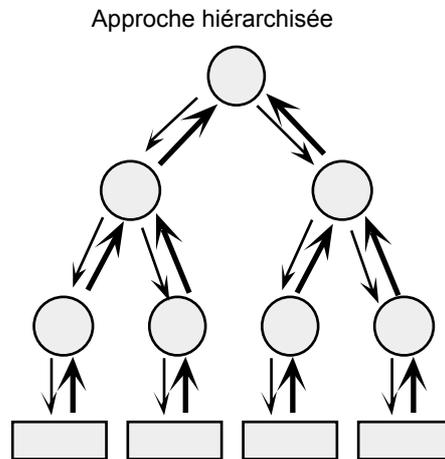


Figure 6: structure hiérarchisée

Généralement ces méthodes proposent une structure de contrôle qui s'établit à la fois au niveau prévisionnel (niveau 1) et temps réel (niveau 2).

Le modèle NBS [MacLean et al. 87] peut-être utilisé pour décrire l'architecture de base de telles structures de pilotage. La notion de niveau d'abstraction permet de modéliser une usine complète. Chaque niveau coordonne les unités de pilotage du niveau inférieur, et ce jusqu'au niveau le plus bas. La relation, à un niveau donné, est donc de dépendance vis à vis du niveau supérieur et de dominance vis à vis du niveau inférieur. Chaque décision est élaborée au niveau où un problème est détecté. Les niveaux inférieurs traitent cette décision comme une contrainte et transmettent en retour une information de suivi au niveau supérieur. La gestion temps réel concerne les niveaux atelier, cellules, postes de travail et automates.

3.2.3. L'approche coordonnée

L'approche coordonnée correspond à un ensemble de structures hiérarchisées où une coopération est possible au sein d'un même niveau (figure 7). Ces structures accroissent la capacité de décision au sein de chacun de ces niveaux.

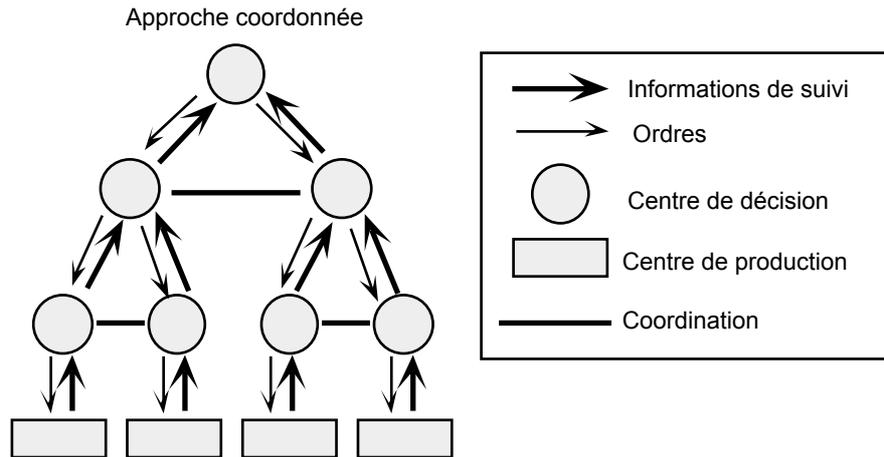


Figure 7: approche coordonnée

Les solutions envisagées pour la réalisation d'un ordonnancement prévisionnel sont fonction de la coopération intra-niveau. Celle-ci doit améliorer la réactivité en proposant une décomposition dynamique du problème au niveau où il apparaît plutôt qu'une décomposition hiérarchique des ordres vers les niveaux inférieurs. Cet aspect dynamique et cette coopération enrichissent la qualité de la décision. Le résultat est une meilleure intégration temps-réel des perturbations.

Cependant, cette coordination intra-niveau nécessite la création d'un contrôle particulier qui, tout en permettant une remise en question plus importante des décisions du niveau 1, doit gérer ces décisions afin d'assurer leur cohérence relative (contradiction, etc.).

3.2.4. L'approche distribuée

L'approche distribuée (figure 8) est fondée sur une distribution totale des capacités de décision parmi un ensemble de ressources pilotant un centre de production [Ferber et al. 88]. La structuration du contrôle d'une telle approche est beaucoup plus complexe que dans le cas d'une structure hiérarchique, plus rigide et plus facilement maîtrisable.

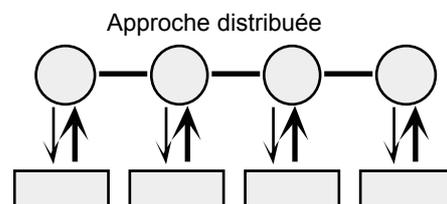


Figure 8: approche distribuée

3.2.5. L'approche distribuée supervisée

Un pilotage distribué supervisé (figure 9) se caractérise par un ensemble d'entités coopérantes sous le contrôle d'une entité superviseur dont le rôle est

d'imposer, de conseiller ou de modifier une décision afin de respecter un objectif plus global. Ainsi, l'entité superviseur possède une vision plus globale du processus de production.

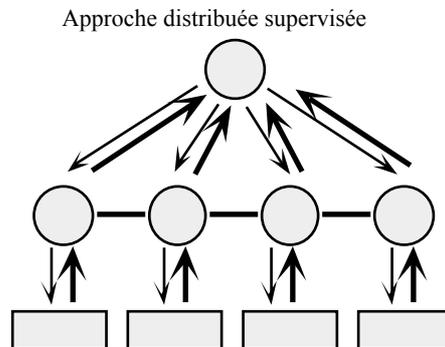


Figure 9: approche distribuée supervisée

4. Pilotage utilisant un ordonnancement prévisionnel total

Nous présentons tout d'abord les fonctions associées à chaque niveau en détaillant plus particulièrement le niveau 2.

4.1. Niveau 1

Les fonctions mises en oeuvre à ce niveau sont (cf. Annexe A):

- planification: élaboration d'un plan directeur de production,
- programmation: calcul d'un programme prévisionnel,
- ordonnancement (ordonnancement prévisionnel total).

4.2. Niveau 2

4.2.1. Objectif

A ce niveau, le but est de réaliser, en évitant dans la mesure du possible de le remettre en question, l'ordonnancement prévisionnel établi par le niveau 1. Ainsi, la responsabilité du pilotage est la réalisation au mieux des différentes tâches décrites. Son efficacité est donc évaluée principalement en regard de sa capacité à gérer les fonctionnements dégradés.

4.2.2. Caractéristiques générales

Les structures de pilotage à ordonnancement prévisionnel total sont principalement de type centralisé. En effet, les contraintes imposées par la réalisation d'un ordonnancement total prévisionnel impliquent un contrôle complexe du système de production. Seule une structure centralisée peut permettre un contrôle suffisant.

Les outils informatiques pour la gestion du niveau 1 en milieu industriel intègrent généralement la fonction "ordonnancement prévisionnel total".

Cependant, très peu d'entreprises informatisent le niveau 2, la responsabilité du pilotage étant assurée par le chef d'atelier. La structure du niveau 2 est donc naturellement centralisée.

4.3. Conclusion

L'ordonnancement pose de nombreux problèmes: d'une part par sa nature même. C'est un problème complexe et NP-complet. D'autre part, il nécessite d'importantes ressources de saisie et de traitement et introduit une étape critique dans le circuit de production.

D'autre part:

- il est difficile de prendre en compte les perturbations potentielles,
- les méthodes de résolution recherchent une optimisation monocritère peu pertinente.

[Parunak 91] caractérise ces difficultés en terme de:

- "desirability": il est difficile de tenir compte à la fois des contraintes réelles de production par une fonction d'évaluation et des souhaits des responsables de la production,
- "stochasticity:" il est difficile de prendre en compte les éventuels aléas puisque le séquençage des tâches anticipe la production effective,
- "tractability": le calcul d'un ordonnancement est NP-complet avec le niveau de complexité,
- "chaos": il est très difficile de prédire les conséquences d'une décision ou d'un événement,
- "decidability": le comportement d'un système de production est complexe et nuit à la compréhension des phénomènes et par conséquent, il nuit à la prise de décision.

Le pilotage est très spécifique à l'environnement de production. Cependant, la capacité de régulation est réduite par la rigidité structurelle.

Ainsi la structure centralisée est par essence peu robuste et ne peut efficacement intégrer la résistance aux perturbations.

5. Structures de pilotage sans ordonnancement prévisionnel

5.1. Niveau 1

Les fonctions associées au niveau 1 sont la planification et la programmation. La fonction ordonnancement n'apparaît plus à ce niveau. Elle est intégrée, sous forme d'un processus d'allocations dynamiques des tâches, dans la fonction

conduite temps réel du niveau 2.

5.2. Niveau 2

5.2.1. Objectifs

Le but est de réaliser le programme prévisionnel défini par la fonction programmation. L'allocation des tâches est réalisé de manière dynamique. L'efficacité est évaluée a posteriori par rapport à la qualité de l'ordonnancement réalisé.

5.2.2. Caractéristiques générales

La problématique de l'ordonnancement temps-réel est liée aux besoins issus du développement des structures de production "flexibles" [Merbaki 95]:

- accroissement de la complexité de la production,
- amélioration des temps de réponse face à la demande, aux perturbations, etc.

Une perte de qualité de l'ordonnancement est admise au profit d'une amélioration de la capacité à absorber les aléas.

Les méthodes de pilotage s'appuient sur un ensemble de structures plus large que précédemment:

- l'approche pilotage centralisé,
- l'approche pilotage distribué,
- l'approche pilotage distribué supervisé,
- l'approche pilotage par Kanban.

5.2.3. L'approche centralisée

Parmi les méthodes pour la réalisation temps-réel d'un ordonnancement, nous pouvons relever:

- L'utilisation d'une méthode multicritère pour la réalisation des tâches les plus contraignantes dans le but de minimiser le temps de cycle (modèle de [Messaoud 86]),
- L'utilisation des règles de priorité pour la gestion dynamique des files d'attente (le modèle de Merbaki [Merbaki 95]),
- L'utilisation de la simulation dans un but d'anticipation [Arai et al. 93], [De Smet et al. 95].

Nous détaillons ci-après les deux premières approches.

5.2.3.1. Le modèle de Messaoud

[Messaoud 86] présente une structure de pilotage temps réel dédiée à un atelier flexible robotisé. Les robots multifonctions permettent de réaliser un ensemble de

tâches. L'objectif est de réduire le temps de cycle total (makespan). Messaoud préconise l'utilisation de méthodes de surclassement multicritère dans le but de prendre en compte un ensemble de critères conflictuels qui représentent la possibilité, pour une allocation, de libérer un maximum de tâches réalisables.

L'algorithme développé est efficace mais il est très sensible à la pondération des différents critères. D'autre part, cette méthode est dédiée à un type d'atelier robotisé et nécessite une généralisation pour être exploitée dans un cadre de production de type job-shop.

5.2.3.2. *Le modèle de Merbaki*

[Merbaki 95] détaille une méthode s'appuyant sur des règles de priorité utilisées dynamiquement en fonction des événements, de l'état courant réel de l'atelier et des objectifs globaux. Une décision porte sur la réalisation d'une tâche présente dans la file d'attente d'une machine. La prise de décision dynamique est réalisée en trois étapes:

- analyse du système: longueur des files, retards, taux d'utilisation, etc.
- diagnostic: goulet d'étranglement, congestion, retard trop important, etc. La détection des symptômes (globaux ou locaux) se fait par règles d'inférence,
- choix de la règle à appliquer pour l'allocation en fonction des objectifs de production.

Une liste de symptômes pré-établis a été élaborée afin de mieux intégrer la notion d'objectifs de production. Un système d'aide à la décision simplifié a pour objet d'aider au choix des règles d'ordonnancement.

Cette méthode ne concerne que la gestion des files d'attente de chaque machine et on suppose que les opérations ne peuvent être effectuées sur des machines différentes. Cette limitation nuit fortement à la généralité de la méthode. Cependant Merbaki propose une méthodologie intéressante pour la gestion temps-réel des files d'attente. Une extension utilise la structuration multi-agents. Ces perspectives vont dans le sens des travaux de Kouiss [Kouiss et al. 95] présentés ci-après.

L'approche centralisée sans ordonnancement prévisionnel met en oeuvre un certain nombre d'outils, plus ou moins performants, pour la gestion des files d'attente et des allocations. L'avantage principal est qu'elle favorise la prise en compte d'un objectif global (niveau de stock par exemple).

Cependant, afin de réduire les temps de réponse, il est nécessaire:

- de simplifier la structure de production: routages, etc.
- de simplifier ou de préparer les calculs avant et pendant le déroulement du

processus de production.

5.2.4. L'approche distribuée

Nous décrivons brièvement quelques exemples de structures distribuées qui ont été proposées.

5.2.4.1. Le modèle de Bakker

[Bakker 88] présente le modèle de pilotage distribué DFMS, Distributed Flexible Manufacturing System (figure 10). Ce système est décomposé en deux sous-systèmes distincts:

- un ensemble de "Station Manager" SM qui gèrent chacune un ensemble de ressources pour: la gestion locale des files d'attente, l'allocation dynamique de tâches en fonction du temps de production restant sur chaque SM, le contrôle de la production, etc.
- un ensemble de "Fonctional Module" FM qui procurent un ensemble de services différents selon le principe client-serveur: gestionnaire d'outils, de programmes numériques, d'interfaçage avec l'opérateur, de base de données technique, de convoyeurs, etc.

Cette structure propose de nombreux avantages: fiabilité, extensibilité et simplicité.

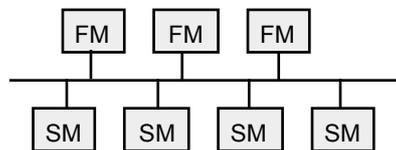


Figure 10: Modèle de Bakker

Les inconvénients sont essentiellement dus à l'absence d'optimisation globale: en effet, la technique d'allocation utilisée est celle du plus faible temps de production restant. Ces optimisations locales selon une fonction mono-objectif ne peuvent conduire à un ordonnancement réellement efficace et encore moins à une optimisation globale de cet ordonnancement.

5.2.4.2. Le modèle de Duffie

Duffie a développé une structure distribuée intégrant l'opérateur humain qui possède les caractéristiques principales suivantes [Duffie et al. 88], [Duffie et al. 94]:

- chaque entité possède le plus haut niveau d'autonomie réalisable,
- il ne peut exister de relation maître/esclave entre ces entités,
- chaque entité doit assumer le fait que la coopération avec une autre entité doit être possible, mais pas forcément nécessaire.

Une telle structure est incompatible avec la notion de globalité en terme d'information, de supervision et de contrôle. Chaque entité est composée d'un sous-système de communication et de contrôle pour la réalisation de l'ensemble des tâches. L'allocation des tâches s'effectue dynamiquement selon le principe de l'appel d'offre et la tâche est allouée à la première entité libre qui répond à cet appel d'offre.

De cette manière, Duffie développe une structure capable d'intégrer efficacement la résistance aux dysfonctionnements.

Il souligne la nécessité:

- d'intégrer l'opérateur humain dans un but de meilleure adaptabilité face à un ensemble d'événements difficilement maîtrisable par le système,
- d'utiliser la simulation en tant qu'outil de prévision,

Il lui semble aussi primordial:

- de réaliser une gestion par anticipation des blocages éventuels,
- de mettre en oeuvre des modules d'optimisation afin d'éviter que l'ensemble des décisions locales ne mènent à un ordonnancement peu efficace.

Cette méthode peut-être largement améliorée en enrichissant notamment le protocole de communication.

5.2.4.3. *Le modèle de Tchako*

Ce modèle sera plus amplement présenté dans le chapitre suivant. Tchako [Tchako 94] propose une méthode de résolution du problème semblable à celle de Duffie et Bakker. Il présente une structure distribuée s'appuyant sur les concepts de l'intelligence artificielle distribuée [Huhns 87]. Il développe ainsi une méthodologie de résolution distribuée de problèmes en particulierisant la notion d'agent définie par Ferber [Ferber et al. 88].

La structure de pilotage se compose d'un ensemble d'agents coopératifs (appelés Station Intégrée de Pilotage) comportant cinq sous-systèmes (décision, contrôle, communication, interface et information). Une SIP assure le contrôle d'une ou d'un ensemble de ressources: production, outillage, etc. Cette méthode est suffisamment générique pour être appliquée à différents types de systèmes, job-shop et ligne de fabrication. L'allocation de tâches et la communication sont considérées comme critiques pour la résolution de l'ordonnancement dynamique et sont étudiées en détail.

L'avantage de cette proposition se situe notamment au niveau de la qualité de la conceptualisation et de la spécification de la structure de pilotage distribuée.

Les caractéristiques principales d'une structure de pilotage distribuée sans ordonnancement prévisionnel sont:

- sa modularité,
- sa maintenabilité,
- son auto-configurabilité,
- la distribution des connaissances.

favorisant la résistance aux perturbations et son autonomie de fonctionnement.

Cependant la délocalisation complète des capacités de décision résultant de la restriction de l'information sur le système de production disponible dans chaque centre de décision induit des prises de décisions basées sur des informations locales. Une optimisation globale n'est donc, en général, pas possible. Un contrôle complexe doit donc être mis en place afin de pallier au mieux à cet inconvénient.

Enfin, l'ordonnancement étant réalisé dynamiquement, une difficulté réside dans la préparation des outils, matériaux, ressources humaines, etc. Ainsi, ce type d'approche ne peut être actuellement envisagé que pour des systèmes particuliers du point de vue de la structure, du type de production ou du mode d'exploitation.

La conclusion de cette étude est que l'approche totalement distribuée n'est pas à ce jour entièrement réalisable sur tous les sites industriels. L'approche distribuée supervisée tente de pallier à cet inconvénient.

5.2.5. L'approche distribuée supervisée

Nous détaillons deux implémentations de cette approche.

5.2.5.1. *Le modèle de Kouiss*

Kouiss [Kouiss et al. 95] présente une approche distribuée dont le niveau de supervision porte sur les règles d'allocation utilisées par chaque entité. Le modèle exploite la notion d'agent définie par Ferber. Chaque agent possède un sous-système de connaissance, d'expertise et de communication. La séquence des tâches de la file d'attente se fait par l'intermédiaire de règles de décision. Cette séquence se déroule en deux étapes:

- détection des symptômes: des symptômes sont pré-définis en fonction de variables d'états sous la forme d'un ensemble de règles. Ces symptômes sont activés lorsqu'une ou plusieurs variables d'état dépassent un seuil donné. La notion de seuil est décrite de manière approfondie,
- application d'une règle de décision: si nécessaire, un agent applique une règle ou une agrégation de règles pour sélectionner une tâche présente dans sa file d'attente [Merbaki 95]. La décision est donc locale et en temps réel.

Chaque agent utilise les règles indépendamment des autres agents. Le superviseur a pour rôle de faire respecter au mieux les objectifs globaux. (les écarts étant évalués par un jeu de symptômes globaux). Ainsi, il peut imposer une règle particulière à un ensemble d'agents dans le but de respecter une partie de ces objectifs.

Cette approche ne propose pas réellement de méthode d'allocation de tâches, mais plutôt une méthode efficace pour la gestion des files d'attente.

5.2.5.2. *Le modèle de Tacquard*

[Tacquard 95] décrit un modèle de pilotage d'un anneau flexible s'appuyant sur une méthode en flux tiré où seuls les besoins de production sont exprimés en terme de ratios. Une structure générique basée sur le modèle agent est développée. Chaque agent possède un sous-système de planification (définissant les buts en terme de sous-objectifs), de réaction (permettant durant la résolution d'un problème de reconsidérer certains choix) et de synthèse (agrégeant les données brutes en grilles de résultats).

Un agent superviseur contrôle le processus global de production. Cette structure de base est alors spécifiée pour un mode de production poussé ou tiré.

Dans le premier cas, le superviseur doit établir les ratios de production (planning) sur plusieurs périodes, proposer des alternatives lorsque les objectifs risquent de ne pas être atteints (réaction) et synthétiser les résultats pour le niveau 1. Les entités supervisées gèrent localement les règles de décision.

Dans le second cas, le superviseur doit établir les ratios de production (planning) pour la période suivante. L'agent superviseur a une responsabilité accrue: la fonction planning est identique, la fonction de réaction comporte la détection des non-respects des objectifs globaux et la prise en compte des aléas locaux au niveau d'une entité supervisée. La synthèse doit détecter les déviations potentielles vis à vis des objectifs globaux.

Une étude menée sur la réactivité face à une série d'aléas montre que l'approche flux tiré est plus efficace que l'approche flux poussé, et ce, en raison de la meilleure intégration du temps-réel. Un parallèle est fait entre la structure à flux tiré et la méthode Kanban que nous présentons ci-après.

L'inconvénient majeur de la méthode de Tacquard est qu'elle est dédiée à la gestion d'un anneau flexible et n'est pas applicable en l'état à un atelier flexible. Ainsi, de nombreuses simplifications portant principalement sur le flux de produits ont été faites.

L'approche distribuée supervisée sans ordonnancement prévisionnel permet de mieux éviter les comportements chaotiques qui peuvent survenir dans l'approche

distribuée, et ainsi d'anticiper des besoins en matériel, matière première, ressource humaine, etc.

Par ailleurs, la présence d'un superviseur apporte une vision globale du processus de production. Une prise en compte des objectifs est alors possible à plus long terme.

Cette approche mixte semble donc établir un compromis intéressant entre l'approche distribuée et l'approche hiérarchisée pour les structures de pilotage temps-réel sans ordonnancement prévisionnel.

5.2.6. L'approche Kanban

La méthode Kanban peut être considérée comme une méthode particulière de gestion distribuée. Les objectifs sont de:

- réguler les cadences de fabrication,
- minimiser les niveaux de stock.

Ceci est effectué par une gestion temps réel couplée à une répartition des capacités de décision au niveau de la fonction conduite. Aucun ordonnancement n'est réalisé. Seuls sont planifiés les besoins en produits finis ou les ratios de produits à réaliser.

Afin de proposer une méthode de gestion efficace, trois conditions assez contraignantes sont nécessaires:

- le flux des produits est simple. La méthode étant particulièrement adaptée dans les cas de processus de production linéaire,
- les demandes sont stables (régulières, constantes ou faiblement variables),
- les fournisseurs adoptent eux aussi une démarche juste à temps afin d'éviter les ruptures de stock et de diminuer le niveau des stocks amont.

Une mise en place efficace d'une telle méthode est conditionnée par la motivation des opérateurs vers la qualité et le progrès permanent et la suppression progressive des aléas: pannes, absentéisme, grève, etc.

6. Conclusion

La principale difficulté au niveau 1 est l'établissement d'un ordonnancement prévisionnel à la fois efficace et robuste. Dans le cas d'une structure de pilotage sans ordonnancement prévisionnel, ce niveau planifie uniquement les besoins.

Pour le niveau 2, les structures centralisées, dans le cas où il est nécessaire de fournir une aide à la décision temps réel, les méthodes de résolution et les modèles du système de production sont fortement simplifiés. Cependant, la centralisation

des données permet une meilleure vue globale de l'atelier et une meilleure anticipation des objectifs de production.

Les structures distribuées présentent un niveau élevé de réactivité, de modularité et de flexibilité, mais il est très difficile d'anticiper les productions effectives, ce qui nuit à leur mise en oeuvre.

Les structures distribuées supervisées semblent fournir un compromis intéressant entre les deux structures précédentes dans le sens où la centralisation du contrôle permet une gestion globale efficace et la distribution des capacités de décision permet une meilleure réaction aux perturbations.

Cependant, quelque soit la structure, l'absence de planification rend impossible la prévision en terme de besoin en ressources humaines et matérielles (outils, matière première, etc.).

Des études précédentes on peut tirer les conclusions suivantes:

L'approche du pilotage avec ordonnancement prévisionnel total ne peut être totalement efficace que dans le cas de systèmes stabilisés ne présentant que peu ou pas d'aléa. Dans le cas contraire, la rigidité du système rend la gestion des perturbations complexe. La réactivité n'est ainsi possible qu'au prix d'une importante baisse d'efficacité, ce qui conduit à ne pas exploiter tous les avantages d'une telle structure.

L'approche du pilotage sans ordonnancement prévisionnel intègre naturellement la robustesse, mais la mise en oeuvre d'une telle approche est complexe: en raison de la gestion temps réel et d'une trop faible capacité d'anticipation.

L'approche à ordonnancement prévisionnel que nous présentons dans la partie suivante cherche à proposer un compromis intéressant.

7. Structures de pilotage utilisant un ordonnancement prévisionnel partiel

Le pilotage à ordonnancement prévisionnel partiel propose de ne planifier qu'une partie des tâches à réaliser. Nous allons caractériser cette approche en spécifiant les rôles des niveaux 1 et 2 puis nous présenterons plusieurs cas pratiques d'application.

7.1. Description fonctionnelle

L'ordonnancement prévisionnel ne concerne qu'une partie des opérations, les autres sont allouées dynamiquement. Deux possibilités s'offrent ainsi à nous:

- soit l'ordonnancement prévisionnel partiel est réalisé par le niveau 1,
- soit il l'est par le niveau 2.

Dans tous les cas, le niveau 2 a pour objectif de minimiser les éventuelles remises en question des décisions du niveau 1. Ce qui revient à dire que le découplage entre les deux niveaux doit être conservé. Or l'objectif de l'ensemble des méthodes à ordonnancement partiel n'est pas d'éviter la remise en question de l'ordonnancement prévisionnel partiel, mais plutôt de l'exploiter en tant que guide qui peut nécessairement être remis en cause. Ainsi, pour préserver le découplage entre le niveau 1 et le niveau 2, la fonction ordonnancement partiel doit être intégrée au niveau du pilotage (niveau 2). Il est donc nécessaire de redéfinir ce niveau. La fonction conduite est enrichie et comporte une fonction de séquençage des tâches dont le rôle est (figure 11):

- d'établir un ordonnancement prévisionnel partiel portant sur une certaine catégorie de tâches,
- d'adapter cet ordonnancement en temps-réel. Une remise en question de celui-ci est possible à chaque instant.

Cette fonction ordonnancement peut-être localisée ou répartie parmi un ensemble de ressources.

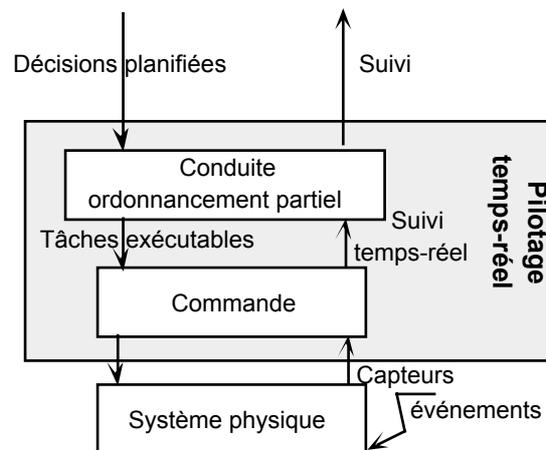


Figure 11: niveau 2: pilotage à ordonnancement prévisionnel partiel

7.1.1. Le niveau 1

Il comporte comme dans le cas de l'approche sans ordonnancement prévisionnel les fonctions de planification (plan directeur) et programmation (programme prévisionnel).

7.1.2. Le niveau 2

L'ensemble des tâches opératoires planifiées peut-être établi sur la base de critères de sélection. L'objectif est d'extraire un sous-ensemble de tâches qui présente des caractéristiques communes. Les critères de sélection sont regroupés en quatre catégories: critères de sélection portant sur les flux, la dimension

temporelle (délai, horizon, etc.), les produits et les ressources. Chacun de ces critères permet d'établir la liste des tâches à planifier.

7.1.2.1. Critères de sélection portant sur les flux

Les critères portant sur les flux peuvent être définis en terme:

- de règle de précedence des produits,
- de ratio temps gamme / temps de passage dans le système de production,
- de degré de liberté: planifier en priorité les opérations les plus contraignantes afin d'éviter les blocages ou les pertes d'efficacité globale et de se concentrer sur l'optimisation du séquençement dynamique des autres opérations.

7.1.2.2. Critères de sélection portant sur la dimension temporelle

Ces critères peuvent être définis:

- en terme de temps de présence sur le site de production: minimiser de manière globale le temps de présence de matières dangereuses, volatiles, instables, etc.
- en terme de délai de livraison: ordres de fabrication à livrer impérativement (contrats à date critique d'annulation de livraison) avec une marge de liberté sur les délais très faible,
- en terme d'horizon temporel: on planifie l'ensemble des opérations ou des ordres de fabrication complets réalisables sur un horizon pré-défini. Cette planification peut-être déclenchée périodiquement ou par une méthode à horizon déroulant.

7.1.2.3. Critères de sélection portant sur les produits

Ces critères s'expriment principalement en terme de coût:

- d'achat de matière première: prix élevés ou qui évoluent rapidement.
- de stockage des produits finis ou semi-finis: prise en compte de la capacité de stockage globale, etc.

Ils peuvent aussi s'exprimer en terme de compatibilité (stockage, production) entre produits.

7.1.2.4. Critères de sélection portant sur les ressources

Les critères de sélection portant sur les ressources sont:

- les ressources critiques décrites en terme de capacité de production: on cherche à détecter et à planifier ces ressources,
- les caractéristiques des ressources (temps de changement de lots, etc.): on exploite au mieux une ou plusieurs ressources dont la(les) caractéristique(s) influent sur les performances du processus de production,
- la logistique associée à une ressource (présence outillage, personnel, etc.)

7.2. Problématique de la structure à ordonnancement partiel

La problématique de ce type de pilotage est liée à celles des deux précédentes approches: d'une part il est nécessaire de planifier et de réaliser la partie prévisionnelle en évitant au maximum une remise en question et d'autre part, il est nécessaire d'effectuer un séquençage efficace des opérations non planifiées, et ce, de manière dynamique. Plusieurs modèles ont été élaborés, nous les présentons ci-dessous.

7.2.1. Approche centralisée

[Sun et al. 94] propose une méthode originale pour le pilotage centralisé s'appuyant sur la notion de vue duale de l'ordonnancement. Les opérations planifiées sont celles qui peuvent être planifiées au mieux dans une période de temps fini. Ainsi, le critère de sélection porte sur la dimension temporelle -horizon déroulant-. La méthode utilisée (méthode "backward") est basée sur une inversion de l'échelle temporelle. Le but de cette méthode est de proposer un pilotage temps réel qui prend en compte les dates de besoins des différents ordres de fabrication. Cette méthode utilisée pour le niveau 1 n'avait pas encore été exploitée au niveau du pilotage.

L'ordonnancement prévisionnel partiel est calculé à l'aide d'une ressource unique qui contrôle à elle seule le processus de fabrication.

L'avantage principal d'une telle méthode est qu'elle prend les dates de besoins des ordres de fabrication comme une contrainte forte.

Cependant, la gestion de cette contrainte est complexe. En effet, en cas de perturbation importante, il est possible qu'une tâche soit planifiée dans le passé. Ce cas n'est que partiellement abordé.

D'autre part, la fonction d'évaluation ne prend pas en compte l'aspect multi-objectifs. Ce qui réduit l'efficacité de l'ordonnancement réalisé.

Enfin, la définition d'horizon temporel a été peu abordée malgré le fait qu'elle conditionne fortement l'efficacité de la méthode.

Les avantages et les inconvénients liés à la structure sont identiques à ceux des structures à approche centralisée avec ordonnancement prévisionnel total: décisions centralisées, ressource unique, etc.

L'intégration d'un ordonnancement prévisionnel partiel implique par ailleurs une complexité accrue de la fonction pilotage. La structure centralisée ne semble donc pas très adaptée à un contexte d'ordonnancement partiel.

7.2.2. Approche hiérarchique

ORABAID [Thomas 80], [Legall 89] est une méthode qui a pour objectif

d'accroître la réactivité du pilotage en utilisant la notion de marge libre et de groupes de tâches permutable pour l'ordonnancement en temps réel. Un ordonnancement partiel est réalisé avec comme critère de sélection des opérations planifiables le degré de liberté. Les informations portant sur la planification ne concernent pas la description des allocations de toutes les tâches, mais seulement la description des allocations de groupes de tâches. Un système d'aide à la décision permet d'accroître la qualité du contrôle et de la réactivité temps-réel. [Billaut et al. 93] développe cette méthode en deux étapes:

- une phase d'analyse: on recherche un sous-ensemble de l'ensemble des ordonnancements possibles à l'aide d'une heuristique. La recherche de l'exhaustivité nuirait fortement à l'efficacité de la méthode en terme de temps de calcul. Les ordonnancements possibles sont alors regroupés par catégories d'ordonnancements selon les degrés de liberté qu'ils autorisent. Deux ordonnancements possédant le même degré de liberté proposent des groupes d'opérations identiques dont la permutation ne modifie pas le séquençement des autres opérations. Ceci est représenté graphiquement par des diagrammes de Gantt étendus: les groupes d'opérations sont agrégés en macro opérations qui n'imposent pas le séquençement des opérations entre deux dates limites, incluant des marges potentielles. Les groupes d'opérations sont alors disponibles au niveau des postes de pilotage local,
- une phase de prise de décision: différents événements déclenchent le besoin de prise de décision (panne, libération de ressource, etc.). Celle-ci doit stipuler plus précisément les opérations à réaliser au sein d'un groupe d'opérations. Le système de pilotage local doit alors détecter les groupes qui ne peuvent respecter les contraintes fortes et par conséquent remettre en question l'ordonnancement. Dans le cas contraire, les groupes d'opérations sont permutés ou les marges sont utilisées et aucune remise en question de l'ordonnancement prévisionnel partiel n'est nécessaire.

L'ordonnancement prévisionnel partiel est réalisé par une seule ressource qui distribue les ordres de fabrication des macro-opérations au niveau des postes de conduite locale. Ces derniers gèrent localement ces macro-opérations et informent la ressource de la nécessité d'une éventuelle remise en question de l'ordonnancement prévisionnel partiel.

L'avantage de cette méthode est qu'elle s'appuie sur l'établissement d'un niveau d'importance pour le séquençement des tâches. La méthode consiste à utiliser les tâches les moins contraignantes pour absorber les perturbations. De plus, la prise de décision est anticipée et l'agrégation en macro-opérations permet de représenter

synthétiquement les solutions envisageables.

La difficulté d'une telle méthode réside dans la difficulté à établir la typologie des situations potentielles et à gérer les cas où une perturbation implique un réordonnement.

L'intérêt d'une approche hiérarchique pour un ordonnancement partiel provient de:

- l'interdépendance très élevée entre niveaux: elle induit une rigidité importante dans la structure de pilotage et permet d'avoir des temps de réponse réduits,
- la phase de conception de la structure de pilotage qui est fortement spécialisée en fonction de la structure de production: ceci implique de très bonnes performances globales,
- les données peuvent être agrégées: ce qui permet d'avoir une vision globale du système par niveau,
- la structure des données est hiérarchisée: ce qui implique une connaissance du système entier la plus large possible. Une optimisation des décisions est dès lors possible.

Mais l'approche hiérarchique pose des problèmes importants qui résultent principalement de l'évolution générale des systèmes de production en terme:

- d'organisation (reconfiguration rapide, augmentation de la complexité des systèmes à gérer, etc.),
- et d'environnement (évolution de la concurrence et des marchés impliquant une variété importante de produits, etc.).

Ceci se traduit par:

- la difficulté à établir une liste d'indicateurs de performances stables, pertinents, non-redondants, exhaustifs, etc. [Bitton 90],
- un manque de flexibilité lorsque l'on désire apporter une modification à la structure du système de production,
- une réactivité faible en raison de la structure même du pilotage. Un niveau de contrôle supplémentaire peut être adjoint en gérant par exemple la relaxation d'un certain nombre de contraintes. Cependant, cette approche pose de nouveaux problèmes: comment reconnaître une perturbation et limiter les conséquences de son occurrence? Comment augmenter le pouvoir de décision du niveau conduite sans diminuer l'efficacité de l'ordonnement? etc.

7.2.3. Approche coordonnée

7.2.3.1. La méthode CODECO

CODECO [Kallel 85], [Pellet 85] peut être considérée comme une extension, par l'intégration de la coopération intra-niveau, de la méthode ORABAID. Cette coopération entre les différentes unités de pilotage cherche à optimiser et à mieux réagir localement en cas de perturbation. Elle permet une meilleure prise en compte de l'imprécision qui peut accompagner les décisions des niveaux supérieurs. L'ordonnancement est décrit partiellement (le critère de sélection des opérations planifiables est celui du degré de liberté) et réalisé par une seule ressource. Celle-ci informe les niveaux inférieurs de ses décisions. Chaque élément du niveau inférieur organise les tâches localement en fonction de ses contraintes et en collaboration avec les autres éléments de même niveau. La difficulté réside dans la détermination du degré de liberté concernant les marges allouées pour chaque poste de conduite locale.

7.2.3.2. La méthode PCS

L'originalité de la méthode PCS [Archimède 91] provient du développement d'une fonction de réactivité basée sur la définition et l'exploitation d'une typologie de perturbation.

La structure hiérarchique décompose périodiquement à chaque niveau le problème de l'ordonnancement en fonction de l'horizon et de la période étudiée (le critère de sélection des opérations planifiables est celui de la planification périodique). La réalisation de l'ordonnancement partiel n'est donc plus du ressort d'une seule ressource, mais d'un ensemble de ressources.

Les avantages liés à l'enrichissement des relations intra-niveaux pour l'intégration d'une fonction d'ordonnancement partiel sont principalement:

- une meilleure intégration de la réactivité au sein de la structure elle-même,
- une autonomie décisionnelle accrue qui permet de diminuer considérablement le nombre de remises en cause de l'ordonnancement prévisionnel partiel.

L'inconvénient majeur est qu'en cas d'incapacité à absorber une perturbation, il est nécessaire de remettre en question les décisions d'ordonnancement.

Nous pouvons par ailleurs ajouter que:

- il est très difficile de déterminer la marge d'autonomie des différents acteurs dans la prise de décision locale,
- la propagation des perturbations et aléas n'est pas maîtrisable facilement.

7.3. Conclusion

Dans la structure à ordonnancement partiel, le niveau 1 planifie un sous-ensemble d'opérations sans s'attacher à la définition d'un quelconque ordonnancement. Cette planification est moins complexe à réaliser qu'un ordonnancement global. Sa seule contrainte est le respect des capacités de l'atelier pour planifier des séquences admissibles par le système de production.

Le niveau 2 établit un compromis entre les contraintes imposées par l'ordonnancement prévisionnel partiel et la gestion temps réel du système de production.

La difficulté de cette approche apparaît lorsqu'une perturbation provoque la remise en question d'une partie de l'ordonnancement prévisionnel. Les solutions sont alors de deux types:

- l'ordonnancement partiel est réévalué (en temps réel ou en temps différé). Dans ce cas, on retrouve une situation similaire au cas des méthodes à ordonnancement prévisionnel. Cependant, la complexité est ici réduite car l'ordonnancement prévisionnel est partiel. La difficulté est d'évaluer ou d'anticiper l'importance d'une quelconque remise en question de l'ordonnancement prévisionnel,
- l'ordonnancement partiel est annulé. Dans ce cas, le niveau 2 doit redéfinir les objectifs et minimiser l'influence de la perturbation.

8. Synthèse

L'étude présentée dans ce chapitre nous amène aux conclusions suivantes:

8.1. Point de vue structurel

Les structures distribuées supervisées proposent de nombreux avantages qui exploitent:

- les caractéristiques de globalisation (en terme d'optimisation ou d'objectif) qui sont celles des approches hiérarchiques: la présence d'un superviseur fournit une vision et une capacité de prise de décision globale,
- les caractéristiques associées aux processus d'allocation dynamique des structures distribuées qui lui octroient une capacité satisfaisante en terme de réactivité et flexibilité.

8.2. Point de vue ordonnancement

La création d'un ordonnancement prévisionnel partiel apporte les avantages suivants:

- elle limite la restriction de la capacité de réaction en planifiant qu'un sous-ensemble des opérations à réaliser,
- elle décrit un ensemble de guides (contraintes) favorisant le respect d'un ou des objectif(s) global(aux) et présente une meilleure viabilité industrielle.

8.3. Proposition

Notre contribution s'inscrit dans le cadre du développement d'une structure de pilotage distribué supervisé intégrant un ordonnancement prévisionnel partiel, ou "pilotage hybride" qui s'appuie sur le modèle distribué de Tchako. Elle porte sur:

- la conception du niveau de supervision,
- le développement d'un modèle de gestion dynamique de contraintes portant sur la réalisation des tâches. L'intégration d'un tel modèle facilite en particulier la détection des charges de chaque ressource de production,
- la conception d'un processus d'échange d'information pour le suivi du processus de fabrication entre agents et le superviseur afin d'améliorer la détection et la gestion de la ressource ordonnancée,
- la définition de la charge potentielle maximale pour chaque ressource,
- la conception du processus d'ordonnancement partiel de la ressource qui présente la charge la plus élevée.

dans le but d'accroître la capacité à intégrer la notion d'objectif global. Elle porte aussi sur l'accroissement de la réactivité:

- en décomposant la structure de communication en deux niveaux hiérarchiques qui prennent en compte les besoins en terme de robustesse. L'intégration de la robustesse est facilitée par le développement d'une méthode qui se base sur une décomposition canonique des différentes étapes de chaque processus d'échange de donnée. Nous proposons selon ces deux niveaux la généralisation de la notion de protocole: protocole de transfert, d'échange d'information, d'ordonnancement et d'allocation,
- en exploitant le modèle de contraintes dynamique dans un cadre réactif. Une telle intégration permet de prendre en compte des conditions de fonctionnement particulières ou fugitives.

Par ailleurs, nous nous proposons de réaliser une étude de faisabilité qui porte sur le développement d'une approche multicritère pour la prise décision au niveau du pilotage. Nous étudions l'influence d'une telle approche au niveau de la réactivité. En outre, la mise en évidence de processus décisionnels peu ou pas structurés au niveau du pilotage hybride nous a poussé à étudier l'intérêt de l'intégration de l'opérateur par l'intermédiaire d'un système interactif d'aide

multicritère à la décision dédié au pilotage hybride.

9. Conclusion

Dans ce chapitre nous avons décrit trois approches de pilotage utilisant ou non un ordonnancement total prévisionnel ou partiel. Nous avons présenté pour ces trois approches des solutions qui ont été apportées au problème du pilotage selon la structure adoptée.

Tenant compte du modèle de Tchako, l'analyse de ces solutions nous a conduit à proposer une structure mixte qui établit, d'une part, un compromis entre les structures centralisées et distribuées et d'autre part, entre les approches à ordonnancement total ou sans ordonnancement.

Le chapitre suivant détaille le modèle de Tchako ainsi que les spécifications de la structure hybride.

Chapitre II

Spécification d'un système de pilotage hybride

1. Introduction

Le but de ce chapitre est de spécifier un système de pilotage hybride, c'est à dire une structure distribuée supervisée avec ordonnancement prévisionnel partiel.

La structure distribuée constitue la base de notre approche. Elle s'appuie sur les concepts de la résolution distribuée de problèmes. Nous présentons ces différents concepts puis la structure distribuée développée par [Tchako 94].

Nous spécifions ensuite la structure de pilotage hybride et la notion d'ordonnancement prévisionnel partiel. L'ordonnancement partiel concerne la ressource la plus contraignante en terme de capacité de production. Nous développons les concepts issus de la théorie du management par les contraintes [Marris 94] et de la gestion des ressources goulets [Goldratt et al. 84]. Nous proposons une définition d'une ressource goulet qui s'appuie sur les concepts de [Marris 94]. La différence se situe au niveau de la prise en compte de la dynamique du processus de fabrication. La notion de charge potentielle maximale d'une ressource nous permet de détecter dynamiquement la ressource la moins capacitaire qui est ordonnancée par le superviseur.

Nous caractérisons alors (selon une optique d'évaluation) l'auto-adaptativité qui représente la capacité pour un système de pilotage à pouvoir gérer les différents événements (aléatoires ou non) qui peuvent survenir durant le processus de fabrication. L'auto-adaptativité est décrite en terme de flexibilité et de réactivité du système de pilotage. L'étude de l'auto-adaptativité et du respect des objectifs globaux nous permet ainsi de mettre en évidence les propriétés de notre modèle.

2. Résolution distribuée de problèmes

La résolution distribuée de problèmes s'attache à appliquer les théories de l'Intelligence Artificielle Distribuée (I.A.D) dans le but de faire interagir un certain nombre d'entités (agents) qui partagent leurs connaissances et leur savoir-faire afin de résoudre un problème général décomposé en un sous-ensemble de tâches.

Dans une structure distribuée, il est nécessaire de spécifier l'organisation interne du système et de proposer un modèle de communication. Nous distinguons ainsi cinq concepts attachés à cette approche: les concepts d'agent, d'objet, de tâche, de contrôle et de communication [Huhns 87], [Rasmussen et al. 91].

2.1. Agent

Un agent est défini de manière conceptuelle comme [Ferber et al. 88]:

"Une entité physique ou abstraite capable d'agir sur elle-même et sur son environnement. Cette entité dispose d'une représentation partielle de cet environnement et peut communiquer avec d'autres agents. Son comportement est la conséquence de sa perception du monde, de ses connaissances et des interactions avec les autres agents".

2.2. Tâche et objet

Le problème à résoudre peut-être décomposé en un ensemble de tâches. Chaque agent contribue à la résolution globale du problème en participant à la résolution (partielle ou totale) d'une ou plusieurs tâches.

Un objet est une ressource physique ou logicielle associée à chaque agent (qui peut en posséder plusieurs). Ces objets permettent la réalisation des tâches.

2.3. Communication

La représentation partielle de l'environnement d'un agent implique nécessairement qu'une information soit échangée dans le but de réaliser l'ensemble des tâches. L'intégration et la spécification du concept de communication en fonction de la typologie de structure adoptée permettent cet échange.

Deux approches sont possibles (figure 1):

- les informations sont échangées indirectement: dans ce cas, un mécanisme de tableau noir est utilisé (Blackboard). Chaque agent vient y chercher les informations qui lui sont nécessaires ou y déposer celles dont il dispose [Hayes-Roth 85],
- les informations sont échangées directement entre agents: dans ce cas, le mécanisme de communication doit être spécifié. Un exemple d'une telle approche est celui de la négociation [Smith 80].

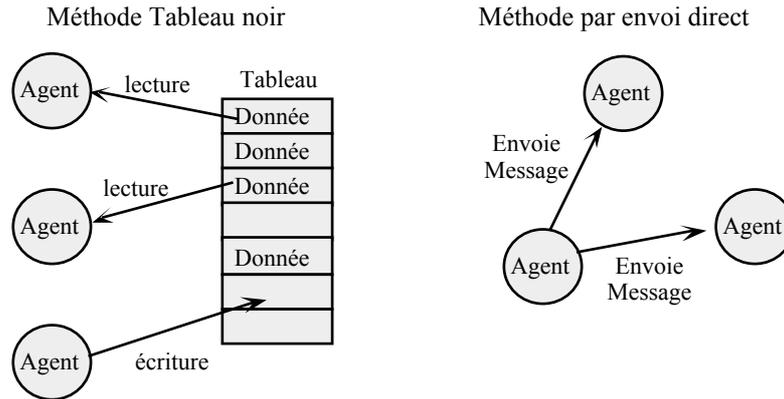


Figure 1: modes de communication entre agents

L'approche et le protocole spécifiés dépendent fortement du contrôle. La communication peut utiliser différents modèles: compétition, appel d'offre, négociation, etc.

2.4. Contrôle

Le contrôle, dont le rôle est de préciser le comportement de chaque agent est défini selon trois dimensions [Erceau et al. 91]:

- le niveau de coopération. Il concerne l'interaction entre tous les agents composant le système. La coopération se décrit par un degré de coopération (agents pleinement coopératifs ou antagonistes) qui influence l'approche adoptée pour la résolution du problème,
- l'organisation du groupe d'agents. Afin de respecter la cohérence globale, le groupe d'agents doit être structuré selon certains critères: fonctionnalité, interactivité, sphère d'influence, etc.
- la dynamique de l'évolution. La structure du groupe d'agents peut, selon les cas, être modifiée de manière dynamique au cours du temps.

Les paradigmes d'organisation et d'évolution spécifient en particulier les possibilités pour un groupe d'agents de modifier l'effectif de l'univers multi-agent (ajout ou suppression d'agent). Ils concernent de manière plus globale la dynamique de la structure.

3. Pilotage de système de production à structure distribuée

L'ensemble des concepts précédents a été appliqué à la modélisation d'une structure de pilotage distribuée [Tchako 94], [Tchako et al. 94]. Le modèle est le suivant:

Les agents (“Station Intégrée de Pilotage”: SIP) sont spécialisés dans les opérations de production et sont de type cognitifs. Ils possèdent un niveau de connaissance suffisant pour permettre une prise de décision.

Le problème global à résoudre est la réalisation d'un ensemble d'ordres de fabrication. Cet ensemble est décomposé en un ensemble d'opérations élémentaires qui sont associées à chaque étape de la production pour un produit donné selon une gamme donnée.

Les objets sont des ressources de type industriel: machines, outillages, convoyeur, etc. Le contrôle est de type pleinement coopératif et la communication est basée sur le paradigme de négociation.

Nous décrivons brièvement ces modèles. Nous nous attacherons plus particulièrement à la structuration des agents et au protocole de communication pour l'allocation de tâches.

3.1. La station intégrée de pilotage: SIP

Le rôle d'une SIP est de:

- gérer localement en temps réel les processus d'allocation de tâches, les files d'attente, etc.
- gérer les actions locales (chargement programme, etc.),
- gérer la disponibilité des éléments nécessaires à la réalisation des tâches opératoires,

Une S.I.P est composée de cinq sous-systèmes:

- le sous-système de décision: ce sous-système est responsable de l'ensemble des décisions prises au cours du processus de résolution du problème global. Les algorithmes dédiés à la prise de décision sont basés sur une méthode à fonction objectif,
- le sous-système de contrôle/ commande: le but de ce sous-système est de commander les ressources de production.

Ces deux systèmes dépendent des sous-systèmes suivants:

- sous-système d'information interne: son rôle est d'assurer la communication entre les sous-systèmes et de gérer des informations locales,
- sous-système de communication: ce système assure la liaison entre la S.I.P et celles qui constituent son environnement,
- sous-système d'interface: il permet les échanges d'informations avec l'opérateur humain (états des sous-systèmes, du processus de fabrication, d'allocation, de gestion de file d'attente).

La structure d'une SIP est représentée sur la figure 2.

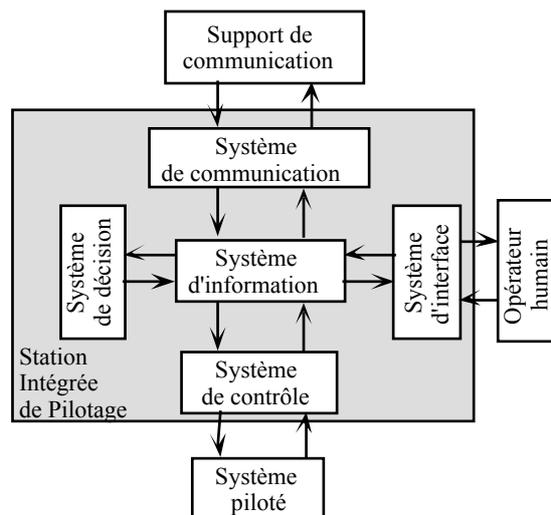


Figure 2: structure d'une SIP

Une SIP particulière, SIP d'entrée, est chargée d'effectuer l'allocation de la première opération de chacun des ordres de fabrication. Elle constitue donc le point d'entrée du système d'agents. Sa fonction n'est donc pas uniquement liée au processus de fabrication, mais aussi à l'intégration initiale des ordres de fabrication au sein de ce processus.

3.2. Protocole de communication

Six primitives de communication ont été développées afin de décrire de manière générique les étapes de tout protocole de communication. Nous allons détailler leur utilisation. Le séquençement des différentes actions menant à un contrat d'échange de données (service) entre un émetteur (demandeur de service) et un récepteur est décrit par les cinq premières primitives. Dans le cas où, suite à l'acceptation du contrat d'échange, les données transmises ne sont pas suffisantes, on peut faire appel à la primitive de transfert. Nous les détaillons ci-dessous:

DEMANDE : une demande est émise lorsqu'il est nécessaire d'obtenir un service de la part de la ressource appelée.

ACCEPTANCE : une acceptation stipule un accord potentiel de service. Le récepteur peut ou non recevoir de demande de service effectif (service activé mais non actif).

RESERVATION : une réservation correspond à une exclusivité de service. Ce dernier est alors actif.

RELAXATION : une relaxation correspond à l'annulation de la demande.

ACQUITTEMENT : un acquittement est émis pour terminer un service.

A ces cinq primitives nous associons une sixième particulière:

TRANSFERT: cette primitive est utilisée pour assurer le transit de toute information indépendamment des cinq primitives ci-dessus. Elle est utilisée lorsqu'un service est actif (une réservation a été envoyée) ou est activé (une

acceptance a été reçue) ou encore de manière totalement asynchrone (cas où l'émetteur ne cherche pas à vérifier la connexion). Elle peut être également associée au message de transfert physique de la pièce.

3.3. Processus d'allocation

Tchako a spécifié et justifié dans le cadre du pilotage temps-réel un agencement particulier de ces primitives afin de décrire un protocole d'allocation fondé sur la négociation (envoi direct de messages), cf. figure 3.

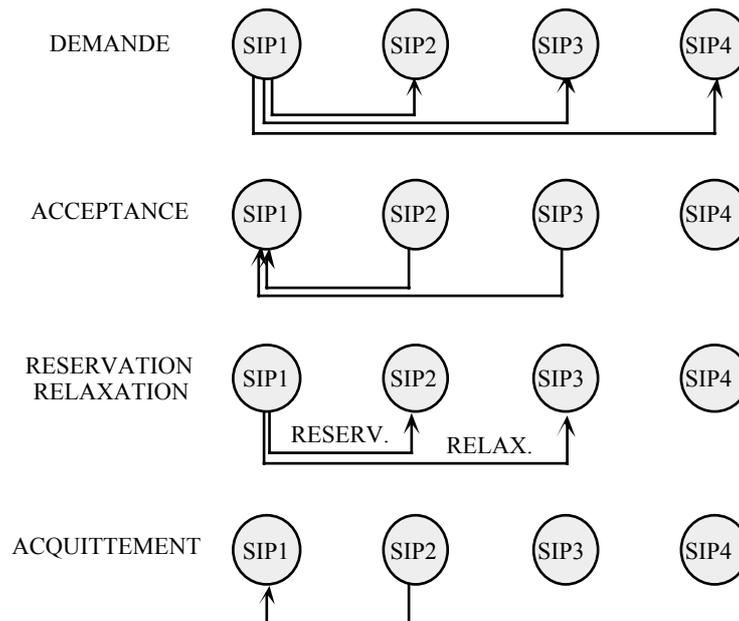


Figure 3: protocole d'allocation par négociation

Un modèle en réseau de Petri coloré temporisé a été élaboré [Trentesaux et al. 94a]. Le processus peut être résumé en quatre étapes:

Etape 1: Lorsqu'une opération sur un produit vient d'être effectuée par un agent, celui-ci vérifie s'il subsiste une opération à réaliser sur ce produit (ou lot de produits). Si oui, une demande est émise à l'ensemble des SIP connectées.

