



**HAL**  
open science

# Contribution à l'étude de l'exploitation des données temporelles en présence d'imperfections

Aymen Gammoudi

► **To cite this version:**

Aymen Gammoudi. Contribution à l'étude de l'exploitation des données temporelles en présence d'imperfections. Autre [cs.OH]. ISAE-ENSMA Ecole Nationale Supérieure de Mécanique et d'Aérotechnique - Poitiers; Institut supérieur de gestion (Tunis), 2017. Français. NNT: 2017ESMA0008 . tel-01551795

**HAL Id: tel-01551795**

**<https://theses.hal.science/tel-01551795v1>**

Submitted on 30 Jun 2017

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

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

# THESE

Pour l'obtention du Grade de  
DOCTEUR DE L'ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET  
D'AEROTECHNIQUE EN COTUTELLE AVEC L'INSTITUT SUPERIEUR DE  
GESTION

(Diplôme National Arrêté du 25 mai 2016)

Ecole Doctorale :  
Sciences et Ingénierie pour l'Information et Mathématiques  
Secteur de Recherche : 911 Informatique et applications

Présenté par :  
GAMMOUDI Aymen

\*\*\*\*\*  
**CONTRIBUTION À L'ÉTUDE DE L'EXPLOITATION DES  
DONNÉES TEMPORELLES EN PRÉSENCE  
D'IMPERFECTIONS**  
\*\*\*\*\*

Directeurs de thèse :

HADJALI Allel  
BEN YAGHLANE Boutheina

\*\*\*\*\*

Soutenu le 17 MAI 2017

devant la Commission d'Examen

\*\*\*\*\*

## JURY

**Président :**

BEN AMOR Nahla, Professeur, Université de Tunis

**Rapporteurs :**

FARAH Riadh, Professeur, Université de Manouba

MAABOUT Sofian, MCF-HDR, Université de Bordeaux

**Membres du jury :**

Daniel ROCACHER, Professeur, Université de Rennes1

Mohamed Salah GOUIDER, MCF-HDR, Université de Tunis

... *Et on ne vous a donné que peu de connaissance.*

A conscience tranquille tout est accessible

Quand il y a la soif d'apprendre  
Tout vient à point à qui sait attendre

Quand il y a le souci de réaliser un dessein  
Tout devient facile pour arriver à nos fins

Malgré les obstacles qui s'opposent  
En dépit des difficultés qui s'interposent

Les études sont avant tout  
Notre unique et seul atout

Ils représentent la lumière de notre existence  
L'étoile brillante de notre réjouissance

Souhaitant que le fruit de nos efforts fournis  
Jour et nuit, nous mènera vers le bonheur fleuri

**A.G**

## Dédicace

*Je dédie cette thèse ...*

*A mon cher père Mokhtar*

Aucune dédicace ne saurait exprimer l'amour, l'estime, le dévouement et le respect que j'ai toujours eu pour toi.  
Rien au monde ne vaut les efforts fournis jour et nuit pour mon éducation et mon bien être.

Tu as su m'inculquer le sens de la responsabilité, de l'optimisme et de la confiance en soi face aux difficultés de la vie.

Tes conseils ont toujours guidé mes pas vers la réussite.

Ta patience sans fin, ta compréhension et ton encouragement sont pour moi le soutien indispensable que tu as toujours su m'apporter.

Je te dois ce que je suis aujourd'hui et ce que je serai demain et je ferai toujours de mon mieux pour rester ta fierté et ne jamais te décevoir.

Que Dieu le tout puissant te préserve, t'accorde santé, bonheur, qui étudie de l'esprit et te protège de tout mal

*A ma très chère mère Rgaya*

Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et de prier pour moi.

Ta prière et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études.

Tu n'as cessé de me soutenir et de m'encourager durant toutes les années de mes études, tu as toujours été présente à mes côtés pour me consoler quand il fallait.

En ce jour mémorable, pour moi ainsi que pour toi, reçoit ce travail en signe de ma vive reconnaissance et ma profonde estime.

Puisse le tout puissant te donner santé, bonheur et longue vie afin que je puisse te combler à mon tour.

*A mes très chers frères Walid et Majdi*

Mes chers qui m'ont le père, les frères et les amis.  
Pour vos encouragements, soutiens moral et surtout vos conseils. Vous avez toujours été à mes côtés pour me pousser vers la réussite.  
Les mots ne suffisent guère pour exprimer l'attachement, l'amour et l'affection que je porte pour vous.  
Que dieu réunisse nos chemins pour un long commun serein et que ce travail soit témoignage de ma reconnaissance et de mon amour sincère et fidèle.

*A mes très chères Soeurs Najia, Amel, Sarra et Olfa*

Votre encouragement et votre soutien étaient la bouffée d'oxygène qui me ressourçait dans les moments pénibles, de solitude et de souffrance.  
En témoignage de l'attachement, de l'amour et de l'affection que je porte pour vous.  
C'est une grande fierté pour moi d'avoir des soeurs comme vous. Je prie dieu le tout puissant pour qu'il vous donne bonheur et prospérité.

*A mes beaux-frères et soeurs : Hsan, Mohamed, Mabrouka et Noura*

En espérant qu'ils y trouvent le fruit de leur soutien, leur aide, leurs conseils et leurs encouragements sans cesse renouvelés et en les exprimant mon profond attachement.

*A mes petits anges Chahd, Minyar, Lojayen, Acil, Adem, Lina, Kodes et Yassine* que dieu les protège et les préserve un avenir radieux.

*A mon cher cousin Aymen GAMMOUDI (Bis)*

Mon cher cousin présent dans tous mes moments d'examens par son soutien moral et ses belles surprises sucrées

*A tous les membres de ma famille, petits et grands*

Veillez trouver dans ce modeste travail l'expression de mon affection

Je cite en particulier oncle **Lazhar** et **Tarek**.

Vous avez toujours été présents pour les bons conseils.

Votre affection et votre soutien m'ont été d'un grand secours au long de ma vie professionnelle et personnelle.

Veillez trouver dans ce modeste travail ma reconnaissance pour tous vos efforts.

A tous mes amis et toutes les personnes qui, à un moment ou un autre, m'ont soutenu et encouragé afin d'atteindre mes aspirations.

## Remerciements

A l'issue de la rédaction de cette recherche, je suis convaincue que la thèse est loin d'être un travail solitaire. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser dans cette phase délicate de «l'apprentichercheur»

En premier lieu, je tiens à remercier mon directeur de thèse, monsieur Allel HADJALI, pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral, pour ses multiples conseils et pour toutes les heures qu'il a consacrées à diriger cette recherche. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et son respect sans faille des délais serrés de relecture des documents que je lui ai adressés. Enfin, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral. Je lui dois ce que je suis aujourd'hui et ce que je serai demain et je ferai toujours de mon mieux pour lui combler à mon tour.

Je tiens à remercier aussi ma directrice de thèse, madame Boutheina BEN YAGHALNE pour la confiance qu'elle m'a accordée en acceptant d'encadrer ce travail doctoral, pour leur encadrement et leur conseil précieux tout au long de mon doctorat.

Je tiens encore à remercier vivement madame Nahla BEN AMOR, qui m'a honoré en acceptant d'être la présidente de jury, ainsi que les rapporteurs, monsieur Imed Riadh FARAH et monsieur Sofian MAABOUT, qui ont bien voulu accepter d'évaluer mon travail. Mes remerciements vont également aux examinateurs, monsieur Daniel ROCACHER et monsieur Mohamed Salah GOUIDER pour avoir accepté de participer à ce jury de thèse.

Je tiens à témoigner toute ma reconnaissance à madame Saoussen KRICHEN, pour leur soutien moral, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral.

Je remercie également tout le personnel de L'ISG et l'ENSMA pour leur aide et assistance dans toutes mes démarches administratives.

Je souhaiterais exprimer ma gratitude aux membres du laboratoire LIAS pour m'avoir donné envie de réaliser la thèse. Je les remercie également pour son accueil chaleureux à chaque fois que je rends visite au laboratoire, ainsi que pour ses multiples encouragements.

Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble des membres du laboratoire LARODEC pour leur soutien moral ainsi que pour la très bonne ambiance que j'ai toujours trouvée au laboratoire.

Je voudrais accorder une place d'honneur dans mes remerciements à ma famille, plus particulièrement, mes parents et mes frères. Leur soutien et encouragement n'a jamais failli au cours de ce long cursus universitaire.

Enfin, je remercie tous mes amis pour m'avoir soutenu durant cette période de thèse.

# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>I État de l'art</b>	<b>6</b>
<b>1 Notions de Base</b>	<b>8</b>
1.1 Introduction . . . . .	8
1.2 Formalismes temporels qualitatifs . . . . .	9
1.2.1 L'algèbre des intervalles . . . . .	9
1.2.2 L'algèbre des points . . . . .	11
1.3 Théorie des ensembles flous . . . . .	12
1.3.1 Compérateurs flous . . . . .	13
1.3.2 Relations temporelles d'Allen floues . . . . .	17
1.3.3 Intervalles flous . . . . .	20
1.3.4 Indices de comparaisons . . . . .	22
1.3.5 Différents types de prédicats flous . . . . .	26
1.3.6 Aperçu du langage SQLf . . . . .	28
1.4 Conclusion . . . . .	30

<b>2 Travaux Connexes</b>	<b>33</b>
2.1 Introduction . . . . .	33
2.2 Travaux en Intelligence Artificielle . . . . .	34
2.3 Travaux dans les Bases de données . . . . .	39
2.4 Conclusion . . . . .	44
<b>II Contributions</b>	<b>51</b>
<b>3 Requêtes en présence des relations temporelles graduelles</b>	<b>53</b>
3.1 Introduction . . . . .	53
3.2 Extensions floues des relations temporelles . . . . .	54
3.3 Approche de Galindo et Medina . . . . .	55
3.4 Approche proposée . . . . .	61
3.4.1 Modélisation . . . . .	61
3.4.2 Langage <i>TSQLf</i> . . . . .	62
3.4.3 Processus de traitement de requêtes <i>TSQLf</i> . . . . .	63
3.5 Exemples illustratifs . . . . .	64
3.6 Conclusion . . . . .	68
<b>4 Requêtes en présence des intervalles temporels flous</b>	<b>69</b>
4.1 Introduction . . . . .	69
4.2 Opérations de dilatation et d'érosion . . . . .	70
4.2.1 Opération de dilatation . . . . .	71
4.2.2 Opération d'érosion . . . . .	72
4.3 Relations d'Allen Tolérantes . . . . .	73
4.3.1 Modélisation . . . . .	74

4.3.2	Indices de comparaison . . . . .	75
4.4	Relations d'Allen entre les intervalles de temps flous . . . . .	79
4.4.1	Ensembles disjonctifs vs ensembles conjonctifs . . . . .	80
4.4.2	Relations temporelles : Vue disjonctive . . . . .	80
4.5	Conclusion . . . . .	83
<b>5</b>	<b>Implémentation et Expérimentation</b>	<b>85</b>
5.1	Introduction . . . . .	85
5.2	Le système <i>Fuzz-TIME</i> . . . . .	86
5.2.1	Architecture du système . . . . .	87
5.2.2	Implémentation et interfaces . . . . .	88
5.3	Exemple Illustratif . . . . .	95
5.4	Conclusion . . . . .	98
	<b>Conclusion générale</b>	<b>99</b>
	<b>Références bibliographiques</b>	<b>103</b>

# Table des figures

1.1	”Égalité approchée et inégalité stricte graduée.” . . . . .	15
1.2	”Au moins approximativement égal à”, ”clairement différent” et ”plus grand (ou plus petit) que”. . . . .	18
1.3	”Intervalles précis et flous”. . . . .	21
1.4	”Après ce soir”. . . . .	21
1.5	”Noyau, Support et Kernel”. . . . .	22
1.6	Certitude que M est strictement plus grand que N . . . . .	23
1.7	Certitude que M est plus grand où égal que N . . . . .	24
1.8	Possibilité que M est plus grand que N . . . . .	24
1.9	Possibilité que M est plus grand où égal que N . . . . .	25
3.1	Exemple 1 des relations d’Allen. . . . .	55
3.2	Exemple 2 des relations d’Allen. . . . .	55
3.3	Trapèze flou . . . . .	57
3.4	Inexactitude au niveau des bornes . . . . .	61
3.5	Processus de traitement de requêtes <i>TSQLf</i> . . . . .	64
4.1	Dilatation et érosion d’un intervalle de temps. . . . .	72