Etape 2: les SIP capables d'effectuer l'opération retournent à la SIP qui a émis la demande une acceptation ainsi qu'un ensemble de données portant sur ses caractéristiques propres (état des file d'attente, etc.). Elles se mettent alors en état d'attente de réservation/relaxation. Une SIP qui reçoit un message de type demande vérifie la concordance entre les contraintes imposées par l'opération et ses propres possibilités. Cette concordance est développée par exemple en terme de compatibilité "type d'opération réalisable par le système de production associé à la SIP" vs. "type d'opération demandée pour la tâche à réaliser".

Etape 3: La SIP émettrice sélectionne la SIP la mieux adaptée et lui envoie une réservation. Elle informe les autres SIP de leur non-sélection en leur envoyant une relaxation. La SIP émettrice se met alors en état d'attente d'acquittement.

L'ensemble des SIP candidates est évalué par un algorithme de recherche opérationnel capable de prendre en compte les règles classiques d'allocation. La fonction objectif peut être alors paramétrée selon des objectifs globaux de production.

Etape 4: la SIP sélectionnée retourne à la SIP émettrice un message de fin de communication (acquiescement) pour l'informer de la réception de la réservation et se place en état d'attente de transfert. Le transfert du produit et de l'ordre de fabrication entre les deux SIP peut avoir lieu.

Nous présentons succinctement les différents types de décision qui ont trait à la réalisation du processus global de fabrication.

3.4. Décisions portant sur la gestion des files d'attente

Un module d'optimisation monocritère (système de décision) assure la gestion locale des files d'attente. Cette fonction peut aussi être paramétrée pour tenir compte, par l'intermédiaire de coefficients de pondération, des différentes règles de décision (SPT, EDD, etc.).

3.5. Décisions portant sur la gestion de la reconfiguration

Cette gestion concerne principalement les ajouts ou les retraits de SIP, ce qui caractérise la notion de flexibilité structurelle.

Cette modification structurelle est automatiquement prise en compte par le groupe des agents. En effet, dans le cas où une SIP est ajoutée, elle sera intégrée au processus de fabrication dès qu'elle émettra sa première acceptation. Dans le cas contraire, où une SIP est retirée, elle ne pourra répondre aux demandes et sera donc naturellement écartée du processus de fabrication. Cette caractéristique fait appel à des procédures locales d'initialisation (réseau, machine, etc.), de déconnexion (réseau), de marche de clôture (procédures d'arrêt machine), etc.

3.6. Décisions portant sur la gestion des aléas

Plusieurs cas d'aléas sont possibles:

- panne d'une SIP ou d'un de ses sous-systèmes: dans ce cas, le processus de fabrication ne stoppe pas, mais la capacité globale est réduite. La panne est plus ou moins critique selon le degré de redondance de la SIP au niveau du système de production. Ainsi une SIP qui possède une spécificité opérationnelle (type d'opération particulier, précision élevée de l'outil, etc.) risque d'altérer de manière plus sensible le fonctionnement du système (voir provoquer un arrêt complet) qu'une SIP dont les particularités sont communes avec un ensemble d'autres SIP. Dans ce dernier cas, les SIP redondantes pourront réaliser les tâches de la SIP déficiente, évitant le blocage du processus de production,

- panne du réseau: la conséquence est l'arrêt à plus ou moins long terme du processus de fabrication entier.

Nous présentons dans la suite de ce chapitre la spécification associée d'une part à la structure de type distribuée et supervisée, et d'autre part à l'ordonnancement prévisionnel partiel de la ressource goulet.

4. Spécification de la structure de pilotage distribuée et supervisée

Nous détaillons dans cette partie les spécifications au niveau de chaque SIP et l'adjonction d'un niveau de supervision.

Le système est composé d'un ensemble de SIP coopératives sous contrôle d'un superviseur. Nous discernons deux types de coopérations (figure 4):

- "verticale" entre l'opérateur et l'agent. L'opérateur humain est responsable des décisions locales. Ce type de coopération est spécifié dans l'optique d'étudier l'influence sur l'auto-adaptativité et sur l'intégration des objectifs de production d'une éventuelle interaction via un système interactif d'aide multicritère à la décision. Ce point sera discuté dans le chapitre suivant.
- "horizontale" qui décrit l'interaction entre agents. C'est à ce niveau qu'intervient la structure de communication robuste. L'intégration d'une résistance aux dysfonctionnements impose d'étendre le protocole d'allocation de tâches de Tchako afin d'intégrer un niveau minimum de robustesse. Par exemple, si le système de communication d'une SIP tombe en panne alors qu'elle vient d'émettre une acceptation, il sera nécessaire dans le cas où une seule acceptation a été reçue par la SIP qui cherche à allouer une tâche, que cette dernière réitère sa demande au bout d'un certain temps.

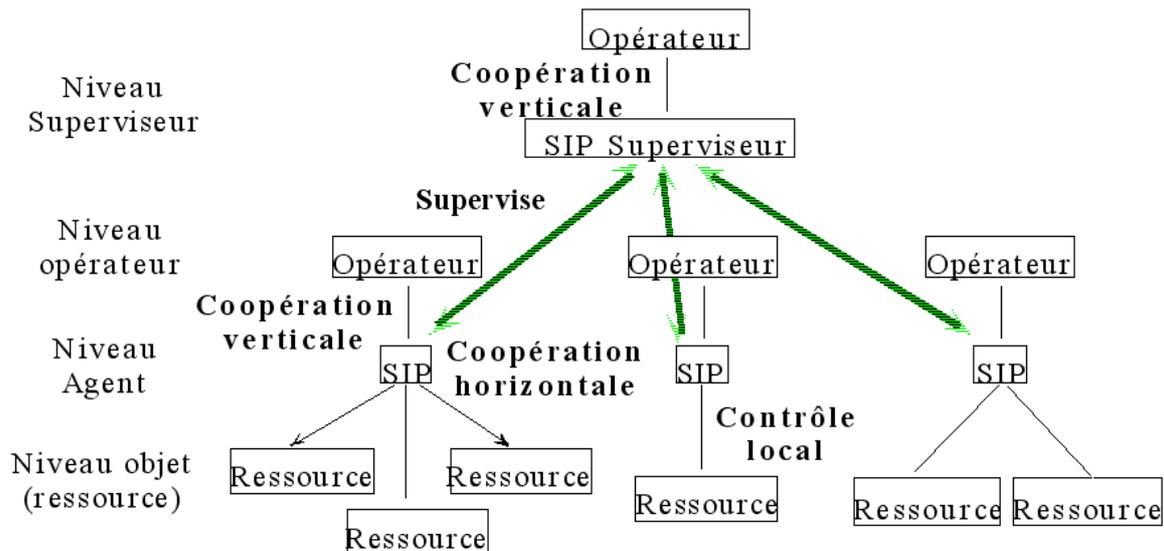


Figure 4: structure distribuée supervisée et types de coopération

La SIP supervision est associée de même à un opérateur. Elle peut contrôler directement un ensemble de ressources et participer ainsi au processus de production. Elle assure en outre les fonctions correspondant à celles de la SIP d'entrée.

Le deuxième aspect qui, concernant la nature du pilotage hybride, est celui de l'ordonnancement partiel développé dans la partie suivante et de l'influence de cette approche sur les fonctions de chaque SIP.

5. Spécification de l'ordonnancement prévisionnel partiel des tâches sur les ressources goulets

L'intégration d'un ordonnancement prévisionnel nécessite le développement de fonctions spécifiques au sein de chaque SIP composant le système (y compris le superviseur). Les spécifications portent sur:

- le développement d'un modèle de gestion dynamique de contraintes,
- la définition d'un protocole d'échange d'information entre le superviseur et chaque ressource,
- la définition de la charge potentielle maximale pour chaque ressource,
- la conception du processus d'ordonnancement partiel de la ressource goulet.

5.1. Modèle de gestion dynamique de contraintes

Le modèle de gestion dynamique de contraintes a pour objet:

- d'établir une liste de contraintes (qui porte sur la tâche à réaliser) à vérifier lors de l'émission d'une primitive de type demande,
- de vérifier si les caractéristiques d'une SIP sont compatibles avec les contraintes imposées par la tâche à réaliser lorsqu'une SIP reçoit une demande.

Ainsi, lors de la réception d'un message de type demande, il est possible de vérifier non seulement la compatibilité d'une tâche avec les caractéristiques de la ressource contrôlée, mais aussi de vérifier certaines contraintes (autres que la nature de la tâche en elle-même) imposées pour sa réalisation. Le message Demande concerne à la fois la détection de la charge potentielle et le processus d'allocation de tâches. Ces contraintes peuvent être déclarées dynamiquement (par l'opérateur par exemple), ces différents points seront détaillés dans le chapitre V.

Ainsi, il est possible d'intégrer à la réalisation d'une tâche des contraintes de type:

- flux: capacité de stockage, possibilité de transport, etc.
- ressource: taux d'utilisation recommandé pour une ressource, qualité du travail à fournir décrite en terme de précision de la machine, niveau d'expérience de l'opérateur, etc.
- produit: règle de précedence imposée par la conception ou par la spécificité du métier concerné, etc.

5.2. Suivi du processus de fabrication par le superviseur

Ce besoin est justifié par la nécessité de faciliter la gestion de la ressource goulet au niveau de la SIP Supervision par la spécification d'une fonction de suivi de production. Cette spécificité porte sur l'évolution des données globales et des ordres de fabrication (historique des allocations).

En même temps qu'une opération est allouée à un agent ou chargée ou déchargée sur une ressource de production, le superviseur doit être informé. Ce processus simplifie la tâche du superviseur qui, lors de la recherche de la SIP goulet focalise l'étude sur les opérations ni allouées et ni chargées.

5.3. Caractérisation du sous-ensemble des opérations planifiables

Le chapitre I a détaillé les différentes approches pour la caractérisation du sous-ensemble des opérations planifiables (ou ordonnançables). Nous avons explicité quatre critères de sélection qui portent sur les flux, la dimension temporelle, les produits ou les ressources. Nous proposons dans cette partie d'exploiter le critère "ressource" et plus particulièrement de planifier les opérations qui sont à réaliser sur une ressource présentant la capacité de production la plus critique par rapport à la demande. Notre démarche qui s'appuie sur la méthode de gestion des ressources goulets développée par Marris [Marris 94] se justifie par:

- la ressource dont la capacité de production est inférieure à celle qui est requise pour réaliser les ordres de fabrication impartis provoque des retards chroniques,

- la synchronisation des flux de production génère une meilleure fluidité et une simplification de la gestion. Cette synchronisation doit être imposée par une ressource sous-capacitaire,
- la planification d'une ressource sous-capacitaire pour synchroniser les flux permet d'anticiper de manière plus fine, sans être forcément exacte, les besoins en composants, matière première, ressource humaine, outillage, etc.

Ces éléments sont issus de la théorie élaborée par Goldratt et spécifiée par la méthode OPT [Goldratt et al. 84]. Cependant [Marris 94] et [Meleton 86] ont relevé de nombreux points sujets à critiques. Marris a proposé une approche beaucoup plus rationnelle et introduit les notions de vue duale et de systèmes de production déséquilibrés sur un horizon long terme.

Mais la réalisation d'un programme de fabrication implique nécessairement de prendre en compte le comportement dynamique du SAP. Par conséquent, notre démarche consiste à modifier l'approche de Marris en introduisant en particulier la notion de charge dynamique potentielle et en définissant la notion dynamique de "goulet probable".

5.4. Pilotage par la gestion prioritaire et dynamique des ressources goulets

Nous considérons que le goulet est détecté en fonction des charges de production, et ce, indépendamment de la méthode de gestion. Dans ce cas, l'objectif est de minimiser les contraintes de production induites par la présence de ce goulet, sans en général pouvoir toutefois le supprimer. Les données prises en compte pour réaliser l'ordonnancement du goulet sont issues de l'état courant et de la charge réelle du système de production.

5.4.1. OPT et la gestion par les contraintes

OPT est une méthode d'ordonnancement prévisionnel total à capacité finie [Lundrigan 86] développée par Goldratt dans une optique juste à temps. Le principe de base est développé autour de ce que Goldratt appelle le but de toute entreprise: maximiser le produit des ventes. Ce but peut être atteint notamment par la maximisation du flux des produits. En effet, c'est ce dernier qui détermine les ventes et seules les ventes permettent de maximiser le gain. Ainsi, Goldratt préconise de porter le maximum d'attention aux ressources qui ne peuvent absorber toute la demande. Ces ressources, appelées ressources goulets, doivent être identifiées et gérées de la manière la plus efficace possible. Pour ce faire, Goldratt préconise l'utilisation de neuf règles dont le but est de maximiser ce flux en se focalisant sur ces ressources [Goldratt et al. 84]. La règle principale est très

connue et conditionne fortement la méthode de gestion: une heure perdue sur une ressource goulet est une heure perdue pour tout le système.

Cette approche a été approfondie par Marris afin de la rendre compatible avec les approches japonaises basées sur la synchronisation, l'amélioration continue (le kaizen), etc. Ainsi, Marris montre qu'OPT n'est pas une méthode de gestion capable de suivre les évolutions technologiques ou encore les évolutions plus générales du marché [Marris 94]. Le management par les contraintes a donné naissance à une méthode proche d'OPT, mais plus robuste (méthode à ordonnancement prévisionnel partiel), plus réaliste, associant la notion de "meilleure situation de production" à celle provenant de la méthode OPT: "gestion par les contraintes de la situation actuelle". Cette approche est regroupée en deux phases:

- analyse de l'existant, : mise en place d'une gestion efficace par les contraintes. La question principale est: "comment gérer l'existant?",
- évolution de l'existant: modification des capacités ressources goulets dans un but d'amélioration. Les questions principales sont: "quelles devraient être les ressources-contraintes?" et "Comment évoluer vers ce nouvel état?".

En outre, Marris préconise un bouclage constant entre les deux phases.

Le thème abordé au sein de ce mémoire concerne uniquement la première phase: spécification, proposition et validation d'une structure de pilotage basée sur l'approche distribuée supervisée à ordonnancement partiel de la ressource goulet. Dans ce but, il nous semble nécessaire de décrire succinctement l'approche de Marris pour établir les différences par rapport à l'approche proposée.

5.4.2. Le management par les contraintes

Nous ne nous intéressons qu'à la phase 1 (gestion du déséquilibre existant). La phase 2 est décrite complètement dans [Marris 94].

5.4.2.1. *Le déséquilibre de la production*

La détection d'un goulet se fait de manière empirique: Marris ne décrit pas précisément une méthode de détection, mais suppose plutôt que l'entreprise sait localiser ses ressources goulets. En effet, les calculs des charges moyennes et des capacités de production moyennes pour tous les produits et toutes les ressources et ce, à long terme sont relativement difficiles à réaliser et souvent spécifiques à l'entreprise.

Le déséquilibre est induit par la présence simultanée de ressources capables de suivre la demande et de ressources goulets. Cette non-linéarité du processus de fabrication doit être exploitée afin de différencier la gestion des ressources goulet des ressources non-goulets. Dans le cas d'une ressource goulet, il est nécessaire de décrire au mieux (suivant une vision globale) le séquençement des tâches afin de

minimiser les risques de retard. En parallèle avec cette procédure il est nécessaire de mettre en place des méthodes de minimisation de temps de non-production et de réduction des tailles de lot par exemple (méthode SMED par exemple). Ces méthodes ne sont pas abordées dans ce mémoire.

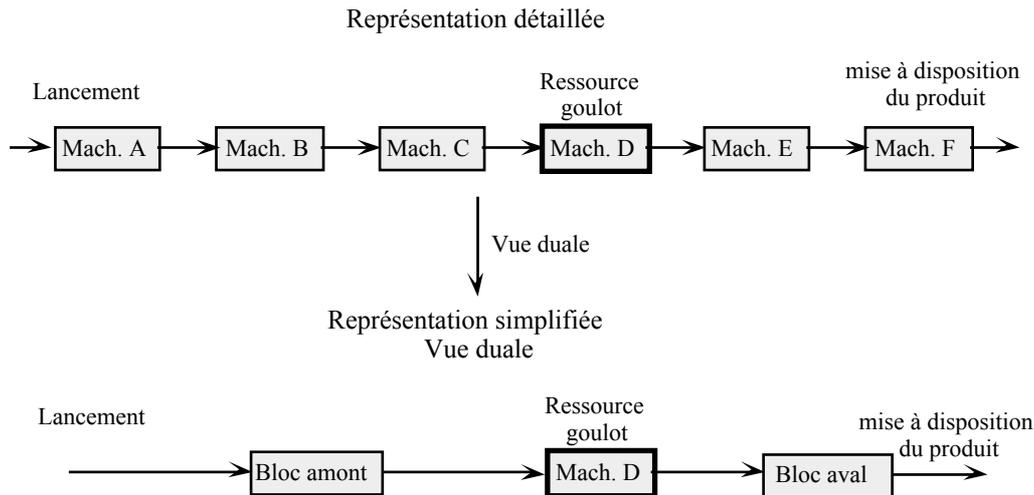
Dans le cas des ressources non-goulets, puisque les ressources sont sur-capacitaires, il n'est pas nécessaire d'imposer des contraintes de production très précises: ces ressources seront facilement en mesure de rattraper un quelconque retard.

L'étude des cas d'aléas est succincte. Il est préconisé de laisser un stock de sécurité en aval des ressources goulet pour éviter de perdre un temps précieux sur une ressource goulet (cf. première règle d'OPT).

D'autre part, la présence potentielle de goulet dynamique (en cas de surcharge locale) est peu abordée, ce qui nuit également à l'intégration du traitement des perturbations au sein du système de pilotage.

5.4.2.2. *Synchronisation*

La synchronisation est induite par l'interaction entre la gestion des ressources goulets et des ressources non-goulets. Marris décrit la notion de "vue duale" qui permet d'obtenir une synchronisation efficace. La vue duale est une méthode qui a pour but de simplifier l'apparence du problème en concentrant l'étude sur les flux de produits schématisés sur les ressources. La figure 5 montre un exemple simple (ligne de production) d'application de la vue duale. Ces flux sont étudiés en détail pour chaque ressource goulet. Chacune d'entre-elles est tout d'abord ordonnancée en juste à temps selon les marges de liberté possible en fonction des dates de besoins. Ces marges, élaborées à partir de la gamme opératoire et des ordres de fabrication, représentent la différence de temps entre la date de besoin de chaque produit et le temps opératoire total après passage sur la ressource goulet. Les opérations sur chaque produit sont planifiées sur cette ressource par marge décroissante. Ceci est réalisé pour chaque ressource goulet. Il est alors possible d'établir les besoins en composants et matériels en étudiant l'amont de chacune des ressources. La planification en besoins est alors évaluée.



Cette synchronisation par vue duale décrit donc totalement l'ordonnement des ressources goulets et succinctement celles des autres ressources.

Le problème est que cette méthode est efficace lorsqu'un, voire deux goulets, sont présents. Si plusieurs ressources ont une capacité inférieure à celle qui est requise, l'ordonnement de ces ressources devient très contraignant et risque de créer d'autres ressources goulets dynamiques (c'est à dire des goulets fugitifs de durée limitée dans le temps) pour respecter les dates de besoins des produits finis. Le cas idéal d'application de cette méthode est une ligne de production linéaire dont le nombre de ressources goulets reste limité (au maximum deux). Cette méthode semble moins adaptée à des structures de production plus complexes.

5.4.2.3. Conclusion

La méthode préconisée par Marris pour la gestion de la phase 1 est une méthode à ordonnancement prévisionnel partiel qui introduit la notion de vue duale. Le but de cette approche est de proposer une méthode de synchronisation des flux de production sur une ou plusieurs ressources goulets. La détection d'un goulet est fondée sur une étude très empirique à long terme du système de production pour faire apparaître la notion de charge moyenne. Ainsi, cette méthode est très peu efficace dans le cas où la présence de perturbations (panne, ordres urgents, etc.) provoque la création de goulets dynamiques. Dans ce cas, si la(les) ressource(s) possède(ent) une(des) capacité(s) de production faiblement inférieure(s) aux besoins exprimés, des goulets dits "baladeurs" sont créés. Ils réduisent fortement la fluidité des flux de produits.

D'autre part, la présence de goulets multiples nuit à l'ordonnement de ces ressources goulets. La complexité croît exponentiellement avec le nombre de goulets. Marris préconise alors de passer en phase 2 afin de modifier le système de production dans le but de ne posséder qu'un ou deux goulet(s).

La méthode développée dans ce mémoire pour l'ordonnancement prévisionnel partiel se base sur celle de Marris, mais augmente considérablement la dynamique de la structure de pilotage en ne contrôlant de manière précise que la ressource goulet la plus contraignante selon le critère de la charge potentielle (ce critère sera défini dans la partie suivante). Pour cela il est nécessaire:

- de redéfinir la notion de ressource goulet et en y intégrant une caractéristique temps-réel,
- de détecter dynamiquement la ressource la plus contraignante,
- de n'ordonner que cette ressource afin d'éviter la multiplicité et la complexité des ordonnancements.

5.5. Détection et ordonnancement du goulet

La détection et l'ordonnancement d'une ressource goulet constitue une démarche événementielle déclenchée par exemple par l'apparition de nouveaux ordres de fabrication.

De cette manière, le sous-ensemble des opérations planifiées correspond aux opérations qui doivent être réalisées sur la ressource goulet durant la période considérée.

5.5.1. Charge potentielle maximale

La charge potentielle maximale est définie pour un ensemble de tâches à effectuer et qui ne sont pas commencées à une date donnée. Le superviseur évalue en premier lieu les tâches qui peuvent être réalisées par chaque ressource en fonction des ordres de fabrication et des gammes opératoires (il doit tenir compte pour chaque ressource des opérations déjà allouées et présentes dans le stock amont de celles-ci). Il évalue ainsi le temps nécessaire à chaque ressource pour réaliser l'ensemble des tâches qui peuvent lui être allouées (et automatiquement celles qui lui sont déjà allouées). Les différents temps caractéristiques du processus de fabrication sont pris en compte (temps opératoires, temps de préparation, temps de changement outil, temps de réglage, etc.). C'est en ce sens que la charge est maximale. Elle est potentielle dans le sens où une opération qui peut être réalisée sous le contrôle d'une SIP ne sera pas nécessairement effectuée sous le contrôle de cette SIP.

Le résultat correspond donc au temps maximum que pourrait passer une ressource sur un ensemble de tâches encore non réalisées. Si une opération est en cours sous le contrôle d'une SIP particulière, alors les temps opératoires ne sont pris en compte que pour cette SIP.

La ressource la plus contraignante a priori est la ressource qui présente la charge potentielle maximale. Cette ressource constitue la ressource goulet jusqu'au

prochain événement qui provoquera la ré-évaluation des charges et l'identification de la ressource goulet.

5.5.2. Ordonnancement de la ressource goulet

Il existe plusieurs méthodes pour l'ordonnancement d'une ressource unique. La performance de l'ordonnancement dépend surtout de la capacité du superviseur à établir un ordonnancement qui tienne compte à la fois des objectifs globaux, des contraintes temps réel représentées par la file d'attente des SIP et des tâches en cours sur chacune d'elles. Ceci conditionne la faisabilité de la réalisation des tâches par la ressource goulet aux dates voulues.

Il est important de noter que l'ordonnancement préconisé doit être considéré non pas comme une contrainte forte, mais plutôt comme un guide qui peut n'être que partiellement respecté. Supposons par exemple, qu'une panne survienne sur la ressource amont à la ressource goulet, alors que cette dernière attend le produit qui était en cours de traitement sur la SIP déficiente. Dans ce cas, si la panne est trop importante, il est possible pour l'opérateur de forcer la réalisation d'une autre tâche sur la ressource goulet avant la date de début de fabrication recommandée afin de ne pas altérer le flux sur cette ressource goulet. En effet le goulet doit être exploité au maximum afin d'éviter une perte d'efficacité trop importante.

5.6. Généralisation de l'approche à une famille de SIP

Cette généralisation permet d'exploiter l'existence d'un ensemble de ressources capables d'effectuer des opérations identiques (familles de ressources). Ces ressources sont regroupées au sein d'un même ensemble appelé famille. La charge potentielle concerne alors la famille, et elle doit être répartie sur chaque ressource de cette famille. Le superviseur peut être confronté aux deux cas suivants:

- une ressource est déclarée goulet (famille réduite à un élément). La méthode présentée dans la partie précédente peut alors s'appliquer,
- une famille est déclarée goulet (après répartition, la charge potentielle est toujours la plus élevée). Il est alors nécessaire d'ordonnancer l'ensemble des ressources composant la famille.

Dans la suite du mémoire nous avons fait les hypothèses suivantes:

- le goulet est une ressource non redondante (ce qui ne sous-entend pas que le système de production ne puisse pas comporter de familles),
- la charge potentielle affectée à chaque ressource de la famille est le rapport de la charge potentielle totale au nombre de ressources de la famille.

Le problème principal est celui où le goulet réel n'est pas forcément celui qui a été déclaré par le superviseur. Ce problème peut survenir par exemple lorsque

plusieurs ressources ont des charges théoriques élevées et très proches, sans que l'on tienne compte des probabilités d'occurrence de pannes. Puisque l'ordonnancement ne concerne qu'une ressource parmi cet ensemble, les autres ressources risquent de perturber cet ordonnancement. Dans ce cas, une attention particulière doit être accordée à ces ressources (ressources "quasi-goulets"). Le rôle de l'opérateur est alors primordial en ce qui concerne l'allocation des tâches et la gestion de la file d'attente lorsqu'il perçoit que la ressource sous son contrôle est quasi-goulet. Ce qui justifie une fois de plus la nécessité et l'intérêt d'un système interactif d'aide à la décision.

La partie suivante spécifie les objectifs de la structure de pilotage hybride en terme de réactivité face aux aléas et en terme de flexibilité.

6. Capacité d'auto-adaptation du pilotage

Nous définissons l'auto-adaptativité d'un système de pilotage comme la capacité à pouvoir intégrer au mieux l'aspect dynamique du processus de fabrication. Cette notion définie dans [Trentesaux et al. 95a] se caractérise par:

- la réactivité ou la capacité à supporter les événement perturbateurs imprévus. Ce sont typiquement les perturbations aléatoires telles que les pannes, les ordres urgents, etc. Cette caractéristique peut être associée à la notion de régulation dans le domaine de l'automatique,
- la flexibilité ou la capacité à supporter les événements perturbateurs prévus. Ce sont des perturbations du type modification de la structure de production (ajout/retrait SIP, etc.), des processus de fabrication (modification de la gamme opératoire, etc.), des caractéristiques ressources (temps de changement outil, taux de production).

On peut considérer la réactivité et la flexibilité selon les trois points de vue ressource, produit et flux.

6.1. Réactivité

6.1.1. Point de vue ressource

La réactivité du point de vue ressource concerne principalement les pannes, c'est à dire:

- pannes SIP (sous-système, ressource contrôlé, etc.) au niveau distribué de la structure , que cette SIP soit ressource goulet ou ressource non-goulet,
- pannes SIP au niveau supervision,
- pannes du réseau de communication, etc.

L'occurrence et la gestion d'une panne d'une SIP au niveau distribué sont traitées de manière approfondie et unifiées pour toutes les ressources dans [Tchako 94]. La gestion est établie en fonction des sous-systèmes et les conséquences sont évaluées selon le niveau de redondance de la SIP au sein du système de production. Le protocole de communication doit être étendu afin d'intégrer ces dysfonctionnements potentiels. Une telle extension peut être réalisée par mémorisation des étapes validées durant le protocole de communication. Des temps d'attente associés à des états d'attente peuvent de cette manière être définis afin de réitérer certaines étapes non validées, et dans certains cas, la totalité d'entre-elles.

Cette approche reste entièrement valable pour une ressource non-goulet. Cependant, si la panne est suffisamment importante, le superviseur peut ré-évaluer l'ordonnancement prévisionnel.

Par ailleurs, la gestion d'une panne sur une ressource goulet ou un élément d'une famille goulet doit être améliorée. Dans les deux cas, lors de la remise en fonctionnement, le superviseur peut exprimer le besoin de ré-évaluer l'ordonnancement.

Le cas de la panne superviseur ne doit pas être nécessairement traitée comme une panne prioritaire. En effet, son rôle principal est l'ordonnancement et les allocations des opérations initiales de chaque nouvel ordre de fabrication. Hormis ces étapes où la ressource est effectivement critique, une panne n'empêche pas nécessairement le processus de fabrication. Il faut dans cette optique conserver un certain niveau d'autonomie du niveau distribué afin de permettre un fonctionnement totalement distribué en l'absence du superviseur pour les ordres de fabrication en cours.

La panne réseau, enfin, en raison de la nature partagée et non redondante de cette ressource est critique pour le système de production. Elle provoque l'isolation des SIP par rupture de tout échange d'information, et a fortiori, les allocations suivantes des tâches en cours seront impossibles à effectuer. Dans ce cas, après la phase de réinitialisation, il sera nécessaire de ré-évaluer l'ordonnancement prévisionnel.

6.1.2. Point de vue produit

La réactivité du point de vue produit concerne les aléas relatifs aux caractéristiques produit. Ces aléas résultent notamment de la modification des gammes opératoires ou des ordres de fabrication qui peuvent être modifiés par le client ou reconsidérés localement.

Le traitement de ces aléas est facilité par la gestion dynamique des événements. Le besoin de ré-ordonnement est guidé par la reconfiguration des gammes opératoires.

6.1.3. Point de vue flux

L'ensemble des événements concernés regroupe notamment les aléas suivants:

- modification des contraintes portant sur les différents ordres de fabrication (intégration au niveau de la structure de nouveaux ordres urgents de fabrication, renégociation des dates de livraison, etc.),
- rupture de(s) stock(s) matière première (fournisseurs) et produits semi-finis (sous-traitants).

Ce type d'aléa est intégré par le niveau pilotage de la même manière que les aléas de type produit. En effet, le premier cas nécessite automatiquement la remise en cause de l'ordonnement partiel afin d'intégrer le nouvel ensemble des ordres de fabrication ou les nouvelles dates de livraison par exemple.

Dans le second cas, un ré-ordonnement est possible en fonction de l'importance de la rupture. Il existe cependant des méthodes permettant d'éviter les ruptures de stock, et ce, spécialement pour les ressources de type goulet. Ces méthodes sont détaillées dans [Marris 94].

6.2. Flexibilité

Dans cette partie, nous étudions les potentialités de la structure supervisée en terme de flexibilité. Les événements concernés sont dus à l'intégration de nouvelles caractéristiques physiques ou à la remise en question des caractéristiques existantes.

6.2.1. Point de vue ressource

Les événements considérés sont:

- la modification de la structure de production (ajout/retrait d'une SIP),
- la modification d'une caractéristique d'une SIP.

L'ajout ou le retrait d'une SIP concerne à la fois le niveau distribué de la structure (les caractéristiques sont identiques à celles présentées dans l'étude de la structure de Tchako) et le niveau supervision. Nous détaillons les conséquences au niveau supervision.

L'ajout d'une ressource augmente le potentiel de production d'une famille. Dans le cas où cette famille est goulet, la ressource ajoutée vient appuyer la famille goulet et l'amélioration concerne l'ensemble du système de production. Dans ce cas, il est nécessaire de ré-évaluer les charges afin de déterminer le nouveau goulet ou la nouvelle famille goulet. Dans le cas contraire, cette ressource vient

augmenter les capacités d'une famille déjà sur-capacitaire (car non-goulet) et le ré-ordonnancement n'est pas nécessaire. L'amélioration de la capacité de production d'une famille déjà sur capacitaire permet par exemple d'absorber une perturbation locale ou une surcharge temporelle.

Dans le cas d'un retrait d'une ressource, plusieurs cas sont possibles (nous supposons qu'un tel retrait ne modifie en rien la diversité des opérations réalisables):

- la ressource supprimée fait partie de la famille goulet. Dans ce cas, il est impératif de ré-évaluer les charges potentielles en sachant qu'un nouvel ordonnancement sera nettement moins performant. Ce type de retrait est évidemment critique pour les performances globales.
- la ressource supprimée fait partie d'une famille sur-capacitaire. Ce retrait peut avoir deux conséquences. Dans le premier cas, la famille reste sur-capacitaire et après évaluation des nouvelles charges potentielles, l'ordonnancement ne sera pas réévalué. Dans le second cas, le retrait rend la famille goulet, le ré-ordonnancement est alors impératif.

Le deuxième type d'événement est la modification d'une caractéristique d'une ressource. Cette modification est guidée par un besoin de modification des possibilités techniques ou de la capacité. Si l'évolution restreint ou accroît les caractéristiques de production, il sera nécessaire de ré-évaluer les charges potentielles pour un éventuel ré-ordonnancement. Le système d'information doit alors intégrer ces modifications et la SIP tiendra automatiquement compte de ces modifications lorsqu'elle retournera sa première acceptance.

6.2.2. Point de vue produit

Les produits sont susceptibles d'être modifiés par le bureau d'études par exemple. Si la structure de production est capable de réaliser la nouvelle gamme opératoire, le superviseur aura à gérer un nouvel ordre de fabrication (associé à une nouvelle gamme opératoire). Il considérera ce dernier de la même manière que l'ensemble des ordres reçus auparavant. La procédure de ré-ordonnancement sera naturellement réalisée. Une telle modification est donc transparente pour le niveau pilotage.

6.2.3. Point de vue flux

L'ensemble des événements touche la réorganisation du flux de production et les objectifs globaux. Notons en particulier:

- la gestion des stocks: réduction des en-cours, prise en compte des contraintes de stockage,

- la synchronisation des flux: modification des cadences, des temps de changement outil, des lots de transformation, des lots de transfert, etc.

La gestion globalisée des stocks est permise par l'intégration au niveau du superviseur d'un ordonnancement prévisionnel. Il est ainsi possible d'ordonnancer la ressource ou la famille goulet en vue de respecter des objectifs de production décrits en terme de flux (cadence, etc.).

D'autre part, il est envisageable d'accroître la responsabilité du superviseur en intégrant des méthodes qui lui permettent d'imposer ou de modifier les critères d'allocation, de gestion des files d'attente, etc. de chaque SIP afin d'intégrer de manière plus efficace un point de vue global. Cette méthode est exploitée par [Kouiss et al. 95] mais suivant une méthode monocritère.

7. Conclusion

La structure distribuée supervisée à ordonnancement partiel qui a été décrite dans ce chapitre doit apporter de nombreux avantages en terme d'intégration d'objectifs globaux et en terme d'anticipation des besoins en matériels, composants et ressources.

Ce chapitre a montré comment les objectifs de réactivité, flexibilité, efficacité de l'ordonnancement, etc. peuvent être intégrés par l'approche proposée dans ce mémoire. Ainsi, nous avons spécifié une structure de pilotage particulière qui tient compte de contraintes telles que les évolutions des objectifs et des technologies, l'amélioration de la résistance aux perturbations, etc.

Le chapitre suivant met en évidence la faible structuration des processus décisionnels, ce qui nous conduit à exprimer et spécifier le besoin d'intégration de l'interactivité avec l'opérateur humain dans le processus de fabrication par un système interactif d'aide à la décision qui tienne compte des contraintes imposées par l'intégration d'un pilotage hybride.

En outre, nous mettons en évidence l'aspect multicritère des décisions associées au niveau du pilotage hybride, ce qui nous pousse à évaluer dans la suite de ce mémoire la possibilité d'une approche de résolution multicritère dans le but d'améliorer l'intégration des objectifs.

Chapitre III

Spécification d'un système interactif d'aide multicritère à la décision distribué

1. Introduction

Dans ce chapitre nous montrons en premier lieu que la plupart des processus de décisions mis en oeuvre au niveau du pilotage se caractérisent par les propriétés développées par [Simon 77] et portant sur la "rationalité limitée". Cette caractérisation peut impliquer le développement d'outils spécifiques pour un pilotage efficace du processus de fabrication afin d'accroître les performances de la structure pilotage. Nous justifions alors la possibilité d'intégration d'un système interactif d'aide multicritère à la décision (SIAMD) distribué au sein de la structure hybride.

Ainsi, après avoir présenté les concepts et les problématiques des approches SIAD et multicritère, nous montrons comment il est possible de développer un SIAMD distribué pour le pilotage hybride en spécifiant une approche interactive et multicritère qui tient compte des contraintes associées à la structure de pilotage introduite dans ce mémoire.

2. Processus de décision

La prise de décision peut difficilement être définie indépendamment de l'environnement dans lequel elle est développée et la décision est elle-même indissociable du contexte global du processus de prise de décision. C'est pourquoi nous nous attacherons particulièrement à l'étude du processus de décision dans sa globalité. Cette optique permet en outre de mettre en évidence l'importance relative des différentes étapes composant ce processus.

De manière générale, un processus de décision consiste à restreindre un ensemble de possibilités à un sous-ensemble strict et à évaluer cette restriction. Simon [Simon 77] détaille ce processus en quatre étapes, non nécessairement séquentielles, cf. figure 1.

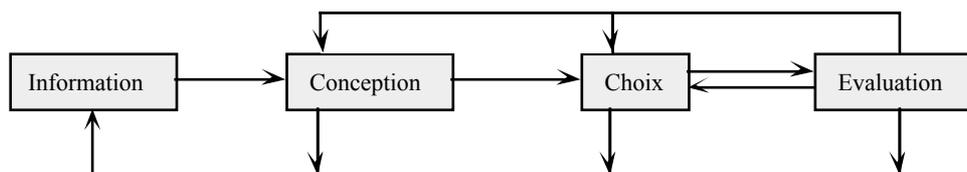


Figure 1: processus de décision selon [Simon 77]

- information: cette étape, avec la suivante, conditionne fortement la qualité de la prise de décision. En effet, elle détermine l'ensemble des données nécessaires (mais pas forcément suffisantes) qui seront utilisées lors des phases suivantes,
- conception: cette phase génère les différentes alternatives qui forment l'ensemble des possibilités. Les différentes solutions sont donc élaborées à ce stade,
- choix: cette phase constitue la phase de décision proprement dite. Elle consiste à restreindre l'ensemble des possibilités au sous-ensemble des possibilités sélectionnées,
- évaluation: cette phase a pour objet d'évaluer la qualité de la prise de décision et peut impliquer si nécessaire un retour à l'une des phases précédentes.

L'acte de décider ne constitue ainsi qu'une partie, indissociable, du processus complet de décision.

3. Processus décisionnel au niveau pilotage

La plupart des approches qui ont trait à une prise de décision sont fondées sur les hypothèses suivantes dites de "stabilité":

- le traitement se fait sur des données stables: les données sont indépendantes du temps, des facteurs externes et des intervenants. C'est le cas si par exemple les différents temps de changement d'outils, de production, les délais d'approvisionnement sont supposés constants. Un autre exemple est celui où l'ensemble des importances relatives des différents critères est supposée être identique pour chaque intervenant lors d'une prise de décision,
- le problème est techniquement stable: il est clairement défini et indépendant des points de vue des intervenants. Par exemple, l'ordonnancement de la production est calculé par optimisation d'une fonction de coût, ou la maximisation de l'utilisation d'une ressource se fait par la minimisation d'une fonction du temps de non-production.

A ces hypothèses de stabilité, s'ajoutent les hypothèses suivantes [Lévine et al. 89]:

- toutes les actions possibles sont identifiées avant le traitement,

- il existe un pré-ordre total sur les actions. Ce pré-ordre s'exprime par la présence d'une fonction d'utilité,
- les intrants (paramètres, données) sont numériques et contiennent toute l'information utile,
- la meilleure décision est celle qui maximise la fonction d'utilité.

Ces hypothèses ont été critiquées par Simon [Simon 77] qui a montré que le processus de décision est parfois plus complexe. Nous présentons ce principe de "rationalité limitée" dans le cadre des processus décisionnels au niveau pilotage.

3.1. Phase d'information

L'approche classique implique que les intrants soient stables, numériques et qu'ils portent toute l'information utile. Malheureusement, dans le cas qui nous concerne, ces hypothèses de stabilité et de numérisation sont peu réalistes.

3.1.1. Stabilité de l'information

La stabilité de l'information est réduite en particulier par la fiabilité difficilement maîtrisable des informations au niveau du pilotage (a fortiori, pilotage distribué). En effet, dans un environnement dynamique, il est préjudiciable de travailler sans une connaissance précise de la qualité de l'information en terme de fiabilité: précision, obsolescence, redondance [Cauffriez 94]. Par exemple, il est possible d'avoir à manipuler une donnée du type "temps d'attente en stock d'entrée ressource". Dans ce cas, il est utile de pouvoir si possible associer un niveau de fiabilité en fonction du temps écoulé depuis la production de cette information.

3.1.2. Accessibilité de l'information

L'information pertinente à la prise de décision peut être difficile à identifier et à localiser. Les notions d'accessibilité et de coût de l'information sont alors à prendre en compte:

- l'accessibilité de l'information: dans un environnement de production, l'information est partagée parmi les ressources. La recherche de l'information parmi celles-ci est donc fondamentale,
- le coût de l'information: la recherche et le stockage de l'information sont coûteux. Le développement de bases de données nécessite de nombreux investissements financiers et en personnel. Par ailleurs, la communication pour l'échange d'informations est également coûteuse. Un compromis doit donc être établi entre la quantité d'information désirée par rapport aux investissements qui lui sont consacrés. Deux solutions sont possibles: décider de ne pas tenir compte de l'information manquante, ou utiliser des méthodes d'estimation pour reconstruire l'information manquante. Dans les deux cas, la qualité de l'information est réduite.

Il semble donc qu'une première difficulté résulte de l'aspect dynamique de la structure de pilotage qui doit tenir compte du manque d'information et de la fiabilité des données.

3.1.3. Fiabilité de l'information

Il nous semble primordial d'étudier plus particulièrement le problème de la fiabilité des données en milieu de production. Les deux aspects coût et accessibilité étant relativement dépendants, nous supposerons, pour la clarté de la discussion, que l'information est toujours disponible pour un coût spécifique (ce qui se traduit par exemple, par un investissement en capteur, en études diverses, etc.).

Les données disponibles lors d'une prise de décision proviennent des différentes fonctions industrielles [Dindeleux 92]:

- gestion des ressources humaines, gestion du personnel: disponibilité opérateur, etc.
- gestion commerciale: délais de livraison client, commandes urgentes, etc.
- gestion prévisionnelle (niveau 1): ordonnancement, programme de production, etc.
- approvisionnements, stocks: quantité disponible, etc.
- méthodes: caractéristiques produit, etc.
- achat en matériel: taille de lot, caractéristiques machine, etc.
- comptabilité, gestion: coûts analytiques, etc.
- suivi de production: indicateurs, ratios, etc.

Les données fournies par les six premières fonctions sont suffisamment fiables pour permettre une qualité de décision suffisante. Or, une partie non négligeable des informations provient des deux dernières fonctions. Malheureusement les outils et les méthodes qui ont pour objectif de fournir un ensemble de données exploitables possèdent un faible niveau de qualité qui réduit sensiblement la qualité des décisions. Par exemple, le coût de production est un critère important. Si ce coût n'est pas estimé correctement, les décisions prises risquent d'altérer l'efficacité du pilotage.

De ce fait, il nous semble nécessaire d'aborder certains points portant sur la comptabilité analytique et l'exploitation d'indicateurs. Nous ne considérons la comptabilité analytique que sous son aspect "outil d'élaboration de données" indépendamment de son utilisation comme outil de diagnostic.

3.1.3.1. Modélisation succincte du système de production

L'approche par indicateurs [Giard 88], [Ancelin 89] et les méthodes de comptabilité analytique s'appuient sur des relations relativement arbitraires, correspondant à un modèle simplifié de la structure de production et du processus de fabrication [Jacot 91]. De ce fait, ils ne prennent en compte que partiellement les relations dites de "cause à effet" et par conséquent ne peuvent fournir que des

estimations très peu précises. Un exemple est celui des coûts de production dont les estimations sont très sensibles au choix de la méthode utilisée.

3.1.3.2. Inadéquation de l'approche analytique

La comptabilité analytique est de plus en plus fortement critiquée [Goldratt et al. 84], [Giard 91], [Jacot 91]. Cette critique porte notamment sur l'inadéquation entre les méthodes de gestion actuelles et les outils de la comptabilité analytique. En effet, l'application des méthodes analytiques requiert au moins deux hypothèses très restrictives:

- une grande inertie (faible variation) des variables d'état. Ce que traduit le principe de stabilité des systèmes de production [Mévellec 93]. Mais la dynamique actuelle des moyens de production va à l'encontre de cette hypothèse. Ainsi, si cette approche est efficace à un niveau élevé d'agrégation comme indicateur de tendance à long terme (niveau stratégique), elle n'est donc pas adaptée aux niveaux tactique et opérationnel de la production.
- les charges directes constituent l'élément prépondérant des charges globales. Par conséquent l'influence des coûts indirects est minimisée (ils sont estimés à partir des charges directes) et les calculs peuvent alors être simplifiés. En effet, les charges directes (par définition) ne sont rattachées qu'au volume de production [Mévellec 93], [Giard 91], introduisant de cette manière une proportionnalité plus facile à cerner. Cette hypothèse est de moins en moins valable car les charges indirectes sont aujourd'hui de plus en plus importantes et faussent les estimations des charges globales [Mévellec 93].

3.1.3.3. Une approche par les activités pour les coûts de production

Les méthodes de calcul de coût par les activités (Activity Based Costing) et plus généralement de management par les activités (Activity Based Management) cherchent entre autres à identifier les activités correspondant à chaque étape de production. Une seconde étape associe à chaque activité un coût [Mévellec 93]. Une hiérarchisation est alors possible sous forme d'agrégation (par regroupement par processus -enchaînement d'activités dépendantes- ou par inducteur de coûts -regroupement d'activités dont le coût dépend d'un même facteur caractéristique-). Dans notre cas, les données brutes suffisent en tant que telles (cf. problème de l'agrégation).

Un des avantages de cette approche est de déterminer de manière très précise une logique de coût commune qui permet d'accroître la cohérence relative des données. De cette manière, des études comparatives sont possibles [Sénéchal et al. 94] et l'intégration de ces données au sein d'un processus décisionnel multicritère est pertinente en terme de qualité de l'information.

3.1.3.4. Ne pas assimiler variable d'état et variable d'action

L'utilisation de données attachées au "temps de cycle produit", au "niveau de stock" ou au "taux d'occupation" permet de constituer une image de l'état du système pour une structure de contrôle particulière. Ce sont des variables d'états (indicateurs) et leur utilisation en tant que donnée pour une prise de décision implique la création d'un bouclage cause - effet qui peut mener à une instabilité à plus ou moins court terme (c'est le cas des logiciels d'ordonnancement qui se basent sur les temps de cycle produit moyen pour établir les dates deancements). Seules les variables d'action doivent être utilisées. Elles sont définies en chaque point de pilotage du processus de production et sont spécifiques à chaque système de production.

3.2. Phase de conception

La méthode de conception est fortement liée à la typologie des actions envisageables. Si l'étude porte sur un ensemble discret à faible cardinal (par exemple, sélection d'une SIP parmi un ensemble), il est envisageable d'explicitier l'ensemble des alternatives. Dans ce cas, l'hypothèse d'exhaustivité des choix est respectée.

Cependant si le domaine d'étude porte sur des alternatives qui sont de nature continue (ordonnancement partiel d'une ressource par exemple), alors il est impossible d'explicitier l'ensemble des solutions, ce qui nuit à la compréhension globale des différents choix possibles.

3.3. Phase de choix

Dans un contexte de pilotage, les problèmes à résoudre ne sont pas techniquement stables. Le décideur est confronté à la multiplicité des points de vue entre responsables de production, magasin, chefs d'atelier, bureau des méthodes, etc.

D'autre part, le problème à résoudre n'est pas toujours clairement défini: le choix d'une re-allocation entre machines par exemple se base sur l'étude des aléas survenus: durée estimée, fréquence, conséquence de la ré-allocation, etc.

3.4. Phase d'évaluation

Comment pouvons-nous évaluer a priori ou en cours de la réalisation du processus de décision l'impact d'une décision dans un environnement dynamique (et a fortiori distribué)? La complexité d'une telle évaluation réduit fortement la possibilité d'évaluation des conséquences. C'est une approche souvent empirique qui, en milieu industriel, permet d'anticiper les conséquences. Ce problème peut être associé à la caractérisation chaotique des ateliers manufacturiers présentée dans [Parunak 91].

3.5. Intégration d'une coopération homme-machine

L'étude précédente a montré que, dans le cas du pilotage, il n'est pas possible d'intégrer toutes les données nécessaires pour prendre une décision satisfaisante. Il peut ainsi être parfois nécessaire de laisser l'opérateur entièrement libre de sa décision lorsque certaines informations ne peuvent pas être formalisées ou prises en considération par le système: préférence d'utilisation pour une ressource particulière, anticipation d'une commande, réaction face à un aléa, etc. L'intégration de l'utilisateur dans processus de décision doit permettre l'accroissement de la réactivité et du respect des objectifs.

Cependant, dans le cas où l'opérateur est intégré au processus de résolution, il faut tenir compte de contraintes relativement restrictives qui sont liées au contexte de production. Relevons principalement:

- les occurrences des tâches décisionnelles qui ne sont pas uniques,
- l'efficacité du processus de fabrication en terme de délai qui dépend du temps de réflexion nécessaire pour chaque processus de décision, et ce même si ceux-ci sont anticipés (pré-allocation, etc.).

Les conséquences sont les suivantes:

- l'opérateur doit en un temps réduit élaborer une décision, et ce, de manière répétitive pour certaines d'entre-elles (gestion des files d'attente, etc.). Il est donc nécessaire de mettre en place des mécanismes de traitement simples et efficaces comme le souligne [Pomerol et al. 93]: "la confiance du décideur se porte naturellement sur les méthodes simples",
- le temps nécessaire au déroulement du processus décisionnel doit être largement inférieur aux caractéristiques temporelles du processus de production (par exemple, temps opératoire moyen par ressource).

Il faut aussi garder à l'esprit que, dans le cas d'un atelier flexible de production, les marges portant sur les délais sont la plupart du temps réduites et le temps moyen entre deux allocations est faible. Ce qui conditionne fortement la méthode de résolution du processus décisionnel dans le cas où l'opérateur humain est pris en compte.

3.6. Vers un système interactif d'aide multicritère à la décision

Nous montrons dans cette partie que l'utilisation d'un système interactif d'aide multicritère à la décision permet l'intégration de l'opérateur humain dans les processus décisionnels au niveau du pilotage.

3.6.1. Typologie des processus décisionnels

Afin de caractériser les processus de décision associés au pilotage nous utilisons la classification développée initialement par Simon et reprise par [Lévine et al. 89]. Elle comporte trois grandes catégories:

- (processus de) décision structuré: le problème est clairement posé en termes techniques, les données sont fiables et numériques, etc. Les modèles associés sont clairement définis. Ils permettent une résolution fiable et algorithmique du processus décisionnel. C'est le cas d'une facturation, d'un achat, etc.
- (processus de) décision peu structuré: le problème peut ne pas être clairement posé, les données sont souvent qualitatives, peu fiables, très peu stables, difficilement accessibles, etc. La résolution est difficilement exprimable sous forme algorithmique. Cette catégorie est constituée de la quasi-totalité des processus de décision dans le cas du pilotage,
- (processus de) décision non structuré: le problème n'est pas clairement posé, le principe de rationalité limitée s'applique à toutes les étapes du processus décisionnel. La décision élaborée est difficilement justifiable de manière rationnelle. C'est le cas par exemple d'une réaction de l'opérateur face à un aléa qui n'a jamais été envisagé.

Selon cette caractérisation, le processus de décision associé à une structure de pilotage à ordonnancement prévisionnel est:

- souvent indissociable de l'opérateur humain car il est seul capable d'intégrer la diversité et la complexité des données dans un cadre décisionnel réactif ou global,
- naturellement peu ou pas structuré, ce qui réduit fortement l'intérêt d'exploiter un ensemble de méthodes purement algorithmiques.

Il est donc justifié d'adopter une démarche décisionnelle coopérative associant l'opérateur humain et un système interactif d'aide à la décision.

Nous montrons en outre dans la partie suivant qu'il est raisonnable de considérer le cadre de décision au niveau du pilotage selon une optique multicritère.

3.6.2. Processus décisionnels multicritères

Les algorithmes d'optimisation imposent d'établir des hypothèses qui réduisent fortement l'adéquation du modèle de décision à la réalité. En effet:

- les calculs sont très complexes (optimisation portant sur de nombreuses variables de nature différente: continue et discrète), et sont d'autant peu efficaces que les données sont réduites et simplifiées,

- l'intégration de contraintes particulières (commande urgente, préférence pour une machine) est difficile à réaliser,
- l'utilisation d'algorithmes d'agrégation (généralement additive) de données de dimensions différentes n'est pas pertinente,
- la création d'une relation de transitivité cardinale entre les différents choix possibles, supprimant par conséquent toute notion d'indifférence ou de seuil de veto, n'est pas représentative de la réalité,
- les caractéristiques propres au système de production sont difficilement prises en compte: usure outil, capacité de stockage, temps de changement outil, procédures de re-initialisation, etc.
- les processus décisionnel au niveau du pilotage résultent d'un compromis entre des critères conflictuels exprimés principalement en terme de coût, délais ou qualité. Chaque processus de décision de pilotage est par essence multicritère: L'allocation d'un tâche par exemple, est un processus décisionnel qui résulte d'une étude portant sur des critères de coût de production, de temps de changement de série, de temps de convoyage, de temps d'attente dans le stock amont, de qualité de production, etc.

L'exploitation de méthodes multicritères permet d'intégrer l'ensemble de ces contraintes, en particulier par le fait que les hypothèses sur lesquelles se basent celles-ci sont plus proches de la réalité que les méthodes d'optimisation.

D'autre part, l'approche multicritère facilite l'intégration de l'opérateur humain au sein du SIAD:

- en proposant une optique de résolution plus adéquate au raisonnement humain: critère, seuil, prise en compte de l'intransitivité de l'indifférence, etc.
- en proposant des méthodes de résolution interactive.

Cependant l'adéquation des modèles à la réalité accroît la non stabilité des données et des traitements:

- les données exploitées correspondent la plupart du temps à une information dont le niveau de stabilité décroît fortement avec la spécificité de la donnée: prise en compte de manière très précise des temps de non production, des coûts indirects, etc. Ce point nous conduit donc à constater une diminution de la stabilité et de l'accessibilité de l'information,
- la multiplicité des points de vue par l'intermédiaire du développement des critères et de leurs pondérations met en évidence la non stabilité du problème à résoudre.

L'adéquation avec la réalité qui accompagne le développement d'une méthode multicritère accroît ainsi la capacité théorique à proposer une prise de décision

efficace. Cependant, elle met en évidence de manière plus flagrante le principe de rationalité limitée, ce qui conduit au besoin d'intégration de l'opérateur humain, seul capable de faire face à la faible structuration des processus décisionnels.

3.7. Conclusion

Il résulte de l'analyse précédente que la conception d'un SIAD pour le pilotage doit rechercher un compromis entre:

- la nécessité d'intégration d'une approche multicritère dans le processus de décision,
- le niveau d'adéquation à la réalité requis pour la méthode multicritère développée,
- les besoins associés à la structure de pilotage:
 - besoin d'interactivité avec l'opérateur qui favorise la réactivité et l'intégration d'objectifs globaux,
 - nécessité de réaliser les différents processus décisionnels selon des temps suffisamment faibles par rapport aux caractéristiques temporelles du système de production,
- l'objectif de limitation des contraintes psychologiques (effet de stress, etc.) auxquelles est soumis l'opérateur humain.

Nous avons justifié l'intérêt de l'intégration au sein de chaque SIP d'un Système Interactif d'Aide Multicritère à la Décision. Les parties suivantes détaillent les concepts associés au SIAD d'une part et les concepts associés à une approche multicritère d'autre part puis présentent une spécification possible d'un SIAMD distribué en fonction du compromis présenté auparavant et appliqué au niveau du pilotage à structure hybride.

4. Système interactif d'aide à la décision et pilotage

Dans cette partie, nous allons décrire le concept de SIAD et ses caractéristiques principales. Nous présenterons ensuite de manière plus précise la problématique de l'intégration d'un SIAD au niveau pilotage.

4.1. Concepts SIAD

L'aide à la décision regroupe un ensemble de domaines assez disparates tels que l'intelligence artificielle, l'ergonomie, et les systèmes de gestion de bases de données (SGBD), etc.

Bonczek [Bonczek et al. 84] décrit un SIAD comme un système informatisé qui utilise ses connaissances sur un sujet particulier afin d'aider le responsable lors de la prise de décision dans une catégorie de problèmes peu ou pas structurés.

Courbon [Courbon 88] le définit comme "un système homme-machine qui à travers un dialogue permet à un décideur d'amplifier son raisonnement dans

l'identification et la résolution de problèmes peu structurés". Il introduit les notions de dialogue homme-machine, notion implicite dans la définition de [Bonczek et al. 84]. Il nous semble important de remarquer que l'efficacité d'un SIAD est fortement conditionnée par la qualité de l'interaction avec l'utilisateur.

Les premiers SIAD sont apparus en 1968 [Courbon 88], [Bonczek et al. 84]: Morton du M.I.T réalisa alors un système d'aide à la planification conjointe, commerciale et industrielle, dans une division de la firme WestingHouse. L'essor des SIAD n'est effectif qu'à la fin des années 1970 où divers outils d'aide à la décision deviennent opérationnels. [Keen 87] estime que cet essor a pu avoir lieu grâce à l'apparition de nouvelles technologies désormais accessibles. Ses conclusions stipulent en outre que la définition et les spécifications d'un SIAD ne peuvent pas être fixées précisément, car il n'est pas possible de trouver une approche résolument générique pour l'ensemble des cas et des domaines possibles. C'est pourquoi les concepts génériques associés à l'approche SIAD sont peu nombreux.

Cependant, il existe une définition conceptuelle généralement admise définissant un modèle systémique générique d'un SIAD [Sprague 87] (figure 2).

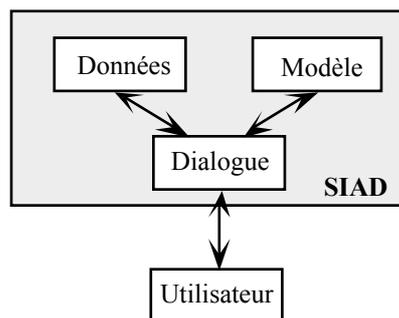


Figure 2: structure générique d'un SIAD selon [Sprague 87].

Ce modèle est assez proche de celui de Bonczek qui propose un modèle comportant trois sous-systèmes: communication, savoir et traitement [Bonczek et al. 80], [Bonczek et al. 84]. Du point de vue fonctionnel, le sous-système de traitement comporte sept activités: collecter des informations, formuler des modèles, gérer (administrer, gouverner), analyser, évaluer, reconnaître les problèmes et les implémenter (figure 3).

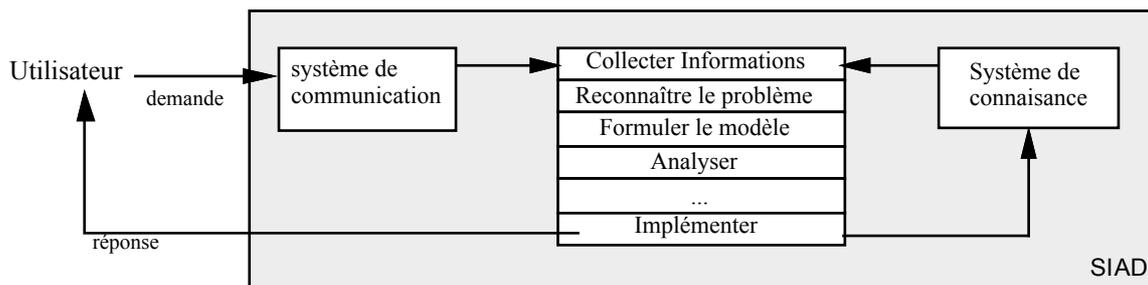


Figure 3: SIAD selon [Bonczek et al. 84]

Cette décomposition fonctionnelle constitue une première étape vers l'intégration du modèle de processus de décision de Simon. En effet, nous constatons l'adéquation entre le processus de décision tel qu'il est décrit par Simon et la structuration du modèle de Bonczek capable d'intégrer une telle méthodologie.

Beulens rappelle qu'un SIAD est une mise en oeuvre particulière d'un système d'informations dédié à l'aide à la décision [Beulens et al. 88] et compare les SIAD à d'autres systèmes d'informations spécifiques. Cette étude met en évidence la spécialisation du SIAD pour traiter des problèmes peu structurés.

4.1.1. Problématique générale des SIAD

Elle est liée principalement à la résolution des problèmes peu structurés. [Hawgood et al. 87]. Un SIAD doit faire face à :

- la complexité du problème,
- l'imprécision ou l'incertitude des objectifs,
- l'aspect conflictuel des préférences,
- le manque d'information valide, dû notamment à la complexité des problèmes à résoudre et l'aspect incertain des données. Lévine relève également le problème des informations fournies en trop grand nombre [Lévine et al. 89]. La conséquence est évidemment une réduction de la pertinence de l'aide apportée aux utilisateurs du système,
- le manque de méthodes précises pour la résolution du problème traité.

Le concepteur d'un SIAD doit s'efforcer de prendre en compte ces différents éléments, mais certains objectifs ne sont pas toujours atteints. [Hawgood et al. 87] constate que beaucoup de SIAD sont :

- trop "artificiels": ils utilisent souvent une forme de dialogue trop abstraite et non adaptée au niveau de compétence de l'utilisateur; ce qui complique la compréhension du problème et des objectifs par l'utilisateur,
- trop "égoïstes": ils prennent rarement en compte les méthodes, les jugements et les préférences des utilisateurs. Ils sont souvent peu interactifs ou coopératifs. Le dialogue entre le système et l'opérateur durant la prise de décision est très restreint ou inexistant, ce qui est contradictoire avec le principe d'interactivité.

La partie suivante traite un aspect particulier de la problématique des SIAD qui est l'intégration de l'opérateur humain.

4.1.2. Intégration de l'opérateur humain

4.1.2.1. *La coopération homme-machine*

L'un des avantages de l'approche SIAD est qu'elle ne supprime pas la part de responsabilité du décideur afin de gérer au mieux la faible structuration des processus décisionnels, comme c'est le cas pour les méthodes classiques d'optimisation qui proposent une solution unique et optimale à partir des hypothèses de stabilité. Ainsi, le décideur reste responsable du choix final. De ce fait, le rôle de l'opérateur évolue de la responsabilité du contrôle à celle du pilotage [Johannsen et al. 94].

L'un des problèmes majeurs dans la conception d'un SIAD est celui de la modélisation de l'opérateur humain. Cette modélisation permet d'améliorer le rendement du système "homme-machine". Cependant, une telle modélisation est complexe et l'on se limite souvent à une étude comportementale [Rasmussen 80].

D'autre part, pour réaliser une intégration efficace de l'opérateur humain, il est nécessaire que ce dernier possède une représentation interne des phénomènes aléatoires, de leurs conséquences, du processus à piloter, etc. Ceci est étroitement lié aux trois comportements spécifiques répertoriés par Rasmussen (comportements basés sur l'habileté, les règles et la connaissance). Il est nécessaire de prendre en compte l'existence de ces trois types de comportement lors de la spécification d'un SIAD.

Le problème inhérent à un échange de données entre deux acteurs qui possèdent chacun une approche différente de la réalité (l'opérateur humain et la machine) est celui de la communication entre deux mondes totalement différents. Cette communication doit s'établir sur une échelle qui peut aller d'une coopération faible voire nulle (l'un ou l'autre exclusivement prend la décision) à une coopération totale entre les deux acteurs. Une telle taxonomie est décrite dans [Levis et al. 94] et sous une forme équivalente dans [Johannsen et al. 94]. Elle généralise l'approche habituelle qui consiste à décrire deux niveaux de coopération (horizontale, verticale) entre l'homme et la machine. Il est évident qu'il existe de nombreuses contraintes quant à une quelconque spécification des niveaux de coopération potentiels: ces contraintes sont dictées par l'application elle-même.

4.1.2.2. *La volonté de coopérer*

Un aspect particulier de l'opérateur humain est qu'il influe très sensiblement sur la qualité de la coopération. Une étude menée par N. Moray a montré que l'efficacité du couple homme-machine dépend de la volonté de coopérer du premier [Moray et al. 94]. Cette volonté dépend du facteur de confiance en soi et de la confiance de l'opérateur envers le système informatisé. Cette caractéristique doit être prise en compte en fonction du niveau de coopération souhaité et en fonction du niveau de responsabilité de l'opérateur.

4.1.3. Typologie des SIAD

Keen propose un regroupement par niveaux d' "Aide interactive" en fonction des désirs de l'utilisateur [Keen 87]:

- SIAD passif: il donne un avis qui ne vient pas empiéter sur l'autonomie de l'utilisateur,
- SIAD traditionnel: ce système est un assistant dont la principale utilisation concerne les interactions du type "What-if?",
- SIAD étendu: il assure les fonctions de consultant et se place au même niveau que l'utilisateur,
- SIAD nominatif: il domine le processus de décision. L'opérateur ne remplit plus qu'un rôle passif dans la prise de décision.

Il est à noter que cette typologie est isomorphe à celle établie suivant les niveaux de coopération. Elle spécifie cependant la manière avec laquelle un SIAD répond à l'attente de coopération.

4.1.4. Système d'aide à la décision distribué

Peu de travaux ont été réalisés dans ce domaine exceptées les applications dans le domaine des systèmes d'aide à la décision de groupe ou SADG (Group Decision Support Systems) [Floyd Lewis 87], [Floyd Lewis et al. 88], [Pollock et al. 93], [Eom et al. 90], [Beauclair et al. 90], [DeSanctis et al. 87].

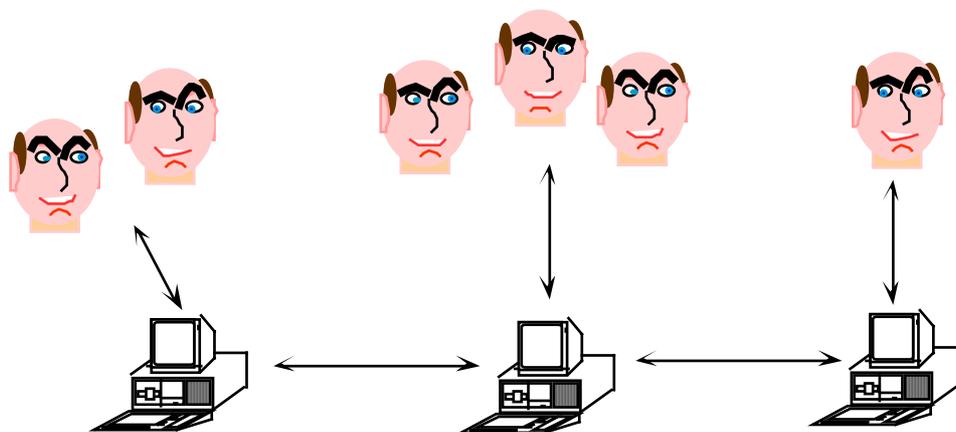


Figure 4: structure générale d'un GDSS

Les SADG sont conçus principalement pour l'aide à la communication au sein d'un groupe de décideurs distribué géographiquement et fonctionnellement (l'horizon décisionnel est en général le long terme), (figure 4). Ces outils constituent en fait une finalisation informatisée de l'approche "séance de créativité" ou "brainstorming" visant à canaliser, synthétiser et organiser les différentes interactions entre les participants dont le but est d'atteindre un objectif global.

Certains points communs existent entre l'approche SADG et SIAD distribués, les SADG particularisant un SIAD distribué pour les décisions d'ordre stratégique.

Cependant, le cadre de décision qui nous concerne est celui du pilotage (niveau opérationnel), et les différents acteurs ne sont plus seulement des êtres humains, mais aussi des ressources informatisées telles que les SIP.

Quelques rares tentatives de formalisation d'un SIAD distribué ont été réalisées [Johannsen et al. 94], [Bonczek et al. 84]. Ce dernier décrit deux approches non nécessairement incompatibles:

- dans un premier cas, le module de connaissances est localisé au sein d'une unique ressource. La recherche de l'information se fait alors par un jeu de requêtes adressées à cette unique ressource,
- une deuxième approche consiste à distribuer les connaissances au niveau de chaque composant du SIAD. Dans ce cas, la recherche de l'information se fait par échange direct entre ressources.

Ces deux approches ont été largement étudiées dans un cadre plus particulier qui concerne les différentes approches de conception du protocole de communication d'une structure multi-agents [Tchako 94].

4.1.5. Conclusion

La conception d'un SIAD est principalement fonction des niveaux de coopération désirés, du niveau de connaissances et du niveau de responsabilité et de compétence de l'opérateur.

[Eom et al. 93] a étudié plus d'une centaine d'auteurs pour déterminer les différents aspects des SIAD, que ce soit en terme de fondement théorique, ou de spécifications selon le domaine d'application. Il décrit notamment le rôle des SADG au sein des systèmes d'informations.

Notre apport se situe dans la spécification d'un SIAD distribué dans un contexte opérationnel (temps-réel). Ce type de SIAD n'a, à notre connaissance, que très peu été étudié et mis en oeuvre.

Nous présentons dans la partie suivante la problématique de l'aide au pilotage de systèmes de production à l'aide d'un SIAD.

4.2. SIAD et pilotage

La présentation de l'état de l'art sur les applications de SIAD au niveau du pilotage et des concepts de l'interaction temps réel entre machine et opérateur nous conduit à définir un ensemble de contraintes à respecter pour la spécification du SIAD appliqué au pilotage hybride.

Cette étude prend en compte la problématique de l'ordonnancement/ ré-ordonnancement prévisionnel.

4.2.1. Apport d'un SIAD pour l'aide au pilotage

D'une manière générale, les décisions au niveau pilotage concernent:

- la réalisation des tâches,
- le suivi des performances,
- l'évaluation du respect des objectifs.

L'intégration d'un SIAD permet de mieux résoudre un problème répertorié par Parunak pour l'ordonnancement de tâches en milieu industriel [Parunak 91]: la désirabilité qui décrit la difficulté à tenir compte à la fois des contraintes réelles de production par une fonction d'évaluation et des souhaits des responsables de la production. L'intégration de l'opérateur humain dans le processus de prise de décision améliore cet aspect et par conséquent l'évaluation de l'impact des décisions a priori et a posteriori et l'évaluation en temps-réel ou a posteriori du respect des objectifs sont nécessaires.

4.2.2. Aide à la décision temps-réel pour le pilotage

L'intégration de la contrainte temps-réel entraîne nécessairement que l'opérateur soit confronté à la réalisation de plusieurs tâches en un temps déterminé [Moray et al. 91].

De ce fait, il est légitime de prendre en compte les situations où l'opérateur peut être en surcharge de travail. Cette surcharge influe sur la qualité de la décision et la performance du système complet. D'une manière duale, l'absence de toute charge nuit à l'intérêt porté par l'opérateur humain au contrôle du système. Un équilibre est donc recommandé pour obtenir l'efficacité dans la prise de décision en temps-réel. Cet aspect, lié à la psychologie du travail, est rarement abordé lors de la spécification ou l'évaluation d'un SIAD dédié au pilotage.

4.2.3. Etat de l'art des SIAD pour l'ordonnancement et le pilotage

Il existe à ce jour peu d'applications des concepts SIAD au niveau pilotage temps-réel, malgré la pertinence d'une telle approche à ce niveau. En revanche, de nombreuses applications ont été développées pour la création d'un ordonnancement prévisionnel partiel ou total. Pour la structure présentée dans ce mémoire, les décisions concernent à la fois l'ordonnancement prévisionnel et la gestion temps-réel. C'est pourquoi il nous semble nécessaire de définir d'une part les cadres de décision prévisionnel et temps-réel, et d'autre part de présenter un état de l'art des méthodes d'aide à la décision pour l'ordonnancement prévisionnel et pour le pilotage temps réel.

Ces systèmes s'appuient principalement sur un ensemble d'approches heuristiques ou empiriques. L'aide est principalement constituée d'un recours à un dialogue du type "What-if?". C'est à dire que l'utilisateur évalue ses décisions à l'aide d'un dialogue plus ou moins complexe basé sur une étude des différents cas de figure. Ces différents cas expriment la plupart du temps le niveau de robustesse vis à vis des aléas (réactivité) ou le respect d'objectifs locaux ou globaux. Le décideur peut alors réitérer l'étape de construction de la décision. Il est évident que

ce schéma correspond à celui de Simon. La qualité globale des différentes mises en oeuvre est évaluée en regard de:

- la qualité de la modélisation du système de production,
- des possibilités d'évaluation (agrégation des résultats, etc.),
- de l'interfaçage (ergonomie, facilité d'apprentissage, interactivité, etc.),
- des possibilités de paramétrage.

Avant de présenter quelques mises en oeuvre existantes, nous présentons les différents approches usuelles de traitement de données utilisées dans les SIAD.

4.2.3.1. Approches de résolution

Le traitement des données peut s'effectuer à l'aide de méthodes et d'outils issus de différents domaines:

- la recherche opérationnelle (méthodes analytiques, méthodes multicritère, etc.),
- la simulation de flux (langages généraux, réseaux de Petri, simulation à événements discrets, etc.),
- l'intelligence artificielle (système expert, programmation par contrainte, et plus récemment les réseaux de neurones, etc.).

Les méthodes multicritères ont été introduites succinctement dans un paragraphe précédent. Nous présentons de manière plus précise cette approche dans la partie suivante.

Il faut noter que la tendance actuelle dans la conception de SIAD est de combiner plusieurs de ces approches (approche multi-modèle).

La recherche opérationnelle

Les méthodes analytiques (réseaux de files d'attente par exemple) sont peu compatibles avec les concepts SIAD. Elles requièrent des hypothèses souvent peu réalistes (probabilité exponentielle, règles FIFO, etc.) et entraînent l'inadéquation du modèle avec la réalité. De ce fait, elles sont généralement exploitées en phase de pré-étude (faisabilité de l'ordonnancement, respect a priori des objectifs globaux, etc.) ou en tant qu'outil de simulation locale (comportement d'une machine, etc.). Elles peuvent cependant constituer un élément d'un système multi-modèle.

La simulation

Les méthodes exploitant les réseaux de Petri sont compatibles avec le concept d'interactivité du SIAD. Ceci est dû à l'aspect visuel du modèle. Cependant, la complexité des graphes s'accroît rapidement avec l'intégration de contraintes particulières visant à rapprocher le modèle de la réalité. L'intégration des extensions (coloration, temporisation, hybridation, etc.) permet d'agréger et donc de simplifier la structure du graphe sans toutefois remédier au problème de lisibilité, car ces extensions conceptualisent les évolutions du graphe. Cette

approche est donc également adoptée dans un but de globalisation ou dans le but d'étudier localement un aspect temporel ou spatial d'une décision. La création d'une structuration hiérarchique des réseaux a été envisagée pour remédier à ce problème de lisibilité. Un premier exemple est proposé par [Trentesaux et al. 94a] où l'utilisation du concept de macro-place a été présentée.

Les modèles analytiques et réseaux de Petri peuvent ainsi être exploités, mais de manière complémentaire avec d'autres modèles. L'objectif est en particulier de proposer des outils de simulation et d'évaluation (l'utilisation de ces outils se situe au niveau de la phase d'évaluation de Simon). La plupart des applications concernent ainsi l'évaluation en terme de performance ou l'étude de viabilité d'un système (blocage, taux d'utilisation moyen, etc.).

La simulation à événement discret utilise des langages dédiés ou des langages de simulation. Les avantages et inconvénients de chacun sont largement connus: l'approche par langage dédié (Fortran, Pascal, C, C++, Lisp, Prolog, etc.) est très souple, très modulaire, flexible mais requiert la présence d'un expert de développement et un temps d'étude qui s'accroît très rapidement avec la complexité de la modélisation. L'intérêt d'une telle approche réside aussi dans les possibilités de paramétrage. De cette manière un modèle très précis peut-être établi, la contrepartie étant la complexité de la gestion d'un tel système. L'approche par langage de simulation (Qnap, Slam, Arena, etc.) est moins fine car elle propose des éléments de modélisation pré-définis. La modélisation est plus rapide mais le niveau de précision est plus faible. Le temps de développement est réduit et il n'est pas forcément nécessaire de faire appel à un expert.

L'intelligence artificielle

Les systèmes experts constituent la forme la plus répandue de l'utilisation de l'intelligence artificielle au sein d'un SIAD. Ils apportent de nombreuses solutions:

- le raisonnement du système expert est "calqué" sur celui de l'être humain, ce qui favorise la compréhension du mécanisme de résolution du système par l'opérateur,
- les méthodes expertes se basent sur un acquis et une expérience élaborée. Leur temps de réponse compatible avec les contraintes temporelles des applications envisagées permet d'augmenter l'efficacité de la prise de décision,
- les potentialités d'apprentissage permettent d'intégrer une meilleure gestion des aléas par mémorisation des décisions et des contextes correspondant,
- il est possible de supporter la nature de certaines données peu précises ou qualitatives,
- la possibilité de l'intégration naturelle du concept "What-if?" accroît naturellement la qualité de la coopération entre l'homme et la machine.

La programmation par contraintes permet de définir un problème en terme de contraintes sans avoir à spécifier la méthode ou le traitement permettant d'obtenir le résultat. L'intérêt provient principalement de l'aspect descriptif et également du fait que la méthode puisse être paramétrée afin de fournir l'ensemble des solutions potentielles qui satisfont les contraintes définies. Le processus de relâchement de contrainte permet en outre d'intégrer aisément différents niveaux d'interactivité.

Ces différents outils ont été diversement exploités. Nous présentons un état de l'art suivant le type d'application du SIAD.

4.2.3.2. *SIAD pour l'ordonnancement prévisionnel*

[Canals 86] propose un système d'aide à l'ordonnancement utilisant la simulation. L'objectif est de découpler les règles de priorité afin de les évaluer indépendamment l'une de l'autre. Une étude précise de l'interaction entre les différentes règles d'allocation permet d'établir une liste de critères dans une optique juste à temps. Un des intérêts de cette approche est qu'elle évite de travailler sur des paramètres qui dépendent de l'ordonnancement contrairement à certaines méthodes d'ordonnancement qui utilisent les délais de production comme paramètres, alors que l'ordonnancement calculé conditionne ces délais.

[Paredis et al. 92] propose l'utilisation de la programmation par contraintes. L'exploitation d'une telle méthode a pour conséquence d'isoler l'utilisateur de toute la problématique technique pour se focaliser sur l'efficacité et la robustesse d'un ordonnancement. Le problème de l'ordonnancement est spécifié exclusivement sous forme de contraintes (précédence, disjonctivité, etc.). Un des avantages est que le système peut fournir un ensemble d'ordonnements possibles. Notons cependant que dans certains cas, le système ne peut fournir qu'une ou pas de solutions. Il faut alors relâcher certaines contraintes. Une étude plus poussée peut alors permettre une comparaison des différentes solutions, cette évaluation se fait par l'étude de l'espace des ordonnancements possibles. Il est possible de forcer des choix et ainsi d'évaluer l'impact des décisions en terme de non-respect de contraintes.

Erschler a montré la compatibilité d'une telle approche avec l'aide à l'ordonnancement [Erschler et al. 93]. Il définit et propose une méthode explicitant la robustesse d'un ordonnancement par contraintes.

4.2.3.3. *SIAD pour le pilotage temps-réel*

[Gravel et al. 92] propose une spécification originale d'un SIAD pour le pilotage d'un système de production géré par Kanbans. Une première étape d'analyse permet de former une base de données statistiques. Cette base de données contient un ensemble de tests élaborés selon le nombre de Kanbans en circulation, les règles appliquées, la taille des lots, etc. Cette base de données est alors exploitée en temps réel pour améliorer un ou plusieurs critères concernant la

production. La résolution utilise une méthode multicritère. Le SIAD permet ainsi à l'utilisateur d'afficher les variables d'état, l'état des stocks, etc. Des simulations à très court terme sont alors possibles. La résolution du choix par l'approche multicritère permet d'ajuster les paramètres pour améliorer l'efficacité globale. L'avantage est la spécificité poussée du SIAD et de la base de données qui permet de réagir en temps réel.

[Zhao 89] propose un système expert d'aide à la conduite fournissant à l'opérateur une assistance experte en vue de saisir synthétiquement l'état du système de production. Le système expert conseille l'utilisateur afin de gérer au mieux une perturbation ou afin de respecter un objectif global. Il décrit la nécessité d'intégrer un filtre pour mettre en évidence les informations pertinentes. Le système est également capable d'auto-apprentissage.

[Grabot et al. 94] discute de la nécessité de structurer les informations vagues ou incertaines au sein d'un système expert d'aide au pilotage. Outre cet aspect, les fonctionnalités sont semblables à celles de Zhao. Le système expert doit pouvoir identifier une situation, proposer différents choix possibles et intégrer un module algorithmique pour l'évaluation des données ainsi qu'une structuration des données floues.

[Dindeleux 92] et [Tahon et al. 92] proposent des spécifications génériques d'un SIAD dédié au pilotage centralisé d'un système de production. Le rôle du pilotage et les fonctions associées sont mis en évidence. Le SIAD est proposé comme outil permettant d'établir un compromis entre la gestion temps-réel et les ordres prévisionnels.

4.2.3.4. Méthodes conjointes d'ordonnancement et de pilotage temps-réel

Dans le premier chapitre, nous avons décrit la méthode ORABAID [Thomas 80], [Legall 89] pour l'élaboration d'un ordonnancement prévisionnel partiel couplé à un module de gestion temps réel. Le système interactif d'aide à la décision exploite les notions de marge libre et de groupes de tâches permutables pour l'ordonnancement en temps réel. Le SIAD est activé en fonction des différents événements (panne, libération de ressource, etc.) [Billaut et al. 93]:

- l'analyse, où l'on recherche à l'aide d'une heuristique un sous-ensemble de l'ensemble des ordonnancements possibles,
- décisions d'ordonnancement élémentaires: ces décisions doivent être prises à chaque fois qu'une ressource devient libre. Elles se basent sur l'ensemble des familles d'ordonnancement élaborées,
- décisions après dysfonctionnement: panne ressource, etc.

En ce qui concerne l'aide à la décision, elle est assez succincte car le modèle de décision utilise exclusivement une approche analytique: un arbre logique détaille

les réactions à envisager et le décideur intervient au niveau de la permutation des familles de produits.

[Tabucanon et al. 94] présente une approche multi-modèle basée sur le développement d'un SIAMD dédié à la sélection d'une ressource de production. La méthode est la suivante (figure 5):

- un système expert réalise la pré-sélection d'un ensemble de ressources admissibles. Cette pré-sélection s'effectue en collaboration avec l'utilisateur qui particularise les caractéristiques de chaque ressource étudiée par le système expert. Ce dernier propose alors au maximum sept ressources potentielles,
- une méthode multicritère basée sur la méthode AHP [Saaty 80] réalise la sélection de la ressource la plus satisfaisante. Les critères sont décomposés par famille (coût, flexibilité, etc.). Le système expert intervient également dans la désignation et la prise en compte de sous-familles de critères et pour l'analyse de sensibilité.

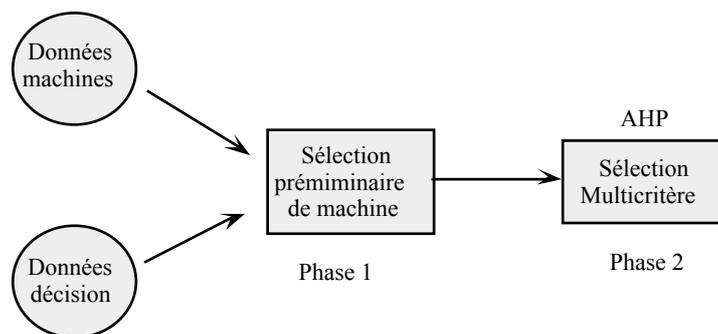


Figure 5: méthode de sélection d'une ressource [Tabucanon et al. 94]

L'intérêt de cette approche est qu'elle est applicable à la fois au niveau prévisionnel et au niveau opérationnel. Dans le premier cas, l'ordonnancement est considéré comme un ensemble d'allocations réalisées dans une optique multicritère. Dans le second cas, les opérations peuvent être allouées dynamiquement. Cette approche est, pour l'allocation dynamique, assez semblable à l'approche que nous avons élaborée.

[Bourgeois 94] propose un SIAD dont le but est d'une part d'assister l'opérateur lors de l'établissement d'un plan de travail réalisable aussi rapidement que possible et d'autre part de réagir face aux aléas de production. Un premier module interactif permet de réaliser un ordonnancement. Celui-ci est ensuite évalué par l'opérateur conjointement avec un système expert. Le module d'évaluation permet la simulation de placement de manière prévisionnelle [Bourgeois 93]. Cette approche permet aussi une gestion temps réel par l'intégration au SIAD d'une gestion de charge dynamique des ressources, de leurs marges capacitaires, etc.

4.3. Synthèse

L'étude précédente a montré la diversité des applications du concept SIAD pour le pilotage ou l'ordonnancement. Cette diversité s'appuie sur la multiplicité des modèles et des structures de pilotage et montre clairement l'adéquation entre la problématique du pilotage et celle des SIAD.

L'approche multi-modèle semble de plus en plus employée car elle permet d'intégrer au sein d'un même SIAD plusieurs outils de nature différente dédiés chacun à une facette du problème de pilotage.

L'intégration de l'opérateur humain est envisageable afin de prendre en compte la multiplicité des situations, l'ordinateur ne pouvant faire face à toutes les situations de manière exhaustive.

Enfin, l'approche multicritère utilisée dans certains systèmes est particulièrement adaptée à l'approche SIAD dans un contexte de production. C'est ce que nous décrivons dans la partie suivante.

5. L'aide multicritère à la décision

Nous allons succinctement décrire les caractéristiques d'un modèle multicritère et présenter ensuite les problèmes liés à son utilisation dans un contexte de pilotage.

5.1. Approche multicritère pour la résolution d'un problème

De très nombreuses références traitent de la théorie de l'approche multicritère et de ses applications industrielles: [Schärlig 85], [Pomerol et al. 93], [Roy 85], [Roy et al. 93], [Vincke 89], [Zeleny 82], etc. Plusieurs synthèses ont été réalisées dans le vue d'élaborer une typologie de problèmes à caractère multicritère (voir par exemple [Korhonen et al. 92]).

Les critiques faites envers les méthodes d'optimisation résultent principalement de l'impossibilité à prendre en compte un modèle adéquat de la réalité pour une prise de décision pertinente. Les hypothèses sont généralement très restrictives et conduisent à élaborer des modèles insuffisamment représentatifs de la réalité.

A l'opposé, une approche multicritère a pour principe de s'appuyer sur des hypothèses plus proches de la réalité, et d'utiliser des méthodes de traitement de données moins complexes [Roy et al. 93]. L'objectif n'est plus de trouver une solution optimale, mais plutôt de proposer un compromis efficace. Ce compromis est réalisé en fonction de l'objectif de résolution qui peut-être principalement [Roy et al. 93]:

- élaborer un choix (problématique de choix),
- élaborer un tri par catégorie (problématique de tri),
- élaborer un classement (problématique de rangement).

Une méthode multicritère se fonde sur une “matrice de décision” qui comporte le maximum d’informations nécessaires au traitement. Elle est élaborée à partir d’un ensemble de choix potentiels (alternatives) qui dans le cadre de cette étude est défini en extension (ensemble à cardinal fini, stable ou non). Chacun de ces choix est évalué en regard d’un critère qui représente le jugement d’une réalité correspondant à un axe de signification précis. On constate qu’une première restriction a lieu, puisque la réalité est restreinte à un ensemble de critères.

Dans le cas des méthodes d’optimisation, on cherche à agréger les données (sous la forme d’un indicateur par exemple) le plus tôt possible dans le déroulement du processus de décision afin d’élaborer une relation transitive entre les différents choix. L’approche multicritère au contraire retarde le plus possible lors de la phase de résolution, l’opération d’agrégation, car cette dernière détruit l’information de base donnée par la matrice de décision. On cherche de cette manière à élaborer en premier lieu les évaluations pour chacun des critères puis à agréger ces données. Deux étapes sont alors définies.

5.1.1. Phase d’identification

L’objectif premier de cette phase est d’obtenir et de structurer la matrice de décision. La difficulté inhérente à cette phase est la prise en compte de la diversité des données et des évaluations. On discerne deux grandes catégories de données:

- les données quantitatives: à relation cardinale (l’ordre indiqué par cette relation est quantifiable, on peut déterminer une distance entre deux données),
- les données qualitatives: relation ordinale (la mesure de la distance est impossible, seul l’ordre est disponible) ou de type linguistique.

A chaque critère correspond un type de donnée. Par exemple, la première catégorie de données donne naissance à un ensemble de critères que l’on peut subdiviser en vrai-critère, quasi-critère, pseudo-critère. Ces notions peuvent être généralisées, l’exemple le plus connu s’appuie sur les concepts de la logique floue [Fodor et al. 94].

La qualité de la prise de décision est fortement conditionnée par le choix des critères pris en considération. [Roy et al. 1993] indique les cinq conditions suivantes portant sur les critères:

- exhaustivité: la liste de critères doit être composée de critères exhaustifs. Cette caractéristique s’exprime par le fait suivant: si pour chaque couple d’alternatives, la même évaluation est faite pour chacun des critères pris en compte, alors il faut conclure à l’indifférence entre ces deux choix. Dans le cas contraire, un critère supplémentaire a été omis et doit être ajouté à la liste,

- cohésion: l'évaluation relative globale pour une alternative vis à vis d'une autre doit être monotone avec l'évolution de chacune de ses évaluations, et ce, pour chaque critère,
- non-redondance: la liste de critères est redondante lorsqu'un critère spécifique peut être supprimé de la liste sans que la nouvelle liste ne remplisse plus la première condition d'exhaustivité,
- séparabilité: l'évaluation globale relative entre deux alternatives peut être influencée par la modification d'une évaluation particulière vis à vis d'un critère spécifique. Dans le cas contraire, le critère est inutile,
- isolabilité: il est important de pouvoir modifier l'évaluation d'une alternative vis à vis d'un critère particulier sans avoir à reconsidérer une évaluation de cette même alternative vis à vis d'un autre critère. Dans le cas contraire, un lien existe entre ces deux critères, et il est conseillé d'élaborer la liste sur des critères plus fondamentaux et éviter ainsi une agrégation anticipée.

En ce qui concerne le niveau d'importance relative des critères, le lecteur pourra se référer aux auteurs cités plus haut pour approfondir cet aspect. Les études réalisées soulignent le fait que la déclaration de l'importance relative entre critères peut-être ou non réalisée par l'intermédiaire de poids. Dans le premier cas, la cardinalisation permet d'élaborer une algorithmique plus simple que dans le second cas où l'aspect qualitatif de l'importance relative prédomine. Par ailleurs, il est clair que la spécification des différents poids est une étape très sensible et peut influencer fortement sur la décision finale.

5.1.2. Phase d'agrégation

Le nombre de méthodes disponibles témoigne de la diversité des différentes approches pour la phase d'agrégation.

Les processus de décision étudiés dans ce mémoire sont associés à la problématique de choix discret (sélection d'une SIP, etc.) ou de rangement discret (séquencement des allocations, etc.).

Un compromis doit être établi entre le niveau d'adéquation de la méthode avec la réalité et les contraintes imposées par l'environnement décisionnel.

Deux grandes écoles coexistent: l'école américaine et l'école européenne. La première cherche à établir un pré-ordre total des alternatives à partir de fonctions d'utilité. Une fonction d'utilité est une fonction de normalisation, croissante, permettant d'agréger (la plupart du temps de manière additive) les différentes évaluations en un critère unique. C'est une méthode d'agrégation complète et transitive qui se rapproche assez des méthodes d'optimisation.

L'école européenne favorise une méthode d'agrégation partielle ou locale et itérative. Elle estime réaliste et plus proche de la philosophie multicritère de conserver au maximum l'information et d'éviter cette agrégation totale. Cette

démarche permet entre autres d'intégrer les notions d'intransitivité des préférences, d'incomparabilité entre actions, d'indifférence, de seuil de veto, etc. L'approche qui permet une agrégation partielle est celle du surclassement qui se décompose en deux étapes:

- évaluation de tous les couples de choix pour chaque critère. On explicite les relations de surclassement partielles (la fonction d'utilité n'existe pas dans cette approche),
- exploitation de ces relations partielles. En ce qui concerne la problématique de choix, un graphe de surclassement partiel est établi en fonction de ces relations. L'étape finale d'agrégation est réalisée en établissant une relation qui permet d'extraire un sous-ensemble d'alternatives possédant certaines particularités. Par exemple, Electre [Roy 85] utilise la notion de seuil (coupure) pour déterminer un noyau de surclassement. En ce qui concerne la problématique de rangement, une fonction de performance (score) établit un pré-ordre total (sans ex-aequo) ou partiel (avec ex-aequo) selon le cas. Prométhée [Brans et al. 85] utilise la notion de flux entrant ou sortant pour déterminer ces pré-ordres.

La phase d'agrégation finale n'est pas systématiquement nécessaire. En effet, il suffit parfois de n'utiliser que les surclassements partiels (qui constituent en eux-mêmes une première étape vers la décision) lorsque la quantité d'information disponible permet la compréhension globale du processus décisionnel.

Il est à noter qu'une phase d'exploitation propre peut-être définie dans le cas où l'exploitation des relations partielles est complexe [Fodor et al. 94]. C'est le cas, par exemple, lorsqu'une méthode utilise abondamment la notion de noyau (problématique de choix) pour établir un sous-ensemble d'alternatives satisfaisantes. Il est alors nécessaire de procéder à une étude complète du graphe des relations afin de pouvoir extraire un unique noyau. La difficulté intervient lorsque ce graphe est cyclique où plusieurs noyaux (voire aucun) peuvent coexister.

L'approche multicritère possède une potentialité élevée en terme de paramétrage. Ces paramétrages concernent outre les différents poids (s'ils sont définis par la méthode), les seuils de veto, des pseudo critères, des quasi critères, des seuils de surclassement, etc. L'effet pervers d'une telle approche est bien sûr le danger de déplacer le problème de la décision en un problème de définition des paramètres. Inversement, il est possible de calculer le paramétrage suffisant pour obtenir un choix particulier. Ce qui n'est pas sans conséquence pour la crédibilité de l'approche multicritère.

Le problème de la subjectivité est généralement peu abordé. Il existe des méthodes qui supportent la notion de subjectivité, de manière statistique, floue,

etc. Le problème sous-jacent est celui de la complexité qui en découle et qui nuit fortement à l'interactivité lorsque l'on considère la contrainte d'une prise de décision temps-réel.

Il nous semble en outre nécessaire de fournir un ensemble de tests de robustesse afin d'évaluer notamment la sensibilité des paramétrages. Cette analyse peut conforter le décideur dans le cas où certaines informations sont peu précises.

Enfin il est intéressant de noter que dans l'optique multicritère, les deux premières phases de Simon sont primordiales: elles conditionnent la qualité de la construction de la solution (agrégation).

5.2. L'approche multicritère en milieu industriel

Plusieurs applications ont été développées en environnement industriel: [Tabucanon 88] détaille l'intégration de l'approche multicritère au sein de plusieurs fonctions industrielles; [Gandibleux 95] pour la conduite de processus continus; [Belton et al. 92] et [Deckro et al. 82] traitent du problème de l'ordonnancement multicritère. [Pomerol et al. 93] fournit de très nombreux exemples divers appliqués à l'industrie.

Cependant, il n'existe encore que très peu d'applications de l'approche multicritère au niveau opérationnel (environnement décisionnel temps-réel). Excepté [Tabucanon et al. 94], nous ne connaissons pas de tentative d'application des méthodes multicritères en environnement de pilotage temps-réel.

[Korhonen et al. 92] a établi une typologie des problèmes multicritère. Selon cette typologie, les processus décisionnels associés au pilotage hybride constituent une problématique discrète, à ensemble d'alternatives défini en extension, à données qualitatives et quantitatives. Le nombre de critères et d'alternatives est suffisamment peu élevé pour permettre à l'utilisateur d'avoir une vue globale du problème.

5.3. Conclusion

L'intégration de l'opérateur humain dans les processus décisionnels au niveau du pilotage permet d'accroître la réactivité du système et le respect des objectifs globaux. Nous avons envisagé une telle intégration. Nous avons par ailleurs montré que les processus décisionnels associés au pilotage sont par nature faiblement structurés. Ce qui nous a conduit à proposer l'intégration de l'opérateur via un SIAD.

D'autre part, nous avons montré que les processus décisionnels au niveau du pilotage sont par nature multicritères. L'intégration d'une méthode multicritère accroît la qualité de l'interaction avec l'opérateur humain mais l'adéquation élevée de l'approche à la réalité diminue de manière sensible la structuration des processus décisionnels en se basant sur des données brutes et temps-réel, ce qui accroît la non stabilité des données.

Cette étude a ainsi montré que, dans le cas de l'intégration de l'opérateur humain, la qualité du processus de fabrication dépend d'un compromis portant à la fois sur la qualité de l'interaction (choix de la méthode et niveau d'adéquation à la réalité, possibilité d'interaction, etc.) et à la fois sur la prise en compte de la contrainte temps réel du pilotage.

Nous spécifions dans la dernière partie de ce chapitre un modèle particulier de SIAMD dédié au pilotage hybride.

6. Spécification d'un SIAMD pour le pilotage hybride

Nous allons tout d'abord spécifier les caractéristiques d'un modèle de SIAMD pour le pilotage hybride et son intégration au sein des sous-systèmes de chaque SIP [Trentesaux et al. 96]. Nous présentons ensuite les caractéristiques du modèle de données ainsi qu'une structuration de la matrice de décision (alternatives, critères, évaluations).

6.1. Structure du SIAMD

Elle doit être cohérente avec la structure fonctionnelle d'une SIP (figure 6).

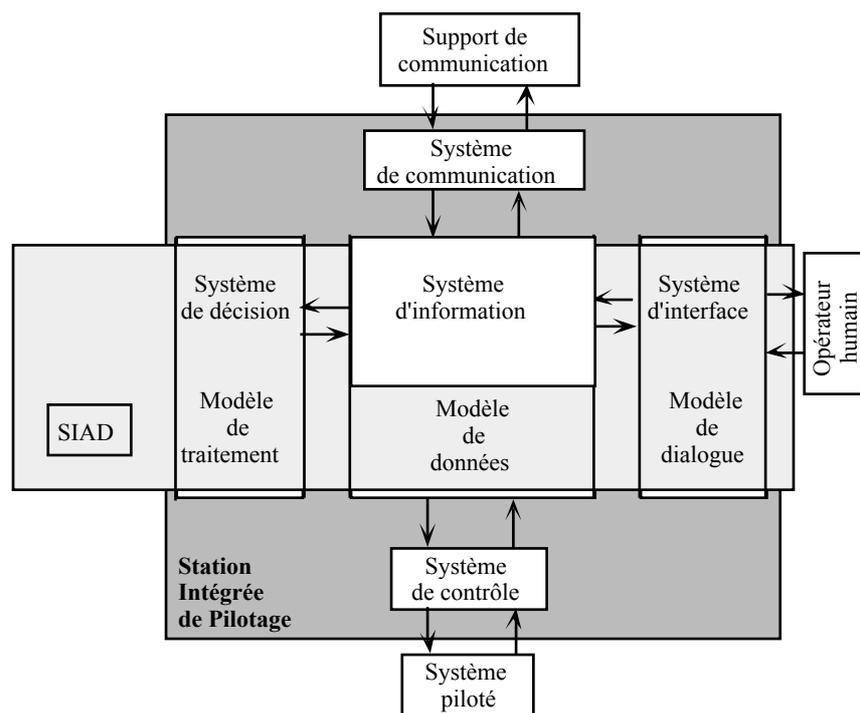


Figure 6: décomposition du SIAMD et intégration dans la SIP

Le modèle de traitement (sous-système de décision) contient l'ensemble des méthodes exploitées par le SIAMD: méthodes multicritère, méthodes analytiques, heuristiques, etc. ainsi que les représentations des données. Le modèle de données contient les informations propres au méthodes de traitement (liste de critères, etc.), il est inclus dans le sous-système d'information. Le modèle de dialogue (sous-système d'interface) contient l'ensemble des outils composant l'interface propre au SIAMD.

6.2. Niveau de connaissance

La responsabilité de l'opérateur est naturellement accrue du fait de son intégration dans le processus décisionnel. Il doit être capable de:

- réaliser, ou faire réaliser chaque opération. Ce rôle est purement technique. Le comportement est basé sur l'entraînement [Rasmussen 83]. Dans ce cas, les processus décisionnels sont généralement bien structurés,
- interpréter les événements qui surviennent de manière périodique (besoin d'allocation, etc.), superviser localement le processus de fabrication. Ces réactions se basent sur des processus décisionnels qui font appel à un comportement de type procédural [Rasmussen 83], l'opérateur prend en compte un ensemble de règles pour identifier la tâche et les procédures à exploiter. Ces événements sont pré-définis, l'opérateur peut ainsi se rattacher à un modèle ou à une démarche. Les processus décisionnels correspondant sont généralement peu structurés, parfois bien structurés mais dans ce cas trop complexes pour permettre à l'opérateur de cerner le processus de décision dans sa globalité,
- réagir face aux événements qui surviennent épisodiquement (aléas, commande urgente, etc.) et qui ne sont pas décrits de manière très précise ou qui n'ont pas été pris en compte. Ces réactions caractérisent un niveau comportemental basé sur la connaissance [Rasmussen 83]. Elles concernent principalement les processus décisionnels non structurés et certains processus peu structurés.

6.3. Niveau de coopération

Le SIAMD doit tenir compte de la multiplicité des types de comportement de l'opérateur mise en évidence et proposer une large échelle de coopération qui permette à l'opérateur:

- d'élaborer sa décision et de la justifier sans l'aide du SIAMD, en se basant uniquement sur les renseignements disponibles (c'est le cas de comportement basé sur la connaissance). Ce mode de coopération est utile dans les cas où le SIAMD ne peut fournir qu'une aide minimale: organisation des données et évaluation succincte de la décision. C'est un mode où le niveau de coopération est peu élevé. Nous le désignons sous le terme "Mode Opérateur",
- d'élaborer conjointement une décision avec le système d'aide. Le rôle du SIAMD n'est plus uniquement de fournir un ensemble de données organisées. L'utilisation de méthodes multicritères permet de fournir une aide en terme de proposition de décision; le choix final de l'opérateur peut naturellement contredire partiellement ou totalement les propositions du SIAMD. Nous désignons ce mode sous le terme "Mode Coopératif". Il est possible de définir plusieurs sous-modes de coopération afin de proposer un niveau de coopération adapté au contexte et à l'opérateur,

- de laisser le SIAMD élaborer les décisions. Le niveau de coopération est dans ce cas minimal. Le SIAMD informe l'opérateur de ses décisions. Nous désignons ce mode sous le terme "Mode Autonome".

La figure 7 résume ces différents modes de coopération et les niveaux d'interaction correspondants.

Mode de coopération	Niveau de coopération
Mode Opérateur	Faible
Mode Coopératif	
Niveau 1	Moyen
Niveau 2	Elevé
Niveau 3	Moyen
Mode Autonome	Faible

Figure 7: modes et niveaux de coopération.

La difficulté majeure est celle du choix du niveau de coopération compte tenu de l'environnement temps-réel. Dans le cas où le temps de réflexion est négligeable par rapport aux temps moyens caractéristiques du processus de production, tous les niveaux de coopérations sont permis. Dans le cas où le temps utilisé par le processus décisionnel conditionne fortement l'efficacité globale, il est peu judicieux de perfectionner la qualité de la décision et une approche non coopérative est nécessaire. Ainsi, le mode autonome doit être favorisé dans le cadre de système de production à temps de réponse très rapide. Une application du mode autonome a été proposée pour l'aide au pilotage d'un système de conditionnement à haut débit et présentée dans [Trentesaux et al. 95b].

L'utilisateur doit sélectionner de ce fait le mode de coopération en fonction de contraintes propres au couple (opérateur, SIAMD):

- la décision à prendre,
- le niveau de connaissance requis,
- le niveau d'interaction nécessaire.

Il doit aussi tenir compte de contraintes externes telles que les contraintes portant sur le temps maximal alloué pour le processus décisionnel.

Le choix du mode de coopération est réalisé de manière empirique. Cependant, il est possible de définir en fonction des processus décisionnels, un mode de coopération par défaut.

6.4. Processus décisionnels du pilotage hybride

Nous détaillons dans cette partie les différents types de processus décisionnels, au niveau distribué et superviseur du pilotage hybride. Les types de processus décisionnels sont présentés suivant la classification définie dans [Lévine et al. 89].

6.4.1. Processus de décision structurés

Trois processus décisionnels sont considérés comme structurés. Ils concernent le processus d'ordonnancement de la ressource goulet:

- détection des charges potentielles maximales par le superviseur,
- choix de la ressource goulet par le superviseur,
- l'ordonnancement partiel de la ressource goulet par le superviseur.

Dans ce cas, le problème d'ordonnancement n'est pas de complexité élevée (faible combinatoire, données exclusivement numériques, etc.), ce qui permet l'exploitation d'une algorithmique d'optimisation sans interaction avec l'opérateur (mode "autonome"). Cependant, il est parfois judicieux de réaliser ces processus décisionnels en mode "coopératif". L'intérêt de cette approche se situe au niveau de la prise en compte de la notion de désirabilité, lors de la réalisation d'un ordonnancement prévisionnel partiel par exemple.

6.4.2. Processus de décision peu structurés

- séquençement des ordres d'allocation: en particulier l'allocation de la première opération de chaque ordre de fabrication par le superviseur. Ce problème se pose aussi pour une SIP qui possède au minimum deux opérations à allouer,
- gestion des files d'attente du stock d'entrée: ce processus décisionnel concerne le choix du séquençement de la réalisation des opérations. Il ne concerne généralement pas la SIP goulet, sauf en cas de ré-ordonnancement local,
- gestion des variables d'action locales: chaque SIP est responsable d'un ensemble de ressources (production, convoyage, etc.). Il est donc nécessaire de spécifier les paramétrages de ces ressources (taux d'occupation, vitesse, etc.),
- gestion de chaque allocation: chaque SIP (y compris éventuellement le superviseur) doit allouer les opérations suivantes à réaliser sur un produit,
- pour la ressource ordonnancée: ré-ordonnancement local d'une partie de l'ordonnancement spécifié par le superviseur.

Pour l'ensemble de ces processus décisionnels, l'intégration de l'opérateur humain permet d'améliorer le pilotage pour le respect des objectifs.

6.4.3. Processus de décision non structurés

- choix de remise en question ou non de l'ordonnancement prévisionnel. Ce choix est fait par le superviseur,

- choix de ré-allocation ou non d'une opération en cours: une SIP peut envisager la ré-allocation d'une opération entamée,
- pour la ressource ordonnancée: choix de la remise en question ou non d'une partie de l'ordonnancement prévisionnel.

La complexité résulte de la difficulté pour le décideur à évaluer dans sa globalité (au niveau du système de production entier) et de manière relative les solutions envisageables.

Dans le cas où le décideur conclut au besoin de ré-allouer par exemple, la ré-allocation est de type peu structuré. Ainsi, dans notre cas:

- choix de la remise en question ou non (ré-allocation, ré-ordonnancement): processus non structuré,
- dans l'affirmative: processus peu structuré ou structuré.

Cette particularisation montre bien la différence, en terme de complexité, par exemple, entre la comparaison de deux ordonnancements prévisionnels (processus peu ou pas structuré) et l'élaboration d'un ordonnancement prévisionnel (processus peu structuré ou structuré).

6.5. Modèle de décision multicritère: spécification de la matrice de décision

Dans cette partie, nous allons nous attacher à la spécification générale d'un problème multicritère appliqué au pilotage hybride, pour les niveaux supervision et distribué. L'objectif est de spécifier un modèle générique applicable à l'ensemble des processus de décision non entièrement structurés [Trentesaux et al. 95c].

6.5.1. Modèle des données

Nous avons montré la nécessité de définir la notion de donnée incertaine. Cette définition permet d'intégrer:

- les contraintes inhérentes à la structure de contrôle distribué. Par exemple, la connaissance des SIP est locale. L'estimation des états des autres SIP doit être possible,
- les contraintes inhérentes au processus de production: données peu fiables, subjectives, etc.