4.2	Indices possibilistes de comparaison des intervalles flous. . . . .	76
4.3	Indice de comparaison $d(A > B)$ . . . . .	78
5.1	Architecture du système <i>Fuzz-TIME</i> . . . . .	86
5.2	Définition d'une condition temporelle floue. . . . .	89
5.3	Définition des intervalles de validité et de tolérance. . . . .	90
5.4	Résultat d'une requête <i>TSQLf</i> . . . . .	91
5.5	Définition des conditions temporelles floues avec des bornes précises. . . . .	92
5.6	Définition des bornes temporelles floues. . . . .	92
5.7	Résultat d'une requête <i>TSQLf</i> 1. . . . .	93
5.8	Résultat d'une requête <i>TSQLf</i> 2. . . . .	93
5.9	Définition des conditions temporelles floues. . . . .	94
5.10	Résultat d'une requête <i>TSQLf</i> . . . . .	95

# Liste des tableaux

1.1	Relations temporelles d'Allen . . . . .	9
1.2	Relations temporelles entre instants . . . . .	11
1.3	Relations d'Allen floues . . . . .	19
1.4	Relation dept . . . . .	29
1.5	Relation emp . . . . .	30
1.6	Résultat de la requête <b>Q1</b> . . . . .	30
2.1	Imperfection dans les unités de temps de base . . . . .	42
2.2	Imperfection dans les relations temporelles . . . . .	45
2.3	Imperfection dans les unités de temps de base et les relations temporelles . . . . .	46
3.1	Les valeurs floues qui peuvent être utilisées dans la base de données floues ou dans des requêtes <i>FSQL</i> . . . . .	58
3.2	Extension des comparateurs du TSQL2 . . . . .	59
3.3	Les comparateurs flous . . . . .	60
3.4	Relations temporelles d'Allen floues . . . . .	62
3.5	Table d'Archéologie. . . . .	66

4.1	Relations d'Allen tolérantes (où $d(\tilde{b}_L > \tilde{a}'_L) = 1 - \sup_{x \leq y} \min(\mu_{\tilde{b}_L}(x), \mu_{\tilde{a}'_L}(y))$ )	79
4.2	Relations d'Allen entre deux intervalles de temps flous. . . . .	82
5.1	Table d'Archéologie. . . . .	96

# Introduction générale

## Contexte et problématique

Une base de données (**BD**) est un ensemble structuré d'éléments d'informations stockées dans un dispositif informatique. Ces éléments sont souvent stockés sous forme de tables selon des critères bien déterminés pour pouvoir les manipuler et les exploiter en cas de besoin d'une information. Plusieurs systèmes de gestion de base de données existants permettent d'organiser et de structurer la base de données de manière à pouvoir aisément manipuler son contenu et stocker efficacement de très grandes quantités d'informations. Les bases de données temporelles sont des bases de données qui gèrent certains aspects temporels (temps, dates ...) en plus des données qu'elles contiennent. Dans ce type de bases, les attributs liés au temps ne sont généralement pas traités comme les autres attributs, bien que tous les deux décrivent les propriétés des mêmes objets. Les attributs liés au temps sont considérés comme éléments ayant un impact sur la cohérence de l'ensemble d'objets modélisés par la base de données.

Les Bases de Données intégrant des aspects temporels, comme dans les Bases de Données historiques (i.e., décrivant des archives), sont de plus en plus sollicitées et utilisées (Nagypal & Motik, 2003). Dans ce contexte, les spécifications temporelles sont souvent incertaines, subjectives ou vagues. Par exemple, les périodes de "*la révolution russe*" et de "*la Renaissance*" sont caractérisées par des débuts et des fins naturellement graduels et mal-définis (comme "*juste avant les années 20*" ou "*vers la fin des années 30*"). L'imprécision ou le flou signifie que les spécifications temporelles ne sont pas rigides mais flexibles. Ceci est dû au fait que notre connaissance sur le temps est exprimée d'une manière linguistique.

Entre autres, nous pouvons définir la notion primitive de dates d'une manière vague ou utiliser des descriptions graduelles pour exprimer des relations linguistiques entre intervalles de temps. Quant à l'imperfection, l'incertitude provient du manque d'information sur l'état du monde. Il est alors impossible de déterminer si certaines connaissances, bien définies sur le monde, sont vraies ou fausses.

Le problème de la représentation des connaissances temporelles et du raisonnement temporel est un problème central en *Intelligence Artificielle*, notamment dans les domaines tels que la planification et l'ordonnancement, le raisonnement causal et le diagnostic, la compréhension du langage naturel, etc. Depuis le début des années 80, le raisonnement temporel a attiré l'attention de très nombreux chercheurs en *Intelligence Artificielle* et plusieurs approches ont été proposées dans ce domaine (Chittaro & Montanari, 2000). La plupart de ces approches se sont focalisées sur l'aspect algorithmique et sur la question de l'efficacité du raisonnement. Relativement, peu de travaux se sont intéressés à la modélisation et à l'expressivité. Actuellement, il n'existe pas un modèle permettant de modéliser les informations temporelles d'une manière fidèle et complètement satisfaisante. De plus, l'interrogation de données temporelles subjectives et empreintes d'imprécision et d'incertitude n'est pas totalement prise en charge par les langages d'interrogation traditionnels.

Par ailleurs, des problèmes similaires restent également posés dans le cadre des Systèmes de Questions/Réponses pour traiter certaines questions temporelles de type "Qui était le président des États Unis d'Amérique durant la guerre froide?". Ce type de question est problématique pour les systèmes traditionnels car, en général, il n'y a pas de date de début et de date de fin acceptées pour la guerre froide.

Cette thèse s'inscrit dans le cadre de la gestion des Bases de Données intégrant des aspects temporels mal-définis. Les défis scientifiques à relever, dans le cadre de cette thèse, sont multiples :

- définir un modèle permettant de représenter, d'une manière fidèle, les informations temporelles (entités temporelles et les relations qui tiennent entre elles) de nature complexes et imparfaites ;
- exploiter et interroger, d'une manière intelligente, ces informations temporelles en développant des mécanismes de raisonnement et d'inférence ;
- proposer des techniques d'optimisation afin de rendre le système plus performant et faire face à l'immensité des données temporelles manipulées.

## Contributions

L'objectif de cette thèse est ainsi de contribuer à l'exploitation des données temporelles avec des spécifications incomplètes. Il s'agit de définir, concevoir et développer un système intelligent pour l'interrogation des bases de données temporelles où la dimension temps est définie d'une manière imprécise et floue. Ce type de base de données est souvent présent dans les applications issues, notamment, des domaines historiques et d'archives.

L'ensemble des travaux présentés dans ce manuscrit se fondent sur des outils théoriques issus de la théorie des ensembles flous et du *soft computing*.

Une première contribution apportée sur la définition d'un nouveau modèle de représentation de données temporelles entachées d'imprécision et d'incertitude. Ce modèle a pris comme point de départ les travaux de (Dubois, Hadjali, & Prade, 2003) sur l'extension floue des relations temporelles d'Allen. Ces premiers travaux ont permis de développer un premier prototype nommé ***Fuzz-TIME***.

La plupart des travaux menés sur l'exploitation et la manipulation des informations temporelles considèrent que les entités temporelles sont définies d'une manière précise (non floue) par exemple, en termes d'intervalles classiques. Or dans le monde des applications réelles, les bornes des intervalles de temps manipulées sont floues. Par exemple, dans l'expression "début des années 20 jusqu'à la fin des années 30" les deux bornes de la période de temps en question sont floues.

Notre deuxième contribution a trait à la définition d'une nouvelle sémantique des relations temporelles d'Allen dans le contexte des intervalles de temps flous. L'élément clé de cette sémantique est une relation de tolérance convenablement choisie. Elle permet notamment d'associer à chaque intervalle de temps flou deux intervalles de temps emboîtés (par application des opérations de dilatation et d'érosion).

Par ailleurs, le calcul sur des relations d'Allen tolérantes a été aussi intégré dans le système ***Fuzz-TIME*** en exploitant les indices de comparaison de deux intervalles temporels flous.

Enfin, notre dernier apport concerne la modélisation d'une nouvelle extension des relations temporelles d'Allen entre des intervalles de temps flous. Cette extension se base sur une interprétation conjonctive des intervalles de temps.

## Organisation du manuscrit

Ce manuscrit est organisé en deux parties.

La première partie est dédiée à l'état de l'art qui est composée de deux chapitres.

Le chapitre 1 présente les différents formalismes utilisés dans le traitement des informations temporelles. Puis, nous présentons les comparateurs flous exprimés en termes de différence de valeurs. Ainsi que la composition des relations floues modélisant des égalités approchées, ou des inégalités graduées, est rappelée. Une deuxième partie de ce chapitre est réservée pour la présentation des intervalles flous, les indices de comparaison et les différents types des prédicats flous. Enfin, nous donnons un aperçu sur le langage *SQLf*. Ces notions liées sont nécessaires à la lecture de ce manuscrit.

Le chapitre 2 passe en revue les différentes approches qui ont abordé la problématique de la gestion et de traitement de l'information temporelle imparfaite. Nous décrivons les deux grandes familles d'approches de la gestion des données temporelles définies d'une manière incomplètes aussi bien en *Intelligence Artificielle* que dans le domaine des *bases de données*.

Dans la deuxième partie, nous décrivons nos contributions. Cette partie est structurée en trois chapitres.

Dans le chapitre 3, nous présentons une première contribution apportée sur la définition d'un nouveau modèle de représentation de données temporelles entachées d'imprécision et d'incertitude. Ce modèle a pris comme point de départ les travaux de (Dubois, Hadjali, & Prade, 2003) sur l'extension floue des relations temporelles d'Allen. Ces premiers travaux ont permis de développer un premier prototype nommé *Fuzz-TIME*.

Le chapitre 4 présente, en premier lieu, la définition d'une nouvelle sémantique des relations temporelles d'Allen dans le contexte des intervalles de temps flous. L'élément clé de cette sémantique est une relation de tolérance convenablement choisie. Elle permet notamment d'associer à chaque intervalle de temps flou deux intervalles de temps emboîtés (par application des opérations de dilatation et d'érosion). En deuxième lieu, l'établissement des indices de comparaisons entre deux intervalles de temps flous. Ces indices permettent d'équiper le système *Fuzz-TIME* par des mécanismes de calcul sur les relations d'Allen tolérantes. Une deuxième partie de ce chapitre est consacré pour notre dernier apport concerne la

modélisation d'une nouvelle extension des relations temporelles d'Allen entre des intervalles de temps flous. Cette extension se base sur une interprétation conjonctive des intervalles de temps.

Le chapitre 5 discute de l'expérimentation menée pour valider les contributions proposées. Nous présentons, en premier lieu, l'architecture du système *Fuzz-TIME* (pour *Fuzzy Temporal Information Management and Exploitation*). Ensuite, nous décrivons comment nous avons implémenté notre système. Enfin, nous donnons un exemple illustratif avec des interfaces systèmes.

Enfin, un bilan global de la thèse est fourni en conclusion générale où nous rappelons nos différentes contributions. Un ensemble de perspectives est également discuté dans ce chapitre.

# **Première partie**

## **État de l'art**



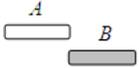
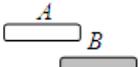
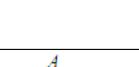
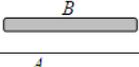
# Notions de Base

## 1.1 Introduction

Ce chapitre se limite à un rappel des notions de base utilisées dans la littérature pour la gestion et le traitement des informations temporelles qui sont essentielles à la compréhension des points développés dans les chapitres ultérieurs. Sa lecture peut donc être omise par un lecteur familier de l'ensemble de ces formalismes. A contrario, le lecteur souhaitant approfondir certains aspects pourra se reporter aux références suivantes : (Chittaro & Montanari, 2000), (Vila, 1994), (Allen, 1983), (Vilain & Kautz, 1986), (Dubois, Prade, & Prade, 2000), (Ohlbach, 2004).