Nous distinguons ainsi trois catégories pour la nature d'une donnée qui peut être:

- quantitative et certaine: nombre d'opérations réalisées, temps écoulé entre deux allocations, etc. Ces données sont de type numérique,

- quantitative et incertaine: temps opératoire, MTBF, MTTR, etc. Ces données sont typiquement des données statistiques (moyenne, écart type, fonction de répartition, etc.),
- qualitative: adéquation entre la qualification de l'opérateur et l'opération à réaliser, qualité générale du produit, etc. Ces données décrivent un ensemble de facteurs plus ou moins subjectifs, difficilement chiffrables, et sont généralement liées à la subjectivité de l'opérateur humain.

Les données quantitatives sont de nature cardinale (numérique), il est possible de travailler sur une échelle chiffrée. Les données qualitatives sont ordinales (non numériques), l'échelle de comparaison n'est pas chiffrable. Seul un ordre peut être établi sur ces données (première, deuxième, etc. ou encore: bon, très bon, excellent, etc.).

Il est donc nécessaire d'utiliser une méthode de modélisation permettant l'utilisation de telles données par le SIAMD. Nous restreignons cependant cette étude de faisabilité aux données numériques (données quantitatives certaines ou incertaines). Un autre avantage à intégrer un tel type de donnée permet d'introduire une marge d'incertitude. Ainsi, il est possible d'étudier l'influence de cette marge sur les décisions potentielles, ce qui permet naturellement une analyse de robustesse.

D'autre part, la modification dynamique de ces données doit également être possible. Nous proposons un processus d'enrichissement qui sera décrit dans le chapitre suivant.

6.5.2. Problématique multicritère des processus décisionnels

L'ensemble des processus décisionnels du niveau pilotage hybride est associé soit à la problématique de rangement discret soit à celle de choix discret, ce qui conduit à deux approches de résolution différentes. Cependant, l'uniformité des approches de résolution réduit la complexité de l'interprétation (interfaçage, dialogue homme-machine) par l'opérateur humain. C'est pourquoi il est judicieux (dans l'optique de développer ultérieurement une interactivité) de proposer une approche de résolution qui puisse être exploitée à la fois par la méthode de résolution de la problématique de choix et à la fois par la méthode de résolution de la problématique de rangement. La différence se situe au niveau de la phase d'agrégation.

6.5.3. L'ensemble des alternatives

L'ensemble des alternatives est à cardinal discret, et ce, pour tout type de décision peu ou pas structurée. De ce fait on peut présenter systématiquement à l'opérateur l'ensemble des choix possibles. De cette manière, la prise de décision est possible en étudiant conjointement tous les choix possibles.

Cependant, il est à noter qu'une limite de saturation mémorielle propre à l'opérateur humain peut réduire l'efficacité d'une telle méthode. Dans le domaine de la psychologie cognitive, il est courant d'utiliser l'axiome empirique qui stipule que l'être humain est, en moyenne, capable de prendre en compte simultanément un nombre d'objets-concepts ("items") compris entre cinq et neuf (en moyenne sept). Dans le cadre qui nous intéresse, le cas le plus contraignant est celui où, à la suite d'une demande d'allocation, plusieurs SIP sont capables de réaliser une même opération. Le nombre limite de SIP est donc compris entre cinq et sept pour une coopération efficace entre le SIAMD et l'opérateur humain. Dans le cas contraire, il sera nécessaire de proposer:

- soit une approche d'aide différente qui intègre la possibilité d'une évaluation portant sur un sous-ensemble de l'ensemble des alternatives,
- soit un regroupement par famille de ressources et allouer la responsabilité de la gestion de cette famille à une seule SIP. Cette généralisation de l'approche distribuée constitue une perspective importante.

Sous réserve du respect de cette condition, on peut uniformiser les schémas de présentation des alternatives en explicitant à chaque fois l'ensemble des cas possibles. Cette uniformisation soulage la tension mentale de l'opérateur. De cette manière, aucun effort de mémorisation des choix possibles ne lui est demandé.

6.5.4. L'ensemble des critères

6.5.4.1. *Elaboration de la liste de critères*

La liste de critères doit être structurée pour permettre une compréhension globale et intuitive de l'opérateur. Pour cette raison, Roy préconise l'exploitation d'une typologie par "axe de signification" [Roy et al. 93]. L'établissement de ces axes a pour objectif de schématiser ou représenter un point de vue du problème. Par exemple, il est possible d'associer à un axe particulier l'ensemble des critères qui ont la même dimension (coût, unité de temps, etc.). Nous détaillons dans le chapitre suivant une liste de critères structurés suivant trois axes.

La diversité des systèmes de production nuit à l'élaboration d'une liste exhaustive de critères. [Trentesaux et al. 95b] présente une liste de critères appliquée à un système de production de type flow-shop. Une liste de critères concernant le pilotage d'un atelier flexible de production est présentée en annexe C. Ces différents exemples montrent la spécificité des critères en fonction de la nature du système de production.

6.5.4.2. *Pondération des critères*

La pondération des critères doit être le reflet des éléments suivants:

- les objectifs globaux,

- les consignes de production locales: éviter les temps d'attente trop longs sur une ressource, équilibrer la charge entre deux ressources, etc.
- les souhaits des responsables concernant une ressource qui possède une caractéristique particulière (moindre coût, etc.),
- les souhaits des responsables vis-à-vis d'un produit: minimiser le temps de cycle global pour un produit particulier, etc.

L'intérêt d'une pondération dynamique réside ainsi dans la possibilité de particulariser chaque choix en fonction d'objectifs locaux, globaux, à caractère permanent ou non. Il est évident que la difficulté se situe au niveau de la particularisation de chaque poids, ainsi qu'au niveau d'incertitude associé à ces poids. Il existe de nombreuses méthodes permettant une telle analyse [Pomerol et al. 93], [Roy et al. 93].

Nous avons décrit un ensemble de spécifications dédiées au pilotage hybride concernant le développement et la gestion de la matrice de décision en fonction des alternatives, des critères et de leur pondération éventuelle. L'étude a porté sur l'application des concepts multicritère à la phase d'information et de conception des schémas de décision.

7. Conclusion

Nous avons justifié la possibilité d'intégrer l'opérateur humain dans les processus décisionnels au niveau du pilotage et montré que ceux-ci présentent un faible niveau de structuration. D'autre part, il est envisageable de proposer une méthode de résolution des processus selon une approche multicritère qui favorise en outre l'intégration de l'opérateur. Nous avons ainsi spécifié, dans un objectif de conception générique, un SIAMD pour le pilotage hybride.

L'approche présentée dans les deux derniers chapitres a ainsi permis de répondre à certains des défis de Parunak présentés dans le premier chapitre et caractérisé en terme de:

- "désirabilité": l'intégration de l'opérateur humain et de méthodes multicritères permet d'accroître fortement la caractéristique de désirabilité,
- "stochasticité": les différents aléas sont intégrés par la nature adaptative de la structure de pilotage,
- "tractabilité": le seul ordonnancement réalisé concerne une seule ressource, ce qui simplifie singulièrement la complexité du problème,
- "décidabilité": l'approche qui consiste à opérer localement par étapes et de manière complémentaire à un horizon plus lointain (superviseur) accroît le niveau de décidabilité,

- “chaos”: la nature du système de production est indépendante du système de commande, et la nature fortement chaotique (peu prévisible) des systèmes de production accentue la difficulté de prévision ou d'évaluation des conséquences. Bien qu'il soit possible d'évaluer localement l'impact d'une décision, il ne sera pas toujours possible de prévoir la conséquence de cette décision à plus long terme. L'aspect chaotique demeure.

Nous présentons dans le chapitre suivant un modèle de donnée certaine et incertaine. Nous montrons alors qu'il est possible d'élaborer un modèle de traitement multicritère qui tienne compte des spécifications présentées dans ce chapitre.

Chapitre IV

Modèle de données et modèle de traitement multicritère

1. Introduction

Nous présentons dans cette partie un modèle de données numériques certaines et incertaines qui a été spécifié dans le chapitre précédent. Ce modèle se base sur la définition d'une entité générique qui peut être associée à une donnée certaine ou à une donnée incertaine. Nous développons le concept d'enrichissement pour une donnée incertaine. Ce concept permet d'améliorer la connaissance sur cette donnée en proposant un processus d'enregistrement dynamique de l'information réelle associée à cette donnée. L'objet de cette proposition est de permettre l'affinement de la connaissance et de la qualité de l'information mémorisée.

Nous présentons ensuite une étude qui porte sur le développement d'une méthode multicritère qui répond au cahier des charges imposé dans le chapitre précédent. Le but de cette étude est de montrer la faisabilité de l'intégration d'une méthode multicritère supportant les données incertaines et les contraintes imposées par le contexte de pilotage hybride. Nous restreignons cependant l'étude à la problématique de rangement.

2. Modélisation des données

2.1. Approche mathématique choisie

Les outils actuels permettant de modéliser le manque d'information portant sur une donnée numérique relèvent principalement des mathématiques statistiques et de la logique floue. Ces outils ont une forte influence sur la simplicité du modèle global et la compréhension par l'utilisateur de la méthode d'aide à la décision.

Ainsi l'apparente simplicité de la modélisation des données en logique floue occulte la gestion complexe des relations entre ces données. En effet, il suffit de considérer:

- le nombre et la complexité des différentes lois de composition interne possibles. Par exemple [Fodor et al. 94] décrit sept formes différentes (les plus courantes) pour représenter la seule notion d'implication,
- le nombre de propriétés diverses qui caractérisent ces données. Par exemple, les sept formes d'implication présentent chacune entre 5 et 12 propriétés spécifiques. Ainsi une étude précise de toutes ces propriétés doit être menée préalablement à toute exploitation.

En outre, les relations de base (et, ou, implication, équivalence) qui constituent l'extension du domaine booléen au domaine continu $[0,1]$, peuvent aussi présenter chacune plusieurs formes généralisées.

La logique floue, a priori outil très intuitif du point de vue modélisation, est ainsi très lourde du point de vue manipulation.

Par ailleurs, les mathématiques des variables statistiques (somme de deux variables aléatoires, inter-dépendance d'événements aléatoires, etc.) présentent elles aussi un certain niveau de complexité au niveau de l'interaction entre variables statistiques. En outre, dans le cas d'événement à caractère aléatoire, elles restreignent souvent le modèle de la réalité à des distributions pré-établies (gaussienne, poisson, etc.). Cependant l'avantage de l'approche par les mathématiques statistiques est l'existence de la notion de probabilité d'occurrence d'un événement, de fonction de répartition, de fréquence cumulée, etc. qui est fortement compatible avec la gestion de l'incertitude associée à une donnée.

Ainsi ces méthodes ne sont pas complètement satisfaisantes dans un objectif d'intégration de l'opérateur humain dans le processus décisionnel.

Nous avons développé une approche mixte qui répond aux spécifications du chapitre précédent. Cette approche s'appuie sur la modélisation floue appliquée au domaine discret et exploite la notion de probabilité et de fréquence cumulée propre aux mathématiques statistiques. Nous montrons que cette approche permet de prendre en compte les contraintes:

- d'accessibilité de coût et de fiabilité de l'information,
- d'évolution dynamique des états du système et des données. La méthode développée nous permet entre autre de définir la notion d'enrichissement dynamique,
- de simplicité de modélisation,
- d'intégration des données certaines et incertaines sous un même formalisme.

Nous présentons la méthode sur un exemple avant de développer le modèle mathématique correspondant.

Soit le temps opératoire pour une ressource pour une opération donnée. Ce temps, qui est en première approximation fixé, évolue en fait en fonction de l'environnement de production (usure de l'outil, opérations de manutention, etc.). Le principe que nous proposons est de relever puis mémoriser les occurrences d'événements de "début opération" et de "fin opération" pour chaque type d'opération pour cette ressource. L'avantage de cette approche est de proposer une image fidèle de la réalité sans la restreindre à un modèle (fonction de répartition gaussienne, etc.). De ce fait, ces données mémorisées (qui sont de nature incertaine) peuvent être exploitées avec un niveau de fiabilité élevé lors d'une prise de décision. Il est ainsi nécessaire de proposer une structuration des données capable de supporter cette approche. Nous avons ainsi développé:

- l'intégration de l'incertitude (et de la certitude) par la notion de structure de donnée générique. Nous montrons qu'une telle donnée peut représenter à la fois une donnée certaine et une donnée incertaine,
- le principe de mémorisation des occurrences par le développement de la fonction d'enrichissement. Nous montrons que la méthode se base sur le principe de l'inversion de la fonction de fréquence cumulé. L'enrichissement peut concerner ainsi les variables d'état (temps opératoire, temps d'attente, etc.) et les variables d'action (rangements issus de la méthode multicritère, résultats de calculs intermédiaires, etc.),
- le principe d'exploitation de ces données durant la résolution des processus de décision (méthode multicritère intégrant la notion de donnée générique).

Ces différents points sont détaillés dans les parties suivantes.

2.2. Modèle des données incertaines

Soit X un ensemble discret de valeurs défini par:

$X = \{x_1, x_2, \dots, x_n\}$ avec $x_k \in \Pi$, un ensemble fini quelconque où il existe une relation. Les éléments de X sont distincts deux à deux et sont indicés de la sorte: $x_i < x_j$ si et seulement si $i < j$.

Soit P une fonction de X dans $[0,1]$ définie par:

$$P: X \rightarrow [0,1]$$

$$x_k \mapsto P(x_k) = p_{k, k \in [0,n]}$$

avec $p_0 = 0$, $p_{k, k \neq 0} \in]0,1]$ et $\sum_{k=0}^n p_k = 1$.

Soit F une fonction de $[0,1]$ sur X définie de la manière suivante:

$$F:]0, 1] \rightarrow X$$

$$p \mapsto F(p) = x_i \text{ et } F(0) = x_1$$

A $p \in]0, 1]$ vérifiant la double inéquation

$$\sum_{k=0}^{i-1} p_k < p \leq \sum_{k=0}^i p_k$$

on associe l'indice i par la fonction F .

On montre facilement que F est définie croissante par escalier (X est discret et possède une relation d'ordre).

Ceci est résumé sur la figure 1.

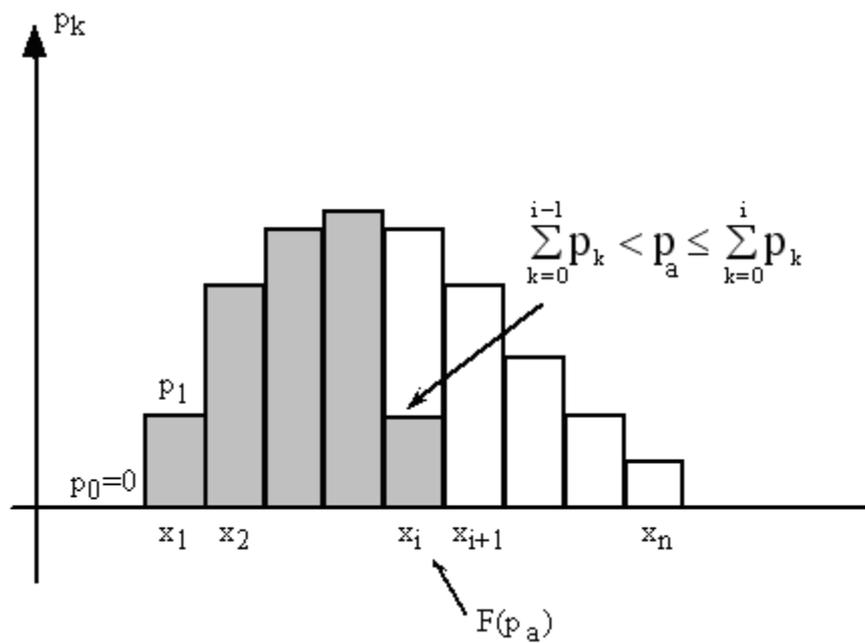


Figure 1: Obtention de $x_i = F(p = p_a)$.

On montre aisément que:

$$F(p_0 = 0) = x_1 \text{ (par définition)}$$

$$F(p_0 + p_1) = x_1$$

...

$$F(p_0 + p_1 + \dots + p_{n-1}) = x_{n-1}$$

$$F(p_0 + p_1 + \dots + p_n = 1) = x_n$$

et de manière plus générale:

$$F\left(\sum_{k=0}^i p_k\right) = x_i \text{ pour } i \in \{1, 2, \dots, n\} \text{ et } F(0) = x_1.$$

F se comporte comme une fonction réciproque de la fonction cumulative croissante des p_k .

D'autre part, la relation:

$$\sum_{k=0}^n p_k = 1$$

permet de spécifier $\phi_W(x_k)$, fonction de répartition discrète (fonction des fréquences cumulées), définie de manière classique par une variable aléatoire discrète W telle que:

Soit P_n l'ensemble discret $P_n = \{p_0, p_1, p_2, \dots, p_n\}$ ($p_i \in \mathfrak{R}$).

alors:

$$\begin{aligned} \phi: X &\rightarrow P_n \\ x_k &\mapsto \phi_W(x_k) = p(W \leq x_k) \end{aligned}$$

On montre que:

$$\phi_W(x_k) = \sum_{l=0}^k p_l$$

Ce qui conduit à définir F comme la pseudo-inverse de $\phi_W(x_k)$ de la manière suivante:

Soit $\alpha_i = \sum_{k=0}^i p_k$ et $A = \{\alpha_0, \alpha_1, \dots, \alpha_n\}$.

Alors si $p \in]\alpha_{i-1}, \alpha_i]$ on a:

$$\begin{aligned} F(p) &= \phi_W^{-1}(\alpha_i) \text{ dans le cas où } p \in A - \{\alpha_0\} \text{ et} \\ F(0) &= \phi_W^{-1}(\alpha_0) \text{ si } p = p_0 \end{aligned}$$

De cette manière F est pseudo-équivalente à ϕ_W^{-1} pour $p \in [0, 1]$.

Cette pseudo-équivalence permet d'associer à F toutes les caractéristiques de l'inverse d'une fonction de répartition discrète, ce qui offre en particulier la possibilité d'utiliser les outils statistiques classiques et les notions sous-jacentes.

Nous définissons un modèle de donnée générique comme un quadruplet (X, P, F, m) . Ce quadruplet est composé d'un ensemble de départ X que nous appellerons dans la suite "ensemble des valeurs possibles", de deux fonctions P et F dont la nature et les propriétés ont été définies auparavant, et d'un paramètre m appelé "coefficient d'enrichissement" dont le rôle sera précisé dans le paragraphe suivant. $P(x_k)$ est la fonction donnant la probabilité d'occurrence associée à la valeur x_k et $F(p) = \phi_W^{-1}(\alpha_i)$ la pseudo-inverse définie auparavant.

L'ensemble des données génériques est noté Ω . Un élément de Ω est noté ω .

On peut alors définir les deux types de données suivants:

- une "donnée numérique incertaine" est un quadruplet (X, P, F, m) où $\text{card}(X) > 1$.

Dans ce cas, la donnée peut prendre au moins deux valeurs,

- une "donnée numérique certaine" est une donnée où $X = \{x_0\}$. Il est évident que dans ce cas $F(p) = x_0, \forall p \in [0, 1]$.

Nous constatons ainsi que notre proposition constitue en quelque sorte une généralisation de la notion de donnée numérique.

2.3. Enrichissement d'une donnée incertaine

Un autre aspect intéressant du modèle proposé est qu'il permet l'enrichissement dynamique d'un élément $\omega \in \Omega$. L'enrichissement d'une donnée générique $\omega \in \Omega$ consiste à mémoriser l'occurrence d'un événement au sein de cette donnée afin de déterminer une fréquence d'apparition pour chaque valeur possible. L'enrichissement s'appuie sur le principe de la récurrence. Nous noterons $\omega_m = (X, P, F, m)$.

La fonction d'enrichissement E_m^k de ω_m sur la valeur x_k est définie de la manière suivante:

$$E_m^k: \Omega \rightarrow \Omega$$

$$\omega_m \mapsto \omega_{m+1} = E_m^k(\omega_m) = (X, P', F', m+1)$$

Nous appelons événement associé à un élément x_k l'accroissement de $p_k = P(x_k)$. Cet accroissement est décrit par un gain linéaire fonction de p_k et de m , $\Delta(p_k, m) = \Delta_m^k$. Il est évident que de nombreuses relations peuvent décrire ce gain. Cependant, pour des raisons qui seront justifiées dans la partie suivante, nous posons:

$$p'_k = \Delta_m^k = \frac{m \cdot p_k + 1}{m + 1}$$

La normalisation modifie les autres paramètres $p_{i, i \neq k}$ de la façon suivante:

$$p'_{i, i \neq k} = \frac{m}{m + 1} \cdot p_{i, i \neq k}$$

La fonction P' peut alors être définie comme suit:

$$P': X \rightarrow [0, 1]$$

$$x_{i, i \neq k} \mapsto P'(x_i) = p'_i = \frac{m}{m + 1} \cdot p_{i, i \neq k}$$

$$x_k \mapsto P'(x_k) = p'_k = \Delta_m^k = \frac{m \cdot p_k + 1}{m + 1}$$

La démonstration de la normalisation $\sum_{i=1}^n p'_i = 1$ est donnée en annexe B.

La fonction F' est la fonction F appliquée à P' .

Le fait de proposer une approche récurrente nécessite de définir l'état initial pour ω : ω est à l'état initial d'indice k si et seulement:

$$\exists! x_k \in X / P(x_k) \neq 0 \text{ et } m=1.$$

La normalisation entraîne nécessairement $P(x_k) = 1$.

Il est clair que l'on peut associer à une donnée générique à l'état initial une donnée numérique certaine ω' définie de la manière suivante: $\omega' = (\{x_k\}, 1, x_k, 1)$.

Cette méthode s'applique dans le cas d'un ensemble X constant. Elle peut cependant se généraliser à l'extension de X par ajout d'une nouvelle valeur x_k . La pré-condition portant sur X ordonné impose à cette nouvelle valeur d'être intercalée correctement. Alors le nouvel ensemble X' se substitue à X . Puisque l'enrichissement concerne x_k , on posera avant enrichissement $P(x_k) = 0$. Naturellement, après enrichissement, $p'_k = \Delta(p_k, m) = \Delta(0, m) = \frac{1}{m+1}$.

Un exemple d'application ainsi que la justification de la structure du gain sont fournis en annexe B.

2.4. Avantages du modèle

Un intérêt de la représentation proposée est qu'elle permet (outre les propriétés de généralisation de la structure d'une donnée et de l'enrichissement dynamique) de retrouver les différentes valeurs (x_i) variant de x_1 à x_n à partir d'un paramètre $p \in [0,1]$ donc inclus dans un domaine constant. Dans le cas où l'étude porte sur l'ensemble des réels, les (x_i) peuvent correspondre à différentes valeurs d'une même donnée. Par exemple, si la donnée est un temps opératoire ($\Pi \equiv \mathfrak{R}$), x_1 correspond au temps minimum relevé et x_n au temps maximum. L'appel de la fonction F avec le paramétrage $p=0$ (resp. 1) retournera la valeur minimale x_1 (resp. maximale x_n). Ainsi, les valeurs de p permettent l'obtention des (x_i) . Dans un cas particulier, il peut être intéressant d'obtenir une approximation de la médiane ($p=.5$) ou de la moyenne, ou encore en restreignant le domaine de définition de p à un sous-ensemble de $[0,1]$.

Nous utiliserons d'autre part la pseudo-équivalence $F(p) = \phi_W^{-1}(\alpha_i)$ qui permet de se placer dans le cadre des mathématiques statistiques. A p nous associons une probabilité et à l'ensemble X une variable aléatoire W . Par conséquent, W peut être utilisée comme une variable aléatoire associée à une répartition uniforme entre 0 et 1. F donne à p l'image x_i considérée de manière aléatoire selon la fonction de distribution discrète associée. Nous préciserons ultérieurement lors de la présentation de l'algorithmique multicritère l'utilisation et les avantages de cette approche.

2.5. Application au cas du pilotage hybride

Cette représentation théorique peut être appliquée aux données propres du niveau pilotage hybride: variables d'état (comme le temps opératoire) et d'action (comme un résultat établi sous forme de rangement).

Nous posons plusieurs restrictions sur l'ensemble X des valeurs possibles:

- le domaine des valeurs possibles est l'ensemble des réels: $\Pi \equiv \mathfrak{R}$. Cette restriction ne réduit en rien les performances de la méthode car dans le cas du pilotage, la quasi-totalité des données génériques sont cardinales et possèdent généralement une unité (coût, temps, vitesse, etc.).
- les valeurs possibles x_k vérifient la condition suivante ($X = \{x_1, x_2, \dots, x_n\}$ est l'ensemble ordonné des valeurs croissantes possibles):

$$\exists \varepsilon > 0 / \forall i \in [1, n-1], x_{i+1} - x_i = \varepsilon$$

Cette condition exprime le fait qu'il existe un écart constant ε entre deux valeurs possibles consécutives. C'est le cas de l'ensemble X de la donnée ω''''_α citée en annexe B où $\varepsilon = 2$. Cette simplification permet de décrire entièrement X à partir de ε , de x_1 et de x_n . Les valeurs extrêmes (x_1, x_n) déterminent les limites acceptables et ε la différence minimale requise permettant de différencier deux valeurs possibles de X . La cohérence de ces trois paramètres dépend donc étroitement des valeurs des données physiques (pour un temps opératoire moyen d'une heure un ε est de l'ordre de la minute, non de la micro-seconde). D'autre part, nous avons montré qu'un enrichissement en dehors des limites (x_1, x_n) est possible, à condition de créer un nombre suffisant de variables intermédiaires pour que la condition d'un écart constant soit respectée. Une étude préalable doit être réalisée pour fixer ε . En effet, seuls les changements d'écart entre les valeurs possibles multiples de ε sont possibles,

- nous supposons que le gain est de la forme $\Delta_m^k = \lambda(m) \cdot \frac{m \cdot p_k + 1}{m + 1}$, (cf. annexe B avec $\lambda(m) = 1$),

- lorsque deux données incertaines sont comparées, on suppose qu'elles sont caractérisées par un même écart ε et qu'elles possèdent au moins une valeur commune (généralement 0). Cette hypothèse réduit considérablement la difficulté d'effectuer une comparaison. En effet, la donnée résultant de la comparaison sera composée de la réunion des deux ensembles X^1 et X^2 , ce qui a un sens car l'écart ε identique pour X^1 et X^2 . En effet, les contraintes précédentes permettent de définir une relation de récurrence qui peut alors décrire les deux ensembles (suite arithmétique définie par une valeur initiale et une raison),
- les enrichissements décrits sont sans ambiguïté sur une valeur possible $x_k \in X$. Si l'enrichissement porte sur une valeur non incluse dans cet ensemble, mais bornée strictement par deux éléments de X , il est nécessaire de l'associer à

une valeur possible existante. La méthode utilisée de “la borne inférieure” est alors la suivante:

L’enrichissement portant sur une valeur possible $x \notin X$ telle que:

$$\exists (x_k, x_{k+1}) \in X^2 / x_k < x < x_{k+1}$$

est réalisé sur $x_k \in X$.

Une autre possibilité (méthode de “la valeur la plus proche”) consiste à associer x à la valeur possible la plus proche définie de la manière suivante:

Soit $D = \{i/i \in [1, \text{card}(X)] \in \mathbb{N}\}$

$$\exists k \in D / \forall i \in D - \{k\}, |x_k - x| < |x_i - x|$$

Si x est équidistant à deux éléments de X , le choix de l’une ou l’autre valeur est possible.

La structure de données proposée constitue ainsi une généralisation de la notion de donnée et supporte l’enrichissement dynamique avec gain non constant. Cette approche permet, en accord avec les spécifications présentées dans le chapitre précédent, de ne prendre en compte que les données réelles (brutes) du système de production. La notion de donnée générique peut aussi être exploitée avantageusement comme élément mathématique issu d’un calcul (rangement intermédiaire, etc.). Ce dernier point sera mis en évidence lors de la présentation de la méthode multicritère.

3. Structure et représentation des matrices génériques de décision multicritère

Dans cette partie nous proposons une structure matricielle générique de décision à laquelle sont associées des données génériques, et comportant trois matrices (critères, alternatives et évaluations).

3.1. Matrice générique des critères

Nous avons retenu une élaboration de listes de critères suivant trois axes de signification reposant sur les notions de coût, de délai et de qualité. L’intérêt d’une telle représentation est triple:

- elle améliore l’adéquation entre la réalité du processus de production et la perception qu’a l’opérateur de la problématique du pilotage,
- elle facilite la tâche de l’opérateur lors de la phase de validation de la liste.

Les poids des critères dépendent de la structure et des objectifs adoptés. D'autre part, il est envisageable de résoudre une partie des problèmes de pondération (sensibilité, adéquation avec les préférences du décideur, etc.) en utilisant des données incertaines comme poids plutôt que des données certaines. Des études ont montré la viabilité de cette approche sans avoir toutefois étudié les propriétés mathématiques du modèle correspondant.

La matrice des critères est composée des éléments suivants:

- nom du critère,
- nature des données associées au critère (certaines ou incertaines),
- poids du critère,
- objectif recherché associé au critère (minimisation ou maximisation).

3.2. Matrice générique des alternatives

La matrice des alternatives est composée de l'ensemble des choix possibles, identifiés de manière unique par un nom et une référence numérique.

3.3. Matrice générique des évaluations

Les évaluations sont composées de données génériques certaines ou incertaines identifiées par un nom pour les données incertaines. La nature des données est la même pour chaque critère.

La représentation générale d'une structure matricielle générique de décision est donnée sur la figure 2.

Matrice de critères		
Nom Critère #1	Nom Critère #c	Nom Critère #nc
Nature des données	Nature des données	Nature des données
Poids critère #1	Poids critère #c	Poids critère #nc
Nature critère #1	Nature critère #c	Nature critère #nc

Matrices des alternatives		Matrice des évaluations		
1	Nom choix #1	Evaluation #1;1		
a	Nom choix #a	...	Evaluation # a;c	
na	Nom choix #na			Evaluation #na;nc

Figure 2: structure matricielle générique de décision.

On note n_a le nombre d'alternatives et n_c le nombre de critères.

La figure 3 décrit un exemple de structure matricielle de décision générique dans le cas d'un processus décisionnel d'allocation.

Matrice de décision générique			
Matrice générique critères			
Tps. Op.	P.T.L	Nb Op.	Préc. outils
Estim.	Estim.	Num.	Num.
2	5	3	3
Min	Min	Min	Min
Matrice générique alternatives		Matrice générique évaluations	
1	SIP 4	[Tps. Op. SIP 4]	[P.T.L SIP 4]
2	SIP 3	1	0.001
3	SIP 5	2	0.001
4	SIP 1	4	0.009
		[Tps. Op. SIP 5]	[P.T.L SIP 5]
		0	0.001
		[Tps. Op. SIP 1]	[P.T.L SIP 1]

Figure 3: un exemple de l'ensemble des matrices de décision génériques

Les critères pris en compte sont respectivement le temps opératoire, la date de disponibilité, le nombre d'opérations effectuées et la précision de l'outil pour l'opération considérée.

Dans le cas présent, la matrice montre que quatre SIP sont capables d'effectuer l'opération en question. La matrice est complétée par les données présentes dans les messages d'acceptances émis par quatre SIP comme les temps opératoires.

4. Choix de la méthode multicritère

Nous avons défini au chapitre précédent les différentes contraintes que la méthode multicritère doit respecter pour permettre:

- que l'ensemble des processus de décisions relèvent de la problématique de choix ou de rangement,
- l'intégration de données de natures différentes (certaines ou incertaines),
- que l'ensemble des alternatives soit à cardinal peu élevé (afin de respecter le principe de "non saturation mémorielle"),
- que l'ensemble des critères utilisés pour un même processus décisionnel discret soit à cardinal peu élevé (idem).

Ces contraintes sont des contraintes fortes qui nous conduisent à développer une méthode multicritère particulière adaptée à la problématique du pilotage. De

plus, le choix de la méthode multicritère est fortement conditionné par le niveau de connaissance des méthodes multicritères que possède l'opérateur.

L'intégration de données incertaines risque de nuire à la compréhension de l'approche par l'opérateur. Voir par exemple, les méthodes AHP [Saaty 80], Qualifex [Paelinck 78], PRIAM [Lévine et al. 86] et Electre III [Roy 78]. Ces méthodes sont, indépendamment de la problématique qui nous concerne, relativement complexes et intègrent des règles particulières pour la gestion de l'incertain comme les relations entre deux données qualitatives.

Aussi, nous arrivons à la conclusion qu'il n'existe pas de méthode présentant toutes les caractéristiques nécessaires pour être appliquée à l'aide au pilotage hybride et que chacune à leur manière propose une solution à une facette de notre problématique [Trentesaux et al. 94b], [Trentesaux et al.??b], [Dindeleux et al. 95]. Ce point est confirmé par le fait qu'à notre connaissance, il n'existe pas de méthodes multicritère appliquée au niveau opérationnel (cf. chapitre précédent). Un compromis portant sur les caractéristiques de ces méthodes a donc été défini afin de tester la faisabilité de l'intégration du multicritère au niveau du pilotage hybride. Notre choix s'est ainsi porté en premier lieu sur la méthode Prométhée de [Brans et al. 85] en raison de sa simplicité, de sa facilité de compréhension et de son aspect modulaire. Nous présentons brièvement cette méthode et décrivons ses avantages et inconvénients.

4.1. La méthode Prométhée

Le lecteur trouvera des présentations de cette méthode dans [Brans et al. 85], [Pomerol et al. 93], [Roy et al. 93]. Nous présentons uniquement les caractéristiques principales décrites dans [Pomerol et al. 93]:

Prométhée est une méthode dédiée à l'analyse de données certaines en environnement discret pour une problématique de rangement.

Pour chaque paire d'actions (a_i, a_k) , on détermine pour chaque critère j la distance entre les évaluations de ces deux actions:

$$d_{ik} = a_{ij} - a_{kj}$$

Une fonction de préférence, relative à ce critère j , $S_j(a_i, a_k) = S_j(d_{ik})$ représente sous la forme d'un vrai ou d'un pseudo critère la préférence (normalisée dans le segment $[0,1]$) de l'utilisateur. Une bibliothèque de fonctions correspondant aux différents cas de préférence est fournie. Ces critères s'expriment en fonction d'un ou plusieurs paramètres (seuil, écart-type, etc.). Par exemple, S peut avoir la forme suivante:

$$S_j(a_i, a_k) = S_j(d_{ik}) = \begin{cases} 0 & \text{si } d_{ik} \leq q \text{ (indifférence)} \\ \frac{d_{ik} - q}{p - q} & \text{si } q \leq d_{ik} \leq p \\ 1 & \text{si } d_{ik} > p \text{ (préférence stricte)} \end{cases}$$

Cette relation est représentée graphiquement sur la figure 4.

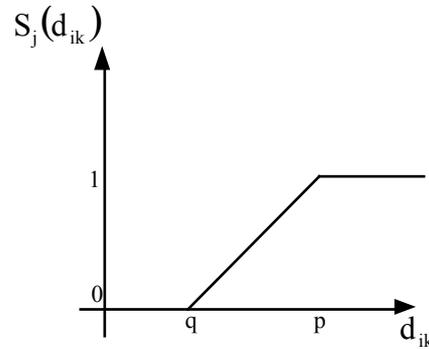


Figure 4: Exemple d'une relation de préférence linéaire avec seuil

D'autres types de relations de préférence sont prédéfinies (six dans la version de Prométhée que nous avons utilisée).

Après la déclaration des préférences de l'utilisateur, un indice de préférence agrège, par addition et pondération (poids des critères w_j normalisés: $\sum_j w_j = 1$), pour chaque critère j , chacune de ces fonctions de préférence relative:

$$c_{ik} = \sum_j w_j \cdot S_j(d_{ik})$$

A partir de ces indices, on détermine de manière additive mais non pondérée, les valeurs dénommées "flux sortant" ϕ_i^+ et "flux entrant" ϕ_i^- pour chaque alternative i :

$$\phi_i^+ = \sum_k c_{ik}$$

$$\phi_i^- = \sum_k c_{ki}$$

Ces deux valeurs sont associées à chaque action et représente l'évaluation, cette fois-ci globale, de chacune de ces actions par rapport aux autres actions.

Ces données sont représentées sous forme matricielle (figure 5).

$$\left(\begin{array}{c|cccc} c_{ik} & \text{Alternative\#1} & \text{Alternative\#2} & \text{Alternative\#a} & \text{Alternative\#na} \\ \hline \text{Alternative\#1} & c_{1,1} & c_{1,2} & c_{1,a} & c_{1,na} \\ \text{Alternative\#2} & c_{2,1} & c_{2,2} & c_{2,a} & c_{2,na} \\ \text{Alternative\#a} & c_{a,1} & c_{a,2} & c_{a,a} & c_{a,na} \\ \text{Alternative\#na} & c_{na,1} & c_{na,2} & c_{na,a} & c_{na,na} \end{array} \right) \left(\begin{array}{c} \phi_i^+ = \sum_k c_{ik} \\ \phi_1^+ \\ \phi_2^+ \\ \phi_a^+ \\ \phi_{na}^+ \end{array} \right)$$

$$\left(\begin{array}{c|cccc} \phi_i^- = \sum_k c_{ki} & \phi_1^- & \phi_2^- & \phi_a^- & \phi_{na}^- \end{array} \right)$$

Figure 5: indices de préférence et flux

Le préordre (total ou partiel) est ainsi établi à partir des flux entrants, flux sortants et du flux net $\phi_i = \phi_i^+ - \phi_i^-$. Ce dernier est représenté sur la figure 6.

$$\left(\begin{array}{c|c} \phi_i = \phi_i^+ - \phi_i^- & \\ \hline \text{Alternative\#1} & \phi_1 \\ \text{Alternative\#2} & \phi_2 \\ \text{Alternative\#a} & \phi_a \\ \text{Alternative\#na} & \phi_{na} \end{array} \right)$$

Figure 6: Flux net pour chaque alternative

Dans le cas de Prométhée II, le préordre de surclassement est établi en fonction de la monotonie croissante du flux net (préordre total).

Une autre possibilité est offerte avec la méthode Prométhée I qui propose une relation de surclassement plus faible en comparant directement les flux entrants et sortants sans tenir compte du flux net. Le résultat est un préordre partiel.

4.2. Inconvénients

Ils proviennent principalement de la création d'un préordre. En effet, la méthode Prométhée est associée à la problématique de rangement et non à la problématique de choix. Prométhée propose, après analyse soit un préordre partiel sur les alternatives (rangement avec ex-aequo), c'est la méthode Prométhée I, soit un préordre total (rangement sans ex-aequo), c'est la méthode Prométhée II. Dans les deux cas, une relation transitive est créée. Cette relation est établie en comparant les flux. Cette méthode n'est donc pas exploitable dans le cas de la problématique de choix.

4.3. Avantages

Les avantages de cette méthode sont:

- le domaine d'étude discret (qui correspond à notre cas d'étude),

- la simplicité et l'aspect intuitif de la méthode (qui accroît pour l'opérateur la facilité de compréhension et accroît la confiance en soi),
- la pré-définition de critères (ce qui simplifie la tâche de l'opérateur),
- l'automatisation aisée (ce qui est important dans le cas où le niveau de coopération est réduit au fonctionnement autonome).

4.4. Proposition

Le processus de décision qui nous concerne relève des deux problématiques de rangement et de choix. Dans une perspective d'intégration de l'opérateur, une approche de résolution similaire pour les deux problématiques doit être envisagée afin de proposer une interaction efficace. Ainsi, la différenciation entre les deux résolutions ne doit donc apparaître qu'en phase d'agrégation. Les étapes précédant cette phase doivent être identiques. C'est pourquoi nous proposons la méthode suivante:

- la méthode Prométhée est étendue pour intégrer les données génériques. Dans tous les cas, Prométhée est exploitée jusqu'à la phase de calcul des indices de préférence.
- après ce calcul, et selon la problématique traitée, deux approches de résolution sont envisageables:

A. Si la problématique est celle de rangement:

1. Une première méthode s'appuyant sur la notion de flux positif, de flux négatif et de flux net est utilisée, ce qui constitue en fait l'exploitation complète de la méthode Prométhée étendue.
2. La seconde méthode est basée sur la notion de degré de dominance. Le rangement peut être réalisé de deux manières différentes, selon la définition des relations de préférence.

B. Si la problématique est celle du choix:

Une méthode de détection de noyau (valué ou non) est développée: la notion de coupure est utilisée dans ce cas. De nombreux outils ont été développés dans ce cadre pour la gestion des relations binaires sous forme de graphes valués ou non et l'obtention des noyaux associés à ces graphes. L'existence des noyaux, qui peuvent être multiples ou impossibles à déterminer, représente une difficulté importante étroitement liée à la présence de cycles dans le graphe. D'autres approches ont été envisagées consistant à relâcher les contraintes fortes d'appartenance au noyau. Ces approches ont mené en particulier à la notion de quasi-noyau qui a pour avantage d'imposer l'existence au minimum d'un noyau.

Le point B ne sera pas étudié dans ce mémoire. Nous n'étudierons que la résolution de la problématique de rangement. Cependant nous pensons qu'il est envisageable en première approximation de l'étude de faisabilité d'exploiter la méthode de résolution de la problématique de rangement dans une optique de choix. En effet, bien que théoriquement les problématiques soient différentes, elles sont beaucoup moins incompatibles si l'on considère le point de vue utilisateur: l'ordre proposé élaboré par une méthode de rangement peut être considéré dans le cas d'une problématique de choix comme une aide à la décision pour l'analyse de sensibilité de paramétrages possibles. C'est cette remarque qui a conduit [Champetier et al. 95] à proposer, indépendamment de nos travaux de recherche, Prométhée comme méthode multicritère dédiée à la problématique de choix. Nous allons montrer sur un exemple utilisant Prométhée II comment la problématique de rangement peut intuitivement contribuer à l'analyse de sensibilité des paramètres dans un cadre de problématique de choix.

Supposons que l'utilisateur auquel la méthode propose un rangement de trois alternatives soit confronté aux deux situations suivantes (figure 7):

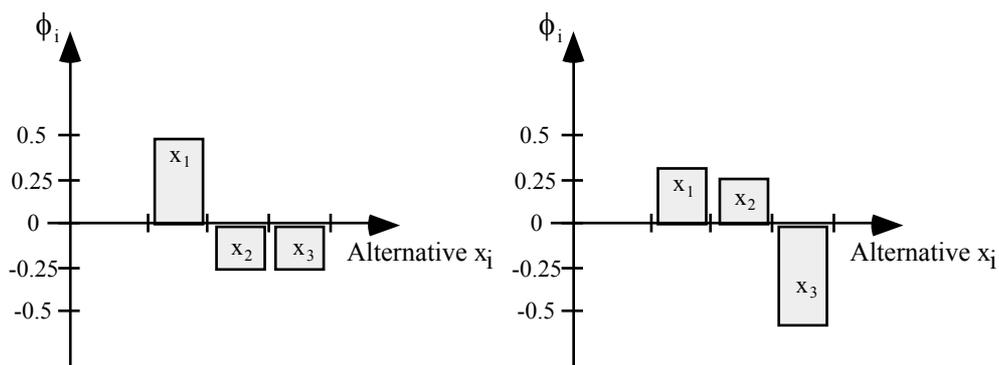


Figure 7: Deux cas possibles de rangement selon le flux net

On voit clairement que les valeurs du flux net apportent une information au niveau du rangement relatif des alternatives. Dans le premier cas, la première alternative est clairement favorisée, alors que dans le second cas, les deux premières alternatives semblent également satisfaisantes. En se plaçant dans la problématique de choix, l'opérateur sera conscient que son choix se portant sur la première alternative est à faible risque dans le premier cas. Dans le second cas, ce choix est peut-être moins judicieux, car fortement lié à la pertinence des paramètres (type de critère, pondération, fiabilité des évaluations, etc.).

Les méthodes qui ont trait à la problématique de choix proposent généralement des analyses de sensibilité de paramétrage, mais la complexité de la méthode rend l'approche très difficilement automatisable sous forme de procédures (informatique, etc.).

Les méthodes proposées ci-dessous possèdent de manière intrinsèque une procédure d'analyse de sensibilité.

5. Première méthode M1

C'est en quelque sorte la généralisation de Prométhée aux données incertaines. La figure 8 représente les différentes étapes de cette méthode.

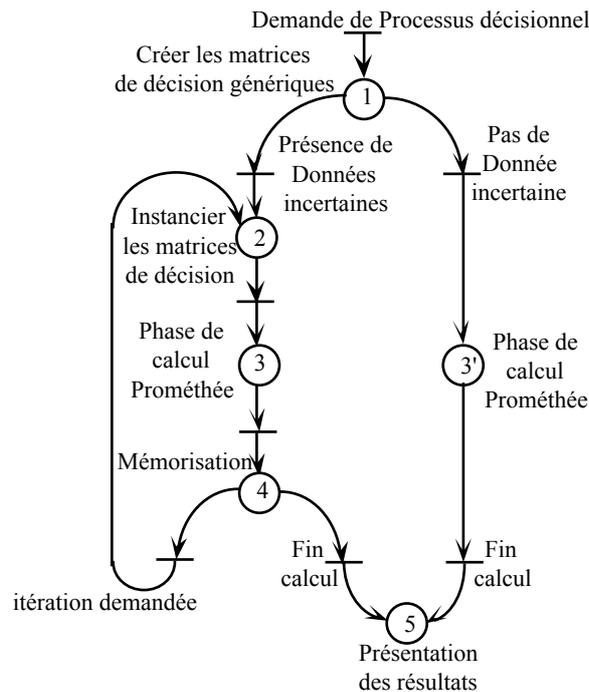


Figure 8: méthode de résolution

Dans le cas où la matrice d'évaluation ne contient que des données numériques certaines, le calcul réalisé par Prométhée n'est itéré qu'une fois (étape 3') et conduit à la proposition d'un pré-ordre (étape 5). Si au moins un critère est associé à des données incertaines, il faut procéder à plusieurs itérations (bouclage des étapes 2, 3 et 4).

5.1. Etape 1: création des matrices génériques

L'étape 1 est une étape préliminaire à la phase de calcul. Elle consiste à utiliser les informations disponibles pour créer les trois matrices génériques. La matrice générique des évaluations peut comporter des éléments numériques certains ou incertains.

5.2. Etape 2: instantiation des matrices génériques

Cette étape est réalisée avant chaque étape de calcul. Elle consiste à créer une matrice de données numériques certaines à partir de la matrice générique des évaluations. La méthode exploite l'algorithme suivant (figure 9).

```

Fonction Instantiation (mg: Matrice Générique): Matrice Numérique
variables
    mn: matrice de données numériques certaines
    mg: matrice générique
    lig, col: entiers
:début
Pour lig variant de 1 à mg.nb_lignes
:début
    Pour col variant de 1 à mg.nb_colonnes
:début
        si mg(lig,col) est une donnée certaine alors
:début
            mn(lig,col)=mg(lig,col)
:fin
        sinon //c'est une donnée incertaine
:début
            mn(lig,col)=mg(lig,col).F(s(p)) // appel de F
:fin
:fin
:fin
retourner mn
:fin
  
```

Figure 9: algorithme d'instanciation des matrices génériques

Cet algorithme crée une matrice composée de données numériques certaines:

- soit en recopiant la donnée si elle est numérique certaine,
- soit en utilisant la fonction F dans le cas d'une donnée incertaine. On substitue à p une fonction s(p). Cette fonction représente la stratégie adoptée. Elle est définie de la manière suivante:

$$s:[0,1] \rightarrow [0,1]$$

$$p \mapsto s(p)$$

Les stratégies les plus courantes sont:

- on étudie le pré-ordre pour les valeurs médianes des données incertaines: s(p)=0.5. Cette approche donne un résultat équivalent à celui d'une matrice d'évaluations composée des valeurs médianes. L'étude est globale.
- on étudie l'influence des données incertaines: s(p)=nombre aléatoire de répartition uniforme dans le domaine concerné. Toutes les données sont concernées et l'étude est également globale.
- on étudie l'influence d'un critère à données incertaines: pour chaque donnée de ce critère, on utilise soit le tirage aléatoire, soit dans le cas où l'on désire évaluer l'influence de ce critère "au pire" les valeurs s(p)=0 (resp. s(p)=1) si le critère est

à minimiser (resp. maximiser). On définit de manière duale l'étude "au mieux". L'étude est "verticale" (par critère), car elle concerne une colonne de la matrice d'évaluation.

- on peut étudier l'influence d'une alternative: l'étude est alors "horizontale" (par alternative). En fonction de $s(p)$ on peut maximiser ou minimiser l'influence de cette alternative qui doit alors posséder des valeurs incertaines. L'étude utilisant les tirages aléatoires et $s(p)=\text{constante}$ est bien sûr possible.

Ces trois études (globale, horizontale -par alternative et verticale -par critère) et la fonction décrite ci-dessus fournissent, dans tous les cas, une matrice de données numériques certaines que l'on est alors capable de traiter à l'aide de la méthode Prométhée.

Voici un exemple qui illustre cette étape. Soit la matrice générique d'évaluation $M_{s(p)}$ indicée par $s(p)$:

$$M_{s(p)} = \begin{pmatrix} \omega''_{\alpha} & 1 \\ \omega'''_{\alpha} & 2 \end{pmatrix}$$

Elle est composée de deux valeurs numériques certaines (1 et 2) et de deux valeurs numériques incertaines introduites dans l'exemple en annexe B. Les données incertaines sont:

$$\omega''_{\alpha} = \left(\{3, 5, 9\}, \left\{ 0, 0, \frac{1}{3}, \frac{2}{3} \right\}, F, 3 \right)$$

$$\omega'''_{\alpha} = \left(\{3, 5, 7, 9\}, \left\{ 0, 0, \frac{1}{4}, \frac{1}{4}, \frac{1}{2} \right\}, F, 4 \right)$$

Il est possible de construire $M_{s(p)}$ avec des données exclusivement génériques en utilisant la généralisation présentée précédemment. Mais cette représentation diminue la lisibilité. Nous préconisons ainsi de conserver les données certaines sous leur forme classique.

On obtient, dans le cas où on applique le même paramétrage:

$$s(p) = 0, 0.2, 0.4, 0.6, 0.8, 1$$

les instanciations suivantes:

$$M_0 = \begin{pmatrix} 3 & 1 \\ 3 & 2 \end{pmatrix}, M_{\frac{1}{5}} = \begin{pmatrix} 3 & 1 \\ 3 & 2 \end{pmatrix}, M_{\frac{2}{5}} = \begin{pmatrix} 5 & 1 \\ 5 & 2 \end{pmatrix}, M_{\frac{3}{5}} = \begin{pmatrix} 5 & 1 \\ 7 & 2 \end{pmatrix}, M_{\frac{4}{5}} = \begin{pmatrix} 9 & 1 \\ 9 & 2 \end{pmatrix}, M_1 = \begin{pmatrix} 9 & 1 \\ 9 & 2 \end{pmatrix}$$

Chacune de ces matrices est alors traitée à l'aide de la méthode Prométhée.

5.3. Etape 3: Utilisation de la méthode multicritère Prométhée

Selon la méthode utilisée (Prométhée I ou II), l'algorithme évalue successivement l'ensemble des préférences relatives puis les flux entrants et sortants. Selon la

méthode, le flux net peut être calculé. De ce fait, un résultat est exprimé sous la forme d'une matrice comportant des données numériques certaines.

5.4. Etape 4: mémorisation des résultats

Il est nécessaire de mémoriser les résultats avant de réitérer l'étape d'instanciation. Cette mémorisation exploite le principe d'enrichissement. On crée autant de données génériques (incertaines) qu'il y a de flux pour chaque alternative. Nous notons pour chaque alternative ω_a^+ , ω_a^- et ω_a^{net} , $a \in [1, na]$ les données génériques flux associées.

5.4.1. Détermination des paramètres des données génériques

Afin d'estimer les écarts et les valeurs extrêmes pour fixer l'ensemble X commun, on peut effectuer l'étape 3 à partir de valeurs particulières des données incertaines. Le principe est le suivant:

Les flux positifs et négatifs sont de signe positif (ceci se vérifie aisément à partir de leur définition). De ce fait, il suffit de définir une borne supérieure pour fixer la valeur à donner à l'écart. Ceci est réalisé par l'intermédiaire du flux positif représentant la qualité avec laquelle une alternative est préférée vis à vis des autres alternatives. On instancie (en vue de minimiser les évaluations) toutes les données avec $s(p)=0$ (resp. $s(p)=1$) si le critère est à minimiser (respectivement à maximiser) sauf une pour laquelle on optimise de manière duale les données instanciées. L'on estime alors le flux positif pour cette donnée. Ceci étant réalisé pour chaque donnée, on est alors en mesure d'évaluer le flux positif maximal.

L'opération est alors inversée pour trouver le flux négatif maximal. La valeur maximale de X est le maximum de ces deux valeurs. L'écart est établi en fonction de la finesse voulue (il représente en quelque sorte un seuil d'indifférence). Le domaine X de la donnée générique ω_a^{net} $a \in [1, na]$ est évalué pour un même écart. L'enrichissement détermine naturellement les valeurs extrêmes pour cette donnée. Nous illustrons cette méthode sur un exemple:

Soit l'ensemble multicritère suivant (nous reprenons la matrice générique d'évaluation $M_{s(p)}$ présentée auparavant):

$$\begin{pmatrix} c_1 & c_2 \\ \hline I & C \\ \hline 2 & 1 \\ \hline m & M \end{pmatrix}$$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} \omega_\alpha'' & 1 \\ \omega_\alpha''' & 2 \end{pmatrix}$$

avec $\omega'' = \left(\{3,5,9\}, \left\{0,0,\frac{1}{3},\frac{2}{3}\right\}, F, 3 \right)$ et $\omega''' = \left(\{3,5,7,9\}, \left\{0,0,\frac{1}{4},\frac{1}{4},\frac{1}{2}\right\}, F, 4 \right)$

Les critères sont notés c_1 et c_2 . Pour simplifier, on les suppose tous deux de nature linéaire $p=4$ sans seuil ($q=0$). Le premier est composé de données incertaines (I), de poids 2 et à minimiser (m). Le second est composé de données certaines (C), de poids 1 et à maximiser. Les alternatives sont notées a_1 et a_2 .

On en déduit, pour le flux positif, les matrices et résultats suivants:

- on favorise a_1 et on minimise a_2 :

$$M_{a_1, a_2} = \begin{pmatrix} 3 & 1 \\ 9 & 2 \end{pmatrix} \text{ et l'on obtient } \phi_{a_1, a_2}^+ = 0.66$$

- de même, on minimise a_1 et on favorise a_2 :

$$M_{a_2, a_1} = \begin{pmatrix} 9 & 1 \\ 3 & 2 \end{pmatrix} \text{ et l'on obtient } \phi_{a_2, a_1}^+ = 0.75$$

$$\max(\phi_{a_1, a_2}^+, \phi_{a_2, a_1}^+) = \phi_{a_2, a_1}^+ = 0.75$$

On montre de même que $\max(\phi_{a_1, a_2}^-, \phi_{a_2, a_1}^-) = 0.75$. Le maximum de ces deux valeurs (0.75) détermine le flux maximal que l'on puisse atteindre. Le flux minimal étant 0, l'écart ε est aisément déterminé. Si l'on désire que X présente un cardinal d'au moins dix valeurs, on se fixe $\varepsilon = 0.075$. Le flux net, de part sa définition, est situé dans le domaine $[-0.75, 0.75]$ pour un même écart. Pour ces trois données, on harmonise les ensembles X en fixant comme valeur commune la valeur 0.

Il est évident que cette étude préliminaire peut constituer une des itérations et être considérée comme un enrichissement à part entière.

5.4.2. Enrichissement des données génériques

A chaque phase de calcul, les trois données sont enrichies des valeurs des flux. La pertinence de cet enrichissement dépend fortement de la stratégie adoptée pour $s(p)$. Il est nécessaire d'apporter une information nouvelle pour que l'enrichissement soit pertinent. Le cas $s(p) = 0.5$ n'apporte qu'une information identique lors des différentes itérations, alors que le choix $s(p) = \text{nombre aléatoire}$ apportera une information pertinente.

Nous allons poursuivre l'exemple précédent. Les paramètres des trois données génériques ont été spécifiés: $\varepsilon = 0.075$ et valeur commune, 0. L'étude porte sur deux des six instanciations présentées auparavant:

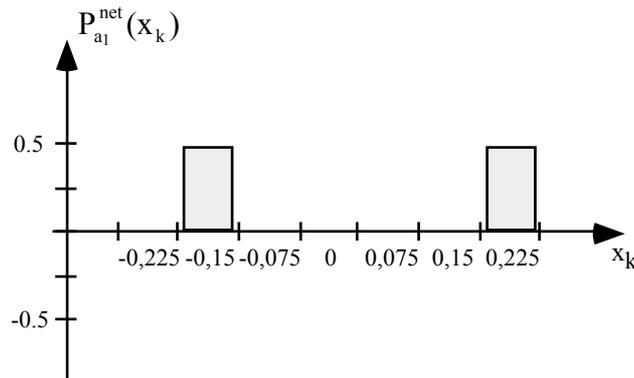
$$M_{\frac{2}{5}} = \begin{pmatrix} 5 & 1 \\ 5 & 2 \end{pmatrix}, M_{\frac{3}{5}} = \begin{pmatrix} 5 & 1 \\ 7 & 2 \end{pmatrix}$$

Pour chacune de ces matrices on relève les trois flux pour chaque alternative (table 1):

Matrice	a_1			a_2		
	$\phi_{a_1}^+$	$\phi_{a_1}^-$	$\phi_{a_1}^{\text{net}}$	$\phi_{a_2}^+$	$\phi_{a_2}^-$	$\phi_{a_2}^{\text{net}}$
$M_{\frac{2}{5}}$	0	0,083	-0,083	0,083	0	0,083
$M_{\frac{3}{5}}$	0,333	0,083	0,25	0,083	0,333	-0,25

Table 1: flux positif, négatif et net pour chaque alternative

Nous donnons l'état d'une seule donnée générique $\omega_{a_1}^{\text{net}}$ après les deux enrichissements présentés (cf. figure 10). Les valeurs -0,083 et 0,25 sont associées aux valeurs possibles -0.15 et 0.225 (méthode de "la borne inférieure").

Figure 10: donnée générique $\omega_{a_1}^{\text{net}}$ après deux itérations.

A ces trois données génériques, nous en associons une quatrième qui concerne le rangement. Cette donnée sera utilisée extensivement car elle renseigne sur les rangements des différentes alternatives selon les itérations.

5.4.3. Donnée générique "rangement"

La donnée rangement $\omega_{a_i}^c$ n'est plus associée à un ensemble X concernant les valeurs possibles portant sur la mesure des flux, mais à un ensemble X_c qui correspond aux différents rangements potentiels. Son cardinal est égal au nombre d'alternatives évaluées:

$X_c = \{i/i \in \mathbb{N}^* \wedge i < n_a\}$, on note $x_k \in X_c$ le $k^{\text{ième}}$ élément de X_c .

Soit la donnée générique $\omega_{a_i}^c$ de l'alternative a_i . L'enrichissement est réalisé sur la valeur $x_k \in X_c$ de $\omega_{a_i}^c$ correspondant à la $k^{\text{ième}}$ place selon le préordre (total ou partiel) établi par Prométhée.

Supposons par exemple que nous utilisions la méthode Prométhée I. Le premier enrichissement présenté correspondant à $M_{\frac{2}{5}}$ propose comme préordre complet

$a_1 \prec a_2$. De cette manière, les enrichissements de la première itération ($m=1$) sont $E_1^2(\omega_{a_1}^c)$ et $E_1^1(\omega_{a_2}^c)$. De même les enrichissements correspondants à la deuxième itération ($a_1 \succ a_2$) sont $E_2^1(\omega_{a_1}^c)$ et $E_2^2(\omega_{a_2}^c)$. Après cette deuxième itération, les deux

alternatives ont été placées chacune une fois première et une fois seconde. Les données génériques $\omega_{a_i}^c$ sont identiques (figure 11).

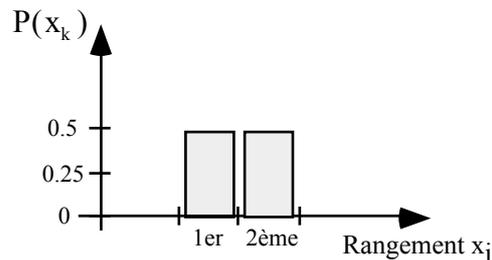


Figure 11: un exemple de rangement

Ces quatre données génériques sont alors exploitées lorsque le nombre d'itération désiré est atteint. Nous allons détailler cet aspect de la méthode.

5.5. Itération et intégration de l'analyse de sensibilité

Si l'utilisateur participe au processus décisionnel (mode coopératif), les itérations sont arrêtées lorsqu'il constate que celles-ci n'apportent plus d'information. Si le processus est réalisé en mode "autonome", une heuristique peut fixer le nombre d'itérations en fonction du nombre de données génériques et de la précision voulue. Si l'on constate par exemple qu'en moyenne elles possèdent dix enrichissements, il est raisonnable de considérer au minimum dix itérations. Les nombreux essais que nous avons effectués ont montré par ailleurs que dix fois le nombre moyen d'enrichissements fournissent une bonne représentativité du système. Ce point est fortement lié à l'analyse de sensibilité si l'on considère le cas où les données génériques incertaines sont instanciées de manière unique. Par exemple avec le paramétrage $s(p)=0.5$, une seule itération est possible (étape 3') et le résultat peut être comparé aux résultats utilisant les données génériques incertaines. Cette comparaison fournit une information similaire à une analyse de sensibilité. Cependant, cette dernière s'appuie sur des variations théoriques, alors que notre approche est fondée sur un ensemble de données réelles. L'analyse de sensibilité est ainsi naturellement intégrée dans la méthode.

5.6. Comparaison avec une méthode d'analyse de sensibilité

[Roy et al. 93] propose une heuristique assez semblable à celle présentée dans ce mémoire mais ne prenant en compte que les valeurs minimales, maximales et moyennes des m données de nature incertaine, ce qui nécessite 3^m calculs.

La limite de cette approche est due à l'explosion combinatoire (3^m calculs) qui n'est pas maîtrisée. Au contraire, dans l'approche que nous préconisons:

- les résultats sont exprimés sous la forme de données génériques, ce qui simplifie la représentation et l'analyse des résultats,

- une méthode heuristique (qui s'appuie dans le meilleur des cas sur le paramétrage aléatoire de $s(p)$ pour les raisons exprimées auparavant) effectue un nombre non pré-défini d'itérations.

Pour dix données incertaines, la méthode proposée par Roy entraîne plus de 50.000 itérations menant à une quantité d'informations très importante. Notre méthode conduit à un nombre d'itérations compris entre 10 et 100.

En conclusion, la méthode de Roy est valable lorsqu'il y a peu de données sujettes à variation: deux ou trois tout au plus. L'information obtenue est exploitée efficacement. La méthode heuristique proposée permet une plus large prise en compte de données incertaines. L'information résultante est moins riche que celle résultant d'une exploitation systématique de tous les enrichissements de toutes les données incertaines (généralisation de la méthode de Roy); mais un choix adapté de $s(p)$ permet de représenter correctement l'influence des valeurs incertaines sur les préordres.

5.7. Etape 5: exploitation des résultats

La donnée rangement $\omega_{a_i}^c$ est exploitable très facilement puisqu'elle donne directement la solution au problème de rangement.

La figure 12 représente un exemple portant sur quatre données rangement (représentées sous forme graphique et numérique) correspondant à une analyse de la matrice générique utilisée dans l'exemple précédent pour 50 itérations. On constate bien que les différents rangements sont issus de l'incertitude qui est associée à chaque donnée.

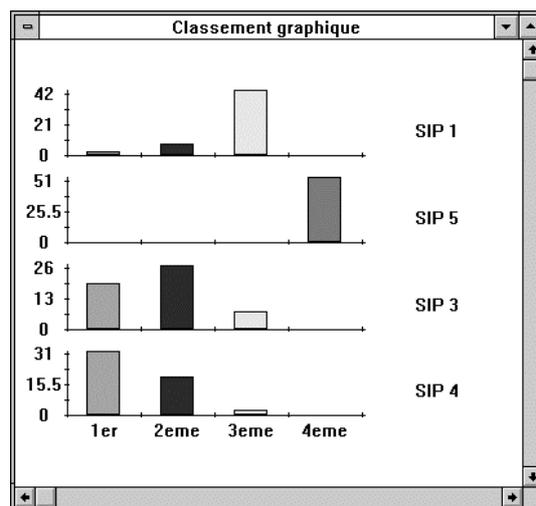


Figure 12: données génériques de rangement

La figure 13 donne ce rangement sous forme numérique (tableau).

Classement				
Name/pos.	1er	2eme	3eme	4eme
SIP 4	30	18	2	0
SIP 3	18	25	7	0
SIP 5	0	0	0	50
SIP 1	2	7	41	0

Figure 13: rangement sous forme de tableau

L'analyse des flux nets peut apporter l'information manquante au niveau des questions que le décideur peut se poser: "Est-ce que cette alternative, souvent première du rangement, se détache nettement des alternatives rangées seconde ou troisième?"

En dernier recours, le décideur peut considérer les flux entrants et sortants sous forme de données génériques. Il doit dans ce cas maîtriser davantage la méthode multicritère.

Une analyse de sensibilité peut être menée pour évaluer l'influence des paramètres qui sont soumis à une variation théorique. Il est possible par exemple d'utiliser des données incertaines particulières au lieu de poids certains. Cet aspect n'est pas abordé dans ce mémoire mais constitue une perspective de développement importante.

Nous présentons dans le paragraphe suivant la deuxième méthode envisagée.

6. Deuxième méthode M2

Cette méthode a pour objectif de fournir une aide complémentaire dans la décision portant sur le rangement.

Nous proposons d'adjoindre à Prométhée une fonction de gestion exploitant directement la matrice de préférence en exploitant la notion de non dominance introduite par Orlovski et présentée dans [Fodor et al. 94]. La méthode, comme la précédente, comporte cinq étapes.

Les étapes 1 (création des matrices génériques) et 2 (instanciation des matrices génériques) sont identiques à celles de la méthode précédente. Nous décrivons les étapes suivantes.

6.1. Etape 3: méthode multicritère basée sur les degrés de non-dominance

A la matrice des indices de préférence $c_{ik} = \sum_j w_j \cdot S_j(d_{ik})$ nous associons une relation binaire valuée R définie de la sorte:

$$R(a_i, a_k) = c_{ik}$$

La relation $R(a_i, a_k)$ décrit la relation de préférence entre a_i et a_k . Plus la valeur de $R(a_i, a_k)$ est élevée, plus a_i est préférée à a_k (indépendamment de tout seuil).

Nous définissons le degré de préférence stricte P de la manière suivante:

$$P(a_i, a_k) = \max\{R(a_i, a_k) - R(a_k, a_i), 0\}$$

Cette relation nous permet de définir le degré de non domination $\mu_{ND}(a_i, R)$ portant sur a_i selon la relation R :

$$\mu_{ND}(a_i, R) = \min_k \{1 - P(a_k, a_i)\}$$

Le rangement (du dernier au premier) est élaboré par valeurs croissantes du degré de non-domination.

Cette méthode est souvent utilisée dans le but d'élaborer un choix, et non plus un rangement [Fodor et al. 94]. Dans ce cas, le choix se porte sur l'alternative qui propose le plus haut degré de non domination. Cependant, l'utilisation de la matrice de préférence élaborée par Prométhée s'appuie sur une relation de préférence non transitive. On montre [Fodor et al. 94] que la valeur maximum (c'est à dire de degré 1) n'est pas forcément atteinte, ce qui peut nuire à la pertinence du choix. Il est conseillé dans ce cas d'utiliser une relation transitive, qui a pour conséquence de fournir au moins une alternative de degré maximal. Le choix de cette alternative est donc sans ambiguïté.

Dans les deux cas (relation transitive ou non), il est pertinent d'évaluer la sensibilité des paramètres. Pour ce faire, on définit l'ensemble D_{ND}^* des éléments non dominés qui correspond au maximum des degrés de non domination:

$$D_{ND}^* = \{a^* / \forall k, \mu_{ND}(a^*) \geq \mu_{ND}(a_k)\}$$

Le cardinal de cet ensemble est au minimum 1. La présence de plusieurs éléments de même degré est possible (cas des ex-aequo des éléments rangés).

On peut définir de même le degré de non dominance:

$$\mu_{Nd}(a_i, R) = \min_k \{1 - P(a_i, a_k)\}$$

On définit de même l'ensemble D_{Nd}^* des éléments non dominants qui correspond au maximum des degrés de non dominance:

$$D_{Nd}^* = \{a^* / \forall k, \mu_{Nd}(a^*) \geq \mu_{Nd}(a_k)\}$$

L'étude peut porter sur les trois ensembles possibles:

- les alternatives fortement recommandées D_{ND}^* ,
- les alternatives très peu recommandées D_{Nd}^* ,
- les alternatives sans recommandation précise.

Reprenons l'exemple étudié dans la première méthode:

$$M_{\frac{2}{5}} = \begin{pmatrix} 5 & 1 \\ 5 & 2 \end{pmatrix}, M_{\frac{3}{5}} = \begin{pmatrix} 5 & 1 \\ 7 & 2 \end{pmatrix}$$

On obtient les deux matrices d'indices de préférence, et l'on indique les degrés de non domination et de non dominance:

$$\begin{pmatrix} c_{ik} & a_1 & a_2 \\ a_1 & 0 & 0 \\ a_2 & 0,083 & 0 \end{pmatrix} \begin{pmatrix} \mu_{ND} \\ 1 \\ 0,917 \end{pmatrix} \text{ pour } M_{\frac{2}{5}} \text{ et } \begin{pmatrix} c_{ik} & a_1 & a_2 \\ a_1 & 0 & 0,333 \\ a_2 & 0,083 & 0 \end{pmatrix} \begin{pmatrix} \mu_{ND} \\ 0,75 \\ 1 \end{pmatrix} \text{ pour } M_{\frac{3}{5}}.$$

$$\begin{pmatrix} \mu_{ND} & 0,917 & 1 \end{pmatrix} \qquad \begin{pmatrix} \mu_{ND} & 1 & 0,75 \end{pmatrix}$$

6.2. Etape 4: mémorisation des résultats

La phase de mémorisation peut être menée en fonction de la problématique:

- la première qui consiste à exploiter uniquement l'ensemble des éléments non dominés D_{ND}^* (cas du rangement),
- la seconde qui consiste à exploiter deux parmi les trois ensembles proposés (cas du choix),
- la troisième qui consiste à exploiter les trois ensembles (cas du choix).

Les données génériques utilisées ont trait à la prise en compte des trois (ou moins) ensembles présentés selon l'approche choisie.

Nous présentons l'intégration des résultats au sein de variables génériques de la forme ω_{a_i} dont l'ensemble des valeurs possibles est naturellement:

$$X = \{x_{ND}\}, X = \{x_{ND}, x_{\overline{ND.Nd}}, x_{Nd}\}, \text{ selon l'approche choisie.}$$

L'enrichissement est alors réalisé au niveau de X en fonction des ensembles de dominance. L'enrichissement porte sur x_{ND} de la donnée générique ω_{a_i} de l'alternative a_i si durant l'itération concernée, $a_i \in D_{ND}^*$. De même pour $a_i \in D_{Nd}^*$ et pour le cas $a_i \notin D_{ND}^* \cap D_{Nd}^*$ si les trois ensembles sont exploités.

Dans l'exemple proposé, et dans le cas où $X = \{x_{ND}\}$, on obtient deux enrichissements différents pour les deux itérations. Le premier (associé à $M_{\frac{2}{5}}$) concerne $a_2 (\mu_{ND} = 1)$: $E_1^1(\omega_{a_2})$ et le second $a_1 (\mu_{ND} = 1)$: $E_1^1(\omega_{a_1})$.

Lorsque le nombre d'itérations est estimé suffisant, l'opérateur peut passer à la phase d'exploitation.

6.3. Etape 5: exploitation des résultats

L'exploitation est réalisée de manière identique à la première méthode. La prise en compte de la valeur possible x_{ND} est évidemment nécessaire. L'opérateur peut alors compléter sa décision en étudiant les deux autres ensembles D_{Nd}^* et $D_{\overline{ND} \cdot \overline{Nd}}^*$ s'il désire exploiter cette méthode pour traiter la problématique de choix et non plus la problématique de rangement.

On ne peut pas assurer a priori que les deux méthodes proposées aboutissent à des conclusions compatibles. Dans une situation où l'incompatibilité est mise en évidence, le système d'aide conclut à la faible validité de ses évaluations et l'opérateur doit alors faire appel à sa propre expertise et à sa connaissance pour lever l'indétermination. La probabilité d'occurrence d'un tel cas de figure est évidemment considérablement diminuée puisqu'une analyse de sensibilité est proposée pour chaque méthode.

7. Conclusion

Ce chapitre avait pour objet de développer en premier lieu un modèle de données incertaines compatible avec les contraintes spécifiées dans le chapitre III. Nous avons développé le processus d'enrichissement d'une donnée et montré qu'une telle méthode possède de nombreux avantages en terme de modélisation générique, d'intégration dynamique des événements et un haut niveau d'adéquation à la réalité.

Nous avons aussi montré qu'il est possible d'envisager une approche multicritère en développant un modèle pour la résolution des processus décisionnels au niveau du pilotage hybride. Celui-ci est dédié à la problématique de rangement.

Le chapitre suivant décrit la structure de communication et l'intégration globale avec les processus décisionnels multicritères au niveau du pilotage hybride.

Chapitre V

Proposition d'une structure de communication et intégration des processus décisionnels

1. Introduction

La résolution des processus décisionnels présentée dans les chapitres précédents est décomposée selon les phases non séquentielles d'information, de conception, de choix et d'évaluation. La recherche de l'information (phase d'information) est réalisée de deux manières complémentaires:

- localement: c'est le système d'information de chaque SIP qui est exploité,
- globalement lorsqu'un processus décisionnel nécessite une information qui n'est pas disponible localement. La structure de communication est utilisée pour accéder à cette information.

De ce fait, l'intégration d'une structure de communication au sein de chaque SIP permet la mise en oeuvre de l'ensemble des processus décisionnels. Elle doit par ailleurs être capable de supporter les ré-itérations envisageables lors de la résolution des processus décisionnels.

La première partie présente la structure de communication et les deux niveaux qui la composent. Les différents protocoles sont définis et modélisés. Nous définissons alors la notion d'état canonique, ce qui nous permet d'intégrer un premier niveau de robustesse selon la hiérarchisation proposée.

La seconde partie définit l'intégration conjointe de la fonction de communication et des processus décisionnels au sein de la structure de pilotage hybride.

2. Structure de communication

Nous utilisons la notion de protocole de communication que nous définissons comme une procédure de transfert d'information, décrite par étapes. De ce fait, c'est une notion très large qui peut s'adresser à plusieurs niveaux d'applications.

Un message est décrit sous forme générique par un triplet $\langle S,C,X \rangle$, il est composé d'un émetteur S, d'un récepteur C et d'un contenu X.

Deux niveaux ont été définis. Le niveau T gère le transfert effectif d'un message et doit s'assurer de la réception sans erreur du message. Le niveau T est lié à la syntaxe de la communication.

Le niveau A traite les informations échangées comme un ensemble de messages $\langle S,C,X \rangle$. Le niveau A est lié à la sémantique de la communication.

La figure 1 détaille la structure hiérarchique des deux niveaux au sein du sous-système de communication.

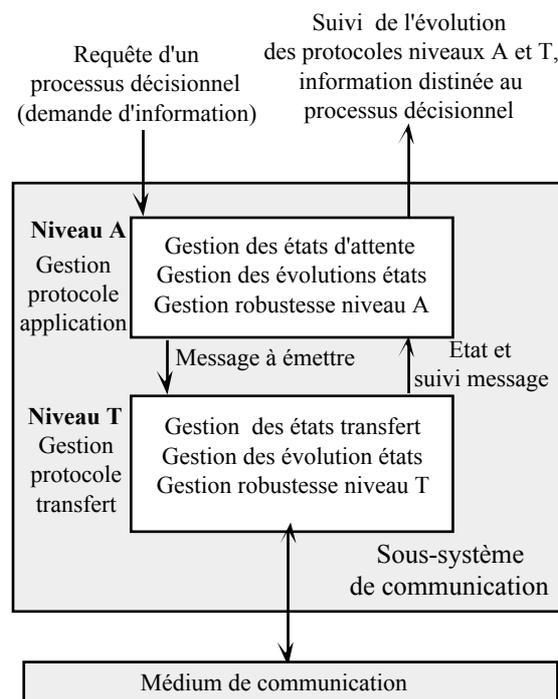


Figure 1: intégration des deux niveaux de communication

Le niveau T correspond à un protocole unique, le Protocole de Transfert et concerne deux agents. Ce protocole est utilisé exclusivement à chaque appel provenant du niveau A concernant au moins deux agents.

Le niveau A regroupe l'ensemble des protocoles, Protocoles Application, utilisés par chaque éléments de la structure de pilotage hybride.

Ces protocoles applications concernent (cf. chapitre III):

- le protocole de détection et d'ordonnancement de la ressource goulet par le superviseur,
- le protocole d'information portant sur la mise à jour du système d'information du superviseur par chaque agent,

- le protocole d'allocation des opérations entre agents et entre les agents et le superviseur.

Ces différents protocoles utilisent les primitives de communication présentées au chapitre I sous la forme de réseaux de Petri colorés. Nous les décrivons dans la partie suivante.

2.1. Modélisation des niveaux T et A

2.1.1. Spécification des systèmes de communication

L'approche distribuée présentée dans les chapitres précédents nous a menés à proposer une structure de communication capable de supporter l'aspect non séquentiel des processus décisionnels. Sur la figure 2, deux agents réalisent un processus d'allocation à des étapes différentes du processus. Ainsi, un agent doit être capable de gérer simultanément:

- une réception multiple de messages,
- une émission multiple de messages.

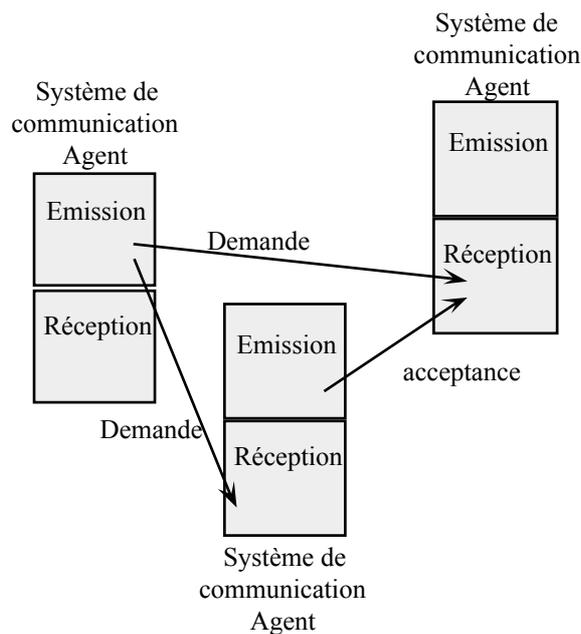


Figure 2: communication inter-agents

La figure 2 illustre deux points importants:

- la nature symétrique de chaque agent. Un agent doit être capable d'émettre et de recevoir de manière indépendante. Chaque sous-système doit donc posséder un canal d'émission indépendant du canal de réception.
- chaque agent est susceptible d'être contacté simultanément par plusieurs émetteurs.

On impose à chaque agent une seule connexion à la fois en tant qu'émetteur, ce qui nécessite d'attendre la libération du canal d'émission.

Ces deux points impliquent d'adopter une structure multi émetteur- récepteur, matérialisée par une structure informatique de réseau multi client - serveur.

2.1.2. Les réseaux de Petri

Pour modéliser les différents protocoles, nous avons retenu les réseaux de Petri [David et al. 89] pour, outre le fait que ce mode de représentation soit largement répandu, leur puissance de modélisation dans les domaines de la gestion de production, de la communication, de l'analyse de performances, de l'étude de robustesse, de la commande ou du suivi de production, etc. [Sahraoui 87].

D'autre part, leurs caractéristiques graphiques proches des représentations de connaissance en intelligence artificielle apportent, dans la mesure où la complexité et la taille du réseau restent raisonnables, un aspect souple, convivialité et lisibilité non négligeable [Simonetti 87].

2.1.3. Les extensions

Quatre extensions ont été apportées à la représentation par réseaux de Petri afin de modéliser les contraintes dynamiques liées aux processus décisionnels et la nature générique de l'information supportée par les primitives (et a fortiori par la primitive Transfert) [Trentesaux et al. 94a]:

- les réseaux colorés sont paramétrés: nous prenons en compte les notations indicielles. Par exemple $\langle C_i \rangle$ symbolisera $\langle C_1 \rangle$, $\langle C_2 \rangle$, $\langle C_3 \rangle$, etc. Ceci permet une meilleure lisibilité et une meilleure gestion des agents et des messages,
- les notations de jetons colorés ont été étendues de telle sorte que les notations de type $\langle C_i, X \rangle$ signifient tout aussi bien $\langle C_i, C_j \rangle$ que $\langle C_i, \langle D_k \rangle \rangle$ ou encore $\langle C_i, \langle D_k, X \rangle \rangle$. Les couleurs peuvent alors être des jetons. La symbolique X est donc réservée. Dans le modèle, cette couleur particulière représentera l'information véhiculée jusqu'aux places qui l'exploitent. La symbolique X n'a pas été spécifiée de manière plus précise afin que l'on puisse lui associer et/ou substituer toute information nécessaire aux heuristiques de calcul,
- les fonctions associées aux arcs gèrent dynamiquement les données. Certaines fonctions d'arc établissent au fur et à mesure les images des jetons concernés, alors que normalement, les images des fonctions sont définies statiquement au niveau de la modélisation. Cette approche se justifie pleinement par l'aspect dynamique de la répartition des opérations. D'autre part, la robustesse aux aléas est une caractéristique primordiale et seule la gestion dynamique des jetons supporte cet aspect. Par exemple, la suppression d'une S.I.P (supposée en panne) peut se modéliser simplement, *en cours de simulation*, par l'interdiction

au niveau de la fonction correspondante de créer des jetons associés à cette S.I.P,

- enfin, la notion de macro-place a été introduite. Cette notion permet de hiérarchiser la structure d'un réseau de Petri afin d'améliorer la structuration et la lisibilité. Il est évident que la seule contrainte qui permet l'utilisation de macro-places est celle de la cohérence des intrants et sortants.

2.1.4. Réseau de Petri associé au niveau T

Parmi les approches envisageables, nous avons opté pour une quasi-synchronisation du protocole entre les agents (figure 3). Seule l'étape de déconnexion (relaxation) est de nature asynchrone. La synchronisation permet d'assurer un transfert correct de l'information et évite une gestion complexe des états des agents (risque d'occurrence d'états bloquants ou d'itérations récursives).

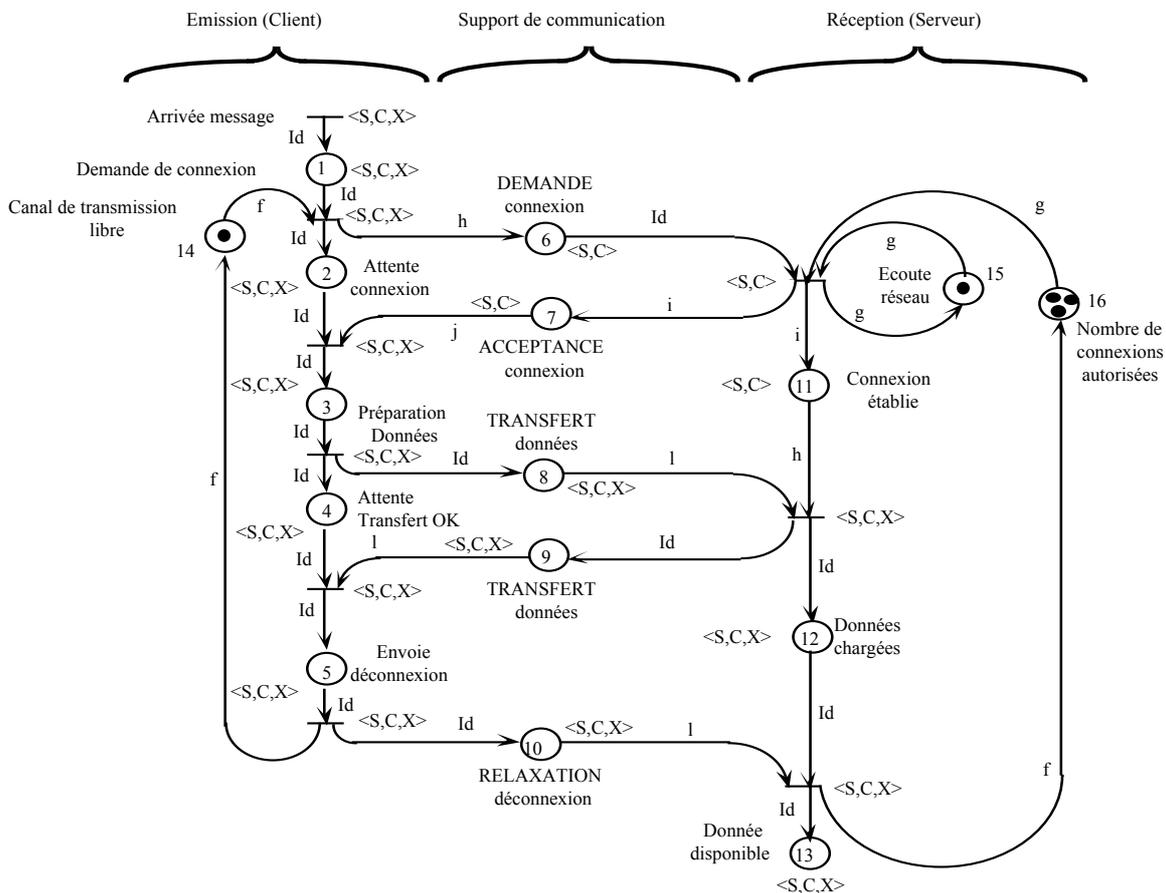


Figure 3: protocole de transfert

Les fonctions sont décrites ci-après:

- $f(\langle S, C, X \rangle) = \langle \bullet \rangle$
- $g(\langle S, C \rangle) = \langle \bullet \rangle$
- $h(\langle S, C, X \rangle) = \langle S, C \rangle$

- $i(\langle S, C \rangle) = \langle C, S \rangle$
- $j(\langle S, C, X \rangle) = \langle C, S \rangle$
- $l(\langle S, C, X \rangle) = \langle C, S, X \rangle$

L'information transmise est associée à X . La première (resp. deuxième) couleur correspond à l'identificateur de l'émetteur (resp. récepteur). La structure de ces deux couleurs est naturellement identique et permet les inversions du type $\langle S, C \rangle$ en $\langle C, S \rangle$.

En remarquant que $j = i \circ h = h \circ l$ et $f = g \circ h$, on simplifie l'ensemble des fonctions du modèle présenté.

L'arrivée d'un message $\langle S, C, X \rangle$ dans la place 1 provoque l'émission d'une demande vers le récepteur si la canal de transmission est libre (une seule connexion est possible du côté client). Si le récepteur (serveur) est à l'écoute du réseau, et si le nombre de connexions autorisées n'est pas atteint, il peut émettre une acceptation. Une fois l'acceptation reçue, les données sont préparées puis émises. Le récepteur reçoit la donnée (objet de la connexion) et retourne cette dernière. Ceci pour deux raisons:

- la première est qu'un tel retour permet la synchronisation du client qui, suite à ce retour, est assuré de la transmission d'une donnée, ce qui lui permet de se déconnecter,
- la seconde est que le retour de la même donnée permet de vérifier de manière simple la qualité de la transmission.

La déconnexion (relaxation) est alors permise et réalisée de manière asynchrone. Ce réseau de Petri décrit donc la syntaxe du transfert de données. Nous supposons en outre que:

- pour des raisons de simplification, les données sont transmises en une seule fois (la transmission de X n'est pas décomposée par paquet). Un niveau intermédiaire, responsable de la décomposition des données d'un côté (client ou serveur) et de la recombinaison de l'autre, doit être défini si l'on désire émettre un message en plusieurs fois (problème de capacité réseau, etc.),
- pour des raisons de compréhension mutuelle, une structure commune pour les messages est adoptée pour l'ensemble des agents.

2.1.5. Réseaux de Petri associé au niveau A

Nous développons les réseaux de Petri des différents protocoles précédemment introduits.

2.1.5.1. Protocole d'information

Ce protocole permet d'informer le superviseur de l'état des ordres de fabrication. Il est purement de nature asynchrone. Le rafraîchissement de la base de données du superviseur se fait de manière fréquente. Il n'est pas nécessaire de le réaliser de manière synchrone. En effet, les évolutions de l'état de chaque ordre de fabrication rendent obsolètes les anciennes mémorisations de cet état dans le système d'information du superviseur. De ce fait, le changement d'état d'un ordre de fabrication non signalé au superviseur à un moment donné lui sera spécifié lors de la prochaine phase d'information concernant cet ordre de fabrication. En même temps que l'opération est allouée ou chargée, l'ordre de fabrication et son état lui sont fournis. L'agent modifie ainsi son historique et informe le superviseur. Le réseau de Petri est présenté sur la figure 4.

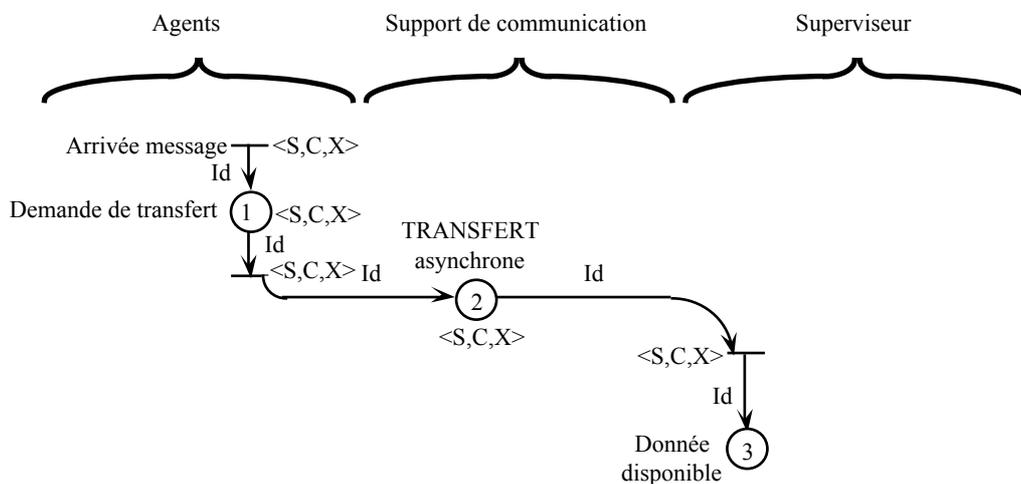


Figure 4: Protocole d'information

La transition "arrivée message" est activée à chaque fois qu'un agent prend en charge ou a terminé une opération d'un ordre de fabrication.

La place 2 correspond ainsi à une macro place qui est associée au niveau T. Notons que les intrants sont de nature $\langle S,C,X \rangle$ et les sortants de nature $\langle C,S,X \rangle$. Le niveau T, en délivrant le message, inverse l'émetteur et le récepteur et conserve l'information X. De ce fait, on peut associer au niveau T la macro-fonction $T(\langle S,C,X \rangle) = \langle C,S,X \rangle$.

Ce protocole peut être généralisé pour l'échange d'informations entre agents et non uniquement entre un agent et le superviseur. Il permet en particulier à un agent d'informer tous les autres de son intégration au sein du réseau en fournissant par exemple son adresse. Il permet enfin d'intégrer la notion de postes de travail (1x8, 2x8 ou 3x8) de manière efficace en créant une information de désactivation et d'activation (ce qui peut être lié à la phase d'intégration) en début

et en fin de chaque poste. L'intérêt dans ce cas est de ne pas provoquer une interrogation de la part des autres agents lorsqu'une fonction opératoire n'est plus disponible au sein du système de production. En outre, dans le cas d'une simulation du système, le superviseur peut avertir les agents du début de la simulation afin de synchroniser au mieux les horloges et les données de simulation.

2.1.5.2. Protocole de calcul de charges et ordonnancement goulet

Le protocole a été spécifié au chapitre II. Nous présentons le modèle dans cette partie.

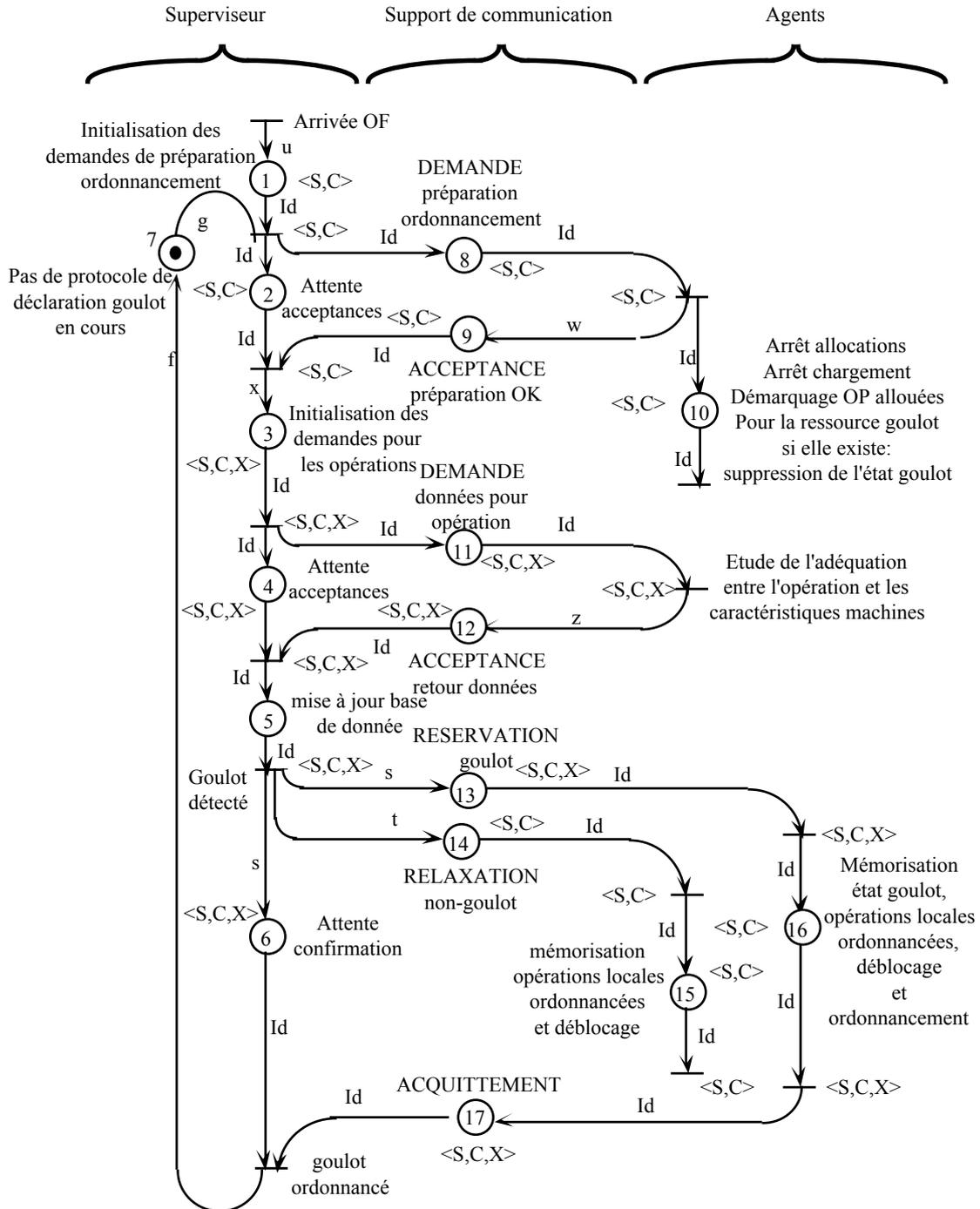


Figure 5: protocole d'ordonnancement

Le modèle est décrit figure 5.

Les fonctions sont définies de la manière suivante:

- $f(\langle S, C, X \rangle) = \langle \bullet \rangle$
- $g(\langle S, C \rangle) = \langle \bullet \rangle$
- $j(\langle S, C, X \rangle) = \langle C, S \rangle$

Ces trois fonctions sont identiques aux fonctions introduites précédemment.

- u est définie de la manière suivante:

$$u(\langle \text{Of} \rangle) = \sum_{a=\text{agents}} \langle S_{\text{superv.}}, C_a \rangle$$

Cette fonction crée autant de demandes qu'il y a d'agents pour informer chacun d'entre-eux de l'initialisation de la phase d'ordonnancement. Le but est de demander à chaque agent de stopper les processus de chargement d'ordre de fabrication et d'allocation.

- $w(\langle S, C \rangle) = \langle S, C \rangle$ indique que le processus d'ordonnancement a été enregistré et que l'agent attend pour répondre aux demandes de faisabilité.

- $x(\langle S, C \rangle) = \sum_{o=\text{opérations}} \sum_{a=\text{agents}} \langle S_{\text{superv.}}, C_a, X_o \rangle$

Pour chaque agent et pour chaque opération à réaliser, on émet une demande de faisabilité.

- $z(\langle S, C, X \rangle) = \langle S, C, X \rangle$ enregistre au niveau de X les données concernant l'opération en question: faisabilité ou non, et si oui, temps nécessaire pour traiter l'opération (donnée estimée ou non).

Les acceptances étant reçues, le calcul des charges potentielles est réalisé. Il tient compte des allocations déjà effectuées (opérations dans les files d'attente). Suite à ce calcul, la ressource goulet est détectée et informée de son état et de l'ordonnancement qui lui est affecté par la fonction s :

- $s(\langle S, C_{\text{goulet}}, X \rangle) = \langle S, C_{\text{goulet}}, X_{\text{ordo}} \rangle$ qui ne concerne que cette ressource goulet.
- la fonction t informe les autres agents de leur état non-goulet:

$$t\left(\sum_{a=\text{agents} \neq \text{goulet}} \langle S, C_a, X \rangle\right) = \langle S, C_a \rangle$$

Les deux fonctions s et t réactivent les processus de fabrication et d'allocation.

2.1.5.3. Protocole d'allocation

Le protocole a été étudié de manière approfondie dans [Trentesaux 94a]. Nous en donnons la représentation fonctionnelle (figure 6).

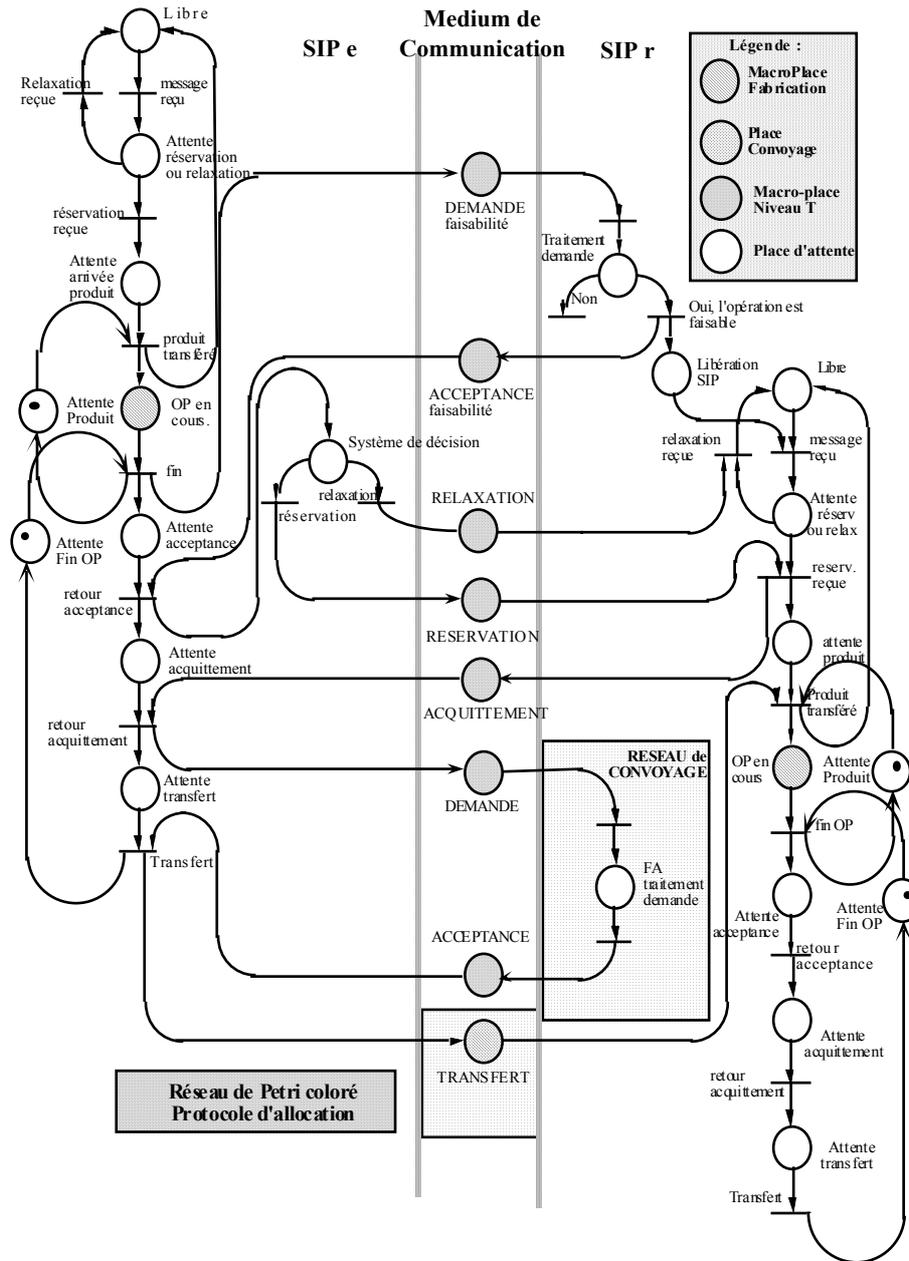


Figure 6: réseau de Petri fonctionnel du protocole d'allocation.

La SIP e cherche à allouer une opération. La SIP r représente l'ensemble des SIP qui reçoivent cette demande d'allocation. Les différentes places gérant l'évolution des états des SIP ont été dupliquées en une SIP e et une SIP r afin d'améliorer la lisibilité du réseau de Petri.

Le processus de ré-allocation est identique à celui de l'allocation. La différence se situe au niveau de la localisation de l'ordre de fabrication dans la file d'attente de sortie pour l'allocation et dans la file d'entrée pour la ré-allocation.

2.2. Intégration de la robustesse au sein des niveaux T et A

Un niveau élémentaire de robustesse peut être introduit dans chaque protocole du niveau T et du niveau A en exploitant la notion de temps d'attente. Ces temps d'attente ont été modélisés sur les différents réseaux de Petri. Ils permettent de répéter certaines étapes non complètement réalisées au bout d'un certain temps. Après plusieurs itérations d'une même étape, l'opérateur est averti de l'existence d'un dysfonctionnement et doit alors exploiter des procédures de réinitialisation particulières.

2.2.1. Intégration d'un premier niveau de robustesse: protocole de transfert robuste

La robustesse au niveau T est nécessaire pour prendre en compte les situations suivantes:

- un agent ne peut poursuivre le protocole de transfert associé à ce niveau (déconnexion soudaine, etc.),
- le réseau de communication ne peut procéder à la transmission d'un message (problème support physique, etc.),
- le réseau de communication transmet un message erroné,
- le réseau de communication est saturé (ce qui augmente sensiblement les temps d'échange et qui peut provoquer une perte d'information par surcharge).

Pour assurer ce niveau de robustesse, le protocole de transfert présenté auparavant a été enrichi. Nous précisons cet aspect dans le cas de l'étape de connexion. Le principe est généralisable aux autres étapes. L'algorithme concerne à la fois le client et le serveur. Il est représenté sur les figures 7 et 8. Il constitue un processus non bloquant activé périodiquement.

Coté client:**initialisation:**

DEMANDE_envoyée = Faux
 canal de transmission = Libre
 ACCEPTANCE_reçue = Faux

Algorithme

:début

Si((DEMANDE_envoyée=Faux) et (canal transmission=Libre)) alors

:début

DEMANDE_envoyée = Vrai

Envoyer demande de connexion

canal de transmission = occupé

temps_attente_ACCEPTANCE = 0

:fin

Si ((DEMANDE_envoyée=Vrai) et (ACCEPTANCE_reçue=Faux)) alors

:début

Si une ACCEPTANCE arrivée alors

:début

ACCEPTANCE_reçue = Vrai

temps_attente_ACCEPTANCE = 0

:fin

Sinon

:début

incrémenter temps_attente_ACCEPTANCE

:fin

Si temps_attente_ACCEPTANCE > Limite_tps_ACCEPTANCE alors

:début

temps_attente_ACCEPTANCE = 0

DEMANDE_envoyée = Faux // ré-initialise protocole

canal de transmission = Libre // idem

:fin

:fin

Si (ACCEPTANCE_reçue = Vrai)

... // phase suivante du protocole

:fin

...

Figure 7: algorithme d'intégration de la robustesse coté client

```

Coté serveur:

intialisation:
canal de transmission = Libre
attente_TRANSFERT = Faux
nb_connexion = 0

Algorithme
:début
Si (écoute réseau) alors
  :début
  Si ((message arrivé) et (message_reçue = DEMANDE)) alors
    :début
    Si ((nouveau client) et (nb_connexion < limite_nb_connexion)) alors
      :début
      référencer client
      envoyer ACCEPTANCE de connexion
      temps_attente_TRANSFERT = 0
      attente_TRANSFERT = Vrai
      incrémenter nb_connexion
      :fin
    Sinon
      :début
      // ne rien faire,
      // on attend la libération d'une connexion,
      // ou c'est un client déjà connecté.
      :fin
    :fin

  Si (attente_TRANSFERT = Vrai)
    :début
    incrémenter temps_attente_TRANSFERT
    Si (temps_attente_TRANSFERT > Limite_tps_TRANSFERT) alors
      :début
      temps_attente_TRANSFERT = 0
      // si cette ré-initialisation intervient
      // un trop grand nombre de fois,
      // avertir l'utilisateur
      :fin
    Si(message arrivé) et (message_reçue=TRANSFERT) alors
      :début
      temps_attente_TRANSFERT = 0
      attente_TRANSFERT = Faux
      :fin
    :fin
  :fin
:fin

```

Figure 8: algorithme d'intégration de la robustesse coté serveur

On peut remarquer que certains processus sont activés de manière événementielle (par exemple, la procédure de détection de demande de connexion activée par l'arrivée d'une demande). Nous les avons cependant décrits

séquentiellement pour faciliter la compréhension de l'algorithme. Cette simplification ne nuit en rien à l'approche.

D'autre part, nous supposons dans ce mémoire que les procédures au niveau du support de communication ne sont pas totalement bloquantes en cas de dysfonctionnement. Il peut exister cependant des cas de figures exclusivement bloquants, et dans ce cas, une procédure complète de réinitialisation est requise. Nous ne prendrons pas en compte cet aspect.

Dans le cas présent, la complexité induite par une telle gestion rend difficile une modélisation efficace par réseaux de Petri. Uniquement pour l'étape de connexion, 14 places et 8 transitions possédant en moyenne 3 arcs entrant ou sortant ont été nécessaires, ce qui rend le réseau de Petri illisible et peu exploitable. C'est pourquoi nous avons présenté une modélisation algorithmique.

Il est possible d'associer ce processus (définition d'un temps d'attente, procédure de ré-initialisation de l'étape du protocole) à chaque étape d'attente. Ce qui a été présenté pour l'étape "émission d'une demande, réception d'une acceptation" peut être appliqué de manière similaire aux étapes "émission et réception d'un Transfert".

2.2.2. Notion d'état canonique

La méthode d'intégration de la robustesse qui a été présentée s'appuie sur la détection des états d'attente. Ces états sont fortement découplés les uns des autres. D'autre part, leur agencement permet de décrire complètement le protocole. Par exemple, le protocole de transfert peut être décrit en deux phases itératives (émission demande, réception acceptation) et (émission transfert, réception transfert). A chacune de ces phases correspondent deux états d'attente qui sont:

- attente d'acceptation,
- attente de transfert.

Chaque état peut ainsi être défini comme canonique. On définit également la notion d'état possible afin de décrire un séquençage de changements d'états. De cette manière, à chaque état correspond un ensemble d'états possibles dont les changements d'état traduisent décrit l'évolution du protocole.

La gestion par les états canoniques simplifie le découplage de chaque protocole, ce qui simplifie l'analyse et l'intégration de la robustesse.

L'approche générale a été présentée et justifiée dans la partie précédente (niveau T). Nous détaillons la méthode complète dans le cadre d'un protocole de niveau A.

2.2.3. Intégration d'un second niveau de robustesse: protocole d'allocation robuste

Nous allons montrer, en étudiant le protocole d'allocation, comment la détection et l'utilisation des états canoniques permet de mieux intégrer la notion de robustesse. La méthode exposée se généralise sans difficulté aux autres protocoles.