Un ensemble de formalismes utilisés dans le traitement des informations temporelles est tout d'abord décrit, puis un aperçu générale sur la théorie des ensembles flous qui constituent la base de la manipulation des relations floues est présenté. Ensuite, La représentation des comparateurs flous exprimés en termes de différence de valeurs (i.e., égalité approchée et inégalité graduée) est discutée, en mettant en évidence des règles d'inférence contenant ces relations floues paramétrisées. Puis, une extension floue des relations temporelles est présentée et une sous-section est consacrée à la définition des intervalles flous et les indices de comparaisons de ces intervalles. Finalement, un aperçu générale sur les différents types de prédicats flous et du langage *SQLf* est présenté.

TABLE 1.1: Relations temporelles d'Allen

Relation	Inverse	Signification	Relations entre les bornes
$A < B$	$B > A$		$b > a'$
$A m B$	$B mi A$		$a' = b$
$A o B$	$B oi A$		$b > a \wedge a' > b \wedge b' > a'$
$A d B$	$B di A$		$a > b \wedge b' > a'$
$A s B$	$B si A$		$a = b \wedge b' > a'$
$A f B$	$B fi A$		$a > b \wedge b' = a'$
$A \equiv B$	$B \equiv A$		$a = b \wedge a' = b'$

## 1.2 Formalismes temporels qualitatifs

De nombreux formalismes temporels qualitatifs ont été proposés dans la littérature (Chittaro & Montanari, 2000) (Vila, 1994). *L'algèbre des intervalles d'Allen et l'algèbre des points* sont sans aucun doute les deux formalismes les plus connus et utilisés.

### 1.2.1 L'algèbre des intervalles

*L'Algèbre des Intervalles (AI) de Allen* (Allen, 1983) est sans aucun doute l'un des formalismes les plus connus et utilisés en raisonnement temporel. Dans le cadre de cette algèbre, les entités temporelles de base sont représentées sous forme d'intervalles de la droite réelle et les relations qualitatives entre ces entités correspondent aux 13 relations possibles et mutuellement exclusives qui peuvent

s'appliquer entre deux intervalles. Ces relations, dites relations d'Allen, sont habituellement dénotées par : *précède* ( $<$ ), *suit* ( $>$ ), *rencontre* ( $m$ ), *rencontré\_par* ( $mi$ ), *chevauche* ( $o$ ), *chevauché\_par* ( $oi$ ), *pendant* ( $d$ ), *contient* ( $di$ ), *commence* ( $s$ ), *commencé\_par* ( $si$ ), *termine* ( $f$ ), *terminé\_par* ( $fi$ ), et *égal* ( $\equiv$ ). Leur signification est illustrée dans le Tableau 1.1 (avec  $A = [a, a']$  et  $B = [b, b']$  deux intervalles temporels).

Ces relations peuvent être définies à partir des trois relations ordinaires  $<$ ,  $=$ ,  $>$  entre les bornes des deux intervalles à positionner l'un par rapport à l'autre. Par exemple, la déclaration  $A$  *chevauche*  $B$  correspond à

$$(b > a) \wedge (a' > b) \wedge (b' > a')$$

comme c'est indiqué dans le Tableau 1.1.

Allen (Allen, 1983) a fourni un ensemble d'axiomes décrivant la composition des 13 relations et un algorithme permettant d'inférer de nouvelles informations. Par exemple,

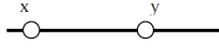
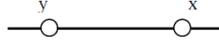
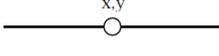
$$\begin{aligned} A < B \wedge B < C &\Rightarrow A < C, \\ A m B \wedge B d C &\Rightarrow (A o C) \vee (A d C) \vee (A s C). \end{aligned}$$

Le dernier exemple montre que Allen était obligé d'introduire des disjonctions de relations primitives pour traiter *l'incertitude* sur les relations, même si la composition est effectuée entre deux relations primitives. Dans de tels réseaux, les variables représentent des intervalles et les contraintes binaires possibles entre chaque paire de variables sont caractérisées à l'aide des 13 relations atomiques ou des relations complexes<sup>1</sup>.

Etant donné un Réseau d'Intervalles (RI), la question principale à laquelle on désire répondre est de savoir si ce réseau est consistant. Allen (Allen, 1983) utilise un algorithme de propagation de contraintes pour détecter l'incohérence d'un RI. Cet algorithme est souvent désigné par la méthode de la chemin-consistance (ou la triangulation d'Allen) qui utilise la table de compositions entre les 13 relations atomiques pour éliminer les contraintes qui ne participent pas à une instantiation consistante du réseau considéré. L'algorithme utilisé est polynomial en temps ( $\theta(n^3)$ ) et en espace ( $\theta(n^2)$ ), où  $n$  est le nombre de variables du RI. Mais il se trouve que cette méthode n'est pas complète. En général, il s'avère que le problème de

1. Les relations complexes sont définies par des disjonctions de relations atomiques. Au total,  $2^{13}$  relations complexes peuvent être formées.

TABLE 1.2: Relations temporelles entre instants

Relation	Symbole	Inverse	Interprétation
$x$ précède $y$	$<$	$y$ suit $x$	
$x$ suit $y$	$>$	$y$ précède $x$	
$x$ égale $y$	$=$	$y$ égale $x$	

la consistance est NP-complet, d'où les nombreuses études proposées pour la recherche des fragments de l'ensemble des relations de l'algèbre des intervalles pour lesquels le problème de la consistance des RI est polynomial (Drakengren & Jonsson, 1997), et en particulier pour lesquels la méthode de la chemin-consistance est complète. A titre d'exemple, l'ensemble des 13 relations atomiques est le plus petit sous-ensemble des relations de l'algèbre des intervalles qui est *traitable*.

### 1.2.2 L'algèbre des points

Le second modèle temporel qualitatif assez usité dans le domaine de l'Intelligence Artificielle est celui introduit par Vilain et Kautz (Vilain & Kautz, 1986), qu'on dénomme par *l'Algèbre des Points* (AP). Les entités temporelles considérées sont les points de la droite des réels (nous parlerons d'instant ou de dates). Les relations temporelles qualitatives qui permettent de comparer deux instants sont au nombre de trois, et correspondent aux situations suivantes : un instant *précède* un autre, un instant *est égal* à un autre, un instant *suit* un autre, voir Tableau 1.2.

Ces relations permettent d'exprimer des situations faisant intervenir des informations *incertaines*, par exemple, "un instant  $t_1$  *précède* ou *égale* un instant  $t_2$ ". Cette relation est notée  $t_1 \leq t_2$  où  $\leq$  représente la disjonction des deux relations atomiques  $<$ ,  $=$ . La composition de deux relations atomiques n'est pas forcément une relation atomique. Par exemple, si  $t_1 < t_2$  et  $t_2 > t_3$  alors  $t_1 r t_3$  où  $r = <, =, >$ . L'algèbre des points comprend au total 8 relations temporelles ( $\emptyset$ ,  $<$ ,  $>$ ,  $=$ ,  $\{<, =\}$ ,  $\{>, =\}$ ,  $\{<, >\}$ ,  $\{<, >, =\}$ ).

A l'instar d'Allen, Vilain et Kautz (Vilain & Kautz, 1986) représentent les informations temporelles entre un ensemble de points par des réseaux de contraintes et utilisent la méthode de la consistance de chemin comme mécanisme de raison-

nement. Dans le cadre de cette algèbre, cette méthode est complète et permet de résoudre le problème de la consistance des réseaux de points en un temps polynomial (de l'ordre de  $\theta(n^3)$  avec  $n$  est le nombre de variables du réseau).

### 1.3 Théorie des ensembles flous

Un ensemble flou  $F$  (Dubois et al., 2000) sur l'univers du discours  $U$  est décrit par une fonction d'appartenance  $\mu_F : U \rightarrow [0, 1]$  où  $\mu_F(u)$  est le degré d'appartenance de  $u$  à  $F$ . Deux ensembles précis sont d'un intérêt particulier lors de la définition d'un ensemble flou  $F$  :

- (i) Le noyau  $C(F) = \{u \in F, \mu_F(u) = 1\}$
- (ii) Le support  $S(F) = \{u \in F, \mu_F(u) > 0\}$ .

Dans la pratique,  $F$  est souvent associée à une fonction d'appartenance trapézoïdale (t.m.f.). Alors,  $F$  est exprimé par le quadruplet  $(A, B, a, b)$  où  $C(F) = [A, B]$  et  $S(F) = [A-a, B+b]$ . Un intervalle régulier  $[A, B]$  peut être considéré comme un ensemble flou représenté par le quadruplet  $(A, B, 0, 0)$ .

Soit  $F$  et  $G$  deux ensembles flous sur l'univers  $U$ , nous disons que  $F \subseteq G$  ssi  $\mu_F(u) \leq \mu_G(u), \forall u \in U$ . Le complément de  $F$ , notée  $F^c$ , est défini par  $\mu_{F^c}(u) = 1 - \mu_F(u)$ . En outre,  $F \cap G$  (resp.  $F \cup G$ ) est défini de telle sorte que  $\mu_{F \cap G}(u) = \min(\mu_F(u), \mu_G(u))$  (resp.  $\mu_{F \cup G}(u) = \max(\mu_F(u), \mu_G(u))$ ).

Normes triangulaires (t-normes en abrégé) et implications floues sont couramment utilisées comme des généralisations logiques floues de la *conjonction* et *l'implication*. Une t-norme  $\mathbb{T}$  est toute application, commutative et associative de  $[0, 1]^2 \rightarrow [0, 1]$  satisfaisant  $\mathbb{T}(1, u) = u$  pour tout  $u$  dans  $[0, 1]$ . Les t-normes les plus populaires sont :

- L'opérateur minimum avec  $\forall (u, v) \in [0, 1]^2, \mathbb{T}(u, v) = \min(u, v)$ .
- Le produit algébrique avec  $\forall (u, v) \in [0, 1]^2, \mathbb{T}(u, v) = u \cdot v$ .
- La t-norme de Lukasiewicz avec  $\forall (u, v) \in [0, 1]^2, \mathbb{T}(u, v) = \max(0, u + v - 1)$ .

Une implication floue  $I$  est toute application  $[0, 1]^2 \rightarrow [0, 1]$  satisfaisant les conditions aux bornes  $I(0, 0) = 1$  et  $I(1, u) = u$  pour tout  $u$  dans  $[0, 1]$ . En outre,  $I$  doit être décroissante sur son premier, et croissante sur son second argument.

Maintenant, soit  $\mathbb{T}$  une t-norme, une implication floue,  $I_{\mathbb{T}}$ , induite par  $\mathbb{T}$  est définie par

$$I_{\mathbb{T}}(u, v) = \sup\{\lambda \in [0, 1] / \mathbb{T}(u, \lambda) \leq v\},$$

pour  $u, v \in [0, 1]$ . Ce type d'implication est généralement appelé l'implication résiduelle (ou R-implication). Les trois R-implications les plus utilisées sont :

- *L'implication de Gödel* :  $I_{Gödel}(u, v) = 1$  si  $u \leq v$ ,  $v$  sinon (avec  $\mathbb{T}(u, v) = \min(u, v)$ ).
- *L'implication de Goguen* :  $I_{Goguen}(u, v) = 1$  si  $u \leq v$ ,  $v/u$  sinon (avec  $\mathbb{T}(u, v) = u \cdot v$ ).
- *L'implication de Lukasiewicz* :  $I_{Luka}(u, v) = 1$  si  $u \leq v$ ,  $1 - u + v$  sinon (avec  $\mathbb{T}(u, v) = \max(0, u + v - 1)$ ).

### 1.3.1 Comparateurs flous

La représentation des comparateurs flous exprimés en termes de différence de valeurs, est tout d'abord discutée. Puis, la composition des relations floues modélisant des égalités approchées, ou des inégalités graduées, est rappelée. Des règles d'inférence contenant ces relations floues paramétrisées sont ensuite établies.