2.2.3.1. Etude du fonctionnement dégradé

Nous supposons le niveau T robuste, c'est à dire qu'un message envoyé est correctement reçu. Nous nous intéressons uniquement aux dysfonctionnements du type "panne sous-système" qui entraînent un blocage du protocole de niveau A (allocation, information bloquée, etc.).

2.2.3.2. Exploitation des réseaux de Petri: détection des états canoniques

Le protocole d'allocation peut être décomposé par la combinaison de trois états canoniques (nous les avons explicitées sur le réseau de Petri du protocole d'allocation, pour le protocole du niveau T, elles ne sont pas indiquées pour des raisons de lisibilité). Nous les désignons par p1, p2 et p3. Les états possibles sont indiqués ci-dessous:

- états possibles pour p1:

{LIBRE, ATTENTE_RESERV_OU_RELAX, ATTENTE_TRANSFERT}

- états possibles pour p2:

{MARCHE, ATTENTE_OF}

- états possibles pour p3:

{ATTENTE_FIN_OF, ATTENTE_ACCEPTANCE, ATTENTE_ACQUIT}

p1 décrit les différents états d'un agent répondant à une demande d'allocation. De manière duale, p3 décrit le comportement d'un agent qui procède à une allocation. L'état p2 décrit le comportement productif d'un agent. On vérifie bien que l'agencement de ces états permet la description complète du protocole. D'autre part, ces états possèdent un sens physique: p1 représente l'arrivée d'une opération (liée à un ordre de fabrication), p2 le processus de production sur une opération et p3 le processus d'allocation d'une opération.

Le séquençement des états est réalisé de la manière suivante (figure 9).

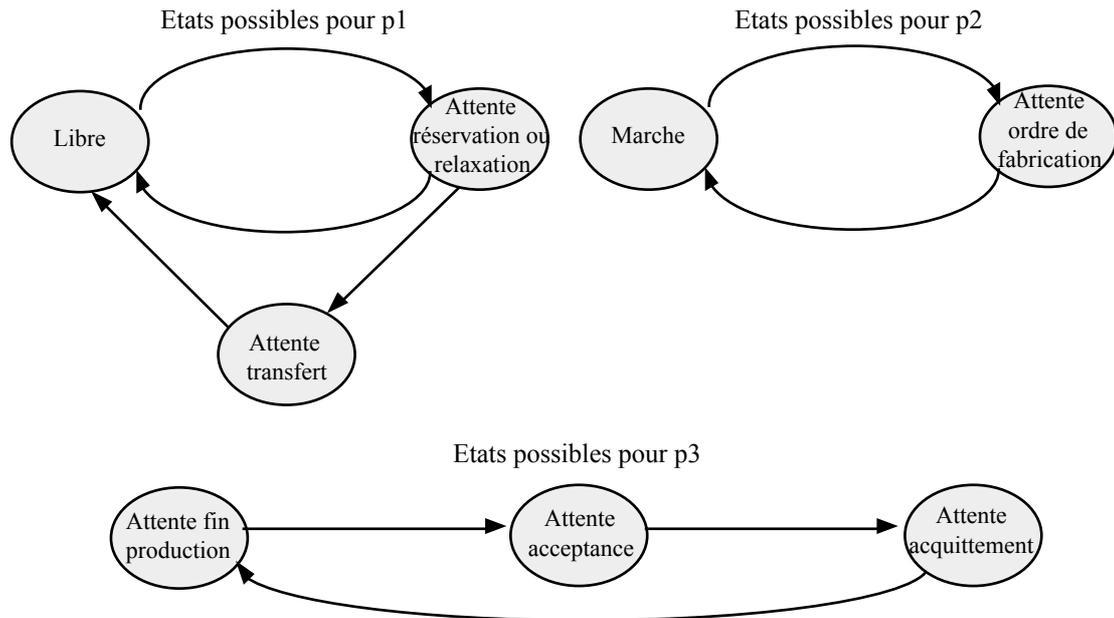


Figure 9: séquençage des états

A partir de ce séquençage, nous pouvons définir le fonctionnement normal, et décrire l'intégration du fonctionnement dégradé.

2.2.3.3. Exploitation des états: détection des étapes d'attente

La première étape consiste à intégrer la notion de temps d'attente maximal pour chaque état correspondant à une attente *exclusivement* au niveau du protocole.

Cette démarche a été introduite dans l'étude de la robustesse du niveau T. Ces états sont `ATTENTE_RESERV_OU_RELAX` et `ATTENTE_TRANSFERT` pour p1, `ATTENTE_ACCEPTANCE` et `ATTENTE_ACQUIT` pour p3. L'état p2 ne concerne que la gestion locale de la production. Elle n'est donc pas exploitée.

Les étapes du type "émission, attente réception" sont associées aux quatre attentes citées précédemment. Elles sont au nombre de deux pour p1:

- "émission acceptance, attente réservation ou relaxation" qui correspond à l'état `ATTENTE_RESERV_OU_RELAX`
- "émission acquittement, attente transfert" correspondant à l'état `ATTENTE_TRANSFERT`.

et au nombre de deux pour p3:

- "émission demande, attente acceptance" qui correspond à l'état `ATTENTE_ACCEPTANCE`
- "émission réservation et relaxations, attente acquittement" qui correspond à l'état `ATTENTE_ACQUIT`.

Ainsi, ces quatre étapes ont pu être identifiées grâce à la représentation canonique. Il est dès lors possible d'intégrer un niveau de robustesse en complétant ces étapes de la manière décrite auparavant:

- adjoindre des temps d'attente bornés,
- réitérer certaines étapes spécifiques en cas de dépassement du temps limite maximal,
- avertir l'opérateur en cas d'un dépassement du nombre d'itérations prévu.

Nous détaillons les procédures de recouvrement pour ces quatre étapes.

Etape "émission demande, attente acceptance"

La procédure de ré-émission des demandes pour une opération donnée est réalisée si aucune acceptance n'est reçue dans les temps impartis. La réception d'une acceptance suffit pour ne pas réitérer l'appel.

Etape "émission acceptance, attente réservation ou relaxation"

La procédure de ré-émission intervient dans le cas où ni une réservation ni une relaxation n'est reçue dans les temps impartis. Ce cas de figure arrive fréquemment dans le cas où l'opérateur humain est intégré dans le processus décisionnel, car il accroît sensiblement le temps nécessaire à la sélection de l'agent le plus satisfaisant. Cependant, la procédure de ré-émission et de re-réception des acceptances est totalement transparente pour l'utilisateur. D'autre part, il est possible d'annihiler cette procédure, ce qui réduit fortement la robustesse de cette étape.

Etape "émission réservation et relaxations, attente acquittement"

Cette étape est peu critique car totalement automatisée et ne dépend que du support de communication. Il est cependant nécessaire de proposer une procédure de ré-émission, car la panne d'un système de communication, par exemple, survenant à ce niveau est toujours possible.

Etape "émission acquittement, attente transfert"

Cette étape est fortement liée à la capacité du réseau de transport physique. Les temps d'attente doivent être conditionnés en fonction des temps de transfert moyens par opération pour chaque couple d'agents.

2.2.3.4. Duplication de messages

L'approche présentée a pour conséquence de dupliquer certains messages, ce qui nécessite l'utilisation d'une interface au niveau de la réception pour la gestion des messages dupliqués dans les deux cas où:

- le message est déjà présent; le message dupliqué est alors détruit. le cas typique est celui de l'agent qui, en attente d'acceptance, ré-envoie des demandes juste avant de recevoir les acceptances retardataires. Il peut constater dès lors la présence d'acceptances dupliquées.
- le message n'est plus présent car il a été pris en compte avant l'arrivée du message dupliqué; cet aspect plus complexe est cependant aisément résolu lorsqu'une trace des étapes précédentes du protocole est conservée. La politique à adopter est alors définie selon le cas d'attente. Par exemple, après avoir envoyé une acceptance, un agent gardant la trace de la demande constate qu'une demande à laquelle il a déjà répondu est arrivée. Il est alors plus judicieux de répondre une nouvelle fois, l'agent demandeur ayant peut-être eu des difficultés pour la réception des acceptances précédentes. D'un autre côté, l'agent demandeur ayant effectué son choix pour l'allocation peut encore recevoir des acceptances concernant cette allocation. En comparant l'opération concernée et les acceptances reçues (trace de l'étape précédente), il est possible de détecter celles qui lui correspondent afin de les supprimer.

2.2.4. Généralisation aux autres protocoles

L'intégration de la robustesse ayant été décrite pour le protocole d'allocation, nous ne la développons pas au niveau des autres protocoles, la méthode présentée s'appliquant à l'ensemble des protocoles.

3. Déroulement des processus décisionnels

L'objectif de cette partie est de montrer les différentes phases de chaque processus décisionnel. Ces phases sont décomposées en deux parties distinctes:

- phase d'information: réalisation des différents protocoles de communication (en fonction du processus décisionnel), recherche locale de données,
- phase de décision: conception, choix et évaluation.

Nous avons montré que ces deux phases peuvent être itérées dans le cas de dysfonctionnements par exemple.

Nous détaillons succinctement chaque processus. Nous supposons que le mode de coopération est fixé par l'opérateur et prenons en considération la nature des processus qui peuvent être:

- non structurés: dans ce cas, l'opérateur est maître de sa décision. Le calculateur peut cependant déclencher ce processus en mode autonome, mais cette approche est parfois déconseillée de par la nature même du processus qui,

résolu de manière automatique sans analyse plus poussée de l'état du système de production, peut mener à une dégradation des performances.

- bien ou peu structurés: dans ce cas il est possible de fournir une aide de nature multicritère à l'opérateur ou de procéder à une prise de décision automatique (mode autonome).

3.1. Processus non structurés

3.1.1. Choix de ré-allocation

Ce processus peut être activé à chaque instant par l'utilisateur et concerne une opération sur un produit qui est dans le stock d'entrée ou en cours de réalisation mais à un niveau d'avancement qui permet une ré-allocation. Le premier cas est pris en compte en considérant l'opération en cours dans le stock de sortie, alors que le second cas nécessite l'intégration de données plus complexes sur les gammes opératoires et les machines. Ce point n'est pas étudié dans ce mémoire.

3.1.2. Choix de non respect d'une partie de l'ordonnancement

En cas de fonctionnement dégradé qui provoque des retards d'opérations planifiées, il est envisageable pour l'opérateur de forcer le lancement d'une opération. C'est un choix très peu structuré qui est difficile à estimer de manière automatique. Cependant, si l'opérateur décide de ne pas respecter l'ordonnancement, il lui est nécessaire d'annuler la planification de l'opération afin de la prendre en charge avant la date prévue. Ceci modifie l'ordonnancement et dans ce cas, il est, la plupart du temps, préférable, suite à cette remise en question, de demander un ré-ordonnancement.

3.1.3. Choix de remise en question de l'ordonnancement prévisionnel

Ce processus concernant la gestion indirecte est activé soit sur demande de l'opérateur, soit automatiquement lorsqu'au moins un ordre de fabrication est ajouté à la pile de tâches. Le processus d'ordonnancement est alors enclenché.

3.2. Processus peu structurés

Dans cette typologie de processus, les méthodes multicritères sont exploitables. De ce fait, nous présentons quelques critères qui permettent d'étayer la décision.

3.2.1. Gestion des files d'attente du stock d'entrée

Ce processus décrit le choix du séquençement des opérations. Il est activé à chaque fois que plusieurs opérations sont présentes dans le stock d'entrée. L'approche par décision multicritère permet d'intégrer les critères suivants:

- temps passé dans la file d'attente (pour améliorer la fluidité des ordres de fabrication),

- marge accordée en terme de temps libre (défini par la date de besoin diminuée du temps total de production restant),
- priorité de l'ordre de fabrication décrit en terme de conséquence en cas de retard (annulation de la commande, etc.).

3.2.2. Séquencement des ordres d'allocation

Le processus associé au séquencement de chaque opération de chaque ordre de fabrication présent est réalisé de la manière suivante:

Chaque fois que plusieurs opérations doivent être allouées simultanément, il peut être judicieux de choisir l'ordre des processus d'allocation. Ce cas se présente notamment dans le cas du superviseur (initialisation du système de production). Le séquencement des allocations est guidé uniquement par la première méthode multicritère qui propose un séquencement (rangement) des allocations. Les critères pris en compte sont par exemple:

- marge accordée en terme de temps libre (défini par la date de besoin diminuée du temps total de production),
- priorité de l'ordre de fabrication décrit en terme de conséquence en cas de retard (annulation de la commande, etc.).

3.2.3. Détection, information et ordonnancement de la ressource goulet

Ce processus est assez complexe. Il englobe le séquencement de:

- la recherche des données nécessaires au traitement,
- le calcul des charges potentielles,
- la détection de la ressource goulet,
- le calcul de l'ordonnancement,
- la déclaration de la ressource goulet et de son ordonnancement.

Dans une première approche, nous supposons que ce processus décisionnel est réalisé indépendamment de l'opérateur. Les phases de recherche des données et d'information sont guidées par les protocoles d'information et d'ordonnancement. La phase de calcul est réalisée par un algorithme d'ordonnancement classique. Un exemple d'un tel ordonnancement sera proposé ultérieurement.

3.2.4. Gestion des variables d'action locales

Les variables d'actions locales concernent les points de pilotage d'un système de production contrôlés par une SIP, par exemple l'ensemble des paramètres de réglages. Les différentes valeurs de ces variables sont généralement fixées par l'expérience ou le cahier des charges. La deuxième catégorie de variables locales concerne les taux de production, les lots de chargement (s'ils sont de taille

variable), etc. Les critères pris en compte sont alors fortement dépendants de la nature du système de production, nous ne présenterons pas ce niveau.

3.2.5. Gestion de chaque allocation

Ce processus constitue le coeur de l'ordonnancement dynamique. Il est réalisé en deux étapes.

3.2.5.1. 1ère étape: satisfaction de contraintes

C'est à ce niveau que les modules de gestion de contraintes sont utilisés:

- par l'agent émetteur qui recherche une SIP pour réaliser l'opération en question (primitive demande). Cette étape est généralement automatique mais peut être forcée par l'opérateur dans le cas, par exemple, d'une ré-allocation.
- par chaque agent récepteur qui doit vérifier la compatibilité entre les contraintes imposées et ses propres caractéristiques. Cette étape est réalisée automatiquement. Une possibilité non envisagée dans ce mémoire consiste à permettre à l'opérateur de pouvoir agir sur l'émission de l'acceptance afin d'empêcher celle-ci (surcharge prévue à plus ou moins long terme, etc.). Mais la SIP qui n'est pas en bonne condition pour réaliser la tâche sera très certainement éliminée au moment de la réservation. Ce processus fait appel à la première partie du protocole d'allocation.

3.2.5.2. 2ème étape: élection de la SIP cible

Les SIP qui retournent une acceptation sont évaluées par l'opérateur et/ ou par le calculateur selon le mode de coopération sélectionné. Les critères pris en compte sont très nombreux et dépendent fortement de la nature du système de production.

Cette deuxième étape intègre la deuxième partie du protocole d'allocation: émission d'une réservation, de relaxations et d'un acquittement.

3.2.6. Ré-ordonnancement local

Le ré-ordonnancement local correspond à un processus peu structuré. Cependant, l'opérateur comme le calculateur peuvent appliquer une heuristique, assez efficace, qui consiste à retenir la prochaine opération à effectuer. Dans le cas où elle n'est pas présente, il faut alors considérer la première opération planifiée présente. Si possible, il faut tenir compte de la date d'arrivée de l'opération manquante pour sélectionner une opération dont le temps de production est le plus faible, la démarche étant identique pour les opérations suivantes. Cependant, nous conseillons de demander le ré-ordonnancement dès que possible. Si une méthode multicritère est considérée plutôt qu'une heuristique, les critères pris en compte sont par exemple:

- la place de l'opération envisagée dans l'ordonnancement prévisionnel,

- le temps nécessaire pour réaliser l'opération envisagée,
- le retard répercuté sur les autres opérations,

4. Conclusion

Nous avons proposé un modèle de communication correspondant aux spécifications présentées dans les chapitres précédents.

Nous avons décrit la notion de protocole de communication et l'avons appliquée à deux niveaux de communication, l'un gérant le transfert effectif de l'information et l'autre permettant la phase de collecte des données réparties parmi les SIP et nous avons retenu une approche de communication quasi synchrone.

Le niveau T peut être généralisé en intégrant une couche intermédiaire responsable de la structuration des messages, en fonction du dimensionnement du support de communication.

L'intégration des protocoles de communication au niveau du pilotage permet de réaliser l'ensemble des processus décisionnels.

Le chapitre suivant propose une mise en oeuvre informatique de l'ensemble des modèles présentés pour la simulation et l'évaluation de ces modèles sur des cas industriels dans un cadre de fonctionnement "autonome".

Chapitre VI

Validation

1. Introduction

Dans ce chapitre nous présentons la structure logicielle que nous avons développée pour l'évaluation des modèles précédents. Les contraintes liées à l'expérimentation nous ont conduits à envisager une évaluation à l'aide de structures mono-poste et multi-postes.

Nous décrivons les modèles informatiques des différents sous-systèmes constituant les SIP, en précisant les éléments spécifiques au superviseur.

Nous présentons ensuite les résultats des évaluations menées en environnement mono-poste pour un objectif de production juste à temps et les perspectives de développement multi-postes. Les données sont issues d'un système de production industriel de type job-shop qui est un atelier de finition de fonderie.

2. Objectifs de l'évaluation

La validation de la structure de pilotage porte sur:

- la structure hybride en fonctionnement "autonome": qualité du pilotage, de l'intégration de la méthode multicritère proposée, de l'ordonnancement prévisionnel partiel de la ressource goulet,
- les protocoles de communication: évaluation de la robustesse aux niveaux A et T,

Cependant, la réalisation de ces objectifs est conditionnée par un ensemble de contraintes expérimentales.

2.1. Contraintes expérimentales

Elles sont de deux types.

2.1.1. Contrainte logicielle

Nous avons retenu pour le développement de la maquette informatique une approche orientée objet. Ce choix se justifie par les nombreux avantages en terme de qualité logicielle [Desfray 92], [Coad et al. 92]:

- en phase d'analyse et de modélisation, l'adéquation entre le modèle informatique et la réalité (ce qui accroît en particulier la facilité de la compréhension globale, la cohérence du modèle, etc.),
- en phase d'implémentation sur ordinateur, la réutilisabilité, extensibilité, maintenance, etc.

Cette approche présente certains inconvénients. Pour l'implémentation d'une structure distribuée seul le mode de communication synchrone est disponible. En effet, les objets sont de nature passive et subissent les requêtes d'activation, de ce fait, les fonctionnements asynchrones ne peuvent être réellement simulés.

2.1.2. Contraintes matérielles

Les contraintes sont liées aux limitations d'un environnement mono-poste et à la difficulté de mettre en place une structure multi-postes:

- cas mono poste - mono processeur: le fonctionnement interne dépend d'une activation par une horloge qui synchronise les processus (le cas mono-poste et multi-processeur est bien sûr envisageable afin d'intégrer une gestion asynchrone, mais peu répandu),
- cas multi postes: le nombre relativement important de ressources nécessaires limite fortement la mise en oeuvre d'une maquette de simulation correspondant à un cas industriel.

L'ensemble de ces contraintes nous a conduit à réaliser l'évaluation à l'aide de deux structures monoposte (mono processeur) et multi-postes (mono processeurs).

2.2. Evaluation mono-poste

La validation mono-poste permet:

- de simuler l'existence et l'interaction d'un nombre significatif d'agents. Une simulation multi-postes est plus difficile à mettre en oeuvre dans la mesure où elle nécessite autant de ressources informatisées que d'agents dans le système,
- d'évaluer les performances du système de pilotage en mode de coopération "autonome": auto-adaptabilité, performance de l'ordonnancement a posteriori,
- de vérifier le bon fonctionnement des protocoles de niveau A et en particulier, leur robustesse.

Nous avons réalisé le développement logiciel mono-poste en tenant compte du besoin de développement d'une structure de base qui doit être capable de supporter les spécifications du SIAD distribué proposé. Ce dernier s'appuie sur la méthode multicritère que nous avons développée.

2.3. Evaluation multi-poste

Le développement d'une structure multi-postes permet:

- de simuler le fonctionnement et la robustesse du protocole de niveau T,
- de développer, de simuler et d'évaluer la qualité de la coopération opérateur - SIAMD dans un environnement multi-utilisateurs (distribué).

3. Structuration logicielle du pilotage hybride

La structure a été développée pour être compatible avec le modèle de structure distribué [Tchako 94]. Le schéma global de l'application distingue l'environnement de production (les SIP, les produits, la structure de production, etc.) et l'environnement de simulation.

De ce fait, nous ne précisons que les extensions (en terme d'héritage, de données et de fonctions membres) effectuées pour modéliser la structure de pilotage hybride. Nous utilisons le modèle classe relation [Desfray 92] comme support de cette modélisation.

3.1. Objets de base

La classe "Objet de base" est une classe générique décrivant notamment les fonctions élémentaires associées à un objet de base: type, comparaison d'objet, etc. Ces fonctions doivent être sur-définies dans les classes héritantes. Nous définissons de même la classe "Attribut" qui associe un nom à chaque classe héritante et gère l'affichage de ce nom, quelque soit le type de la donnée.

Les différentes listes d'objets sont modélisées par:

- une liste simple entrée pour modéliser une file d'attente FIFO,
- une liste double entrée pour modéliser une file d'attente FIFO avec possibilité d'introduire la notion de priorité,
- un vecteur pour modéliser une liste d'objets sans ordre particulier.

Une liste d'objets est utilisée pour modéliser par exemple:

- les produits en stock d'entrée/ sortie,
- certaines informations (principalement les files d'attentes d'émission de messages).

3.2. Structure logicielle SIP niveau distribué

Les modifications concernant les sous-systèmes par rapport au modèle de la structure distribuée portent sur:

- l'intégration d'une structure logicielle qui permet le développement d'un SIAD distribué. Cette modification concerne les sous-systèmes de décision, d'interface et d'information,
- le développement logiciel de la structure de données (certaines, incertaines) et de la méthode multicritère introduites dans le chapitre IV,
- l'intégration des différents protocoles de niveau A (nous verrons ultérieurement l'intégration du niveau T).

3.2.1. Sous-classe sous-système d'information

Elle possède toutes les informations susceptibles d'être exploitées par au moins deux sous-systèmes. Dans le cas contraire, chaque information est exclusivement propre à une classe. Par exemple le sous-système de communication possède deux listes d'attente de messages, considérées comme données mais non incluse dans le sous-système d'information.

3.2.1.1. Structuration Base de donnée

La structure du sous-système est donnée sur la figure 1.

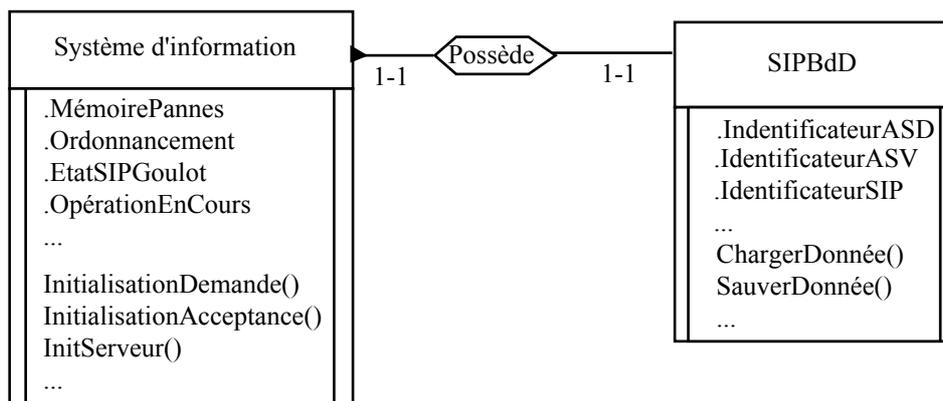


Figure 1: Sous-système d'information

L'environnement de production est constitué de:

- la classe "SIPBdD". Elle gère les accès (lecture, sauvegarde, effacement, etc.) à la base de données locale. Les données mémorisées concernent les caractéristiques SIP. C'est le cas du type d'opération, des temps de préparation, changement outil, coefficient opératoire, etc.
- la classe Ordonnancement (vecteur d'opérations et de dates de début au plus tôt et de fin au plus tard) et de la variable EtatSIPGoulet qui sont exploitées dans le cas où la ressource est goulet,
- OpérationEnCours, OrdreFabEnCours, PourcentageProcessEnCours, StockEntréeOf (aspect information), StockSortieOf (aspect information), MémorisationOf, etc.

- les données portant sur le type de poste de la SIP. Ceci permet en particulier d'anticiper les demandes afin de ne pas débiter une opération avant la fin de la phase de travail.

L'environnement de simulation est constitué principalement d'un vecteur MemoirePannes qui mémorise toutes les dates d'occurrence de pannes. Les pannes sont utilisées en simulation afin de vérifier la robustesse de la structure, des protocoles, etc.

Les données mémorisées sont structurées suivant le modèle de donnée générique présenté dans le chapitre IV. Nous donnons la structure informatique qui correspond à ce modèle.

3.2.1.2. La structure de données

La structuration des données est représentée sur la figure 2 (ne sont représentées que les données et les fonctions utiles à cette partie).

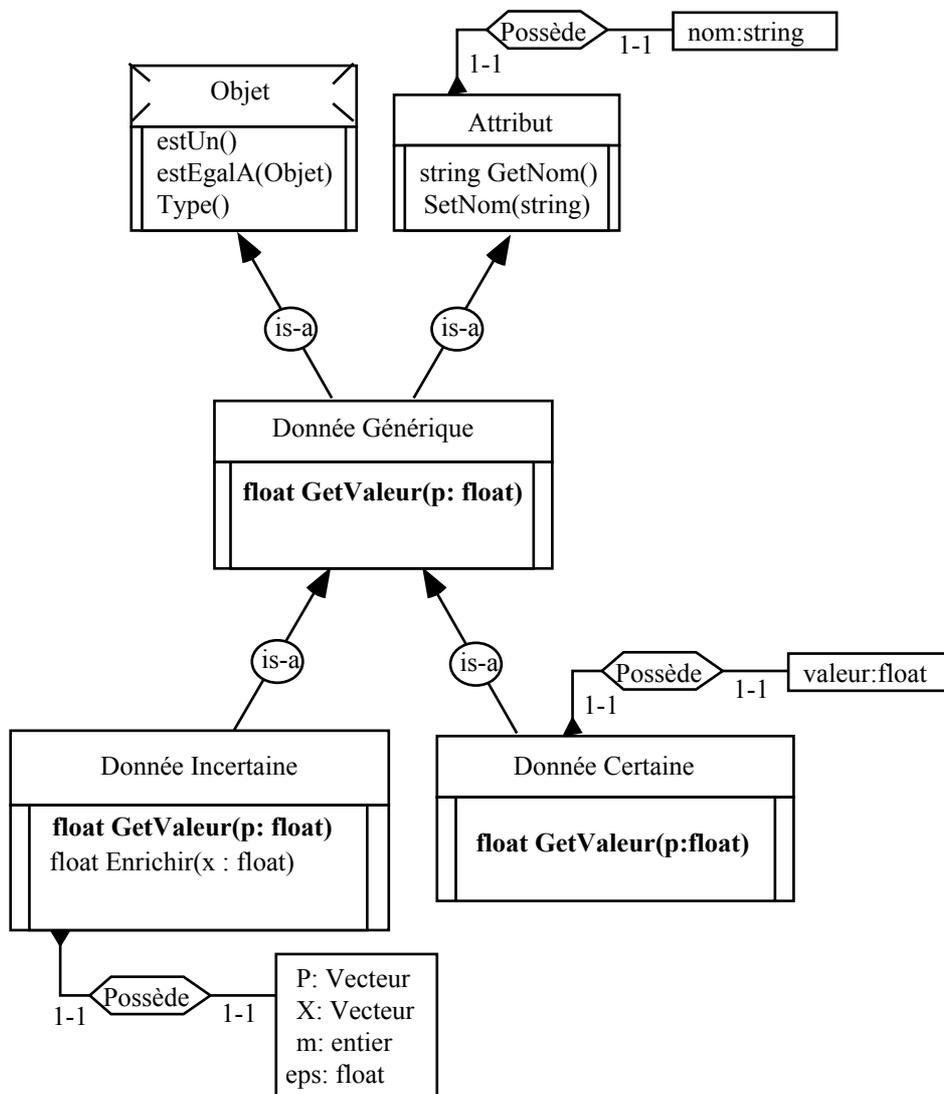


Figure 2: structure de donnée numérique/statistique

La classe “Donnée Générique” constitue la classe de base pour les deux types de données: certaines ou incertaines. La fonction membre `GetValeur()` de “Donnée Générique” permet d’accéder (via l’attribut “type” de la classe “Objet”) aux données des classes “Donnée Incertaine” et “Donnée Certaine” par sur-définition.

3.2.1.3. *Structure des données certaines*

Une donnée numérique certaine est représentée par la classe “Donnée Certaine”. La valeur numérique “valeur” est un attribut membre de la classe “Donnée Certaine”. La fonction `GetValeur()` fournit ce paramètre. Le paramètre `p`, utile pour la classe “Donnée Incertaine”, est alors ignoré.

3.2.1.4. *Structure des données incertaines*

Une donnée incertaine est représentée par la classe “Donnée Incertaine”. Cette classe possède deux vecteurs `X` et `P`, un entier `m` et un réel `eps` (ces trois variables ont été définies dans le chapitre IV) ainsi qu’une fonction pour l’enrichissement.

La fonction `GetValeur(p: float)` est sur-définie et a pour but de calculer une valeur fonction du paramètre `p`. Elle est associée à la fonction $F(p)$, cf. chapitre IV. De ce fait, elle est utilisée pour accéder à la donnée générique associée à l’instance de classe par l’intermédiaire du paramètre `p`.

3.2.2. Sous-système de décision

Nous avons étendu les fonctionnalités du modèle initial en ajoutant:

- le modèle de contraintes,
- les deux méthodes multicritères.

dont nous présentons les structures associées dans la suite de ce chapitre.

3.2.2.1. *Module de satisfaction de contraintes*

Aux deux types de données issus d’une donnée générique, nous ajoutons les types élémentaires objet d’entier, de caractère, de chaîne de caractères, de booléen, etc. Nous ne détaillerons pas ces classes qui héritent elles aussi des classes “Objet de base” de “Attribut”.

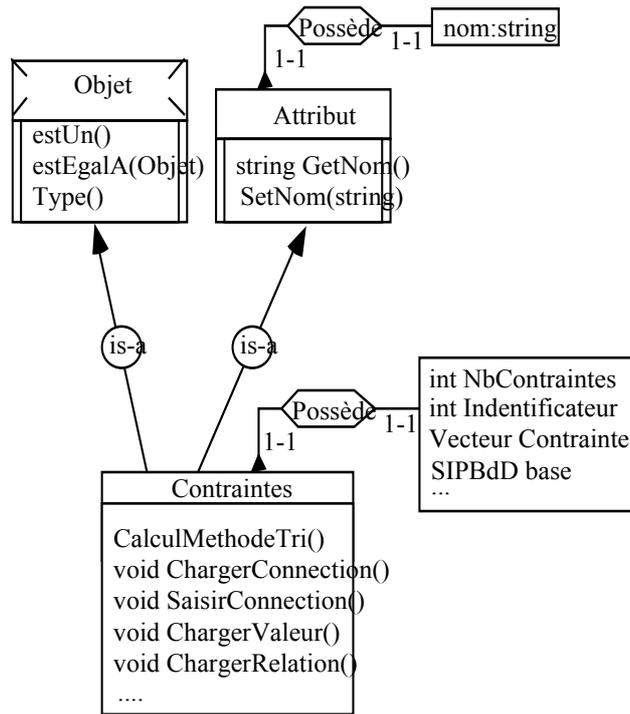


Figure 3: modèle classe-relation du gestionnaire de contraintes.

Les spécifications, le modèle et la structuration (figure 3) du gestionnaire de contraintes sont détaillés en Annexe D.

3.2.2.2. Matrice de décision multicritère

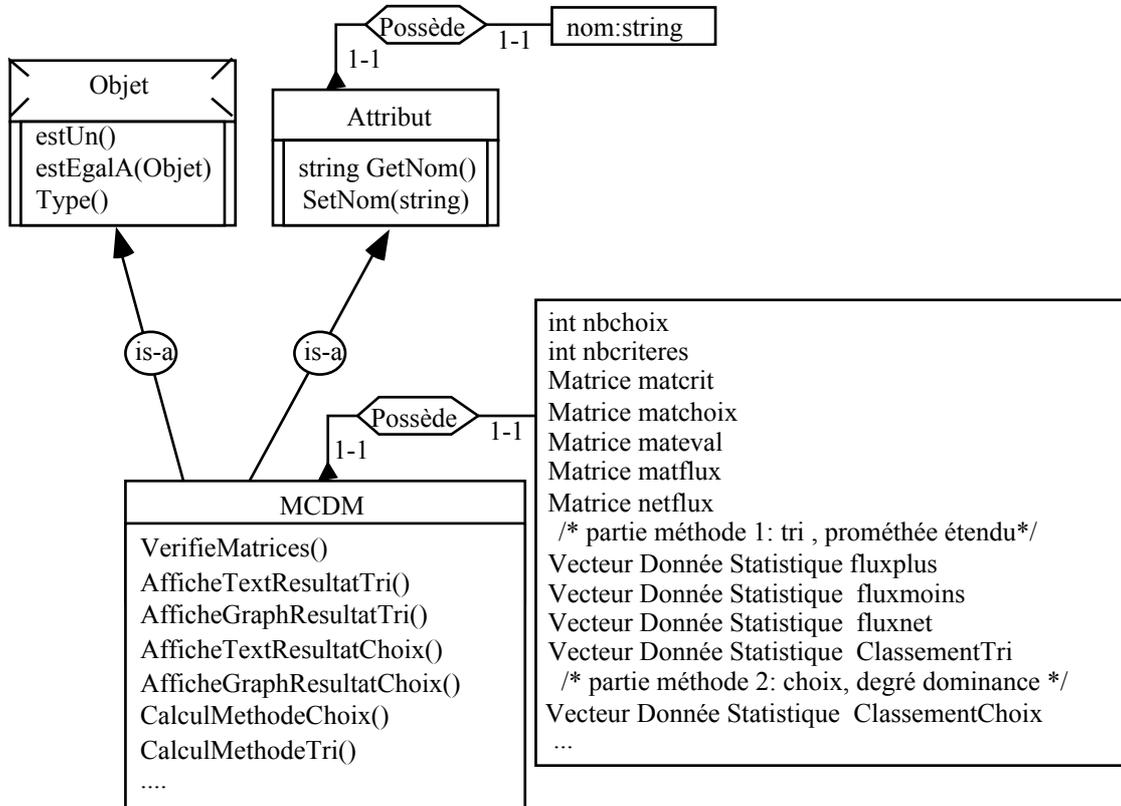


Figure 4: modèle classe-relation du modèle de décision multicritère

Les méthodes et les données membres intègrent les deux méthodes multicritère présentées dans le chapitre précédent au sein de la classe “MCDM” (figure 4). Cette classe est intégrée dans le sous-système de décision.

Les différentes matrices sont composées d’objets de type “donnée générique”. Le type précis (héritage) de ces données est fourni par la fonction `isA()`, ce qui permet d’accéder à la nature de la donnée (caractère, entier, donnée statistique, etc.).

3.2.3. Sous-système de communication

Le sous-système de communication n’est concerné que par l’environnement de simulation. Ce dernier contient:

- une liste (double entrée) d’entrée et de sortie des messages. Les demandes possèdent une priorité normale, les autres messages une priorité élevée,
- la mémorisation des différents états de chaque protocole du niveau A,
- la gestion des temps d’attentes associés aux états canoniques (cf. partie intégration de la robustesse, chapitre V),
- d’un gestionnaire de messages qui scrute chaque arrivée de message, chaque demande d’envoi de message et qui redistribue ces messages en fonction de leur nature. De manière générale, on définit les fonctions `TraiterMArrivée()` et `EnvoieM()` où M est la nature du message à gérer, par exemple `TraiteAcceptanceArrivée()` et `EnvoieReservation()`. Ceci s’applique aussi dans le cas où la ressource est goulet, ce qui implique l’arrivée de messages du type: `TraiteReservationArrivée()` pour la déclaration de l’état goulet. Il est donc nécessaire d’associer la nature même du message à l’état dans de la SIP en ce qui concerne le(s) protocole(s), en effet, plusieurs messages de type RESERVATION peuvent arriver par exemple. Les procédure `EnvoieM()` font appel aux procédures `InitialisationM()` de la classe sous-système d’information afin de paramétrer le message selon la nature du message et le protocole (cf. classe Message).

Ce gestionnaire a pour rôle:

- de scruter chaque liste de sortie du système de communication (dans lequel le gestionnaire de niveau A dépose les messages à envoyer) et de gérer le transport effectif du message vers son destinataire via le réseau. C’est la fonction client,
- la fonction Serveur qui doit pouvoir répondre à plusieurs clients qui essayent de se connecter,
- d’assurer le codage/ décodage du contenu d’un message X de $\langle S,C,X \rangle$ pour le rendre compatible avec le support de communication (réseau).

Afin de réaliser ces fonctions, il est nécessaire pour le serveur de posséder les identificateurs de connexion de chaque client et, pour chaque client, l'identificateur (unique) de connexion vers le serveur. Les états et les temps d'attente associés sont alors définis.

Le schéma (figure 5) décrit dans le cas multipostes la structure logicielle du système de communication (Dans le cas mono-poste, le gestionnaire de niveau T est inexistant).

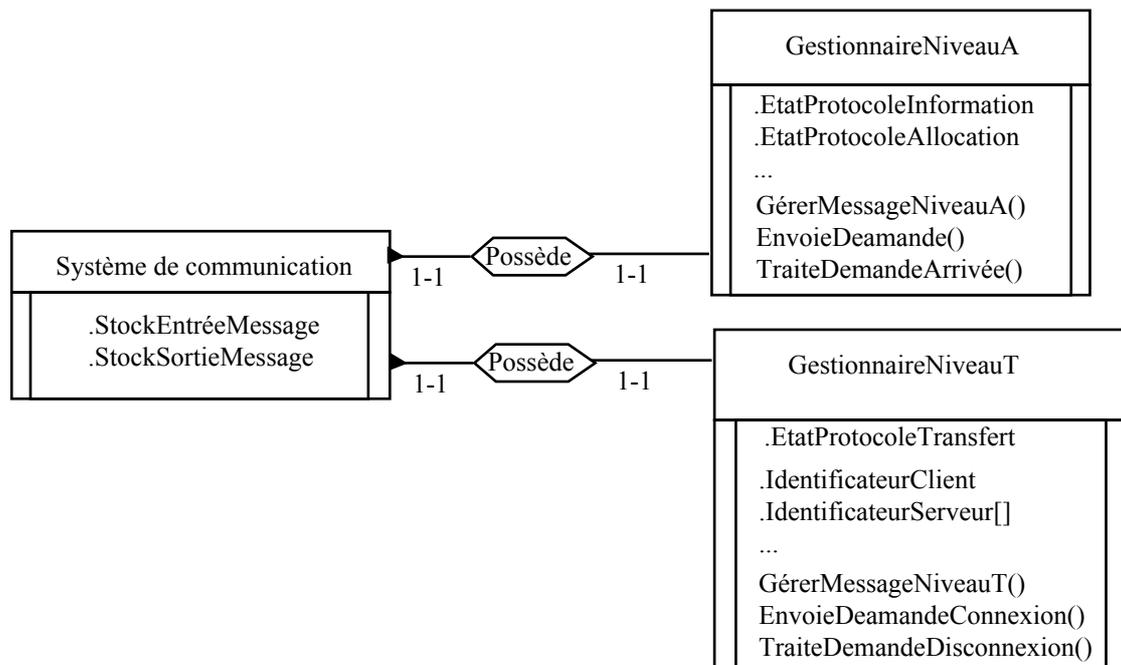


Figure 5: structure du système de communication

Un message est composé de quatre éléments:

- le type du message (en fonction du protocole adressé),
- l'identificateur de la source,
- l'identificateur de la cible,
- le contenu X qui est fonction du protocole. Par exemple, lors de l'émission d'une demande, c'est l'ensemble des contraintes imposées et les identificateurs des données à retourner en cas d'acceptance. Le cas du protocole d'ordonnancement peut correspondre à l'ordonnancement des tâches à réaliser pour la ressource goulet, etc.

Selon les cas, les identificateurs sont de nature différente; dans le cas mono-poste, c'est le nom (chaîne de caractère) de la SIP (source ou cible); dans le cas multi-postes, c'est l'adresse réseau de la ressource SIP. Des fonctions dédiées permettent alors de retrouver le nom associé à cette adresse.

3.2.4. Sous-système de contrôle

Le sous-système de contrôle est associé exclusivement à l'environnement de production. Ce dernier contient:

- le modèle du stock d'entrée/ sortie de produit, produit en cours, etc. le tout référencé en lien avec la liste d'entrée/ sortie d'ordres de fabrication et ordre de fabrication en cours du sous-système d'information, etc.
- les méthodes principales attachées au processus de fabrication: gestionnaire de production: ChargementOf(), DéchargementOf(), etc. L'extension apportée est le lien entre ces méthodes, les processus décisionnels (sous-système de décision) et l'opérateur (sous-système d'interface).

3.2.5. Sous-système d'interface

Les fonctions de l'interface sont liées au suivi du processus de fabrication, à l'élaboration et au déroulement des processus décisionnels, et à la surveillance des protocoles de communication.

3.2.5.1. *Information sur le processus de fabrication*

Les informations caractérisant le processus de fabrication sont:

- synoptique décrivant la réalisation dynamique des tâches. La meilleure approche consiste à présenter un diagramme de type Gantt dont l'échelle de temps évolue en fonction du temps de simulation,
- informations caractéristiques sur l'état de l'ordonnancement (ressource goulet): respect, avance, attente opération, etc. Elles peuvent figurer sur le synoptique décrivant la réalisation des tâches et constituent également une aide empirique lors du choix de remise en question de l'ordonnancement,
- un synoptique décrivant les entrées et sorties des produits (ordres de fabrication) afin de vérifier entre autres la bonne réception ou encore le bon déroulement du protocole d'allocation,
- un synoptique (cas monoposte) portant sur les transferts d'ordre de fabrication via le réseau simulé.

3.2.5.2. *Elaboration des processus décisionnels*

Un ensemble de synoptiques ont été développés, ils permettent de représenter:

- les niveaux de coopération,
- les matrices de décision, les listes de critères, les résultats intermédiaires et définitifs suivant la méthode (M1 ou M2) utilisée,

3.2.5.3. *Surveillance des protocoles de communication*

L'interface fournit:

- une trace des protocoles en cours du niveau A,

- une trace des protocoles en cours du niveau T (dans le cas de la structure multi-postes).

La nature et le contenu des messages échangés sont alors utiles pour fournir des informations en cas de problème de communication. Ceci constitue une aide empirique pour la remise en question de certains choix pendant le déroulement des protocoles de communication.

3.3. Structure logicielle superviseur

Le superviseur n'hérite pas de la classe agent. L'héritage se situe au niveau des sous-systèmes, ce qui permet de sur-définir certaines fonctions et de considérer un nouvel ensemble intégrant les fonctionnalités et les responsabilités spécifiques au superviseur. La structure complète est donnée figure 6 (seules figurent les éléments essentiels comme la classe Ordonnanceur définie ci-après).

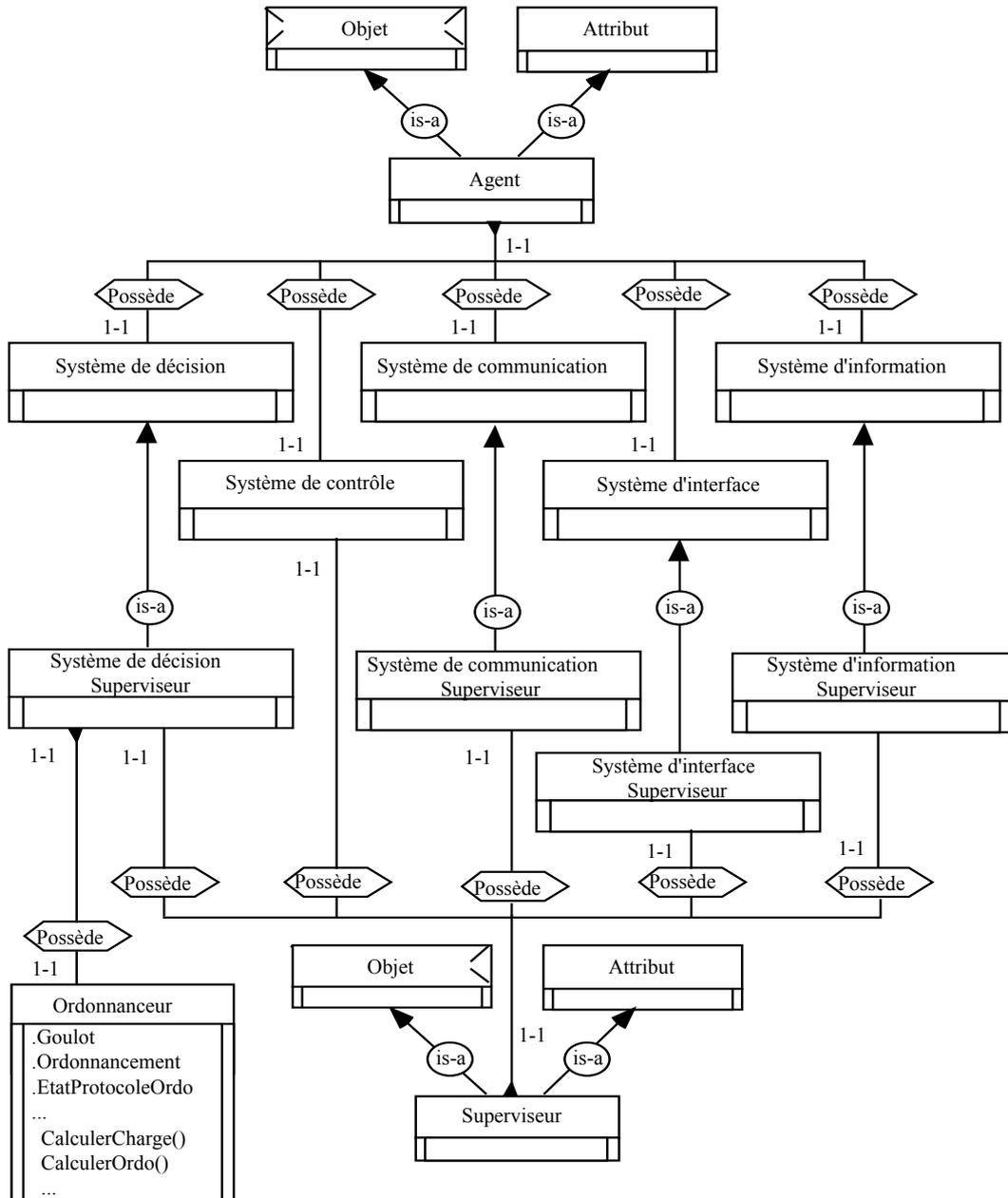


Figure 6: structures agent, superviseur et sous-systèmes

3.3.1. Sous-système de décision

Outre les fonctions de la structure distribuée, les processus décisionnels ont été étendus en ce qui concerne l'élaboration de l'ordonnancement: calcul des charges potentielles, détection de la ressource goulet et calcul de l'ordonnancement. En particulier, La classe "Ordonnanceur" a été définie pour gérer la réalisation du processus d'ordonnancement.

3.3.2. Sous-système de communication

D'autres données et fonctions membres ont été adjointes afin d'intégrer:

- l'aspect communication du processus d'ordonnancement: protocole d'ordonnancement,
- l'extension du protocole d'information dans le but d'informer chaque agent du début de simulation.

3.3.3. Sous-système d'information

Il contient outre les informations communes à toutes les SIP, les informations propres aux processus décisionnel du niveau Superviseur: liste dynamique des ordres de fabrication, liste des opérations planifiées (après processus d'ordonnancement), ordonnancement planifié en cours, etc.

3.3.4. Sous-système de contrôle

Ce sous-système est identique à celui d'une SIP de niveau distribué.

3.3.5. Sous-système d'interface

Outre les fonctionnalités précédentes, un synoptique pour l'affichage des données centralisées par le Superviseur a été développé. Il apporte à l'opérateur de supervision une aide potentielle pour la remise en question de l'ordonnancement par exemple.

4. Mise en oeuvre, cas mono-poste

Une classe théorique sans équivalent physique a été définie. Cette classe "Atelier" intègre l'ensemble des agents qui ne communiquent entre-eux que par l'intermédiaire d'un adresse sous forme de chaîne de caractère et qui représente l'adresse "réseau". Elle contient également le simulateur de réseau (modélisé à cette étape par un retard pur sur chaque message et gérés suivant la règle FIFO).

Cette structuration est résumée sur la figure 7.

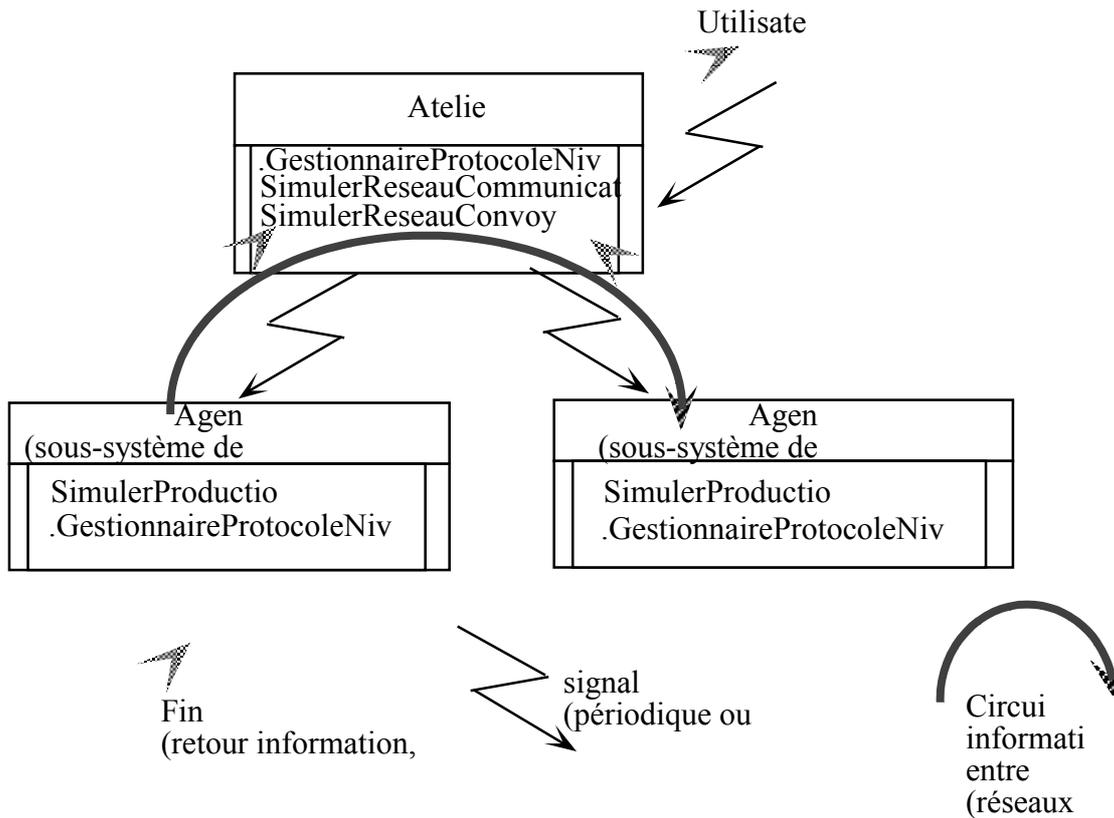


Figure 7: simulation mono-poste

La classe “Atelier” est associée à un environnement de simulation qui porte notamment sur le fonctionnement des réseaux de communication. La simulation du réseau de communication consiste à scruter chaque point d’entrée et de sortie du sous-système de communication SIP afin de délivrer (ou d’extraire) les messages aux SIP concernées (ou de charger ceux-ci dans le réseau). Ces deux réseaux sont modélisés par deux files d’attente à retard pur. Une modélisation plus fine n’est pas nécessaire car dans l’optique multi-postes, la structure distribuée fait directement appel à un réseau existant.

5. Evaluation

5.1. Cas d’étude industriel

Le cas d’étude porte sur un atelier flexible de finition d’une usine spécialisée dans la fonderie de précision.

La structure de production globale comporte deux parties (figure 8): une partie présentant des caractéristiques flow shop et une seconde job-shop.

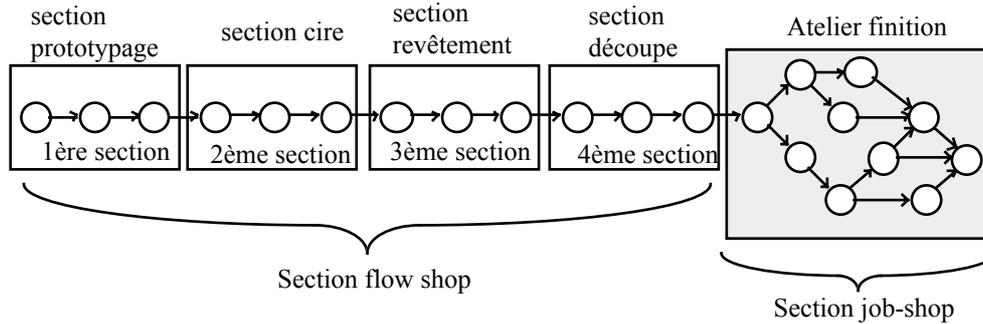


Figure 8: représentation schématique de l'usine

La méthode pour obtenir les pièces de fonderie s'appuie sur le principe de la cire perdue. Ce principe est le suivant:

- section flow shop: un prototype de la pièce est réalisé. Ce prototype donne naissance à quelques moules réalisés à l'aide de ressources physiques telles que l'électro-érosion. La seconde section réalise autant de pièces spécifiées par le cahier des charges (inclus les rebuts éventuels) à partir des moules. Les pièces sont réalisées à partir d'une cire spéciale. Ces pièces sont alors fixées en grappe. Les grappes passent ensuite dans la section revêtement où elles sont enduites de différentes couches de céramiques. Une fois sèches, ces grappes sont cuites dans des fours thermiques. La cire est détruite et ne reste que la couche en céramique cuite. De l'acier en fusion (nuances) est alors versé dans ces grappes. Une fois refroidies, ces grappes sont détruites et sont tronçonnées. Les pièces sont alors prêtes à entrer dans l'atelier de finition.
- section job-shop: les produits semi-finis sont alors ébavurés, sciés, soumis à traitement thermique, grenillés, contrôlés (contrôle visuel et qualité), calibrés, emballés, expédiés. Certaines étapes peuvent être répétées plusieurs fois selon la gamme opératoire.