#### A. Égalité approchée et inégalité graduée

Une *égalité approchée* entre deux valeurs, représentant ici des dates, modélisée par une relation floue E avec  $\mu_E$  sa fonction d'appartenance (E signifie "Égal"), peut être basée sur une distance telle que la valeur absolue de la différence, c'est-à-dire,

$$\mu_E(x, y) = \mu_L(|x - y|)$$

où  $L$  est un ensemble flou qui modélise la disparité floue maximale entre deux valeurs qui sont estimées approximativement égales. Une égalité approchée peut être représentée par :

$$\forall x, y \in \mathbb{R}$$

$$\mu_E(x, y) = \mu_L(|x - y|) = \max(0, \min(1, \frac{\delta + \varepsilon - |x - y|}{\varepsilon})) =$$

$$\begin{cases} 1 & \text{si } |x - y| \leq \delta \\ 0 & \text{si } |x - y| > \delta + \epsilon \\ \frac{\delta + \epsilon - |x - y|}{\epsilon} & \text{sinon} \end{cases}$$

avec  $\delta$  et  $\epsilon$  sont respectivement des paramètres positifs et strictement positifs, qui modulent l'égalité approchée. Ici  $L$  est un ensemble flou centré en 0. Voir Figure 1.1 où  $L = (-\delta, \delta, \epsilon, \epsilon)$ . L'égalité classique peut être obtenue pour  $\delta = 0$  et  $\epsilon \rightarrow 0$ . L'égalité approchée entre les quantités  $a$  et  $b$  (au sens de E) peut donc être écrite sous la forme :

$$a - b \in L \Leftrightarrow b - a \in L \Leftrightarrow a \text{ E}(L) b,$$

ce qui signifie que les valeurs possibles de la différence  $a - b$  sont restreintes par l'ensemble flou  $L$ . En particulier,  $a \text{ E}(0) b$  signifie que  $a = b$ . D'une façon similaire, une inégalité plus ou moins forte peut être modélisée par une relation floue  $G$  ( $G$  signifie "Grand") de la forme :

$$\mu_G(x, y) = \mu_K(x - y).$$

Dans ce qui suit, on considère :

$$\forall x, y \in \mathbb{R}$$

$$\mu_G(x, y) = \mu_K(x - y) = \max(0, \min(1, \frac{x - y - \lambda}{\rho})) =$$

$$\begin{cases} 1 & \text{si } x > y + \lambda + \rho \\ 0 & \text{si } x \leq y + \lambda \\ \frac{x - y - \lambda}{\rho} & \text{sinon} \end{cases}$$

On suppose  $\rho > 0$ , i.e.  $G$  renforce l'idée de "strictement plus grand que". On doit aussi avoir  $\lambda \geq 0$  afin de garantir que  $G$  soit une relation en conformité avec l'idée de "plus grand que". La relation "strictement plus grand que" ( $>$ ) est obtenue pour  $\lambda = 0$  et  $\rho \rightarrow 0$ . Une contrainte plus ou moins forte de type "a est plus grand que b" (au sens de  $G$ ) peut donc être écrite sous la forme

$$a - b \in K \Leftrightarrow a \text{ G}(K) b,$$

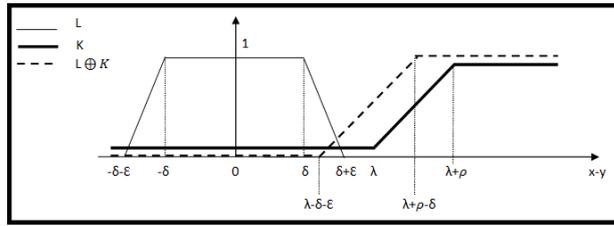


FIGURE 1.1: "Égalité approchée et inégalité stricte graduée."

où  $K$  est un ensemble flou de valeurs positives avec une fonction d'appartenance strictement croissante entre  $\lambda$  et  $\lambda + \rho$ . Voir Figure 1.1 où  $K = (\lambda, \lambda + \rho, +\infty, +\infty)$ . Selon les valeurs des paramètres  $\lambda$  et  $\rho$ , la modalité qui indique de combien  $a$  est plus grand que  $b$  peut être exprimée dans un contexte donné par les termes linguistiques tels que "légèrement", "beaucoup", etc.  $G(0)$  signifie '>'. Des relations floues "plus petit que" peuvent être aussi définies. Notons que  $a G(K) c \Leftrightarrow a - c \in K \Leftrightarrow c - a \in K^{ant} \Leftrightarrow c S(K^{ant}) a$  où  $S$  signifie "plus petit que",  $K^{ant}$  est l'antonyme de  $K$  défini par  $\mu_{K^{ant}}(d) = \mu_K(-d)$ . Ainsi, si  $a$  est beaucoup plus grand que  $c$ ,  $c$  est beaucoup plus petit que  $a$ .

### B. Composition des relations floues

La composition  $G(K) \circ E(L)$  de deux relations floues  $G(K)$  et  $E(L)$  est définie par :

$$\forall x \in X, \forall z \in Z$$

$$\mu_{G(K) \circ E(L)}(x, z) = \sup_{y \in Y} \min(\mu_G(x, y), \mu_E(y, z)) =$$

$$\sup_{y \in Y} \min(\mu_K(x - y), \mu_L(y - z))$$

$$= \mu_{K \oplus L}(x - z)$$

où on reconnaît l'expression de l'addition étendue  $\oplus$  des ensembles flous  $K$  et  $L$ . Dans la Figure 1.1,  $K \oplus L = (\lambda - \delta - \epsilon, \lambda + \rho - \delta, +\infty, +\infty)$ . Si on sait que " $a$  est approximativement égal à  $b$ " (i.e.,  $a E(L) b$ ) et que " $b$  est plus grand que  $c$ " (i.e.,  $b G(K) c$ ), alors on peut déduire que  $a - c \in K \oplus L \Leftrightarrow a G(K \oplus L) c$ , en utilisant la formule de composition ci-dessus. Le résultat est représenté en Figure 1.1.

### C. Règles d'inférence paramétrisées

En exploitant le fait que la composition des relations floues  $E(L)$  et  $G(K)$  se réduit à de simples opérations arithmétiques sur les paramètres  $K$  et  $L$  sous-

jacents, précisant les sémantiques de E et G, des règles d'inférence paramétrisées peuvent être obtenues. Ainsi, un ensemble de règles suivantes, décrivant les comportements des comparateurs flous E et G, a été établi (Dubois, Hadjali, & Prade, 2004),(Dubois, Hadjali, & Prade, 2003).

- **Propriétés de base de E et G**

$$R_1 : a \leq c \leq b \text{ et } a \text{ E(L) } b \Rightarrow a \text{ E(L) } c \quad \text{Convexité}$$

$$R_2 : a \text{ E(L) } b \Leftrightarrow b \text{ E(L) } a \quad \text{Symétrie}$$

$$R_3 : a \text{ E(L) } b \Leftrightarrow a + c \text{ E(L) } b + c \quad \text{E-Compatibilité avec l'addition}$$

$$R_4 : a \text{ G(K) } b \Leftrightarrow a + c \text{ G(K) } b + c \quad \text{G-Compatibilité avec l'addition}$$

- **Règles de fermeture**

$$R_5 : a \text{ E(L) } b \text{ et } b \text{ E(L')} c \Rightarrow a \text{ E(L} \oplus \text{L')} c \quad \text{E-Transitivité}$$

$$R_6 : a \text{ G(K) } b \text{ et } b \text{ G(K')} c \Rightarrow a \text{ G(K} \oplus \text{K')} c \quad \text{G-Transitivité}$$

$$R_7 : a \text{ E(L) } b \text{ et } b \text{ G(K) } c \Rightarrow a \text{ G(K} \oplus \text{L) } c. \quad \text{E-G-Composition}$$

La règle  $R_5$  exprime l'affaiblissement de la transitivité pour l'égalité approchée. Au contraire,  $R_6$  exprime un renforcement pour l'inégalité G. Ainsi, les règles  $R_1$  à  $R_7$  nous permettent de calculer formellement les paramètres flous des relations dans un processus d'inférence, et de les interpréter. A partir de ces règles de base, d'autres règles peuvent être établies :

- **Stabilité par rapport à la somme**

$$R_8 : a \text{ E(L) } b \text{ et } c \text{ E(L')} d \Rightarrow a + c \text{ E(L} \oplus \text{L')} b + d$$

$$R_9 : a \text{ E(L) } b \text{ et } c \text{ G(K) } d \Rightarrow a + c \text{ G(L} \oplus \text{K) } b + d$$

$$R_{10} : a \text{ G(K) } b \text{ et } c \text{ G(K')} d \Rightarrow a + c \text{ G(K} \oplus \text{K')} b + d$$

En effet, considérons la règle  $R_8$  par exemple. L'application de  $R_3$  permet d'écrire  $a \text{ E(L) } b \Rightarrow a + c \text{ E(L) } b + c$  et  $c \text{ E(L')} d \Rightarrow b + c \text{ E(L')} b + d$ , alors par  $R_5$  on prouve  $R_8$ . Il peut être montré qu'en fait  $R_8$  est équivalente à  $R_5$  puisque  $R_8$  implique  $R_5$  (en posant  $b = c$  et en appliquant  $R_3$ ). Maintenant, puisque  $a > b$  est équivalent à  $a \text{ G(0) } b$  où  $k = 0$  correspond au cas où  $\lambda = 0$  et  $\rho \rightarrow 0$  dans (2). Les propriétés intuitives suivantes de la relation floue G peuvent être aussi dérivées en utilisant la règle  $R_6$  (puisque  $K \oplus 0 = K$ ) :

$$R_{11} : a > b \text{ et } b \text{ G(K) } c \Rightarrow a \text{ G(K) } c,$$

$$R_{12} : a \text{ G(K) } b \text{ et } b > c \Rightarrow a \text{ G(K) } c.$$

On peut encore montrer que :

$$R_{13} : a + b \text{ G(K) } c + d \text{ et } c \text{ E(L) } a \Rightarrow b \text{ G(K} \oplus \text{L) } d, \text{ (en utilisant } R_3 \text{ deux fois et } R_7)$$

$R_{14} : a + b \text{ E}(L) \text{ c et } c \text{ G}(K) a \Rightarrow b \text{ G}(L \oplus K) 0. \text{ (en utilisant } R_9 \text{ et } R_4)$

Des règles contenant des durées peuvent être aussi dérivées. Par exemple, Si  $a \text{ G}(K) b$  et  $b - c \in D$  alors  $a \text{ G}(K \oplus D) c$ , où  $D$  représente l'information floue sur le temps passé entre  $b$  et  $c$ .

### 1.3.2 Relations temporelles d'Allen floues

Cette section discute l'extension des relations d'Allen, basée sur des relations d'égalité approchée et d'inégalité stricte graduée. En utilisant les règles d'inférence établies précédemment, on montre que la composition des relations d'Allen ordinaires peut être facilement étendue, en pratique, en ajoutant au calcul classique la manipulation arithmétique des paramètres flous.

#### A. Modélisation

En utilisant les comparateurs flous paramétrisés  $\text{E}(L)$  et  $\text{G}(K)$ , on peut définir des contreparties floues des relations d'Allen. L'idée est que les relations qui peuvent être vérifiées entre des bornes d'intervalles peuvent ne pas être décrites en termes précis. Par exemple, on s'exprimera en termes d'égalité approchée (au sens de  $\text{E}$ ) plutôt qu'en termes d'égalité précise afin d'éviter les discontinuités brutales entre le cas d'une relation de "parfaite" rencontre et le cas d'une relation *précède* quand la borne supérieure du premier intervalle est proche de la borne inférieure du second intervalle.

Ainsi, en termes approximatifs, étant donné deux dates  $a$  et  $b$ . Elles peuvent être "approximativement égales" au sens de  $\text{E}(L)$  ou,  $a$  peut être "clairement différent de"  $b$  au sens de non  $\text{E}(L)$ . Cette dernière relation correspond à "beaucoup plus grand que" au sens de  $\text{G}(K)$  ou à "beaucoup plus petit que" au sens de  $\text{S}(K^{ant})$ . Par conséquent, les paramètres  $L$ ,  $K$  et  $K^{ant}$  sont des éléments d'une partition floue (comme indiqué dans la Figure 1.2). C'est-à-dire,

$$\forall d \in \mathbb{R}, \mu_K(d) + \mu_{K^{ant}}(d) + \mu_L(d) = 1,$$

puisque  $K$  et  $K^{ant}$  sont obtenus par une opération de complémentation floue appliquée sur  $L$ . Chaque paramètre peut être obtenu à partir de chaque autre pa-

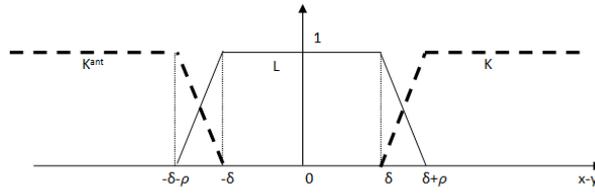


FIGURE 1.2: "Au moins approximativement égal à", "clairement différent" et "plus grand (ou plus petit) que".

ramètre de la manière suivante (où  $\bar{L}$  désigne le complément de l'ensemble flou  $L$ ) :

$$\begin{aligned} K &= \bar{L} \cap [0, +\infty) \text{ noté } L_+^c, \\ K^{ant} &= \bar{L} \cap (-\infty, 0] \text{ noté } L_-^c. \end{aligned}$$

Réciproquement  $L$  peut être obtenu à partir de  $K$  comme suit :  $L = \overline{K \cup K^{ant}}$ , noté  $K^c$ . En cas de représentation trapézoïdale, on a :

Si  $K = (\gamma, +\infty, \rho, +\infty)$  alors  $K^c = (-\gamma + \rho, \gamma - \rho, \rho, \rho)$ ,

Si  $L = (-\delta, \delta, \rho, \rho)$  alors  $L_+^c = (\delta + \rho, +\infty, \rho, +\infty)$  et  $L_-^c = (-\infty, -\delta - \rho, +\infty, \rho)$ .

Ainsi, puisque  $L$ ,  $K$  et  $K^{ant}$  sont étroitement liés, il suffit d'avoir un seul paramètre sous-jacent pour ces relations. Ce qui a été adopté dans la définition des relations d'Allen floues (introduites ci-dessous). Il apparaît ainsi que  $E(L)$ ,  $G(K)$  et  $S(K^{ant})$  avec  $L$  et  $K$  interdéfinis comme expliqué ci-dessus, sont des contreparties floues des relations classiques =, > et < dans le sens non flou.

Soient  $A = [a, a']$  et  $B = [b, b']$  deux intervalles de temps. Les relations d'Allen floues peuvent être maintenant définies comme illustré dans la Table 1.3.

Dans la modélisation proposée, il convient de noter :

- que la nature graduelle des relations temporelles est clairement prise en compte puisque la satisfaction de chacune d'elle est maintenant une question de degré. De plus selon la sémantique associée au paramètre flou  $L$  (et donc au paramètre flou  $L_+^c$ ), et par exemple dans le cas de la relation  $fb(L)$ , un agent peut distinguer entre la situation "bien avant" et la situation "juste avant".
- qu'une relation temporelle floue permet de recouvrir plusieurs situations correspondant à différentes relations d'Allen ordinaires ; par exemple,  $F-rencontre(L)$  recouvre la situation "rencontre" ordinaire et aussi les situations comme "précède légèrement" ou "chevauche légèrement". Cependant ici, le paramètre flou  $L$  contrôle à quel point on peut s'éloigner du

TABLE 1.3: Relations d'Allen floues

Relation d'Allen floue	Définition	Label
$A F\text{-précède}(L) B$ $B F\text{-suit}(L) A$	$b G(L_+^c) a'$	$fb(L)$ $fa(L)$
$A F\text{-rencontre}(L) B$ $B F\text{-rencontré\_par}(L) A$	$a' E(L) b$	$fm(L)$ $fmi(L)$
$A F\text{-chevauche}(L) B$ $B F\text{-chevauché\_par}(L) A$	$b G(L_+^c) a \wedge a' G(L_+^c) b \wedge b' G(L_+^c) a'$	$fo(L)$ $foi(L)$
$A F\text{-pendant}(L) B$ $B F\text{-contient}(L) A$	$a G(L_+^c) b \wedge b' G(L_+^c) a'$	$fd(L)$ $fdi(L)$
$A F\text{-commence}(L) B$ $B F\text{-commencé\_par}(L) A$	$a E(L) b \wedge b' G(L_+^c) a'$	$fs(L)$ $fsi(L)$
$A F\text{-termine}(L) B$ $B F\text{-terminé\_par}(L) A$	$a G(L_+^c) b \wedge a' E(L) b'$	$ff(L)$ $ffi(L)$
$A F\text{-égal}(L) B$ $B F\text{-égal}(L) A$	$a E(L) b \wedge a' E(L) b'$	$fe(L)$ $fe(L)$

”rencontre” ordinaire, et peut fournir des bases pour la sémantique de ce que ”légèrement” signifie dans les expressions ci-dessus. De la même manière, on peut voir que  $F\text{-égal}(L)$  peut recouvrir la situation ordinaire exprimée par ”presque contient” ou ”presque pendant”.

## B. Raisonnement

Comme dans le cas classique, on peut raisonner sur la base des relations d'Allen floues établies par le calcul de la fermeture transitive des relations temporelles floues en utilisant les règles d'inférence données en Section 1.3.1. Par exemple, l'axiome suivant est établi par application de la règle ( $a > b$  et  $b G(K) c \Rightarrow a G(K) c$ ) et de  $R_6$  :

$$A \text{ fb}(L_1) B \wedge B \text{ fb}(L_2) C \Rightarrow A \text{ fb}(L_1 \oplus L_2) C.$$

D'autres axiomes avec différents paramètres flous peuvent aussi être établis, comme :

$$\begin{aligned} A \text{ fo}(L_1) B \wedge B \text{ fm}(L_2) C &\Rightarrow A \text{ fb}(\mathbb{L}_1) C, \\ A \text{ ff}(L_1) B \wedge B \text{ fd}(L_2) C &\Rightarrow A \text{ fd}(\mathbb{L}_2) C, \\ A \text{ fs}(L_1) B \wedge B \text{ ff}(L_2) C &\Rightarrow A \text{ fd}(\mathbb{L}) C, \end{aligned}$$

où  $\mathbb{L}_1 = ((L_1)_+^c \oplus L_2)^c$ ,  $\mathbb{L}_2 = (L_1 \oplus (L_2)_+^c)^c$  et  $\mathbb{L} = ((L_1 \oplus (L_2)_+^c) \cup ((L_1)_+^c \oplus L_2))^c$ . L'ensemble complet des axiomes de transitivité est disponible dans (Dubois et al., 2004), (Dubois, Hadjali, & Prade, 2003).

Les inférences qui peuvent être dérivées de la table de composition, conduisent à des relations dont les paramètres flous ont les formes suivantes :

- $(L_1 \oplus L_2)$  quand les relations initiales sont paramétrisées par  $L_1$  et  $L_2$ . Puisque  $(L_1)_+^c \supseteq (L_1 \oplus L_2)_+^c$  et  $(L_2)_+^c \supseteq (L_1 \oplus L_2)_+^c$  (resp.  $L_1 \subseteq L_1 \oplus L_2$  et  $L_2 \subseteq L_1 \oplus L_2$ ) alors, la relation temporelle floue obtenue, qui est définie seulement sur la base de l'inégalité G (resp. l'égalité approchée E), est renforcée (resp. affaiblie).
- $((L_1)_+^c \oplus L_2)^c$  quand les relations initiales sont paramétrisées par  $L_1$  et  $L_2$ . Puisque  $(L_1)_+^c \subseteq (L_1)_+^c \oplus L_2$  (resp.  $L_1 \supseteq ((L_1)_+^c \oplus L_2)^c$ ) alors, la relation temporelle floue obtenue, qui est définie seulement sur la base de l'inégalité G (resp. l'égalité approchée E), est affaiblie (resp. renforcée).

Il est important de noter que l'itération des axiomes de transitivité peut conduire à des effets de dégradation dans les relations temporelles floues inférées. En effet, les relations symboliques obtenues pourraient ne pas être en complète conformité avec les sémantiques intuitives sous-jacentes des relations temporelles auxquelles elles réfèrent. Ceci est essentiellement dû au fait que quand le paramètre flou  $L$  (resp.  $K$ ) devient trop permissif (resp. proche de 0), on s'éloigne des sémantiques intuitives de l'égalité approchée (resp. beaucoup plus grand que) exprimées par la relation E (resp. G).

### 1.3.3 Intervalles flous

Les intervalles flous (Ohlbach, 2004) sont généralement définis par leurs fonctions d'appartenance. Une fonction d'appartenance associe un ensemble de base

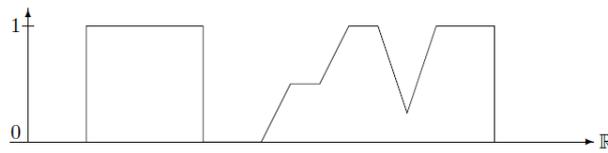


FIGURE 1.3: "Intervalles précis et flous".



FIGURE 1.4: "Après ce soir".

à un nombre réel compris entre 0 et 1. Cette "valeur floue" désigne une sorte de degré d'appartenance à un ensemble flou  $S$ . L'ensemble de base pour des intervalles de temps flous est l'axe temporel, représenté par l'ensemble  $\mathbb{R}$  des nombres réels. Les nombres réels nous permettent de modéliser les flux en temps continu que nous percevons dans notre vie.

Une fonction d'appartenance floue est une fonction totale  $f : \mathbb{R} \mapsto [0,1]$  qui n'a pas besoin d'être continue, mais il doit être intégrable. L'intervalle floue  $i_f$  qui correspond à une fonction d'appartenance floue  $f$  est

$$i_f \stackrel{\text{def}}{=} \{(x, y) \subseteq \mathbb{R} \times [0, 1] \mid y \leq f(x)\}.$$

Étant donné un intervalle flou  $i$  nous écrivons généralement  $i(x)$  pour indiquer la fonction d'appartenance correspondante.

Soit  $F_{\mathbb{R}}$  l'ensemble des intervalles de temps flous. La figure 1.3 présente la définition des intervalles précis (uniques ou multiples). Les intervalles flous peuvent également être infinis (voir figure 1.4). Par exemple, le terme "Après ce soir" peut être représentée par une valeur floue qui se soulève de valeur floue 0 à 18 pm jusqu'à la valeur floue 1 à 20 pm, puis reste 1 à l'infini. Les intervalles de temps flous peuvent être de structures assez complexes avec différentes caractéristiques. Les plus simples sont le *noyau* et le *support* ;

- Le *noyau*  $C(i)$  est la partie de l'intervalle  $i$  où la valeur floue est 1,
- Le *support*  $S(i)$  est le sous-ensemble de  $\mathbb{R}$  où la valeur floue est différent de zéro.

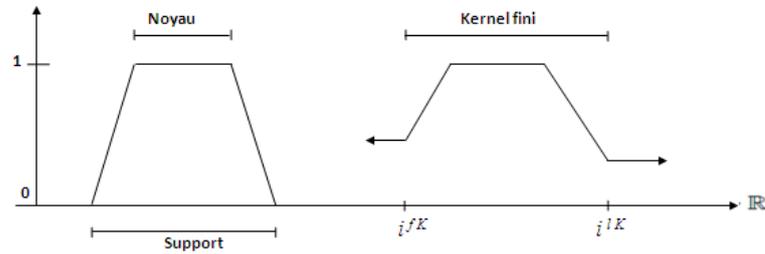


FIGURE 1.5: "Noyau, Support et Kernel".

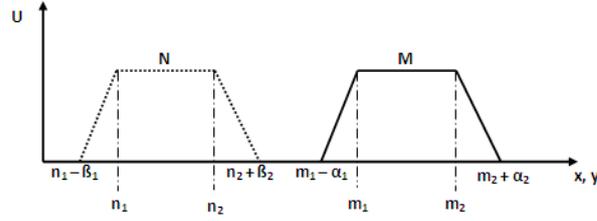
De plus, nous définissons le kernel comme la partie de l'intervalle où la valeur floue est non constante à l'infini. C'est-à-dire le kernel est le plus petit sous-ensemble de  $\mathbb{R}$  tel que  $i(x)$  est constante pour toutes  $x$  à l'extérieur de  $K(x)$ . Les intervalles de temps flous avec un kernel fini sont particulièrement intéressants, car bien qu'ils puissent être infinis, ils peuvent facilement être implémentés avec des structures de données finies.  $K(i)$  peut être vide, fini ou infini. Si  $K(i) = \emptyset$  alors  $i$  est non plus vide ou infini et précis (voir figure 1.5).

Les intervalles de temps flous qui sont, en fait, des intervalles précis peuvent maintenant être caractérisés, très facilement, comme des intervalles où le noyau et le support sont les mêmes. Ces intervalles sont appelés intervalles *précis flous*. Très souvent, nous venons de parler des intervalles précis, cependant, nous voulons parler des intervalles précis flous. Il devrait être clair à partir du contexte ce qui est vraiment destiné.