L'étude porte sur onze ressources. Les ordres de fabrication ont été élaborés à partir de plusieurs gammes de fabrication réelles.

5.2. Objectifs

L'objectif de l'entreprise est de travailler en juste à temps. De ce fait, nous nous attachons à l'intégration de cet objectif global. La définition de la caractéristique "juste à temps" d'un ordonnancement est proposée dans [Trentesaux et al. 95a].

Cette caractérisation se base sur le retard moyen et l'écart type des retards de l'ensemble des ordres de fabrication (annexe G).

6. Support matériel pour la mise en oeuvre de la maquette

La maquette a été développée sur les matériels et logiciels suivants:

- matériels: ordinateurs de type PC 80486 (calculateurs mono-processeur),
- logiciels: Microsoft Windows 3.1, Borland C++ 3.1, Winsockets, gestionnaire réseau TCP/IP.

6.1. Efficacité de l'ordonnement a posteriori

L'étude porte sur la charge globale de la structure de production pour un paramétrage fixe de la méthode multicritère. L'efficacité des ordonnancements a posteriori est évaluée en les comparant à ceux calculés, par une heuristique, de manière prévisionnelle par un logiciel d'ordonnement (SavePlan). Cette évaluation n'a de sens que dans le cas d'un fonctionnement non dégradé (pas de panne, etc.). On compare ainsi:

- une approche de pilotage hybride par ordonnancement prévisionnel partiel. L'ordonnement est calculé selon la méthode présentée en annexe H,
- une approche de pilotage distribué sans ordonnancement prévisionnel.

Nous présentons les résultats qui sont décrits en détail dans [Trentesaux et al. ??a].

L'étude comporte six simulations différentes. Chaque simulation consiste en la réalisation de 10 ordres de fabrication (composés en moyenne de huit opérations chacun). La charge globale est décroissante avec le numéro de la simulation. Les résultats sont indiqués sur la table 1.

Numéro Simulation	Pilotage distribué		Pilotage Hybride		SavePlan	
	Moyenne retard	Ecart type	Moyenne retard	Ecart type	Moyenne retard	Ecart type
1	50.25	244.58	81.91	34.37	117.00	241.25
2	56.87	212.17	86.72	71.12	138.00	282.56
3	58.57	263.35	77.28	48.20	128.00	278.52
4	-87.86	288.17	68.35	56.67	23.00	155.28
5	-157.62	273.73	76.63	67.14	55.00	175.52
6	-159.49	348.31	23.10	11.56	-8.00	7.89

Table 1: résultat de simulation, fonctionnement non dégradé

Ces résultats sont illustrés sur la figure 9.

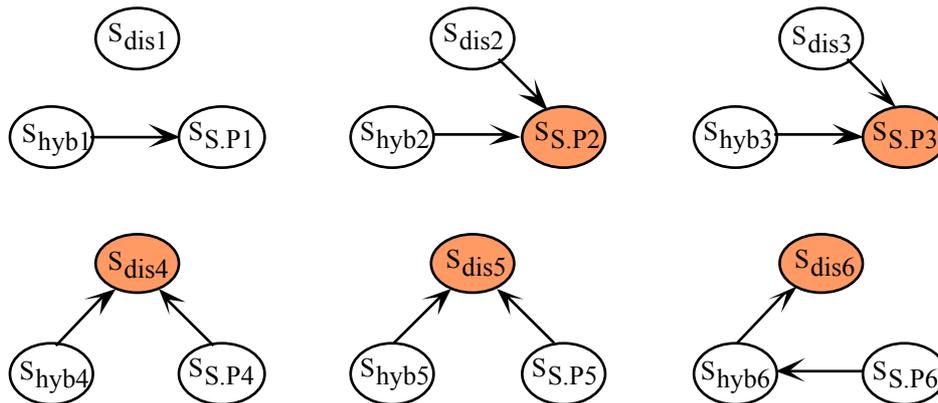


Figure 9: surclassement des ordonnancements

Un ordonnancement surclassé par les deux autres (c'est à dire "moins juste à temps que" les deux autres) est grisé. La flèche indique les surclassements entre deux ordonnancements (une flèche de a vers b signifie que a surclasse b). Les relations d'indifférence ne sont pas indiquées.

On constate que:

- moins la charge globale est élevée, meilleur est l'ordonnancement proposé par SavePlan (cf. les simulations 4, 5 et 6). Le moins efficace est celui proposé par la structure distribuée.
- plus elle est élevée, moins il est possible de discerner un surclassement (cf. simulation 1). L'ordonnancement proposé par SavePlan perd de son efficacité à cause de sa méthode d'ordonnancement qui consiste à planifier un maximum d'ordres de fabrication dans les délais, au détriment des autres. Ce qui se répercute au niveau de la moyenne des retards.

Les conclusions de cette étude sont:

- l'efficacité de la structure distribuée n'est pas prouvée. L'ordonnancement réalisé est souvent surclassé par les deux autres à la fois. Nous concluons à la mauvaise intégration de l'objectif global de production juste à temps. Ceci s'explique par le fait suivant: si la charge de travail est élevée, la structure est confrontée comme les deux autres approches à un problème insurmontable. Ses résultats ne peuvent être mieux. Si la charge est faible, les tâches sont réalisées en flux poussées et sont terminées en avance, réduisant ainsi les performances globales.
- la structure hybride gère efficacement les différents cas de charge de travail. En effet, l'ordonnancement réalisé n'est jamais le meilleur, jamais le moins bon, et ce, quelque soit la charge (très élevée ou très faible). Le pilotage hybride propose donc un compromis efficace en terme de gestion juste à temps.

Cet exemple illustre la possibilité du pilotage hybride à prendre en compte des objectifs globaux. La méthode peut ainsi être qualifiée de robuste en terme de charge atelier.

6.2. Intégration de l'auto-adaptation et la robustesse

L'étude a porté sur l'intégration de la flexibilité et de la réactivité, et en particulier sur l'intégration de la robustesse au niveau A.

Pour chacun de ces deux aspects, nous avons étudié extensivement les trois typologies (produit, ressource et flux). Nous présentons quelques résultats extraits de [Trentesaux et al. ??a].

6.2.1. Réactivité - point de vue flux-

Dans cette étude, cinq ordres de fabrication (avec un nombre moyen d'opérations de sept) sont disponibles à la date 0. A la date 200, deux autres ordres de fabrications sont fournis. De même à la date 320. A partir de cette date, neuf ordres de fabrication sont en cours dans l'atelier. Le superviseur réordonne le goulet à ces deux dates. On suppose également que tous les paramétrages portant sur la méthode multicritère sont fixés constants.

Les deux ordonnancements réalisés sont représentés sur les figures 10 et 11. Le symbole triangle blanc qui figure sur l'ordonnancement réalisé par le pilotage hybride représente la date de début de fabrication d'une opération planifiée.

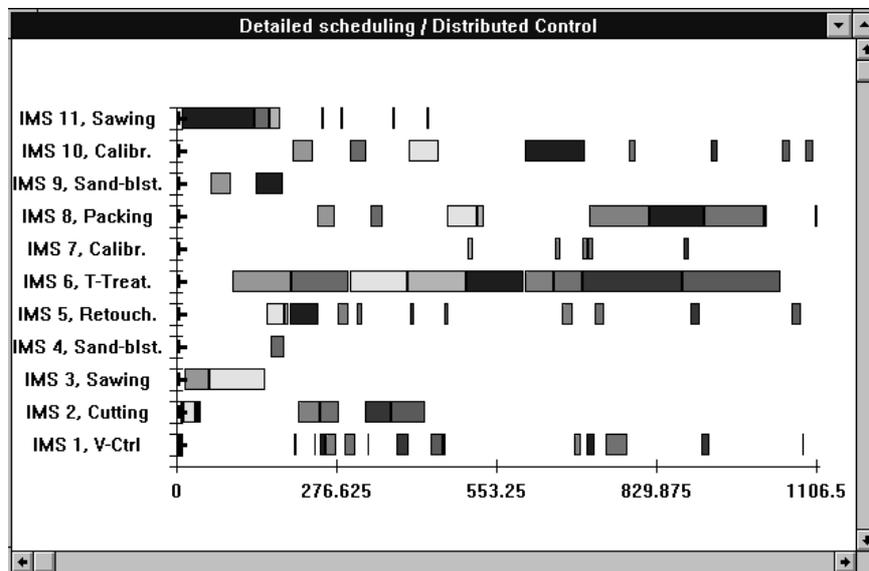


Figure 10: pilotage distribué

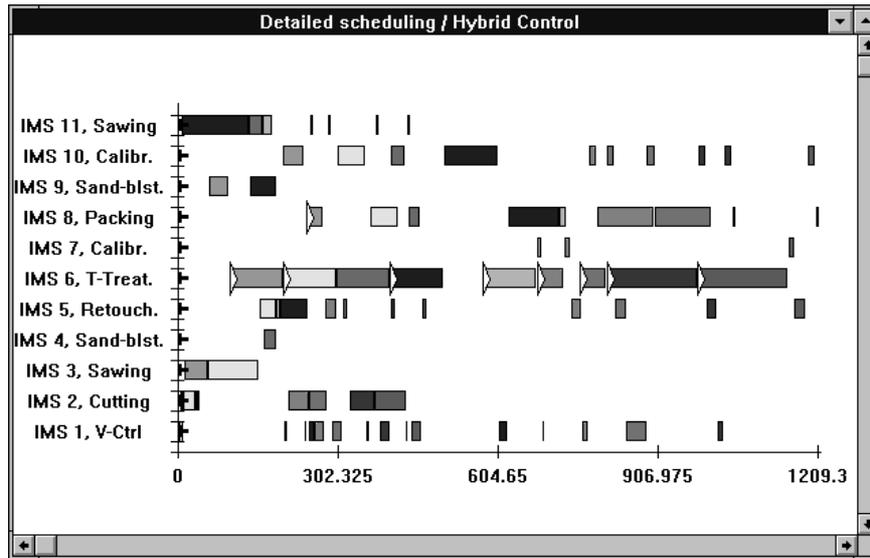


Figure 11: pilotage hybride

En ce qui concerne le pilotage hybride, de la date 0 à la date 200, l’agent 6 est goulet. La modification de la charge réelle provoque un changement de ressource goulet: de la date 200 à la date 320, c’est l’agent 8 qui est goulet. Lors du deuxième ajout d’un ordre de fabrication, l’agent 6 redevient à nouveau ressource goulet.

Les résultats sont indiqués dans la table 2:

Numéro ordre fabrication	Date arrivée	Date besoin	Pilotage distribué		Pilotage hybride	
			Date Fin réelle	Retard/ Avance	Date Fin réelle	Retard/ Avance
#1	0	250	272.8	+22.8	272.8	+22.8
#2	0	700	916.4	+216.4	716.6	+16.6
#3	0	400	413.4	+13.4	413.4	+13.4
#4	0	450	456.4	+6.4	456.4	+6.4
#5	0	700	532.9	-167.1	729.6	+29.6
#6	200	900	822.5	-77.5	894.7	-5.3
#7	200	1000	1020.7	+20.7	1002.6	+2.6
#8	320	1000	1025.7	+25.7	1049.5	+49.5
#9	320	1100	1094.2	-5.8	1204.9	+104.9

Table 2: retard/ avance des ordres de fabrication

Les valeurs en caractères gras correspondent à l’avance et au retard maxima pour les deux approches de pilotage. On obtient:

- $(m_d, \sigma_d) = (6.1; 101.4)$ pour la structure distribuée,
- $(m_h, \sigma_h) = (26.7; 33.4)$ pour la structure hybride.

D’après la définition donnée en annexe G pour la qualité d’un ordonnancement juste à temps, on remarque que $\underline{S}_d \approx \underline{S}_h$. C’est à dire que l’on ne peut établir une préférence.

Cependant, en remarquant que les retards sont négligeables par rapport aux dates de besoin (ils représentent au pire 5% de la moyenne des dates), on constate

une différence importante au niveau des écarts type: σ_d est élevée et réduit l'efficacité du pilotage concernant les ordres de fabrication 2 et 5. La différence se situe au niveau de la capacité pour le pilotage hybride à intégrer ce type d'aléa en exploitant la notion de goulet.

6.2.2. Réactivité - point de vue produit-

Une évaluation de l'influence des paramétrages de la méthode multicritère a été réalisée pour deux critères qui sont présentés en annexe C en ce qui concerne les processus d'allocation, mode "autonome". Cette étude met en relief la commandabilité du processus d'allocation par le poids des critères. Les trois stratégies sont les suivantes:

- favoriser le critère D1 (temps opératoire) par rapport aux autres critères,
- favoriser le critère D4 (prochain date de disponibilité) par rapport aux autres critères,
- favoriser à la fois D1 et D4 par rapport aux autres critères.

Les résultats sont données sur la table 3.

Ordre de fabrication	Date de besoin	Pilotage Hybride			Pilotage distribué		
		$\underline{S}_{h/D1}$	$\underline{S}_{h/D4}$	$\underline{S}_{h/D4,D1}$	$\underline{S}_{d/D1}$	$\underline{S}_{d/D4}$	$\underline{S}_{d/D1,D4}$
		Retard D1	Retard D4	Retard MIX	Retard D1	Retard D4	Retard MIX
1	250	-0,4	+32,3	+11,6	+282	+32,3	+41,7
2	700	+73,3	+85,3	+238,6	+281	+232,2	+132
3	450	+35,4	+47,4	+63,3	-82,8	+47,4	+433,4
4	300	+35,4	+55,8	+55,8	+318	+55,8	+47,4
5	500	+16,6	+32,5	+32,5	-101,6	+32,5	-67,7
6	880	+101,9	+113,3	+162,9	-96,6	-146	-246,2
7	850	+27,6	+39,6	-5,3	+37,7	-11,7	-111,9
8	950	+36,9	+48,9	+97,4	+54,7	-7,6	-7,6
9	1000	+124,7	+132,6	+132,6	+164	+101,7	+101,7
	(\bar{m}, σ)	(50,2; 41,2)	(65,3; 36,6)	(87,7; 78,8)	(95,2; 171,7)	(37,4; 100,4)	(35,9; 189)

Table 3: influence du critère

Le graphe de surclassement est le suivant (figure 12):

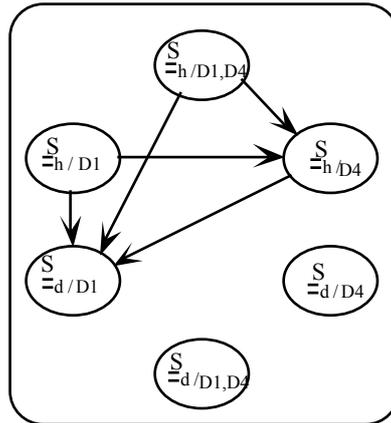


Figure 12: relation de surclassement

Outre le fait que le pilotage hybride semble globalement plus performant que l'approche centralisée, cet exemple met en évidence la capacité d'intégration des préférences sous la forme de critère.

6.2.3. Réactivité - point de vue ressource-

Nous donnons une série d'exemple qui montre que la réalisation du processus de fabrication n'est pas contraint par l'occurrence de pannes (paramétrage multicritère fixe). Une panne est représentée par un rectangle hachuré (figures 13, 14 et 15). Le type de panne concerné est celui d'une panne de l'outil de production qui peut être associée à une panne du système de communication.

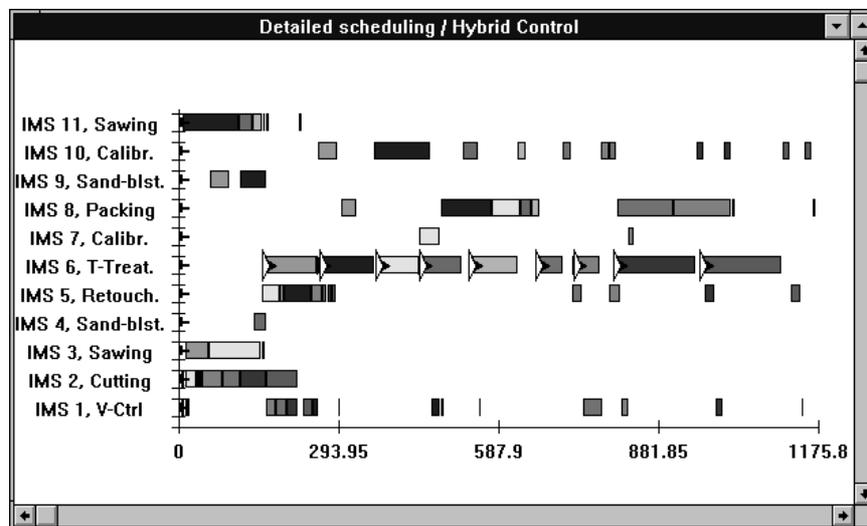


Figure 13: pilotage hybride sans occurrence de panne

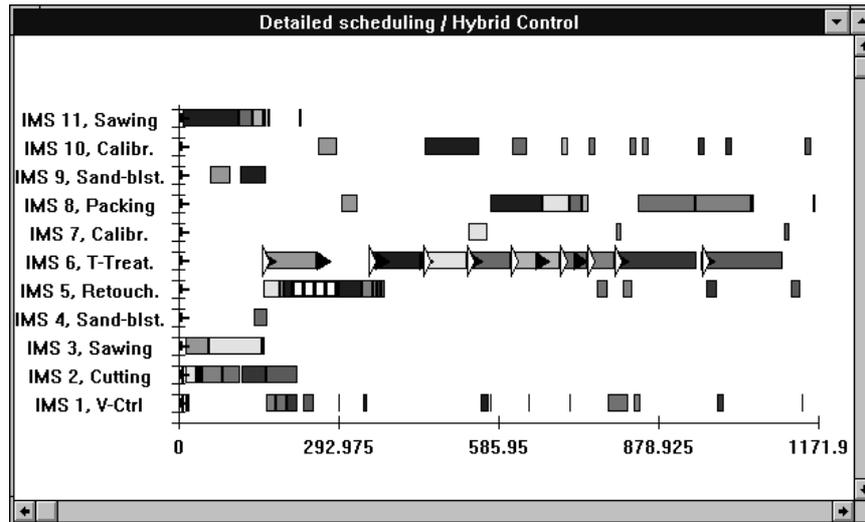


Figure 14: pilotage hybride avec une occurrence de panne (SIP 5)

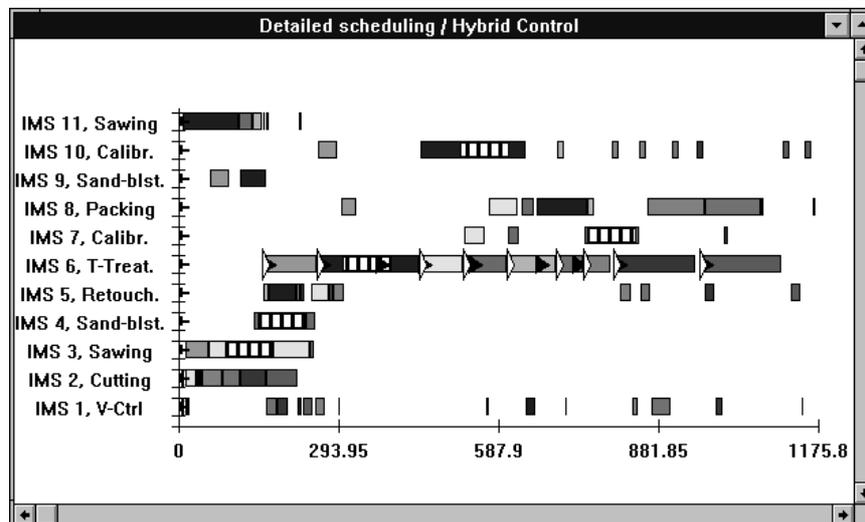


Figure 15: pilotage hybride avec cinq occurrences de panne (SIP 3, 4, 6, 7 et 10)

Il nous n'est pas possible d'évaluer la qualité de l'ordonnancement en fonction des pannes car elles n'ont pas la même influence sur le pilotage. Cet exemple met en évidence les capacités à intégrer les aléas de fonctionnement.

6.2.4. Flexibilité - point de vue ressource-

Une autre étude concerne l'intégration d'une ressource particulière dans le but d'absorber une surcharge de production (paramétrage multicritère fixe). Cette surcharge est relativement importante.

Dans l'exemple étudié, l'agent 12 est intégré à la date 240.

Ces exemples détaillent l'intégration de la ressource 12 dans les deux approches de pilotage.

6.2.4.1. Pilotage distribué

L'influence de l'adjonction de la SIP 12 est montrée sur les figures 16, 17 et 18

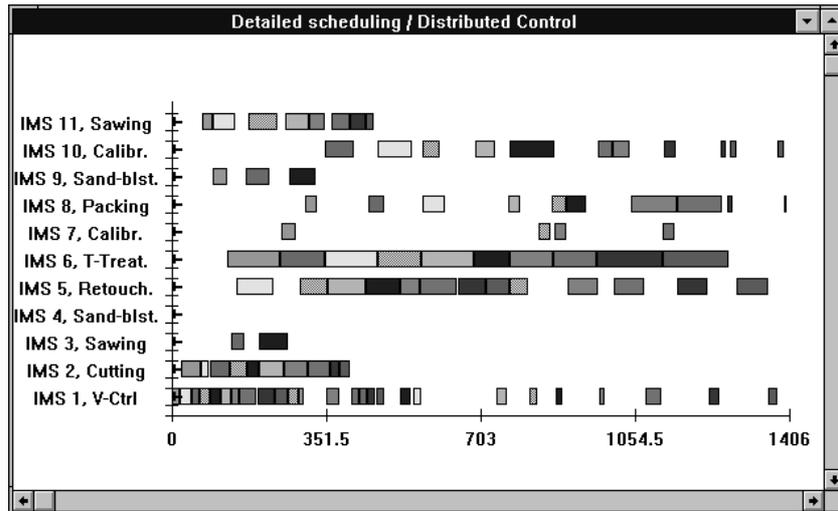


Figure 16: pas de modification structurelle

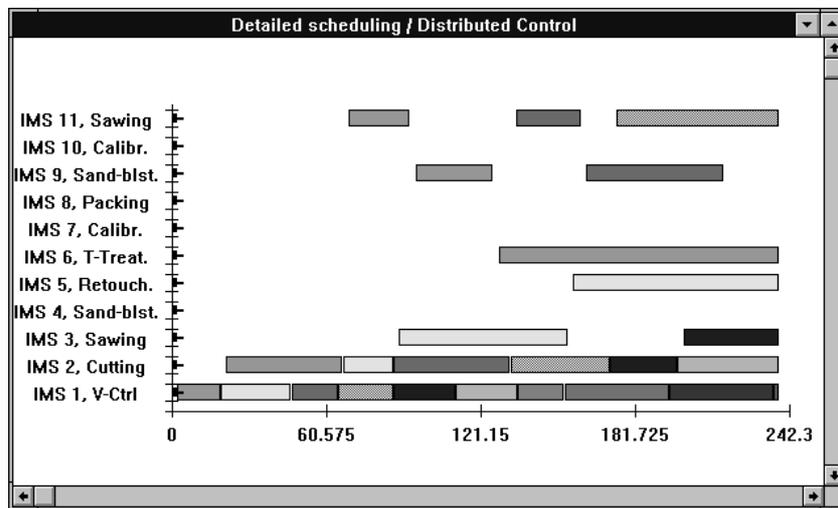


Figure 17: 11 agents sont disponibles jusqu'à la date 240

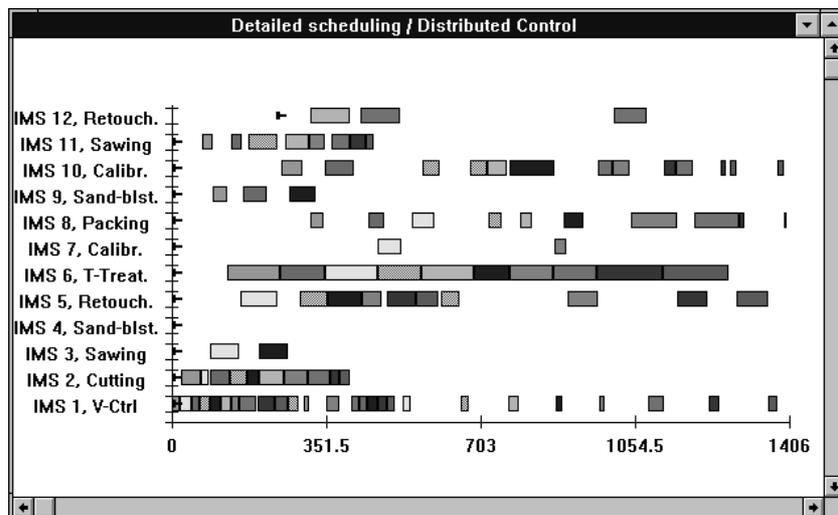


Figure 18: IMS 12 est intégrée à la date 240

6.2.4.2. Pilotage hybride

L'ordonnancement est remis en question dès l'intégration de l'agent 12. Un triangle noir représente la date planifiée, un triangle blanc, la date effective de début de travail (figures 19, 20 et 21).

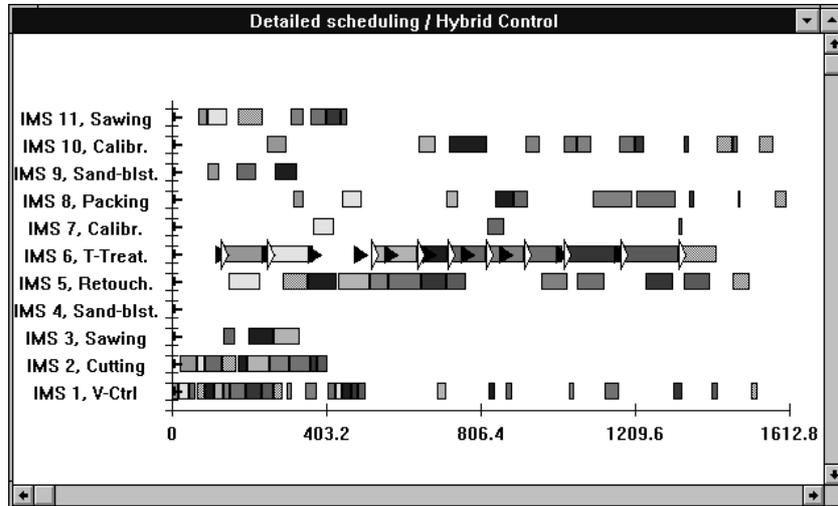


Figure 19: pas de modification structurelle

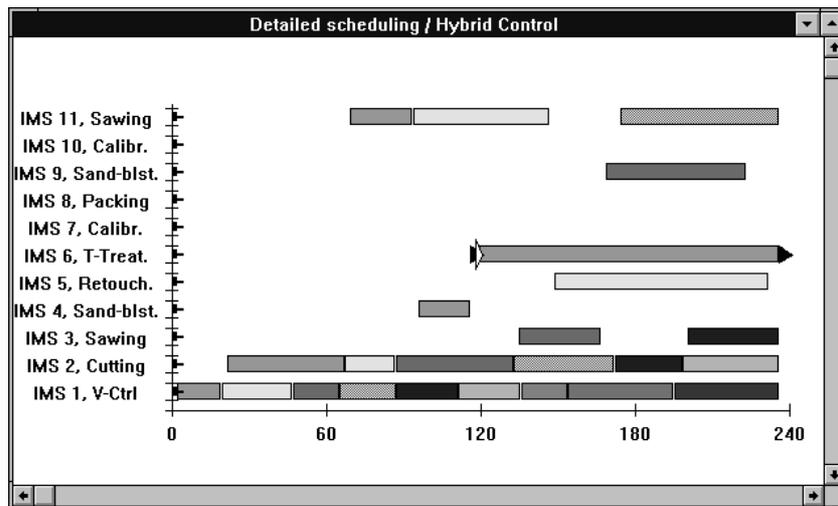


Figure 20: jusqu'à la date 240, 11 agents sont disponibles

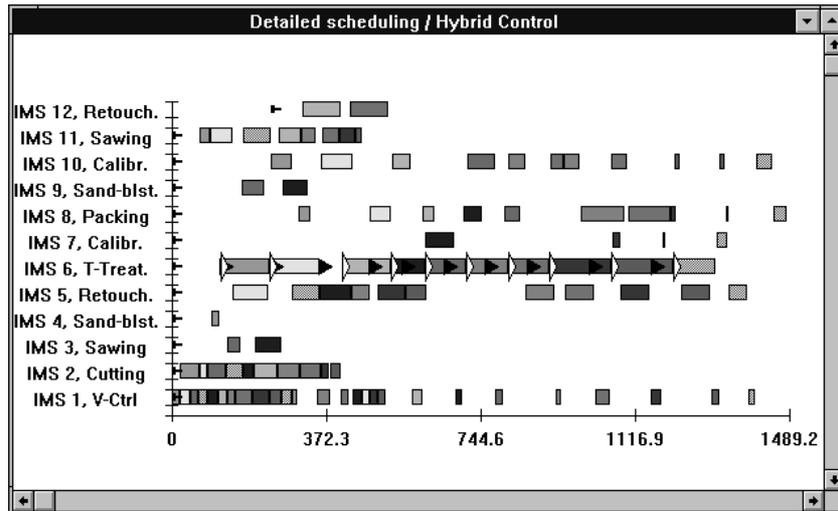


Figure 21: IMS 12 est intégrée à la date 240

On constate la difficulté du système de pilotage à absorber la charge dans le premier cas. On remarque que de nombreuses opérations n’ont pas pu être débutées à la date de planification. L’adjonction de l’agent 12 permet d’absorber une partie de cette surcharge et améliore de manière significative l’efficacité de l’ordonnancement. Les résultats sont les suivants (table 4):

	Pilotage distribué		Pilotage hybride	
	11 Agents	12 Agents	11 Agents	12 Agents
(m, σ)	(85,4;245,9)	(79;288,3)	(208,9;96,4)	(112,6;54,6)

Table 4: retard et écart type selon les deux structures de pilotage

La structure hybride intègre plus efficacement l’ajout d’un agent que la structure distribuée dans l’objectif d’absorber une surcharge supplémentaire, malgré l’absence de surclassement. Ceci est justifié par le faible écart type qui est conservé: cela signifie que le pilotage hybride réalise les tâches au mieux compte tenu des capacités des ressources.

6.3. Conclusion

L’étude présentée montre que la structure hybride supporte d’une manière plus efficace qu’une structure distribuée un objectif global de production juste à temps. L’efficacité de l’intégration de l’objectif juste à temps résulte principalement de la réalisation d’un ordonnancement prévisionnel partiel et de la notion de charge potentielle.

En terme d’auto-adaptabilité, nous avons montré sur de nombreux exemples les possibilités de la structure de pilotage.

Par ailleurs, l’intégration de la robustesse du niveau A a pu être vérifiée notamment par la simulation de pannes au niveau de certains agents.

Un autre cas d'étude concerne un système de production de type flow-shop. La structure étudiée est une chaîne d'embouteillage de la société Perrier. L'étude est développée dans [Trentesaux et al. 95b]. Elle porte notamment sur la pertinence de l'intégration d'une méthode multicritère (plus précisément la méthode M1).

7. Evaluation multi-postes

Chaque ressource informatisée (PC) correspond à un agent (figure 22). On constate que la gestion du niveau T est désormais du ressort de chaque agent. De ce fait, la structure logicielle est légèrement différente. Il faut noter que cette structuration a été réalisée rapidement, ceci grâce à la spécification et la modélisation orientée objet qui permet l'extensibilité et la généricité du modèle.

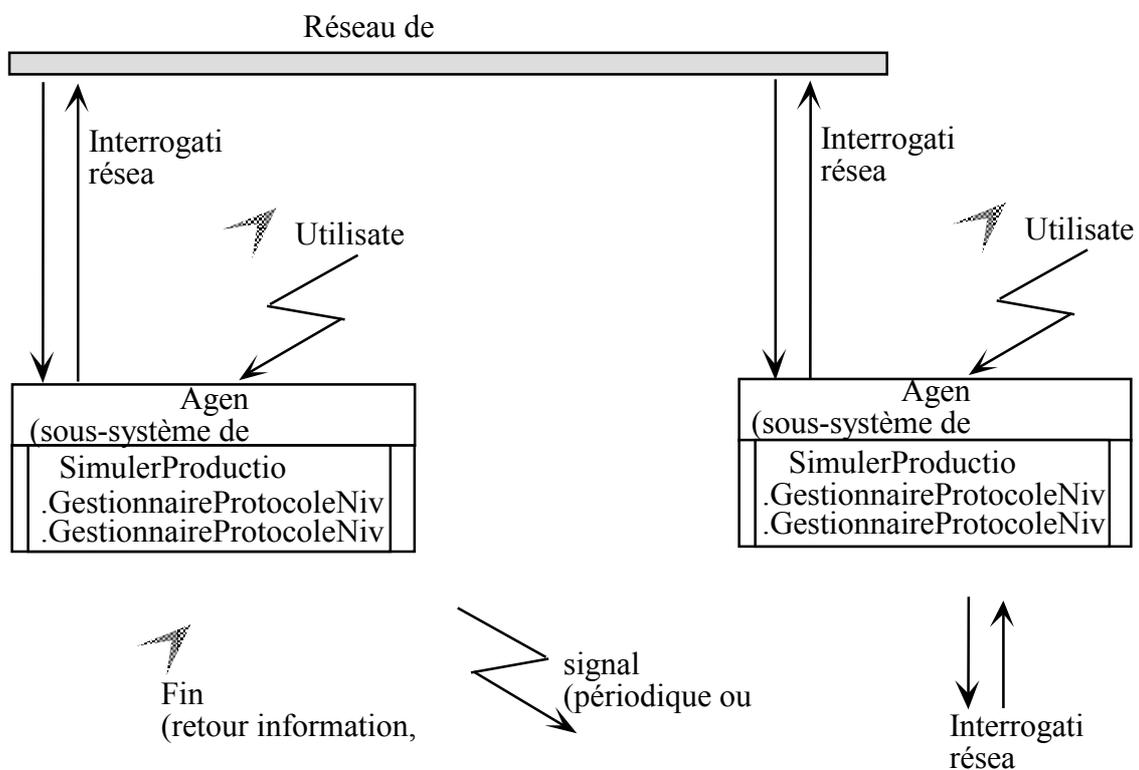


Figure 22: prototype réellement distribué

Le développement d'une simulation multipostes a été réalisée sur trois ordinateurs de type PC via une connexion sur InterNet par une bibliothèque de fonction de Socket pour Windows.

Ce développement a permis de vérifier le bon fonctionnement du protocole de niveau T en mode normal et dégradé. Elle permet par ailleurs de prendre en compte les contraintes réseau telles que le nombre de connexions possibles, la taille maximale des données envoyées, etc. Ce prototype multi-poste est développé en particulier pour intégrer et évaluer la notion de SIAMD distribué introduite dans ce mémoire. Ce type de validation n'a pas pu être menée dans le cadre des travaux

présentés dans ce mémoire et constitue une des principales perspectives de recherche.

8. Conclusion

Dans ce chapitre nous avons montré une structuration logicielle correspondant aux méthodes et modèles présentés dans les chapitres précédents.

Nous avons ensuite montré la validité des concepts proposés en décrivant les résultats obtenus sur un cas de processus industriel de type job-shop (atelier de finition fonderie). Ces résultats portent sur:

- l'évaluation relative des ordonnancements a posteriori réalisés par les structures de pilotage hybride et distribué afin d'évaluer la réactivité des deux méthodes. Ce point met en outre en évidence la qualité de la résolution multicritère,
- l'évaluation conjointe des ordonnancements a posteriori réalisés par les structures de pilotage hybride et distribué avec un ordonnancement réalisé a priori par un logiciel SavePlan. Cette étude avait pour objectif de mettre en évidence les capacités des deux structures à intégrer de manière statique un potentiel de charge différent,
- la mise en évidence de l'influence des paramétrages de la méthode multicritère dans l'optique d'intégrer un SIAMD distribué,
- la mise en évidence de la qualité (point de vue réactivité) du pilotage dans différents cas d'étude.

Nous avons montré que la structure de pilotage hybride et multicritère propose un compromis efficace en terme de réactivité, flexibilité et d'intégration d'objectifs globaux.

Conclusion générale

Nous avons proposé dans ce mémoire un modèle de pilotage hybride qui s'appuie sur:

- une structure distribuée, supervisée et multicritère élaborée à partir des concepts de l'intelligence artificielle distribuée, et appliquée au pilotage,
- la détection dynamique et l'ordonnancement d'une ressource goulet. Le principe de détection se fonde sur le concept de charge potentielle qui détermine, à un instant donnée, la charge probable de chaque entité du processus de pilotage. L'ordonnancement concerne la ressource goulet présentant la charge maximale,
- la spécification d'un système interactif d'aide multicritère à la décision. L'intégration d'un tel système est justifiée par la mise en évidence d'un ensemble de processus décisionnels peu structurés. Ceci nous a conduit à proposer l'intégration de l'opérateur dans la prise de décision, ce qui doit assurer ainsi une meilleure prise en compte de l'adaptabilité du processus de pilotage, le développement d'une méthode multicritère facilitant la mise en oeuvre de la notion d'interactivité,
- le développement d'une structure de communication décomposée en deux niveaux. Le premier niveau concerne le déroulement des processus décisionnels. Le second niveau gère la transmission de l'information elle-même.

Nous avons montré sur un cas industriel que ces modèles apportent un compromis efficace en terme de flexibilité et de réactivité temps-réel tout en permettant l'intégration d'objectifs globaux tels qu'assurer une politique de production juste à temps.

Perspectives

De nombreuses perspectives tant sur les aspects recherche qu'applicatif sont à envisager. Elles ont été introduites tout au long de ce mémoire. Elles concernent la structure de pilotage, les processus décisionnels, les protocoles de communication et le développement d'un SIAMD distribué.

Structure de pilotage

Les perspectives portent sur:

- la définition d'un regroupement des ressources par famille. La réalisation de l'opération allouée est de la responsabilité de la famille. Cette approche introduit un deuxième niveau de décision pour l'allocation de cette opération au sein de la famille.
- La définition d'une structure de supervision dynamique. Chaque agent est capable d'assumer les responsabilités du superviseur, mais un seul agent supervise à un instant donné le processus de fabrication. En cas de panne de cet agent, le rôle de superviseur est alloué à un autre agent. L'intérêt est d'éviter de rendre une ressource critique. La définition d'une telle structure est cependant très complexe.
- l'étude de différents critères de sélection des opérations ou des ressources planifiables. Nous avons montré dans ce mémoire que la gestion par goulet constitue une approche générique efficace dans un cadre d'objectif global de production juste à temps. Il est cependant envisageable d'étudier, en fonction des objectifs ou de la structure de production, la pertinence du choix des critères.
- la conception d'une structure hybride aux niveaux de décision tactique et stratégique. Cette approche permettrait d'intégrer par exemple la notion d'usine comme ressource-SIP. Cette approche fait référence aux notions d'entreprise virtuelle ou multi-sites.
- le développement de la notion de Kanban généralisé. Nous avons en effet considéré la gestion par Kanban comme une approche distribuée du pilotage. En outre, nous avons développé une structure de communication (dans un contexte de pilotage distribué) qui nous pousse à considérer l'échange d'information par la notion de protocole d'application comme une généralisation de l'approche Kanban.

Processus décisionnels

Suite à la mise en évidence dans ce mémoire de la faisabilité de l'approche multicritère (problématique de rangement) dans un cadre de pilotage hybride, les perspectives portent en particulier sur:

- le développement de la problématique de choix et de l'étude de faisabilité associée,

- le développement d'autres méthodes multicritères dans un cadre d'évaluation relative afin d'améliorer les performances du pilotage et l'intégration de l'opérateur,
- l'intégration et l'évaluation d'une méthode d'utilisation des données incertaines pour définir la notion de poids des critères incertains,
- l'intégration du modèle d'Activity Based Management dans la phase d'information afin d'améliorer la qualité des processus décisionnels.

Protocoles de communication

Il est nécessaire d'évaluer différents protocoles de communication (asynchrone ou synchrone) permettant de mettre en évidence la robustesse et l'extensibilité du système de pilotage.

Une autre perspective porte sur la modélisation d'un niveau intermédiaire entre les niveaux A et T, et assurant la structuration (décomposition, agrégation, etc.) des messages afin de les rendre compatibles avec les contraintes imposées par le support de communication.

SIAMD distribué

Ce point constitue la perspective majeure de ce mémoire. Il n'existe actuellement pas de spécification ou de modèle générique de SIAMD distribué au niveau opérationnel. Il nous semble important de pouvoir définir une structuration générique permettant l'application à divers domaines, dans un environnement temps-réel et distribué.

Le projet système coopératif d'aide à la décision pour les systèmes à événements discret (SCAD) mené dans le cadre d'un programme CNRS de coopération SHS-SPI porte notamment sur cet aspect.

Bibliographie

[Ancelin 89] Ancelin B., Quels critères de performance pour les nouveaux ateliers?, *Revue française de gestion industrielle*, n°1, 1989, pp. 67-79.

[Arai et al. 93] Arai E., Amnuay S. T. et Uchiyama N., Real time simulation and monitor forecast for dynamic scheduling in distributed production systems, *Proceedings of JSPE-IFIP WG 5.3 Workshop DIISM'93*, Japon, 1993, pp. 137-147.

[Archimède 91] Archimède B., Conception d'une architecture réactive distribuée et hiérarchisée pour le pilotage des systèmes de production, *Thèse de doctorat*, université de Bordeaux I, Bordeaux, France, 1991.

[Bakker 88] Bakker H., DFMS : A new control structure for FMS, *Computers In Industries*, vol. 10, 1988, pp. 1-9.

[Belton et al.92] Belton V. et Elder M. D., Can Multiple Criteria Methods Help Production Scheduling?, *10th international conference on M.C.D.M TAIPEI*, Taïwan, 1992, pp. 341-349.

[Beauclair et al. 90] Beauclair R. A. et Straub D. W., Utilizing GDSS technology: final report on a recent empirical study, *Information and Management*, vol. 18, 1990, pp. 213-220.

[Beulens et al. 88] Beulens A. J. M. et Van Nunen J. A. E. E., The use of expert system technology in DSS, *Decision Support System*, vol. 4, 1988, pp. 421-431.

[Billaut et al. 93] Billaut J.C., Roubellat F., A decision support system for real-time production scheduling, *International conference on Industrial Engineering and Production Management*, Mons, Belgique, Juin 1993, pp. 449-462.

[Bitton 90] Bitton M., Ecograi: méthode de conception et d'implantation de systèmes de mesures de performances pour organisations industrielles, *thèse de doctorat*, université de Bordeaux I, Bordeaux, France, 1990.

[Bonczek et al. 80] Bonczek R. H., Holshapple C. W., Whinston A. B., Future directions for developing decision support systems, *Decision Science*, vol. 11, 1980, pp. 616-631.

[Bonczek et al. 84] Bonczek R. H., Holshapple C. W., Whinston A. B., *Developments in Decision Support Systems*, Advances In Computers, vol. 23, 1984, pp. 141-175.

[Bourgeois 93] Bourgeois S., Decision support system for dynamic job-shop scheduling, *Revue des systèmes de décision*, vol. 2, n° 3-4, 1993, pp. 277- 292.

[Bourgeois 94] Bourgeois S., Contribution à la conception de systèmes interactifs d'aide à l'ordonnancement dans les ateliers manufacturiers, *thèse de doctorat*, université de Valenciennes, France, 1994.

[Brans et al. 85] Brans J. P. et Vincke P., A preference ranking organization method- The Prométhée method, *Management Science*, vol 31, pp. 647-656, 1985.

[Canals 86] Canals D., Ordonnancement d'atelier par simulation: étude des règles de priorité et aide au lancement, thèse de docteur-ingénieur, *université de Toulouse*, France, 1986.

[Carlier et al. 82] Carlier J. et Chrétienne P., Un domaine très ouvert: les problèmes d'ordonnancement, *RAIRO Recherche opérationnelle*, vol. n° 16, n° 3, Août 1982, pp. 175-217.

[Cauffriez 94] Cauffriez L., Contribution à la mesure en temps-réel des performances de production d'ateliers manufacturiers dans les applications de supervision distribuées, *Thèse de doctorat*, Université de Valenciennes, France, 1994.

[Champetier et al. 95] Champetier M., Loppinet T., Barra C., Cerezuela C. et Kieffer J. P., Un système d'aide à la décision pour la conception de produits complexes dans un contexte d'ingénierie concurrente- application à la conception de systèmes électriques câblés, *Congrès International de Génie Industriel*, Octobre, Montreal, pp. 1491-1500.

[Coad et al. 92] Coad P. et Yourdon E., Analyse orientée objet, *Masson, MIPS*, Paris, 1992.

[Courbon 88] Courbon J. C., Les S.I.A.D: outil, concepts et mode d'action, *Afcet/interfaces*, n° 9, Juillet 1988, pp. 30-36.

[David et al. 89] David R. et Alla H., Du Grafcet aux réseaux de Petri, *Hermès*, Paris, 1989.

[Deckro et al. 82] Deckro R. F, Herbert J. E. et Winkofsky E. P., Multiple criteria job-shop scheduling, *Computer & operational research*, vol. 9, n° 4, 1982, pp.279-285.

[Desfray 92] Desfray P., Ingénierie des objets; approche classe-relation; application à C++, *Masson, MIPS*, Paris, 1992.

[DeSanctis et al. 87] DeSanctis G. et Gallupe B., A foundation for the study of Group Decision Support Systems, *Management science*, vol. 33, n° 5, 1987, pp. 589-609.

[De Smet et al. 95] De Smet O., Abou-Kandil H., Ordonnancement temps-réel pour des systèmes flexibles de production sujets à pannes, *revue d'automatique et de productique appliquées*, vol. 8, n° 2-3, 1995, pp. 291-296.

[Dindeleux 92] Dindeleux E., Proposition d'un modèle et d'un système interactif d'aide à la décision pour la conduite d'atelier, *thèse de doctorat*, université de Valenciennes, Valenciennes, France, 1992

[Dindeleux et al. 95] Dindeleux R. et Trentesaux D., Intégration de l'aide multicritère à la conduite dans le système de pilotage en vue de l'amélioration des processus de production, *42ème journées du groupe de travail européen "aide multicritère à la décision"*, Namur, Octobre, 1995.

[Doumeingts et al. 83] Doumeingts G., Breuil D., Pun L., La gestion de production assistée par ordinateur, *Hermès*, Paris, France, 1983.

[Doumeingts 90] Doumeingts G., Méthodes pour concevoir et spécifier les systèmes de production, CIM'90: productique et intégrations, *éd. productic A.*, Marseille, France, juin, 1990, pp. 89-103.

[Duffie et al. 88] Duffie N.A. et Piper R.S., Fault tolerant heterarchical control of Heterogeneous manufacturing system entities, *Journal Of Manufacturing systems*, vol. 7, n°4, 1988, pp. 315-327.

[Duffie et al. 94] Duffie N. A. et Prabhu V. V., Real-time distributed scheduling of heterarchical manufacturing systems, *Journal of Manufacturing Systems*, vol. 13, n° 2, pp. 94-107.

[Eom et al. 90] Eom H. B., Lee S. M. et Suh E. H., Group Decision Support Systems: an essential tool for resolving organizational conflicts, *International journal of information management*, vol. 10, 1990, pp. 215-227.

[Eom et al. 93] Eom H. B., Lee S. M. et Kim J. K., The intellectual structure of Decision Support Systems(1971-1989), *Decision Support Systems*, vol. 10, 1993, pp. 19-35.

[Erceau et al. 91] Erceau J. et Ferber J., L'intelligence Artificielle Distribuée, *La recherche*, n° 233, Juin 1991, pp. 750-761.

[Erschler et al. 93] Erschler J., Fontan G. et Mercé C., Approche par contraintes en planification et ordonnancement de la production, *APII*, vol. 27, n° 6, 1993, pp. 669-695.

[Ferber et al. 88] Ferber J., Ghallab M., Problématique des Univers Multi Agents Intelligents, *Proc. Journées nationales du PRC/IA Cepadues-Edition*, Toulouse, France, mars 1988.

[Floyd Lewis 87] Floyd Lewis L., A Decision Support System for face-to-face groups, *Journal of Information science*, vol. 13, 1987, pp. 211-219.

[Floyd Lewis et al. 88] Floyd Lewis L, Keleman K. S., Issues in Group Decision Support System, *Journal of Information science*, vol. 14, 1988, pp. 347-354.

[Fodor et al. 94] Fodor J. et Roubens M., Fuzzy preference modelling and multicriteria decision support, *Theory and decision library*, n° 14, Kluwer academic publishers, 1994.

[Gandibleux 95] Gandibleux X., 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*, université de Valenciennes, France, 1995.

[Giard 88] Giard V., Gestion de la production, 2nde éd., *Collection gestion*, Paris, France, 1988.

[Giard 91] Giard V., Une comptabilité de gestion en crise, ECOSIP, *Economica*, Paris, France, 1991, pp. 135-149.

[Goldratt et al. 84] Goldratt E. et Cox J., The goal: excellence in Manufacturing, *North River Press*, Croton-on-Hudson, New-York, USA, 1984.

[Gonthier 90] Gonthier A., De l'optimisation des coûts de changement defabrication vers l'atelier décentralisé d'ordonnancement, *Thèse de doctorat*, Institut National Polytechnique de Grenoble, Grenoble, France, 1990.

[Grabot et al. 94] Grabot B. et Geneste L., Un système expert gérant les informations vagues et incertaines pour l'aide au pilotage d'atelier, *revue des systèmes de décision*, vol. 3, n° 2, 1994, pp. 127-155.

[Gravel et al. 92] Gravel M., Martel J. M., Nadeau R., Price W. et Tremblay R., A multicriterion view of optimal resource allocation in job-shop production, *European journal of operational research*, vol. 61, 1992, pp. 230-244.

[Hawgood et al. 87] Hawgood J. et Humphreys P., Effective Decision Support Systems, *Technical Press*, 1987.

[Hayes-Roth 85] Hayes-Roth B., A blackboard architecture for Control, *Artificial intelligence*, vol. 26, 1985, pp. 251-321.

[Huhns 87] Huhns M. N., Distributed Artificial Intelligence, *research notes in artificial Intelligence*, vol. 1, ed. Pitman, Londres, Angleterre, 1987.

[Jacot 91] Jacot J.H., A propos de l'évaluation économique des systèmes intégrés de production, ECOSIP, *Economica*, Paris, France, 1991, pp. 61-70.

[Johannsen et al. 94] Johannsen G., Levis A. H. et Stassen H. G., Theoretical problems in Man-Machine Systems and their experimental validation, *Automatica*, Vol. 30, n° 2, février 1994, pp. 217-231.

[Kallel 85] Kallel G., Proposition d'une conduite décentralisée coordonnée pour un atelier de fabrication, *thèse de doctorat*, Institut national polytechnique de Grenoble, Grenoble, France, 1985.

[Keen 87] Keen P. G. W., Decision Support Systems: The Next Decade, *Decision Support Systems*, n° 3, 1987, pp. 253-265.

[Korhonen et al. 92] Korhonen P., Moskowitz H. et Wallenius J., Multiple criteria decision support- A review, *European Journal of Operational Research*, n° 63, 1992, pp. 361-375.

[Kouiss et al. 95] Kouiss K., Pierreval H. et Mebarki N., Towards the use of a multi-agent approach to the dynamic scheduling of flexible manufacturing systems, *International conference on Industrial Engineering and Production Management, Marrakech*, Maroc, avril 1995, pp. 118-125.

[Legall 89] Legall A., Un système interactif d'aide à la décision pour l'ordonnancement et le pilotage temps réel d'atelier, *Thèse de doctorat*, université de Toulouse, Toulouse, France, 1989.

[Lévine et al. 86] Lévine P. et Pomerol J. C., PRIAM, an interactive program for choosing among multiple attribute alternatives, *European Journal of Operational Research*, n° 25, pp. 272-280, 1986.

[Lévine et al. 89] Lévine P. et Pomerol J. C., Systèmes interactifs d'aide à la décision et systèmes experts, *Hermès*, Paris, 1989.

[Levis et al. 94] Levis A. H., Moray N. et Hu B., Task decomposition and allocation problems and discrete event systems, *Automatica*, vol. 30, n° 2, 1994, pp. 203-216.

[Lundrigan 86] Lundrigan R., Qu'appelle-t-on OPT?, *Revue française de Gestion Industrielle*, vol. 5, n°3, 1986, pp. 27-38.

[MacLean et al. 87] MacLean C.R. et Brown P.F., The architecture of the NBS factory automation research testbed, *proceedings IFAC 10th trienal world congress*, Munich, Allemagne, 1987.

[Marris 94] Marris P., Le management par les contraintes en gestion industrielle, *Les Editions d'Organisation*, Paris, France, 1994.

[Meleton 86] Meleton M. P. Jr., OPT-fantaisie ou révolution?, *Revue française de Gestion Industrielle*, vol. 5, n°3, 1986, pp.17-26.

[Merbaki 95] Merbaki N., Une approche d'ordonnancement temps-réel basée sur la sélection dynamique de règles de priorité, *thèse de doctorat*, université Claude Bernard Lyon I, Lyon, France, 1995.

[Mercé 87] Mercé C., Cohérence des décisions en planification hiérarchisée, *Thèse de doctorat d'état-sciences*, université Paul Sabatier, Toulouse, Toulouse, France, 1987.

[Messaoud 86] Messaoud D., Problèmes de décision dans une cellule de production flexible utilisant la coopération entre robots, *thèse de doctorat*, université de Lille Flandres Artois, Lille, France, 1986.

[Mévellec 93] Mévellec P., Outils de Gestion, La pertinence retrouvée, éd. *comptables malesherbes*, Paris, France, 1993

[Moray et al. 91] Moray N., Dessouky M. I., Kijowski B. A. et Adapathya R., Strategic behavior, workload and performance in task scheduling, *Human factors*, vol. 33, n°6, 1991, pp. 607-629.

[Moray et al. 94] Moray N., Lee J. D., Muir B. M., Trust and human intervention in automated systems, éd. *Hoc, Cacciabue&Hollnagel*, Expertise and technology cognition and human Computer interaction, 1994.

[Noyes 87] Noyes D., Approches méthodologiques pour l'aide à la conception et à la conduite des systèmes de production, *Thèse de doctorat d'état-informatique/productique*, Institut national polytechnique de Toulouse, Toulouse, France, 1987.