On peut transformer le noyau, le support et le kernel en intervalles précis flous :  $C^\square(i)$  est l'intervalle précis flou avec  $S(C^\square(i))=C(i)$ .  $S^\square(i)$  est l'intervalle précis flou avec  $S(S^\square(i))=S(i)$ .  $K^\square(i)$  est l'intervalle précis flou avec  $S(K^\square(i))=K(i)$ .

### 1.3.4 Indices de comparaisons

Soit deux intervalles de temps flous  $M$  et  $N$  exprimés sous formes de quadruplets de la forme  $(m_1, m_2, \alpha_1, \alpha_2)$  et  $(n_1, n_2, \beta_1, \beta_2)$  représentés respectivement par deux fonctions trapézoïdales. Fondamentalement, selon (Dubois & Prade, 1983b) il y

FIGURE 1.6: Certitude que  $M$  est strictement plus grand que  $N$ 

a quatre indices pour interpréter à quel point  $M$  est plus grand que  $N$ .

$$d(M > N) = \inf_{x,y} \{ \max(1 - \mu_M(x), 1 - \mu_N(y)) : x \leq y \} \quad (1.1)$$

$$d(M >^+ N) = \inf_x \sup_y \{ \max(1 - \mu_M(x), \mu_N(y)) : x \geq y \} \quad (1.2)$$

$$d(M >^- N) = \sup_x \inf_y \{ \max(\mu_M(x), 1 - \mu_N(y)) : x < y \} \quad (1.3)$$

$$d(M \geq N) = \sup_{x,y} \{ \min(\mu_M(x), \mu_N(y)) : x \geq y \} \quad (1.4)$$

L'équation (1.1) exprime la certitude que  $x$  est plus grand que  $N$ , sachant que  $x$  est  $M$ . La Figure 1.6 montre que cette équation est la relation la plus forte entre deux intervalles ; en terme de variables, elle signifie que  $M$  est nécessairement plus grand que  $N$ . Elle peut être exprimée en fonction d'un degré de nécessité de la proposition que  $M$  est strictement supérieur à  $N$  comme suit ;

$$\mathfrak{N}_M([N, +\infty)) = 1 - \sup_{x \leq y} \min(\mu_M(x), \mu_N(y)) \quad (1.5)$$

L'équation (1.5) se réfère au degré d'inclusion de l'ensemble flou  $M$  dans  $]N, +\infty)$ . Sachant que ;

$$\mathfrak{N}_M([N, +\infty)) = \text{Ness}(x > N \mid x \text{ est } M)$$

L'équation (1.2) exprime la certitude que  $x$  est plus grand où égal que  $N$ , sachant que  $x$  est  $M$ . Cette relation est relative au plus petite et au plus grande valeur possible de  $M$  et  $N$  respectivement (voir Figure 1.7). Elle peut être exprimée en fonction d'un degré de nécessité de la proposition que  $M$  est supérieur ou égal à  $N$  comme suit ;

$$\mathfrak{N}_M([N, +\infty)) = \inf_x \sup_{y: y \leq x} \max(1 - \mu_M(x), \mu_N(y)) \quad (1.6)$$

L'équation (1.6) se réfère au degré d'inclusion de l'ensemble flou  $M$  dans  $[N, +\infty)$ . Sachant que ;

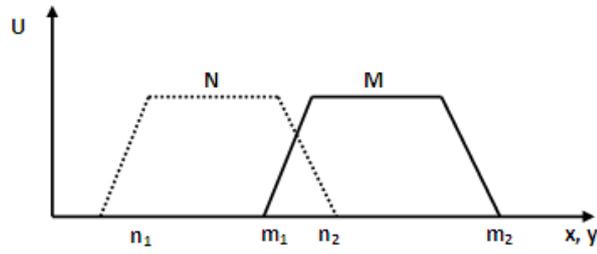


FIGURE 1.7: Certitude que M est plus grand ou égal que N

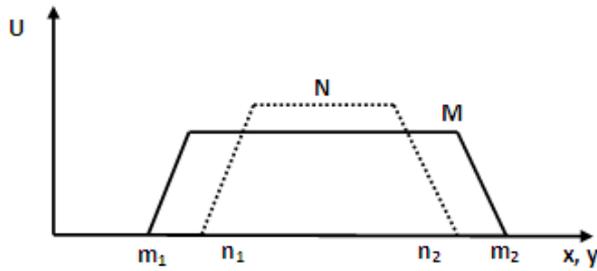


FIGURE 1.8: Possibilité que M est plus grand que N

$$\mathfrak{N}_M([N, +\infty)) = \text{Ness}(x \geq N \mid x \text{ est } M)$$

L'équation (1.3) exprime la possibilité que  $x$  est plus grand que  $N$ , sachant que  $x$  est  $M$ . Cette relation est relative au plus petite et au plus grande valeur possible de  $M$  et  $N$  respectivement (voir Figure 1.8). Elle peut être exprimée en fonction d'un degré de possibilité de la proposition que  $M$  est strictement supérieur à  $N$  comme suit ;

$$\prod_M([N, +\infty)) = \sup_x \inf_{y: y \geq x} \min(\mu_M(x), 1 - \mu_N(y)) \quad (1.7)$$

L'équation (1.7) se réfère au degré de correspondance partielle de l'ensemble flou  $M \cap ]N, +\infty)$  des nombres strictement supérieur à  $N$ , étant donné qu'ils sont limités par  $M$ . Sachant que ;

$$\prod_M([N, +\infty)) = \text{Poss}(x > N \mid x \text{ est } M)$$

L'équation (1.4) exprime la possibilité que  $x$  est plus grand ou égal à  $N$ , sachant que  $x$  est  $M$ . La Figure 1.9 montre que cette relation est relative au plus petite et

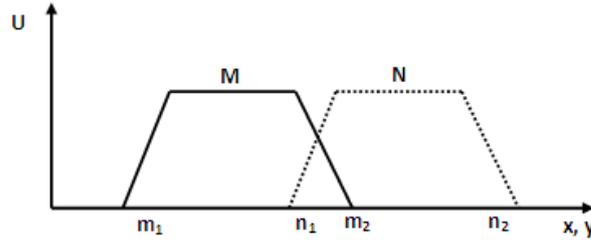


FIGURE 1.9: Possibilité que M est plus grand où égal que N

au plus grande valeur possible de  $M$  et  $N$  respectivement. Elle peut être exprimée en fonction d'un degré de possibilité de la proposition que  $M$  est supérieur ou égal à  $N$  comme suit ;

$$\prod_M([N, +\infty)) = \sup_{x,y: x \geq y} \min(\mu_M(x), \mu_N(y)) \quad (1.8)$$

L'équation (1.8) se réfère au degré de correspondance partielle de l'ensemble flou  $M \cap [N, +\infty)$  des nombres supérieur ou égal à  $N$ , étant donné qu'ils sont limités par  $M$ . Sachant que ;

$$\prod_M([N, +\infty)) = \text{Poss}(x \geq N \mid x \text{ est } M)$$

D'une façon similaire, nous pouvons définir les équations de comparaison entre deux intervalles précis. Soient  $M = [m_0, m_1]$  et  $N = [n_0, n_1]$ , deux intervalle fermés restreignant les variables  $X$  et  $Y$  respectivement. Il y a quatres façons d'interpréter le fait que  $M$  soit plus grand que  $N$  :

$$M > N \text{ si et seulement si } \forall x \in M, \forall y \in N, x > y, \text{ c'est-à-dire } m_0 > n_1. \quad (1.9)$$

$$M >^+ N \text{ si et seulement si } \forall x \in M, \exists y \in N, x > y, \text{ c'est-à-dire } m_0 > n_0. \quad (1.10)$$

$$M >^- N \text{ si et seulement si } \forall y \in N, \forall x \in M, x > y, \text{ c'est-à-dire } m_1 > n_1. \quad (1.11)$$

$$M \geq N \equiv \neg(N > M) \text{ si et seulement si } \exists (x, y) \in M \times N, x \geq y, \text{ soit } m_1 \geq n_0. \quad (1.12)$$

L'équation (1.9) est la plus forte entre deux intervalles ; en termes de variables, elle signifie que  $X$  est nécessairement plus grand que  $Y$ , tandis que l'équation

(1.12) est la plus faible, exprimant seulement que  $X \geq Y$  est possible. (1.10) et (1.11) sont des énoncés intermédiaires qui sont relatifs aux plus petites et aux plus grandes valeurs possibles de  $X$  et  $Y$  respectivement. En effet, (1.9) implique à la fois (1.10) et (1.11); et chacun de ces trois énoncés implique (1.12).

### 1.3.5 Différents types de prédicats flous

Comme il a été dit précédemment, les ensembles flous visent à étendre les ensembles ordinaires en permettant notamment d'exprimer des classes ou ensembles aux frontières mal définies. La transition entre l'appartenance complète et la non-appartenance est graduelle et non plus brutale. De même qu'il existe un lien fort entre ensemble et prédicat booléen, ensembles flous et prédicats flous sont étroitement associés. Un prédicat flou  $F$  s'applique à un ensemble (flou ou non) et prend une valeur de vérité dans l'intervalle unité  $[0, 1]$  traduisant l'adéquation vis-à-vis du concept  $F$ . Une requête faisant intervenir des conditions vagues représentées par des ensembles flous peut être définie à l'aide de différents constructeurs qui sont présentés brièvement ci-après.

Un prédicat atomique est défini par un ensemble flou sur un (ou plusieurs) domaine(s)  $D_i$ , i.e., par une fonction d'appartenance de  $D_1 \times \dots \times D_n$  dans  $[0, 1]$ ; des exemples de tels prédicats sont "grand", "jeune", "important", etc. En pratique, des fonctions de forme triangulaire ou trapézoïdale sont souvent utilisées. Il est clair que la représentation choisie dépend à la fois du contexte ("grand" ne signifie pas la même chose suivant qu'il s'agit d'une personne ou d'une maison) et de la subjectivité de l'utilisateur.

Un modificateur est représenté par une fonction de  $[0, 1]$  dans  $[0, 1]$  et s'applique à la fonction d'appartenance d'un ensemble flou de façon à modéliser un adverbe du langage naturel tel que "très", "plus ou moins", "relativement", etc. Les fonctions de modification les plus utilisées sont de la forme :

- $\mu_{modC}(x) = (\mu_C(x))^n$  avec  $n > 1$  pour un concentrateur et  $n < 1$  pour un dilatateur ;
- $\mu_{modC}(x) = op^n(\mu_C(x))$  où  $op$  est une norme (resp. co-norme) *non idempotente* pour un concentrateur (resp. dilatateur) ;
- $\mu_{modC}(x) = \mu_C(x \pm a)$  qui effectue une translation telle que suggérée dans (Meunier & Yao, 1992); ce mécanisme suppose disponible un ensemble de termes flous de type croissant-décroissant constituant une partition floue

du domaine concerné.

Ces approches offrent une vision *compositionnelle* de la définition du terme flou "mod C", connaissant les fonctions caractéristiques associées au prédicat flou C et au modificateur. Il reste toujours possible de définir *directement* la fonction voulue pour "mod C" si aucune composition standard ne convient. Le choix d'une interprétation est évidemment fonction de la sémantique désirée pour l'adverbe dans un contexte particulier, et donc des propriétés souhaitées (modification ou non du support et/ou du noyau du prédicat, etc.).

**Exemple :** Soit l'ensemble flou "jeune" défini comme suit :

$$\begin{aligned}\mu_{jeune}(x) &= 1 && \text{si } x \leq 25 \\ \mu_{jeune}(x) &= -x/20 + 9/4 && \text{si } 25 < x < 45 \\ \mu_{jeune}(x) &= 0 && \text{si } x \geq 45\end{aligned}$$

Avec le modificateur "très" défini par  $\mu_{très P}(x) = (\mu_P(x))^2$ , on obtient :

$$\mu_{très\ jeune}(30) = 0.56.$$

Avec le modificateur "très" défini par  $\mu_{très P}(x) = \top_L(\mu_P(x), \mu_P(x))$  où  $\top_L(x, y) = \max(x + y - 1, 0)$  est la norme de Lukasiewicz, on obtient cette fois :

$$\mu_{très\ jeune}(30) = 0.5.$$

Dans les deux cas, le résultat trouvé traduit un effet de renforcement (concentration) puisque  $\mu_{jeune}(30)$  vaut 0.75.

La négation se traduit habituellement, quant à elle, par le calcul du complément à 1 du degré. Conformément à l'intuition, l'élément x satisfait d'autant plus le prédicat "non P" que x satisfait peu P.

Enfin, des prédicats peuvent être combinés dans le cadre de propositions quantifiées, comme dans la requête visant à trouver les employés satisfaisant presque toutes les conditions :

- être bien payé ;
- être moyennement âgé ;
- exercer une activité spécialisée ;
- travailler dans une grande compagnie ;

généralisant le cas des quantificateurs existentiel et universel. Ce type d'usage a été initialement suggéré par (Kacprzyk & Ziolkowski, 1986) et permet d'obtenir

un résultat graduel fonction du quantificateur linguistique employé et de la valeur de vérité de chacun des prédicats impliqués. Plusieurs méthodes ont été proposées pour déterminer la valeur de vérité de telles expressions.

### 1.3.6 Aperçu du langage SQLf

Si le point de vue algébrique est intéressant pour définir un cadre formel d'interrogation, le langage *SQL* est le langage d'interrogation utilisé dans la pratique. C'est la raison pour laquelle différents auteurs se sont attachés à en proposer une extension permettant de formuler des requêtes graduées. Le langage appelé *SQLf*, décrit dans (Bosc & Pivert, 1995), peut être vu comme une synthèse et une extension de propositions plus anciennes parmi lesquelles (Bosc, Galibourg, & Hamon, 1988), (Kacprzyk & Ziolkowski, 1986), (Tahani, 1977). Le principe sous-tendant la définition de ce langage réside dans l'introduction de la gradualité partout où elle présente un intérêt. Cette section vise à donner quelques points-clés du langage *SQLf* dont on trouvera une description plus complète dans (Bosc & Pivert, 1995).

La structure en trois clauses : "**select**", "**from**" et "**where**" du bloc de base *SQL* est conservée en *SQLf*. La clause "**from**" ne subit aucune modification et les différences concernent essentiellement deux points : i) *le calibrage du résultat* qui peut se traduire par un nombre de réponses désirées (noté  $n$ ) ou un seuil qualitatif (noté  $t$ ) agissant comme une  $\alpha$ -coupe sur la relation résultante, ou les deux, et ii) *la nature des conditions autorisées*.

En conséquence, en *SQLf*, la formulation du bloc de base est l'expression ci-après :

**select** [**distinct**] [ $n$  |  $t$  |  $n, t$ ] attributs **from** relations **where** cond-floue

où "cond-floue" peut contenir des conditions booléennes et graduées. Cette expression s'interprète comme la restriction (ici floue) du produit cartésien des relations impliquées, suivie d'une projection (aux doubles près qui sont éliminés avec conservation du degré maximum si la clause "**distinct**" est spécifiée) sur les attributs de la clause "**select**", puis du calibrage des réponses ( $n$  meilleurs éléments et/ou ceux dépassant le seuil  $t$ ).

Dans la clause "**where**", peuvent apparaître les divers types de constituants d'une condition graduelle, ainsi que des conditions de jointure du type " $r.A \theta$

TABLE 1.4: Relation dept

d-num	d-nom	budget (k€)	lieu
1	Réseaux	5000 (1)	Lyon
3	Composants	1500 (0.65)	Lyon
5	Bureautique	400 (0.1)	Paris

s.B” où A et B sont des attributs respectifs des relations r et s, et  $\theta$  est un opérateur de comparaison graduelle ou non. Les prédicats et opérateurs paramétrables intervenant dans une requête sont supposés définis au moyen d’une interface appropriée.

**Exemple** : Soit les relations emp de schéma EMP(emp, e-nom, salaire, emploi, âge, ville, dep) et dept de schéma DEP(dep, d-nom, budget, lieu) où les attributs ont leur sens canonique.

**Q1** : ”trouver les employés (numéros et noms) satisfaisant au delà du degré 0.6 la condition : être jeune et travailler dans un département à gros budget”

s’écrit :

**select** 0.6 emp, e-nom

**from** emp E, dept D

**where** E.dep = D.dep and E.âge = ”jeune” and D.budget = ”gros”.

Supposons que la fonction d’appartenance du prédicat flou ”jeune” prenne la valeur 1 pour tout âge inférieur ou égal à 25 et décroisse linéairement jusqu’à valoir 0 pour un âge de 45 ans. De même supposons que la fonction d’appartenance du prédicat ”gros” passe de 0 à 1 entre 200 k€ et 2.2 M€. Considérons les extensions suivantes des relations emp et dept où les degrés obtenus pour chaque prédicat flou apparaissent entre parenthèses :

la table 1.6 présente le résultat de la requête précédente, puisque seuls ces deux couples se voient associer un degré supérieur au seuil exigé (0.6).

Illustrons maintenant le cas d’une condition de sélection complexe où l’opérateur de combinaison n’est plus une conjonction mais un *quantificateur flou*.

**Q2** : ”trouver les 5 meilleurs employés satisfaisant la plupart des conditions : vivre à Rennes, être bien payé, avoir un emploi hautement qualifié, être plutôt jeune et travailler dans un département à gros budget”.

Celle-ci s’écrit :

TABLE 1.5: Relation emp

e-num	e-nom	salaire	emploi	âge	ville	d-num
17	Dupont	3500	Ingénieur	51 (0)	Lyon	3
76	Martin	3000	Ingénieur	40 (0.25)	Paris	5
26	Durant	2000	Secrétaire	24 (1)	Lyon	3
12	Dubois	2500	Technicien	39 (0.3)	Lyon	3
55	Lorant	3500	Comptable	30 (0.75)	Lyon	1

TABLE 1.6: Résultat de la requête Q1

e-num	e-nom	degré
26	Durant	0.65
55	Lorant	0.75

**select** 5 e-num, e-nom

**from** EMP E, DEP D

**where** E.d-num = D.d-num and la-plupart (E.ville = "Rennes", E.salaire = "bien payé", E.emploi = "hautement qualifié", E.âge = "plutôt jeune", D.budget = "gros");

## 1.4 Conclusion

L'objectif principal du présent chapitre a été l'introduction des principaux concepts liés à notre travail de recherche, à savoir les formalismes utilisés dans le traitement des informations temporelles et la théorie des ensembles flous qui constitue la base de la manipulation des relations floues.

Les informations temporelles sont présentées généralement par deux manières (à savoir, l'algèbre des points et l'algèbre des intervalles). Pour exprimer la situation faisant intervenir deux points temporelles, trois relations ordinaires sont utilisées (à savoir, <, =, >). Les relations d'Allen (Allen, 1983) sont utilisées pour identifier la position d'un intervalle de temps par rapport à un autre. Nous

avons également décrit les différentes étapes et techniques usuellement utilisées dans le processus de comparaison des entités temporelles floues. Pour ce faire, nous avons commencé par donner un aperçu générale sur la théorie des ensembles flous. Puis, nous avons présenté les comparateurs flous exprimés en termes de différence de valeurs et la composition des relations floues modélisant des égalités approchées, ou des inégalités graduées.

Une phase de raisonnement sur les relations temporelles d'Allen basée sur les règles d'inférence contenant ces relations floues paramétrisées est ensuite discutée. Par la suite, nous avons rappelé la définition des intervalles flous et les indices de comparaisons de ces intervalles.



## Travaux Connexes

### 2.1 Introduction

La représentation des informations temporelles est utilisée dans plusieurs applications du monde réel (telles que la médecine, l'histoire, les applications criminelles et financières, etc.). L'information temporelle est souvent perçue ou exprimée d'une manière vague et imprécise.

Plusieurs études ont été proposées par (Freksa, 1992), (Ligozat, 1991), (Khatib, 1994), (Dubois, Hadjali, & Prade, 2007), (Dubois, Hadjali, & Prade, 2003), (Dubois & Prade, 1989), (Dubois, Lang, & Prade, 1991), (Qian & Lu, 1994), (Barro, Marin, Mira, & Paton, 1994), (Vila & Godo, 1994), (Godo & Vila, 1995), (Ortega-Binderberger, Chakrabarti, & Merhotra, 2002), ... pour la gestion de l'incertitude et de l'imprécision des informations temporelles dans le domaine d'Intelligence Artificielle. Dans le contexte des bases de données, il n'existe pas beaucoup des travaux pour traiter les données temporelles imparfaites à l'exception de (Billiet, Pons, Matthé, Tré, & Capote, 2011), (Tudorie, Vlase, Nica, & Muntranu, 2012), (Billiet, Pons, Capote, & Tré, 2012),... qui ont abordé cette question.

Dans ce chapitre, un état de l'art sur la gestion des informations imparfaites dans les bases de données temporelles est présenté (pour situer le contexte du travail introduit dans ce manuscrit). Dans un premier temps, nous allons introduire quelques travaux qui ont abordé la question de la gestion et de traitement de l'information temporelle imparfaite en Intelligence Artificielle. Puis, nous

donnons une revue critique des travaux proposés dans le contexte des bases de données.

## 2.2 Travaux en Intelligence Artificielle

Plusieurs généralisations de l'approche d'Allen ont été développées dans la littérature. Citons par exemple l'étude de Freska (Freksa, 1992) qui a développé une généralisation, basée sur les semi-intervalles, de l'approche d'Allen permettant de traiter des informations grossières et incomplètes. Voir aussi les travaux de Ligozat (Ligozat, 1991) et ceux de Khatib (Khatib, 1994) qui ont adapté l'approche d'Allen à des objets temporels généralisant les intervalles.

L'utilisation des ensembles flous (et la théorie des possibilités) en raisonnement temporel a été motivée par trois besoins essentiels (Dubois et al., 2007), (Dubois, Hadjali, & Prade, 2003) :

- La représentation des connaissances temporelles quantitatives entachées d'imprécision et d'incertitude.
- Le traitement de la flexibilité dans les approches (fondées sur les contraintes) pour le raisonnement temporel.
- La modélisation de la description graduelle et linguistique des relations temporelles.

### **Représentation des connaissances temporelles quantitatives et imparfaites ;**

L'information temporelle est souvent perçue ou exprimée d'une manière floue. Par exemple, "l'anniversaire de Lisa débutera en début d'après midi" ou "le trajet pour l'aéroport prend à peu près 30 minutes". (Dubois & Prade, 1989) ont proposé une approche, fondée sur la théorie des possibilités, pour représenter l'imprécision ou l'incertitude des connaissances temporelles quantitatives. Voir aussi (Dubois et al., 1991) pour un cadre logique et unifié permettant de traiter simultanément l'incertitude et le temps.

### **Modélisation de la flexibilité dans les contraintes temporelles ;**

Les ensembles flous qui jouent un rôle clé dans la modélisation des contraintes flexibles, ont été aussi utilisés dans différentes approches basées sur les contraintes pour le raisonnement temporel. Voir (Qian & Lu, 1994) (Barro et al., 1994) (Vila & Godo, 1994) (Godo & Vila, 1995) (Ortega-Binderberger et al., 2002) (Dubois, Hadjali, & Prade, 2003) (Aigner, Miksch, Thurnher, & Biffel, 2005) (Asmussen,

Qiang, DeMaeyer, & DeWeghe, 2009) (Tré, de Weghe, de Caluwe, & Maeyer, 2006).

**Introduction de la gradualité dans la description des relations temporelles symboliques** ; Les informations temporelles sont souvent de nature graduelle et exprimées d'une manière linguistique et qualitative (l'action  $A$  est bien avant l'action  $B$ ). La théorie des ensembles flous offre des outils appropriés pour modéliser tous ces aspects. Dans ce contexte, la plupart des travaux ont conduit à des extensions de l'approche d'Allen pour la représentation des relations temporelles entre intervalles. Des relations d'Allen étendues (ou floues) ont été proposées dans les cas des intervalles temporels classiques (non flous) et des intervalles temporels flous.

(Guesgen, Hertzberg, & Philpott, 1994) ont proposé des relations d'Allen floues vues comme des ensembles flous de relations d'Allen ordinaires en prenant en compte la notion de voisinage conceptuel introduite par (Freksa, 1992). Par exemple, la relation d'Allen floue  $\tilde{r}$  d'une relation d'Allen  $r$  est définie par  $\tilde{r} = \{(r', \mu_{\tilde{r}}(r')) \mid r' \in A_R\}$  où  $A_R = \{<, m, o, fi, di, si, \equiv, s, d, f, oi, mi, >\}$  et  $\mu_{\tilde{r}}(r) = 1$ . Pour définir  $\mu_{\tilde{r}}$ , six niveaux d'appartenance  $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\}$  compatibles avec le voisinage conceptuel de Freksa sont requis et vérifient la contrainte  $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \alpha_4 \geq \alpha_5 \geq \alpha_6 \geq 0$ . Le cas particulier où  $\alpha_i = 0$ , pour  $i = 1 \dots 6$ , correspond à une relation classique d'Allen. La composition des relations floues a été définie en utilisant le principe de min/max combinaison et l'algorithme d'Allen a été étendu à ces relations.