[Paelinck 78] Paelinck J., Qualifex, a flexible multiple criteria method, *Economic Letters*, n° 3, 1978.

[Paredis et al. 92] Paredis J. et Van Rij T., Simulation and constraint programming as support methodologies for job-shop scheduling, *Revue des systèmes de décision*, vol. 1, n° 1, 1992, pp. 59-77.

[Parunak 91] Parunak V.D.H., Characterizing the manufacturing scheduling problem, *Journal of manufacturing systems*, vol. 10, n° 3, 1991, pp. 241-252.

[Pellet 85] Pellet X., Sur la hiérarchisation des décisions. Application à la conduite d'ateliers, thèse de doctorat, *Institut national polytechnique de Grenoble*, Grenoble, France, 1985.

[Pollock et al. 93] Pollock C. et Kanachowski A., Application of theories of Decision Making to Group Decision Support Systems, *International journal of Human-Computer Interaction*, vol. 5, n° 1, 1993, pp. 71-94.

[Pomerol et al. 93] Pomerol J. C. et Barba-Romero S., Choix multi-critère dans l'entreprise, collection Informatique, *Hermès*, Paris, 1993.

[Pujolle et al. 92] Pujolle G. et Horlait E., Architecture des réseaux informatique, tome 1, les outils de communication, 2nde éd., *Eyrolles*, Paris, 1992.

[Rasmussen 80] Rasmussen J., The Human as a System Component, éd. *Smith&Green*, Human Interaction with Computer, *London Academic Press*, Londres, 1980.

[Rasmussen 83] Rasmussen J., Skill, Rules and knowledge: signals, signs and symbols and other distinctions in human performance models, *IEEE SMC*, vol. 13, n° 3, 1983, pp. 257-266.

[Rasmussen et al. 91] Rasmussen J., Leplat J. et Brehmer B., Distributed Decision Making; cognitive models for cooperative work, *New Technologies and Work*, a Wiley Series, Engleterre, 1991.

[Roy 78] Roy B., Electre III: un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples, *Cahiers du CERO* 20, pp. 3-24, 1978.

[Roy 85] Roy B., Méthodologie multi-critère d'aide à la décision, *Economica*, Paris, 1985.

[Roy et al. 93] Roy B., Bouyssou D., Aide multi-critère à la décision: Méthodes et cas, *Economica*, Collection Gestion, Paris, 1993.

[Saaty 80] Saaty T. L., The Analytic Hierarchy Process, *MacGraw Hill*, New-York, 1980.

[Sahraoui 87] Sahraoui A., Contribution à la commande et à la surveillance d'un atelier flexible, *Thèse de troisième cycle*, LAAS, université de Toulouse, 1987.

[Sénéchal et al. 94] Sénéchal O., Lenclud T. et Tahon C., Production costs identification using value engineering and simulation in a concurrent engineering process, *international conference IFIP Feature modelling and recognition in advanced CAD/CAM systems*, valenciennes, France, Mai, 1994, pp.583-602.

[Schärlig 85] Schärlig A., Décider sur plusieurs critères - Panorama de l'aide à la décision multicritère, *Presse polytechniques et universitaires romandes*, Col. diriger l'entreprise 1, Lausanne, Suisse, 1985.

[Simon 77] Simon H. A., The new science of management decision, *Prentice-Hall*, New-Jersey, USA, 1977.

[Simonetti 87] Simonetti- Barbalho D., Conception et mise en oeuvre de la fonction coordination pour une commande distribuée d'atelier, *Thèse de troisième cycle*, LAAS, université de Toulouse, France, 1987.

[Smith 80] Smith R.G., The contract net protocol: High level communication and control in a distributed problem solver, *IEEE Transactions on Computer*, vol. C29, n°12, 1980, pp. 1104-1113.

[Sprague 87] Sprague R. H., DSS in context, *Decision Support System*, n°3, 1987, pp. 197-202.

[Sun et al. 94] Sun D., Lin L., A dynamic job-shop scheduling framework: a backward approach, *international journal of production research*, vol. 32, n° 4, 1994, pp. 967-985.

[Tabucanon 88] Tabucanon M. T., Multiple criteria decision making in industry, *Studies in production and engineering economics*, n° 8, *Elsevier*, 1988.

[Tabucanon et al. 94] Tabucanon M. T., Batanov D. N. et Verma D. K., Decision support system for multicriteria machine selection for flexible manufacturing systems, *Computers in industry*, vol. 25, n° 2, 1994, pp. 131-143.

[Tacquard 95] Tacquard C., Baptiste P. et Manier H., A behavioral model to schedule and control a flexible manufacturing ring: application to the push and the pull flows, *International conference on Industrial Engineering and Production Management, Marrakech, Maroc*, avril 1995, pp. 127-136.

[Tahon et al. 92] Tahon C., Soënen R., Dindeleux E. et Gorthcinsky G., Decision support system in CIM: an integrated production activity control system, Human aspects in computer integrated manufacturing, *IFIP TC5/WG 5.3, eighth international Prolamat conference, Man in Cim, Tokyo, Japon*, 1992, pp. 475-493.

[Tchako 94] Tchako J.F.N., Contribution à la conception d'un système de pilotage distribué pour les systèmes automatisés de production, *Thèse de doctorat, université de Valenciennes, Valenciennes, France*, 1994.

[Tchako et al. 94] Tchako J.F.N., Beldjilali B., Trentesaux D. et Tahon C., Modelling with coloured Petri nets and simulation of a dynamic and distributed management system for a manufacturing cell, *International Journal of Computer Integrated Manufacturing*, vol. 7, n° 6, 1994, pp. 323-339.

[Thomas 80] Thomas V., Aide à la décision pour l'ordonnancement en temps réel d'atelier, *Thèse de doctorat, université de Toulouse, Toulouse, France*, 1980.

[Trentesaux et al. 94a] Trentesaux D. et Tahon C., Modèle de communication inter-agents pour une structure de pilotage temps réel distribuée, *Revue d'Automatique et de Productique Appliquées*, vol. 7, n° 6, 1994, pp. 703-727.

[Trentesaux et al. 94b] Trentesaux D., Dindeleux R. et Tahon C., A MultiCriteria Decision Support System for Dynamic task Allocation in a Distributed Production Activity Control Structure, *European Workshop on integrated Manufacturing Systems Engineering, IMSE'94, INRIA, Grenoble, France*, 1994, pp. 383-393.

[Trentesaux et al. 95a] Trentesaux D. et Tahon C., DPACS: a self-adaptative production activity control structure, *INRIA/IEEE conference on emerging technologies and factory automation, Tome II, Paris, France*, 1995, pp. 543-551.

[Trentesaux et al. 95b] Trentesaux D., Tchako J.F.N. et Tahon C., Distributed and multicriteria management tools for integrated manufacturing, Life cycle modelling for innovative products and processes, *Chapman&Hall, Berlin, Allemagne*, 1995, pp.576-588.

[Trentesaux et al. 95c] Trentesaux D. et Tahon C., Dynamic and Distributed Production Activity Control: a multicriteria approach for task allocation problematic, *International Conference on Industrial Engineering and Production Management, Marrakech, Maroc*, 1995, pp. 137-154.

[Trentesaux et al. 96] Trentesaux D. et Tahon C., HPACS: une approche égalitaire supervisée et interactive pour le pilotage adaptatif d'un système de production, *Journées Ordonnancement Coopératif, Bordeaux*, 1996.

[Trentesaux et al. ??a] Trentesaux D., Tahon C. et Ladet P., Hybrid production activity control: a JIT distributed-based structure for FMS, *soumise à Artificial Intelligence Engineering, ???*

[Trentesaux et al. ??b] Trentesaux D., Dindeleux R. et Tahon C., A MultiCriteria Decision Support System for Dynamic task Allocation in a Distributed Production Activity Control Structure, *soumise à IJCIM, ??*

[Vincke 89] Vincke P., L'aide multicritère à la décision, ed. Université de Bruxelles, *ed. Ellipses*, Belgique, 1989.

[Zhao 89] Zhao X., Contribution à la construction d'un système expert d'aide à la conduite d'ateliers manufacturiers, *Thèse de doctorat*, université de Valenciennes, Valenciennes, France, 1989.

[Zeleny 82] Zeleny M., Multiple Criteria Decision Making, *Mc Graw-Hill*, New-York, 1982.

Annexe A

Gestion de production et fonctions prévisionnelles

Système de production

L'acte de produire consiste en une opération de transformation dans laquelle un ensemble de matières premières ou de produits semi-finis sont convertis en produits finis. De cette manière, Mercé définit un système de production comme l'ensemble des éléments qui interviennent dans cette transformation, ainsi que l'ensemble des relations entre ces éléments [Mercé 87].

Il existe plusieurs classifications pour l'ensemble des systèmes de production. Nous retenons la plus connue qui établit une typologie selon la complexité des contraintes imposées au flux de produit au sein du système [Carlier et al. 82]:

- flow-shop: l'ordre de passage des pièces sur les machines est imposé et est le même pour chaque pièce,
- job-shop: cet ordre reste imposé mais peut différer selon les pièces,
- open-shop: aucun ordre n'est imposé.

Gestion de production

Le rôle de la gestion de production est, selon Doumeingts, d' "organiser et piloter le fonctionnement des processus physiques mis en oeuvre dans l'entreprise, afin d'assurer une meilleure utilisation des moyens humains, physiques et technologiques disponibles, et de satisfaire au mieux l'objectif global de production défini en terme de quantités à fabriquer avec une qualité demandée, et des délais à respecter" [Doumeingts et al. 83].

Cette définition présente la gestion de production comme une fonction qui doit tenir compte de nombreuses contraintes et objectifs très diversifiés selon leurs natures dans le but de piloter l'ensemble des moyens de production:

- contraintes décrites en terme d'utilisation efficace des moyens disponibles: ressources humaines ou matérielles,
- objectifs définis en terme global de délai de qualité et de coût.

L'automatisation est le moyen le plus courant qui permette d'obtenir une gestion performante de la production.

Fonctions prévisionnelles de gestion de production

Les fonctions classiques associées au niveau prévisionnel (figure 1) sont les suivantes [Doumeingts 90]:

- planification: cette fonction définit le plan directeur qui constitue un compromis entre les objectifs commerciaux, financier et de production,
- programmation: cette fonction élabore à partir du plan directeur un programme prévisionnel de production et calcule l'ensemble des besoins en fonction des en-cours et des stocks,
- ordonnancement: cette fonction définit l'ordonnancement prévisionnel total dans le cas où celui-ci est nécessaire, dans le cas où aucun ordonnancement n'est requis, cette fonction est inexistante). Elle décrit ainsi l'utilisation des ressources matérielles et humaines. Les besoins en approvisionnement sont estimés de manière plus fine.

L'agencement de ces fonctions permet de définir la gestion prévisionnelle.

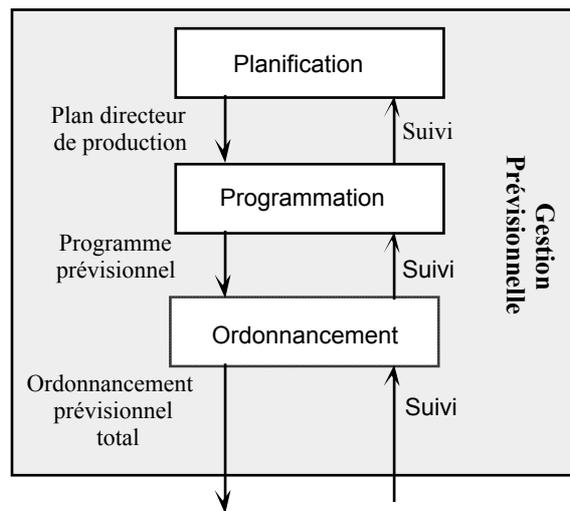


Figure 1: Fonction de gestion prévisionnelle

Annexe B

Compléments sur la notion d'enrichissement

Nous donnons dans cette annexe la démonstration de la normalisation du processus d'enrichissement. Un exemple est fourni quand au processus d'enrichissement. Enfin, nous justifions la structure du gain qui est associée à la procédure d'enrichissement.

démonstration de la normalisation

La démonstration suivante a pour objet de prouver que l'enrichissement décrit au chapitre IV est une opération qui ne détruit pas la normalisation. Nous reprenons les notations développées dans ce chapitre.

En développant $\sum_{i=0}^n p'_i = p'_0 + p'_1 + \dots + p'_k + \dots + p'_{n-1} + p'_n$

on introduit $p'_k = \Delta_m^k = \frac{m \cdot p_k + 1}{m + 1}$:

$$\sum_{i=0}^n p'_i = \frac{m}{m+1} \cdot p_0 + \frac{m}{m+1} \cdot p_1 + \dots + \frac{m \cdot p_k + 1}{m+1} + \dots + \frac{m}{m+1} \cdot p_{n-1} + \frac{m}{m+1} \cdot p_n$$

$$\sum_{i=0}^n p'_i = \frac{m}{m+1} \cdot p_0 + \frac{m}{m+1} \cdot p_1 + \dots + \frac{m}{m+1} \cdot p_k + \frac{1}{m+1} + \dots + \frac{m}{m+1} \cdot p_{n-1} + \frac{m}{m+1} \cdot p_n$$

Ce qui donne après factorisation:

$$\sum_{i=0}^n p'_i = \frac{m}{m+1} (p_0 + p_1 + \dots + p_k + \dots + p_{n-1} + p_n) + \frac{1}{m+1}$$

or

$$\sum_{k=0}^n p_k = 1$$

D'où

$$\sum_{i=1}^n p'_i = \frac{m}{m+1} + \frac{1}{m+1} = 1, \text{ cqfd.}$$

Exemple d'application d'enrichissement

Nous prendrons dans cet exemple $\Pi \equiv \mathbb{N}$. Soit l'ensemble des valeurs potentielles $X_\alpha = \{x_1 = 3, x_2 = 5, x_3 = 9\}$. L'étude portera sur une seule donnée

générique définie par son état initial d'indice 2: $\Omega = \{\omega_\alpha\}$. De cette manière, $\omega_\alpha = (X, P, F, m)$ avec $P(x_2) = 1$, $P(x_0) = P(x_1) = P(x_3) = 0$, $F(p) = x_2$ et $m=1$.

Evaluons l'enrichissement récurrent de ω_α sur la valeur x_3 défini par $\omega'_\alpha = E_1^3(\omega_\alpha)$.

Les équations d'enrichissement P' appliquées au cas présent donnent:

$$P': \{x_1, x_2, x_3\} \rightarrow [0, 1]$$

$$x_1 \mapsto P'(x_1) = p'_1 = \frac{m}{m+1} \cdot p_1 = 0$$

$$x_2 \mapsto P'(x_2) = p'_2 = \frac{m}{m+1} \cdot p_2 = \frac{1}{2}$$

$$x_3 \mapsto P'(x_3) = p'_3 = \Delta_1^3 = \frac{m \cdot p_3 + 1}{m+1} = \frac{1}{2}$$

Si on note ω_α sous une forme explicitée $\omega_\alpha = (\{3, 5, 9\}, \{0, 0, 1, 0\}, F, 1)$, on obtient $\omega'_\alpha = \left(\{3, 5, 9\}, \left\{0, 0, \frac{1}{2}, \frac{1}{2}\right\}, F, 2 \right)$. On montre de même que:

$$\omega''_\alpha = E_2^3(\omega'_\alpha) = \left(\{3, 5, 9\}, \left\{0, 0, \frac{1}{3}, \frac{2}{3}\right\}, F, 3 \right).$$

On vérifie bien que $\sum_{k=0}^n p_k = \sum_{k=0}^n p'_k = \sum_{k=0}^n p''_k = 1$.

Supposons qu'à cette étape un enrichissement porte sur une valeur $x=7$. Dans ce cas l'ensemble X_α est modifié pour devenir $X'_\alpha = \{x_1 = 3, x_2 = 5, x_3 = 7, x_4 = 9\}$.

L'enrichissement porte nécessairement sur la valeur introduite et $\omega'''_\alpha = E_3^3(\omega''_\alpha) = \left(\{3, 5, 7, 9\}, \left\{0, 0, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right\}, F, 4 \right)$.

Ceci est résumé sur les figures 1 et 2:

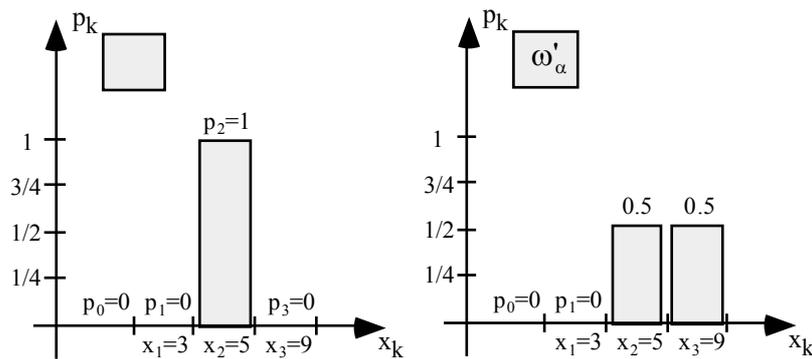


Figure 1: deux premiers enrichissements

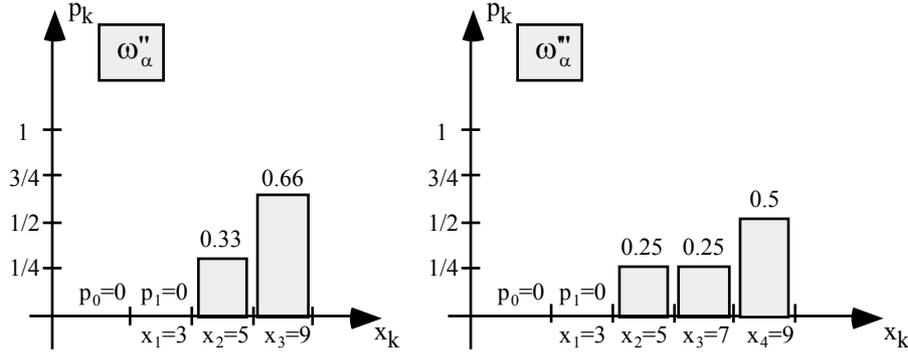


Figure 2: troisième et quatrième enrichissements

On en déduit par ailleurs facilement pour ω'' les valeurs de F suivantes:

$$F\left(\frac{1}{10}\right) = 3, F\left(\frac{1}{4}\right) = 3, F\left(\frac{1}{3}\right) = 5, F\left(\frac{1}{2}\right) = 7 \text{ et } F(1) = 9.$$

Justification de la structure du gain

La relation $p'_k = \Delta_m^k = \frac{m \cdot p_k + 1}{m + 1}$ correspond à l'enregistrement du $(m + 1)^{\text{ième}}$ enrichissement pour un gain de une unité pour une valeur possible x_k par rapport à m unités présentes. Ceci est clairement illustré par l'équation suivante:

$$\sum_{i=0}^n p'_i \equiv 1 = \frac{m}{m+1} \cdot p_0 + \frac{m}{m+1} \cdot p_1 + \dots + \frac{m \cdot p_k + 1}{m+1} + \dots + \frac{m}{m+1} \cdot p_{n-1} + \frac{m}{m+1} \cdot p_n$$

qui peut s'écrire sous la forme:

$$m \cdot p_0 + m \cdot p_1 + \dots + (m \cdot p_k + 1) + \dots + m \cdot p_{n-1} + m \cdot p_n \equiv m + 1$$

où le facteur $(m \cdot p_k + 1)$ montre bien l'accroissement de 1 unité de la valeur $m \cdot p_k$, la normalisation n'intervenant qu'après cette étape. Cette structure induit donc un gain linéaire avec le nombre d'enrichissement. Cette propriété s'exprime différemment de la manière suivante:

L'influence d'un enrichissement est constante dans le temps. Le gain est calculé pour uniformiser chaque enrichissement de telle sorte qu'après l'enrichissement k , l'élément $\omega \in \Omega$ soit composé de $(k+1)$ enrichissements avec un poids identique pour chacun de $\frac{1}{k+1}$.

Autrement dit, le gain non constant est calculé pour que tous les enrichissements aient le même poids. Il est évident que d'autres structures peuvent être définies. C'est le cas par exemple d'une structure de gain avec facteur d'oubli $\lambda(m)$. Un exemple d'une telle structure de gain est $\Delta_m^k = \lambda(m) \cdot \frac{m \cdot p_k + 1}{m + 1}$. La fonction $\lambda(m)$ est une fonction qui peut-être de la forme exponentielle par exemple.

Il est à noter qu'il n'y a pas mémorisation du séquençement des différents enrichissements. Ce qui implique que la structure de gain ne pourra agir que sur l'enrichissement en cours: le réduire si les premiers enrichissements sont considérés comme plus important, ou inversement. Une seconde extension possible consiste à attribuer à un enrichissement un paramètre de crédibilité qui influe sur l'importance (gain) accordé à cet enrichissement. Cette optique permet donc d'intégrer les concepts de fiabilité de l'information.

Annexe C

Exemple d'une liste de critère appliquée à un processus de fabrication de type Job-shop. Application au pilotage hybride

Nous avons décomposé chacun des trois axes en critères qui vérifient au mieux les contraintes fixées par Roy sur la qualité de la liste de critères. Nous présentons une liste de critères selon les axes coûts, délais et qualité.

Axe de signification: le coût (unité de mesure monétaire)

Critère n°C1: le coût de production

Nous découplons les critères qui sont naturellement explicités en terme de coût des autres critères. Il est clair que pour un industriel, une perte de temps se décrit en perte d'argent. C'est une approche par agrégation transitive (plus une machine est en retard sur son programme, plus les coûts sont élevés) qui réduit fortement l'intérêt d'intégrer une approche multicritère. C'est pourquoi ce premier critère doit être exempt de toute interprétation en terme d'unité de temps ou en terme de qualité.

Axe de signification: le délai (unité de mesure temporelle)

Critère n° D1: temps opératoire d'une ressource pour une opération

Le temps opératoire représente le temps nécessaire à une ressource pour effectuer une tâche particulière. Les temps de préparation, de changement outil, de préchauffage, etc. sont exclus.

Critère n° D2: temps de préparation d'une ressource pour une opération

Ce temps représente l'ensemble des opérations de préparation sur une machine pour une tâche donnée. Il est possible d'agrèger les temps de changement outil, de

préchauffage, de manutention, etc. En effet, cette agrégation ne nuit en rien à la spécification du critère.

Critère n° D3: temps de transfert potentiel

Ce critère indique les temps de transfert entre ressources lorsqu'il est nécessaire. Ce critère mesure le temps écoulé entre le stock de sortie de la SIP émettrice et l'arrivée du produit dans le stock d'entrée de la SIP réceptrice.

Critère n° D4: prochaine date de disponibilité

La prochaine date de disponibilité libre correspond à la date à partir de laquelle la SIP sera capable de réaliser, en théorie, l'opération donnée. Ce critère ne prend pas en compte les éventuelles modifications de cette date liées par exemple aux allocations ou ré-allocations potentielles.

Critère n° D5: marge potentielle libre du produit

Ce critère représente la marge libre pour un produit entre un instant t et la date de besoin du produit. Ce critère peut prendre les deux formes suivantes:

- “marge minimum” (en fonction de la date au plus tôt) qui décrit le temps libre en fonction de la date courante, de la date de besoin au plus tôt et des tâches qui restent à réaliser,
- “marge potentielle” qui décrit la capacité d'absorption d'un aléa de production sans provoquer de retard. Ce critère peut expliciter un phénomène largement usité dans le milieu industriel qui est celui de la priorité. Cependant, la notion de règle de priorité a pour inconvénient de provoquer la perte de l'information du retard car elle exprime une notion ordinale généralement par l'intermédiaire d'un nombre entier. Nous lui préférons ainsi celle de marge potentielle.

Les données utilisées pour les critères D1 à D5 peuvent être quantitatives. Cependant, il nous semble plus réaliste de les modéliser sous forme de données incertaines, puisqu'elles concernent un temps anticipé probable.

Axe de signification: la qualité (avec ou sans unité)

Critère n°Q1: taux d'usure

Ce critère regroupe l'ensemble des taux d'usures dont le décideur doit tenir compte lors d'un processus décisionnel. En effet, la qualité de production peut dépendre de l'état d'usure d'un outil.

Critère n°Q2: caractéristique outil

Ce critère permet de prendre en compte le fait qu'un outil soit considéré comme investissement à rentabiliser.

Critère n°Q3: niveau de spécialisation

Il est utile également de pouvoir prendre en compte la spécialisation de l'opérateur (critère N°Q3a) et de la machine (critère N°Q3b). Ce critère permet ainsi d'utiliser les compétences respectives. Une attention particulière doit être donnée à l'intégration du critère Q3a qui nécessite implicitement la diffusion parmi l'ensemble des SIP d'une évaluation de l'opérateur. Cet aspect psychologique doit être scrupuleusement étudié.

Les critères Q1, Q2 et Q3 sont représentés sous forme de pourcentage. Le critère Q3 est un critère difficilement chiffrable dont la nature est sans dimension.

Validité de la liste des critères retenus

La liste de critères présentée nous semble suffisante pour constituer une base de critères qui prend en compte la plupart des cas industriels d'atelier job-shop. Il peut cependant être nécessaire de définir d'autres critères en fonction de la spécificité du système de production ou des objectifs de production. Par exemple, dans le cas d'un équilibrage de charge, un critère prenant en compte le nombre d'opérations réalisées ou le temps total productif pour chaque ressource peut-être utilisé. Un autre exemple concerne le cas où un des objectifs globaux est de réduire le niveau de stock moyen. Un critère particulier doit être élaboré pour représenter chaque aspect.

C'est au niveau de la spécification en fonction des contraintes de gestion qu'il faut établir un ensemble de listes et ainsi vérifier les conditions de qualité de la liste explicitée par Roy. Cette liste est donc constituée d'un sous-ensemble de la liste présentée.

Nous allons proposer un ensemble de sous-listes en fonction du processus décisionnel concerné.

Gestion du stock d'entrée d'une SIP

Le critère C1 est inutile car la tâche est déjà allouée. Il en est de même pour les critères D1, D3, D4, Q1, Q2 et Q3. Ainsi seuls les critères D2 (temps de préparation) qui stipule qu'il peut-être intéressant d'éviter de modifier la configuration machine en réalisant une tâche de la même famille et D5 qui tient compte du temps disponible pour la réalisation de la tâche subsistent.

Séquencement des allocations

Que ce soit au niveau superviseur ou au niveau distribué, lorsque deux opérations ou plus sont en stock de sortie, le séquencement des allocations influe sur les performances globales. Cette influence est encore plus sensible lors de la phase initiale d'allocation par le superviseur.

Le critère D5 est prépondérant pour déterminer l'ordre avec lequel les allocations sont à réaliser. Le critère D3 est utilisé, ainsi que les critères Q1, Q2 pour prendre en compte des souhaits de rentabilité.

Allocation de tâche

Tous les critères explicités auparavant excepté le critère n°D5 (qui n'est pas lié aux capacités des ressources) sont utilisés. Ces sept critères montrent la complexité du processus d'allocation. La qualité des allocations conditionne fortement les performances d'un système de production distribué non supervisé. Ainsi, une allocation peu judicieuse peut causer des conséquences critiques sur l'ensemble du processus de production. Nous avons pu constater que l'approche par supervision permet de diminuer cette influence, et ainsi amoindrir l'importance de cette phase.

Gestion des variables d'action locales

L'ensemble des décisions qui ont trait à la gestion locale des ressources sont concernées. Dans le cas d'une ressource de type machine, ces variables s'apparentent aux vitesses (taux) de production, à la taille des lots, aux capacités de stockage, etc. Pour une ressource de type convoyeur, elles s'apparentent à la vitesse de convoyage.

Re-ordonnement local (modification de l'ordonnement superviseur)

Il est possible pour une ressource ordonnancée de remarquer un retard par exemple dans l'arrivée d'une tâche qui devrait être en cours de réalisation selon l'ordonnement prévisionnel instauré par le superviseur. Dans le cas où une remise en question est décidée, il est nécessaire de trier les tâches en attente. Ce problème est donc typiquement celui de la gestion du stock d'entrée.

Processus non structurés

Choix de remise en question de l'ordonnement prévisionnel

Ce processus décisionnel concerne le niveau de supervision. Il est clair qu'une telle décision requiert une expérience et une connaissance du système de production approfondie pour évaluer l'impact d'un événement au niveau global. Une première approche consiste à établir une typologie des cas qui nécessitent automatiquement une remise en question de l'ordonnement. Certains événements s'intègrent aisément dans cette typologie. C'est le cas à chaque fois qu'une commande est lancée en production, car elle modifie le processus entier de fabrication. Mais la question se pose réellement dans le cas d'une ou plusieurs pannes machines. Il faut alors être capable d'évaluer l'impact de cette panne sur le

séquencement des opérations pour justifier ou non du besoin de ré-ordonnancement. Cette évaluation est difficile à réaliser et fait explicitement appel à la connaissance et à l'expérience de l'opérateur humain. Ce qui caractérise un processus décisionnel mal normalisé.

Choix de re-allocation d'une opération en cours

Le choix de réallocation peut-être réalisé à deux moments distincts:

- la tâche à ré-allouer n'a pas encore été commencée, ce qui correspond à la présence de l'opération dans le stock d'entrée. Ce cas de figure est possible si le moyen de production est en fonctionnement dégradé durant un certain temps. Dans ce cas, on peut décider de ré-allouer si possible les tâches en attente. Un module de gestion particulier doit prendre en compte une ré-allocation de ce type (l'opération passe alors du stock d'entrée au stock de sortie). La difficulté intervient lors de l'évaluation de la nécessité de la ré-allocation. Ceci dépend principalement du temps de non-disponibilité de la ressource de production.
- la tâche à allouer est en cours de production. Ce processus décisionnel nécessite un niveau de supervision pour intégrer la modification sur la tâche à réaliser qui stipule que l'opération a été partiellement réalisée. Cela sous-entend qu'une telle réalisation partielle puisse être possible à la fois de la part de la machine qui arrête sa production et à la fois de la part de celle qui reprendra ce processus de production. Dans tous les cas, la nature de la décision est identique au cas précédent et est associée à un processus décisionnel non normalisé.

Choix de remise en question de l'ordonnancement partiel

Toutes les opérations qui correspondent à un produit présent dans le stock d'entrée de la ressource goulet sont planifiées. Le choix de remise en question doit être envisagée par exemple, dans le cas d'un retard. Une évaluation de celui-ci permet de réaliser une tâche (sous la réserve qu'il y en ait une disponible) dont la durée est inférieure à la durée de retard estimé (une autre alternative consiste à ne pas tenir compte du retard et commencer l'opération suivante planifiée et présente sans attendre).

De manière générale, il est évident qu'une telle remise en question nuit fortement au respect des objectifs globaux car elle modifie le séquencement des opérations planifiées, et ce, bien qu'une modification de l'ordonnancement puisse éviter de rendre non-productive une ressource goulet. Une méthode systématique consiste à demander un ré-ordonnancement à chaque modification, mais cela imposerait au superviseur de procéder à ce ré-ordonnancement plus fréquemment.

Processus structurés

Ces processus sont:

- calcul des charges potentielles,
- choix de la ressource goulet,
- élaboration de l'ordonnancement.

La catégorie des processus normalisés peut-être gérée sans intervention de l'opérateur humain par une méthode algorithmique. De plus, une approche par optimisation est possible car l'ordonnancement n'est plus NP-complet. En effet, il concerne une ressource au lieu de la totalité.

Cependant, de nombreux exemples récents (voir par exemple [Bourgeois 94], [Billaut 93], etc.) ont montré l'intérêt de l'interactivité dans le calcul et l'évaluation d'un ordonnancement. L'intégration du multicritère est nécessaire [Belton et al. 92]. Un autre intérêt de cette méthode est par exemple de permettre à l'utilisateur de trancher entre plusieurs goulets potentiels ou de décider lui-même sans l'intervention du SIAMD quelle est la ressource goulet. Dans ce cas, la prise en compte d'événements spécifiques est possible (anticipation de commandes, etc.).

Annexe D

Modèle fonctionnel et structure logicielle du gestionnaire de contraintes

Modèle fonctionnel

Spécifications

Le gestionnaire de contraintes a été développé pour améliorer la gestion des allocations en prenant en compte un ensemble de contraintes portant sur:

- le type de l'opération à effectuer,
- la précision désirée,
- la qualification requise, etc.

Les fonctions du gestionnaire concernent à la fois:

- la fonction création d'une contrainte: une SIP doit associer à la primitive Demande du protocole d'allocation un ensemble de contraintes portant sur l'opération à réaliser.
- la fonction satisfaction de contraintes: une SIP qui reçoit cette Demande doit pouvoir vérifier si les contraintes imposées sont vérifiées ou non.

Modèle

Nous définissons une contrainte comme un ensemble de conditions qui sont associées chacune à une variable particulière. Une contrainte peut porter ainsi sur plusieurs variables. Chaque condition s'exprime sous la forme:

Condition := <Expression> [et / ou <Expression>]

Une expression est une relation binaire établie sur une donnée. La syntaxe est la suivante:

Expression := <Donnée> R <Donnée>

Dans le cadre de ce mémoire, nous restreignons l'étude d'une expression à la forme suivante:

Expression := <Variable> R <Valeur>

où Valeur est une constante dont le type est compatible avec la Variable.

Exemples:

temps_transfert < 10.0

type_opération = TOURNAGE

La variable est déclarée par un type de l'ensemble des types possibles: entier, réel, chaîne, énumération, etc. Elle est identifiée par un nom qui est commun pour chaque agent. Dans le cas où le type est une énumération, il est nécessaire de définir l'énumération complète (dans le cas précédent, type_opération est une énumération).

La relation R est définie parmi les relations possibles incluses dans l'ensemble $\{>, <, =, \leq, \geq\}$. Dans le cas d'une variable de type numérique certaine, les relations sont définies de manière univoque. Pour les données incertaines, il est nécessaire de redéfinir cet ensemble. Ce point n'est pas abordé dans ce mémoire car il concerne la gestion de contraintes floues. Nous supposons dans une première approche que dans le cas d'une donnée incertaine, la comparaison porte sur la valeur possible médiane. Cette restriction ne gêne en rien la viabilité de l'approche, mais réduit significativement le réalisme de la relation.

Structure Logicielle

Plusieurs fichiers sont créés au niveau de chaque agent pour la gestion des contraintes:

- un fichier de données "ASD.DTA" (pour ASsociation de Données) qui contient la structure des données (organisation, type et nom de chaque donnée). Par exemple, la donnée TypeOpérationSIP est de type énumération "Type_Operation" et le fichier indique comment localiser cette donnée dans le fichier de base de données,
- un fichier d'énumération "ASV.DTA" (pour ASsociation de Valeurs) qui contient les énumérations définies en extension. Par exemple, le type "Type_Operation" est une énumération, il est donc ainsi défini en extension dans ce fichier de la manière suivante:

\$Type_Operation

NULLE

CONTROLE_VISUEL

ARASAGE

TOURNAGE
 USINAGE_CU
 etc.

Ce dernier fichier est commun à toutes les SIP afin d'uniformiser les énumérations. Le fichier "ASD.DTA" a pour seule contrainte de fournir les mêmes noms pour chaque donnée, et ce, pour chaque SIP. La structuration n'est pas nécessairement fixée, cependant, pour des raisons de simplicité, il est fortement conseillé d'adopter la même structure pour chaque SIP.

- un fichier "SIP.TXT" qui contient les données propres à chaque SIP (si nécessaire, les énumérations sont déclarées dans le fichier précédent "ASV.DTA"). Ce fichier est structuré selon le fichier "ASD.DTA". Ce fichier contient aussi les adresses de chaque SIP répertoriée (adresse réseau cas multi-postes ou nom SIP dans le cas mono-poste). Ce dernier champ peut être enrichi ou diminué selon l'adjonction/ suppression de SIP du système de production, cf. le protocole d'information défini dans le chapitre précédent.

Ces fichiers concernent la base de données qui est utilisée par la classe "Contraintes", cf. chapitre VI. Il existe un lien entre la classe "Contraintes" et la classe "SIPBdD" définie dans le système d'information afin que ce dernier puisse accéder à la fois aux données locales de la SIP et au gestionnaire de contraintes. Nous donnons les procédures principales qui ont trait à la création d'une contrainte:

Contrainte.ChargerValeur(Identificateur Donnée, Valeur, Type Donnée, Numéro Relation)

Contrainte.ChargerRelation(Identificateur Donnée, Type Relation, Numéro Relation)

Contrainte.ChargerConnexion(Identificateur Donnée, Type Connexion)

SystèmeInformation.VérifierContrainte(Contrainte)

Les paramètres utilisés sont:

- "Identificateur Donnée", qui correspond au champ associé à la donnée étudiée dans le fichier "ASD.DTA",
- "Valeur", qui correspond à la valeur avec laquelle la donnée est comparée,
- "Type Donnée" qui indique le type de la donnée en question. Ce type est défini dans "ASD.DTA" mais il permet au gestionnaire de s'assurer de la validité du type de la donnée afin d'éviter toute erreur,
- "Numéro Relation" qui décrit à quelle membre est associé la valeur (membre gauche: Numéro Relation = 1 ou droite: Numéro Relation = 2),

- “Type Relation” décrit le type de la relation pour le membre en question: par exemple, “inférieur_égal” pour la relation correspondante,
- “Type Connexion” décrit la nature de la connexion qui lie les deux membres (lorsque le deuxième est effectivement décrit). Si un deuxième membre est déclaré, alors Type Connexion peut être égal à “Connexion OU” ou “Connexion ET”. Si le membre de droite est inexistant la valeur par défaut “Connexion NO” est utilisée,
- la fonction “VérifierContrainte” du sous-système d’information traite la contrainte selon sa base de données caractéristiques “SIP.TXT”.

Il faut noter qu’un squelette de compilateur a été élaboré afin de vérifier la cohérence des bases de données “SIP.TXT” selon la structure “ASD.DTA” et les types décrits dans “ASV.DTA” et les contraintes créées (respect des définitions des connexions, adéquation entre type donnée et type valeur, etc.). Il est alors possible d’élaborer un éditeur de contraintes, mais cette possibilité n’a pas été étudiée dans ce mémoire.

Nous donnons un exemple de création. Nous supposons que la donnée identifiée 0 selon “ASD.DTA” est le type opératoire de la SIP “TypeOpération”, l’énumération étant décrite dans “ASV.DTA”.

```
MaContrainte.ChargerValeur(0, TOURNAGE, énumération, 1);
MaContrainte.ChargerRelation(0, Egal, 1);
MaContrainte.ChargerValeur(0, USINAGE_CU, énumération, 2);
MaContrainte.ChargerRelation(0, Egal, 2);
Macontrainte.ChargerConnexion(0, Connexion_OU);
Ceci représente la contrainte suivante:
```

```
TypeOpération=TOURNAGE ou TypeOpération=USINAGE_CU
```

Cette contrainte représente le fait qu’une opération puisse être réalisée à la fois sur une SIP qui contrôle un tour ou sur une SIP qui contrôle un centre d’usinage. C’est le cas notamment d’une opération de type “contournage” sur centre d’usinage (nommée “tournage” sur tour) qui est réalisable techniquement par les deux ressources.

Annexe E

Parallèle entre les niveaux A et T et la norme OSI de l'ISO

La typologie d'architecture de réseau auquel est associé notre proposition se situe au niveau intra-entreprise, ce qui correspond à un réseau local (Local Area Network: LAN) qui a la particularité de proposer un débit facilement élevé, ce qui nous permet en particulier de spécifier des protocoles assez complexes sans nous soucier à première vue du problème de mise en oeuvre de notre approche. Les normalisateurs de l'ISO ont proposé une décomposition d'une telle architecture en sept couches (niveaux) non exclusivement nécessaires à la bonne intégration du réseau [Pujolle et al. 92]. Cette décomposition porte le nom d'OSI pour Open Systems Architecture. Nous la décrivons très succinctement:

Le niveau 0 correspond au support physique.

Le niveau 1 correspond à la couche physique et a pour responsabilité le transport de l'information brute sous forme binaire. Sont inclus les interfaces de connexion.

Le niveau 2 (couche liaison) fournit les moyens fonctionnels et procéduraux nécessaires à l'établissement, au maintien et à la libération des connexions entre entités et réseaux. Ce niveau a pour but de corriger les erreurs de transmission (le taux d'erreur est négligeable, mais existe).

Le niveau 3 (couche réseau) a pour rôle de fournir les outils pour établir, maintenir, etc. les connexions réseaux. Elle est responsable de ce fait en particulier de l'acheminement physique (via les passerelles par exemple) des données (routage).

Le niveau 4 (couche transport) assure le transfert de données formatées (de manière à être transparent pour l'utilisateur). C'est une couche charnière entre l'aspect physique du réseau et l'aspect logiciel de l'application qui permet d'introduire une qualité logicielle en terme de genericité d'application au point de vue implémentation réseau.

Le niveau 5 (couche session) fournit les moyens pour synchroniser et organiser le dialogue de deux entités. Ce niveau fournit les outils pour ouvrir, maintenir et libérer une connexion.

Le niveau 6 (couche présentation) se charge de la syntaxe des informations. De ce fait, son rôle est primordial dans un environnement hétérogène. Il doit posséder des méthodes de traduction.

Le niveau 7 (couche application) donne le moyen d'accéder à l'environnement OSI. Elle gère l'aspect sémantique de la communication.

De cette manière, il est possible d'établir un lien fonctionnel entre ces description et la décomposition en deux niveau A et T.

Le niveau T s'intègre à la fois au niveau 4 et au niveau 5. En effet, son rôle est de fournir les moyens pour établir/ maintenir/ fermer une connexion entre un client et une serveur (niveau 5) et de formater le message <S,C,X> pour être compréhensible par les deux entités, ce qui est de la responsabilité du niveau 4. D'autre part, le niveau T est spécifié indépendamment de la structure et des caractéristiques réseau, il ne s'adresse pas au niveau 3. Il peut à l'extrême être associé au niveau 6 dans le cas où le système distribué est composé d'éléments hétérogènes. Cependant, nous supposons que tous les éléments sont de même nature (en terme de système d'exploitation par exemple).

Le niveau A s'établit très clairement au niveau 7. Il s'intègre naturellement à ce niveau de par son aspect logiciel qui exploite les fonctionnalités du réseau sans se préoccuper de la procédure d'échange même de données qui est réalisée par les niveaux 6 et inférieurs.

Annexe F

Liste des processus décisionnels et fonctions d'interface

Cette description porte sur les potentialités d'interaction de chaque mode de coopération (modification de la liste de critères, etc.) en fonction des processus décisionnels. Sur les tables suivantes nous résumons l'ensemble des paramètres entrants/ sortants, de déclenchement (conditions), etc. nécessaires à un processus décisionnel complet. Il est évident qu'une équivalence peut être établie entre ces tables et un diagramme SADT, mais étant considéré la complexité de l'ensemble des processus décisionnels, il nous semble plus pertinent de les présenter de la sorte.

La première table prend en compte les processus non structurés.

	Problématique selon la décision	Pré-condition d'activation	Mode de coopération possible	Déclenchement	Paramètres
Processus non-structurés					
Choix ré-allocation	Choix booléen (oui/non)	Présence opération sur machine, dans stock d'entrée	Opérateur	décision opérateur (panne, etc.)	Indéterminés (non structuré)
Choix de violation ordonnancement	Choix booléen (oui/non)	Présence opération non planifiée dans stock entrée	Opérateur	décision opérateur (panne, etc.)	Indéterminés (non structuré)
Choix remise en question ordo prévisionnel	Choix booléen (oui/non)	Il existe un ordonnancement prévisionnel	Opérateur	décision opérateur agent goulet (remise en question de l'ordonnancement)	Indéterminés (non structuré)
			Autonome	Arrivée ordre fabrication superviseur	Aucun

Table 1: processus non structurés

La seconde table concerne les processus peu structurés.

	Problématique selon la décision	Pré-condition d'activation	Mode de coopération possible	Déclenchement	Paramètres
Processus peu ou non structurés					
Gestion file d'attente produit en entrée	rangement des opérations ou choix d'une opération	Présence de plus d'un produit dans stock entré, ressource non-goulet	Opérateur	Décision opérateur, besoin chargement produit	Indéterminés (pas de coopération)
			Autonome	besoin chargement produit	- Temps passé en file d'attente entrée, - Marge, - Priorité, etc.
			Coopératif	Décision opérateur, besoin chargement produit	- Temps passé en file d'attente entrée, - Marge, - Priorité, etc.
Séquencement des ordres d'allocation	rangement des allocations ou choix d'une allocation	Présence de plus d'un produit dans stock sortie	Opérateur	Décision opérateur, besoin allocation produit	Indéterminés (pas de coopération)
			Autonome	besoin allocation produit	- Temps passé en file d'attente sortie, - Marge, - Priorité, - opération planifiée, etc.
			Coopératif	Décision opérateur, besoin allocation produit	- Temps passé en file d'attente, - Marge, - Priorité, - opération planifiée, etc.
Processus d'ordonnancement goulet	Réalisation d'un ordonnancement	Structure hybride activée, Présence d'ordres de fabrication, protocole d'ordonnancement en cours	Autonome	Choix affirmatif au processus de choix de remise en question de l'ordonnancement	- charges potentielles maximales, - acceptances SIP - nom de la ressource goulet, etc.
Gestion des variables locales	Choix pour chacune des variables	Spécifique à chaque variable (présence produit, etc.)	Autonome Coopératif	Décision opérateur, besoin modification valeurs	Spécifique à chaque variable (présence produit, etc.)
			Autonome	Besoin modification valeurs	Spécifique à chaque variable (présence produit, etc.)
Processus d'allocation	Allocation d'opération	Choix/ rangement séquencement allocation effectué ou une opération à allouer, protocole d'allocation en cours, acceptation(s) reçue(s)	Opérateur	Décision opérateur, opération présente dans stock sortie	Indéterminés (pas de coopération)
			Autonome	opération présente dans stock sortie	- Critères portant sur les coûts, délais, qualité, etc.
			Coopératif	Décision opérateur, opération présente dans stock sortie	- Critères portant sur les coûts, délais, qualité, etc.
Processus de ré-ordonnancement local	Choix/ rangement des prochaines opérations à réaliser	choix de ré-ordonnancement local positif, ressource goulet	Opérateur		Indéterminés (pas de coopération)
			Autonome		- le retard répercuté sur les autres opérations, etc.
			Coopératif		- le retard répercuté sur les autres opérations, etc.

Table 2: processus mal structurés

Ces tables permettent de définir les caractéristiques fonctionnelles de l'interface.

.

Annexe G

Modélisation générique d'un ordonnancement. Caractérisation d'un ordonnancement juste à temps

Soit R l'ensemble des ressources de production:

$$R = \{r_i | i \in [1, m]\}, m \text{ est le nombre de ressources,}$$

Soit P l'ensemble des tâches à réaliser:

$$P = \{p_j | j \in [1, n]\}, n \text{ est le nombre de tâches.}$$

Soit M l'ensemble des ordres de fabrication:

$$M = \{m_k | k \in [1, l]\}, l \text{ est le nombre d'ordres de fabrication.}$$

Soit B , E et D les fonctions qui donnent les dates de début et de fin de fabrication et la date de besoin d'un ordre $m_k \in M$:

$$B: M \rightarrow \mathfrak{R}$$

$$m_k \mapsto B(m_k) = b(k)$$

$$E: M \rightarrow \mathfrak{R}$$

$$m_k \mapsto E(m_k) = e(k)$$

$$D: M \rightarrow \mathfrak{R}$$

$$m_k \mapsto D(m_k) = d(k)$$

Ces fonctions sont exploitées lors de la caractérisation de l'aspect juste à temps.

La tâche $p_j \in P$ peut être réalisée sur la ressource $r_i \in R$. La fonction T donne le temps de réalisation de p_j par:

$$T: R \times P \rightarrow \mathfrak{R}$$

$$(r_i, p_j) \mapsto T(r_i, p_j) = d(i, j)$$

On pose $d(i, j) > 0$ si r_i peut réaliser p_j , 0 sinon.

Soit R_s et P_s sous ensembles de R et P , $n_s = \text{card}(P_s)$ et $m_s = \text{card}(R_s)$. Les éléments de P_s sont à ordonnancer, les éléments de R_s sont les ressources ordonnancées pour ces tâches.

Soit S la fonction ordonnancement:

$$S : R_s \times P_s \rightarrow \mathfrak{R}$$

$$(r_i, p_j) \mapsto S(r_i, p_j) \equiv s(i, j)$$

et \underline{S} la matrice d'ordonnancement associée à $s(i, j)$. $s(i, j)$ est égal à la date de début de fabrication de la ressource r_i pour la tâche p_j , sinon $s(i, j) = 0$.

Soit C la fonction qui intègre les trois types de contraintes:

- contraintes de disjonction,
- contraintes de capacité,
- contraintes associées au processus de fabrication.

C ne peut être décrite de manière générique. Nous pouvons seulement définir son domaine d'étude comme suit:

$$C : R^m \times P^n \times \overbrace{\mathfrak{R} \dots \times \mathfrak{R}}^{n_s \cdot m_s \text{ fois}} \rightarrow [0; 1]$$

$$\left((r_1, r_2, \dots, r_m), (p_1, p_2, \dots, p_n), s(\alpha_1, \beta_1), \dots, s(\alpha_{m_s}, \beta_{n_s}) \right) \mapsto C \left(\begin{array}{l} (r_1, r_2, \dots, r_m), \\ (p_1, p_2, \dots, p_n), \\ s(\alpha_1, \beta_1), \dots, s(\alpha_{m_s}, \beta_{n_s}) \end{array} \right) \equiv C(\underline{r}, \underline{p}, \underline{S})$$

Le n-uplet $(\underline{r}, \underline{p}, \underline{S})$ satisfait les contraintes si $C(\underline{r}, \underline{p}, \underline{S}) = 1$

Définition: la matrice \underline{S} est un ordonnancement satisfaisant si $C(\underline{r}, \underline{p}, \underline{S}) = 1$

Soit S^* l'ensemble des ordonnancements satisfaisants:

$$S^* = \{ \underline{S} / \underline{S} \text{ application} \wedge C(\underline{r}, \underline{p}, \underline{S}) = 1 \}$$

Définition: soit $(\underline{S}_1, \underline{S}_2) \in S^{*2}$ deux ordonnancements satisfaisants. Soit (m_1, σ_1) (resp. (m_2, σ_2)) le couple (retard moyen, écart type des retards) de \underline{S}_1 (resp. \underline{S}_2) défini comme suit:

$$m = \sum_{k=1}^1 \frac{e(k) - d(k)}{1}$$

$$\sigma = \sqrt{\frac{1}{1-1} \sum_{k=1}^1 (e(k) - d(k) - m)^2}$$

\underline{S}_1 est un ordonnancement plus juste à temps que \underline{S}_2 (noté $\underline{S}_1 \triangleright \underline{S}_2$) lorsque:

$$(1) |m_1| < |m_2| \text{ et}$$

$$(2) \sigma_1 < \sigma_2.$$

Si (1) ou (2) n'est pas vérifiée, on ne peut conclure. On pose $\underline{S}_1 \approx \underline{S}_2$.

Cette définition n'est certainement pas générique. Elle nous permet cependant d'avoir une image correcte de la caractéristique juste à temps d'un ordonnancement. Cette problématique est en fait multicritère: dans le cas $\underline{S}_1 \approx \underline{S}_2$, il est nécessaire de définir d'autres critères pour élaborer une préférence.

Annexe H

Calcul des charges potentielles Ordonnancement juste à temps de la ressource goulet

Charge potentielle

Soit Ω la fonction qui informe de l'état d'une tâche p_j sur une ressource r_i :

$$\Omega: R \times P \rightarrow \mathfrak{R}$$

$$(r_i, p_j) \mapsto \Omega(r_i, p_j) = \omega(i, j) = \begin{cases} 1 & \text{si } p_j \text{ a été réalisée sur } r_i \\ & \text{ou ne peut être réalisée par } r_i. \\ \text{OU} \\ \text{le pourcentage de temps passé par } r_i \text{ sur la tâche } p_j \\ & \text{si } p_j \text{ est en cours de réalisation par } r_i. \\ \text{OU} \\ 0 & \text{si } p_j \text{ peut être réalisé par } r_i. \end{cases}$$

Définition: la charge potentielle w_i d'une ressource $r_i \in R$ est définie par:

$$w_i = \sum_{j=1}^n d(i, j) * (1 - \omega(i, j))$$

Définition: le goulet \tilde{r} est la ressource r_i dont la charge potentielle est maximale pour l'ensemble des ressources.

L'ordonnancement n'est défini que pour cette ressource (cf. annexe G): $R_s = \{\tilde{r}\}$, $m_s = \text{card}(R_s) = 1$. Les éléments de P_s sont les tâches qui sont planifiées sur \tilde{r} . la fonction d'ordonnancement S est simplifiée, elle ne concerne que \tilde{r} :

$$S: \{\tilde{r}\} \times P_s \rightarrow \mathfrak{R}$$

$$(\tilde{r}, p_j) \mapsto S(\tilde{r}, p_j) \equiv s(j)$$

On définit alors \underline{S} le vecteur des tâches $s(j)$ planifiées sur la ressource \tilde{r} .

Ordonnancement juste à temps de la ressource goulet

L'ordonnancement est le suivant:

Début
 Soit w_i les charges potentielles
 Déterminer la ressource goulet \tilde{N}
 Pour chaque tâche p_j d'un ordre de fabrication $m_k \in M$ faire
 Début
 Si p_j doit être réalisée sur \tilde{N}
 Soit req_j = date de besoin de p_j - somme des temps de
 production en aval de \tilde{N} pour l'ordre de fabrication $m_k \in M$
 Fin
 Soit sch = tableau d'enregistrement du type (date, compteur)
 Trier les dates req_j de manière croissante dans $sch(cpt).date=req_j$ avec
 $sch(cpt).compteur=j$ et incrémenter cpt
 Pour $cpt=1$ à $n_s = \text{card}(P_s)$
 Début
 Planifier \tilde{N} pour tâche $p_{sch(cpt).count}$ à la date $sch(cpt).date - T(\tilde{N}, p_{sch(cpt).count})$
 Si ce n'est pas possible, planifier \tilde{N} au plus tôt après cette date
 Fin
 Fin

Cet algorithme recherche en premier lieu \tilde{N} et ordonnance cette ressource selon un critère qui favorise les tâches qui présentent les marges de temps libre après goulet qui sont les plus faibles. Une marge de temps libre après goulet correspond à la date de besoin moins le temps nécessaire à la réalisation des tâches en aval du goulet. Si une opération ne peut pas être réalisée à une date planifiée, elle est planifiée au plus tôt après cette date.