(Badaloni & Giacomini, 2000, 2002, 2006) ont développé une extension floue de l'algèbre des intervalles, appelée  $IA^{fuz}$ , où des degrés de préférence (appartenant à  $[0, 1]$ ) sont attachés à chaque relation atomique entre deux intervalles temporels classiques. Ce type d'approche permet, par exemple, de représenter des contraintes temporelles de la forme "l'activité  $A_1$  doit être disjointe de l'activité  $A_2$ , et il est préférable que  $A_1$  soit avant  $A_2$ ". Dans ce cadre, une relation  $R$  entre deux intervalles  $I$  et  $J$  s'écrit  $I R J$  avec  $R = \{r_{\alpha_1}^1, \dots, r_{\alpha_{13}}^{13}\}$  et  $r^p \in A_R$  et  $\alpha_p \in [0, 1]$  pour  $p = 1 \dots 13$ . Par exemple,  $I(<_{0.3}, m_{0.7})J$  exprime que la relation  $<$  (respectivement  $m$ ) entre les intervalles  $I$  et  $J$  est préférée avec un degré 0.3 (respectivement 0.7) ; toutes les autres relations ont un degré de préférence égale à 0. En s'appuyant sur des résultats obtenus dans le cadre des FCSP<sup>1</sup> (Dubois, Fargier, & Prade, 1996), Badaloni et Giacomini ont étendu la plupart des algorithmes de raisonnement proposés dans le contexte de l'algèbre  $AI$  (comme l'algorithme de vérification de la consistance d'un réseau temporelles). On peut

1. FSCP pour Fuzzy Constraint Satisfaction Problem.

voir cette approche comme une généralisation de l'approche de (Guesgen et al., 1994) où le nombre des niveaux de préférences utilisés est égal à 6. Voir aussi l'extension de cette approche (aux contraintes temporelles quantitatives et floues) proposée dans (Badaloni, Fade, & Giacomini, 2004). (Qian & Lu, 1994) ont étudié plusieurs stratégies de propagation pour le traitement de réseaux de règles temporelles floues ; (Barro et al., 1994) ont proposé une généralisation, basée sur les ensembles flous, de la notion de contrainte temporelle et utilisé des mesures de possibilité pour vérifier le degré de cohérence d'un réseau de contraintes temporelles floues (voir aussi (Vila & Godo, 1994) et (Wainer & Sandri, 1998)) ; (Godo & Vila, 1995) ont défini une logique temporelle approchée sur la base de l'intégration des contraintes floues dans un langage logique. Le système d'inférence est basé sur des règles spécifiques traitant des contraintes floues, et de propositions ayant des degrés de certitude. Voir aussi les travaux récents de (Cárdenas, Marin, & Navarette, 2001). (Dubois et al., 1991) ont proposé une logique temporelle possibiliste où une formule logique est associée avec un ensemble flou de dates où la formule est plus ou moins certainement vraie. La plupart des travaux en raisonnement temporel adopte une représentation faiblement expressive de l'incertitude. En effet, l'incertitude temporelle est modélisée au moyen d'une disjonction de relations primitives possibles entre deux entités temporelles. Par exemple, si un agent n'est pas certain de la relation, parmi les relations ">" et "=", qui existe entre deux dates  $t_1$  et  $t_2$ , il exprime que  $t_1 \geq t_2$  (où  $\geq = \{>, =\}$ ). Plus la disjonction est large, plus la relation temporelle décrite est incertaine. L'inconvénient majeur de cette approche est qu'elle ne permet pas de quantifier l'incertitude inhérente à la connaissance temporelle disponible. Une seconde limitation non négligeable de cette approche survient quand un ordre est requis entre les relations de chaque disjonction. L'approche ne permet pas donc d'exprimer qu'une relation est plus plausible qu'une autre.

Très peu d'études ont été menées sur la représentation et la gestion de l'incertitude en raisonnement temporel dans le cas où les entités temporelles sont des dates. (Ryabov & Puuronen, 2001) ont proposé un modèle probabiliste permettant de traiter des relations incertaines entre dates. Ils définissent une relation incertaine entre deux dates  $a$  et  $b$  comme toute disjonction possible entre les trois relations temporelles primitives (i.e.,  $<$ ,  $=$ ,  $>$ ). Cette disjonction est une relation composée de type  $\leq = \{<, =\}$ ,  $\geq = \{>, =\}$ ,  $\neq = \{<, >\}$ , ou l'ignorance totale  $? = \{<, =, >\}$ . L'incertitude est modélisée par un vecteur  $P_{ab} = (e^<, e^=, e^>)_{a,b}$  où  $e^<_{a,b}$  (respectivement  $e^=_{a,b}$ ,  $e^>_{a,b}$ ) est la probabilité que  $a < b$ , i.e.,  $a$  précède  $b$ , (respectivement  $a = b$ ,  $a > b$ ), et la somme  $e^< + e^= + e^> = 1$ . Supposons que la relation temporelle  $r$  entre les deux dates  $a$  et  $b$  est "<" (précède) avec une

probabilité de 0.7, ou ”=” (*égal*) avec une probabilité de 0.3. Cela signifie que  $r$  est ”>” (*suit*) avec une probabilité de 0. Le vecteur d’incertitude associé à  $r$  est donc (0.7, 0.3, 0).

(Ryabov & Puuronen, 2001) ont proposé des opérations, qui supposent préserver la sémantique des probabilités, permettant de gérer la propagation de l’incertitude quand deux relations sont composées, ou quand deux informations temporelles portant sur des mêmes dates sont fusionnées. Le point problématique majeur de cette approche concerne la façon avec laquelle l’état de l’ignorance totale est traité<sup>2</sup>. Quand aucune information n’est connue sur la relation entre deux dates quelconque, Ryabov et Puuronen suggèrent l’utilisation des valeurs de probabilités du domaine lié à l’application considérée. Ces valeurs, dénotées par  $e_D^<$ ,  $e_D^=$  et  $e_D^>$ , représentent les probabilités des trois relations primitives entre deux dates dans cette situation. Cette proposition a du sens uniquement si une distribution de probabilité à priori est disponible. Cependant, en pratique, il est difficile de produire de telle probabilité à priori. Quant aux règles d’inférences probabilistes utilisées sont de nature *ad hoc* et posent de sérieux problèmes comme c’est démontré dans (Dubois et al., 2007). En particulier, l’espace probabiliste sous-jacent tous les calculs effectués n’est jamais clarifié et le mécanisme d’inférence appliqué n’est pas complet.

(Ryabov & Trudel, 2004) utilise le même modèle probabiliste pour représenter l’incertitude dans les relations entre deux intervalles temporels. Dans ce cas, une relation temporelle incertaine est modélisée comme une disjonction de relations d’Allen où chaque relation est associée à une probabilité. Ryabov et Trudel définissent ainsi des réseaux d’intervalles temporels probabilistes où les contraintes entre les variables (i.e., les intervalles) sont les relations incertaines introduites. Ils proposent aussi une adaptation de l’algorithme standard de la chemin-consistance dans le cadre de ces réseaux temporels incertains (plus de détails sont aussi donnés dans (K. Zhang & Trudel, 2006)). Notons que cette approche est quelque peu similaire à celle proposée par (Badaloni & Giacomini, 2000, 2002, 2006) où les degrés attachés aux relations temporelles sont interprétés d’une manière non-probabiliste. Un autre modèle de réseaux temporels probabilistes combinant les réseaux Bayésiens et l’algèbre d’intervalle d’Allen, a aussi été étudié dans (Santos & Young, 1999). Ce modèle a été

---

2. Ce qui n’est pas surprenant car l’approche se fonde sur le modèle probabiliste pour représenter l’incertitude et il est bien connu que ce modèle ne permet pas de modéliser l’ignorance (ou l’ignorance partielle) d’une manière naturelle et consistante. Rappelons aussi que ce modèle est très exigeant en termes de données car le codage probabiliste de la connaissance est purement numérique.

appliqué dans le contexte des relations de causalité dépendantes du temps où l'hypothèse des probabilités à priori subsiste. Voir aussi l'étude de (Petridis, Paliouras, & Perantonis, 2010) sur une autre extension probabiliste des relations d'Allen.

Toutes les approches ci-dessus considèrent les relations temporelles floues seulement entre les intervalles de temps réguliers (à savoir, leurs bornes inférieures et supérieures sont des instants précis). Alors que dans les applications du monde réel, les intervalles de temps sont souvent décrits par des bornes floues afin de mieux saisir l'imprécision inhérente aux informations temporelles. Dans ce contexte (par exemple, les bornes d'intervalles temporels sont mal définies), très peu de travaux existent dans la littérature. Dans le cadre d'une application utilisant des informations temporelles historiques, (Nagypal & Motik, 2003) ont défini un modèle temporel basé sur les ensembles flous. Ce modèle permet d'étendre les relations d'Allen aux intervalles temporels flous (*ITFs*). Un *ITF* désigne un intervalle de temps classique où les bornes sont définies d'une manière imprécise (par exemple, "la période allant de la fin des années 20 au début des années 30" est un *ITF* avec la sémantique suivante (1928, 1933, 2, 2)). Nagypal et Motik ont introduit un ensemble d'opérateurs auxiliaires sur les intervalles comme, par exemple, l'opérateur prenant un intervalle  $I$  et construit l'intervalle contenant tous les instants qui *précèdent* l'instant initial de  $I$ . Puis, des contreparties floues à ces opérateurs ont été définies sur les *ITFs*<sup>3</sup>.

Des relations d'Allen étendues aux *ITFs* ont été introduites en utilisant ces opérateurs flous. Notons que la composition de ces relations n'a pas été abordée par les auteurs. (Ohlbach, 2004) a aussi proposé une extension similaire des relations intervalles à intervalles dans le cadre des *ITFs*.

(Schockaert, Cock, & Kerre, 2005 ; Schockaert & Cock, 2008) ont proposé aussi une généralisation des relations d'Allen aux *ITFs*. Cette généralisation permet de traiter des relations classiques entre événements imprécis (comme, "*Roosevelt est décédé avant le début de la guerre froide*"), et aussi des relations de nature imprécise (comme, "*Roosevelt est décédé juste avant le début de la guerre froide*"). Le concept clé utilisé dans cette approche est celui des ordres flous sur les instants (comme, par exemple, l'ordre flou qui exprime à quel point un instant  $a$  est beaucoup plus petit que l'instant  $b$ ). Ces ordres représentés par des relations floues paramétrisées sont appliqués aux bornes graduelles des *ITFs* pour définir, par exemple, le degré  $bb^?(A, B)$  (exprimant à quel point le début défini d'une

3. Ces opérateurs auxiliaires flous conduisent à des ensembles flous exprimant les ensembles des instants *possiblement/nécessairement après/avant* une date  $a$  définie par la d.p.  $\pi_a$ . Notons que ces opérateurs ont déjà été discutés dans (Dubois & Prade, 1989).

manière imprécise d'un *ITF A* est *avant la fin* définie d'une manière imprécise d'un *ITF B*). La table de transitivité des relations temporelles introduites a également été discutée dans (Schockaert, Cock, & Kerre, 2008 ; Schockaert & Cock, 2009). La consistance d'une base de formules temporelles (correspondant à des relations temporelles floues entre *ITFs*) a aussi été étudiée dans (Schockaert & Cock, 2008) ainsi que le problème de déduction.

## 2.3 Travaux dans les Bases de données

Peu de travaux ont été proposés pour traiter les données imparfaites dans les bases de données. (Billiet et al., 2011) ont proposé une approche qui intègre les classifications bipolaires pour déterminer le degré de satisfaction des enregistrements en utilisant les deux préférences imprécises et éventuellement temporelles positives et négatives. En effet, les utilisateurs expriment leurs préférences avec deux énoncés positifs et négatifs : ils expriment ce qu'ils désirent avec des déclarations positives et ils expriment ce qu'ils veulent rejeter avec des déclarations négatives. Dans ce cas il est nécessaire d'utiliser l'interrogation bipolaire ("technique d'interrogation qui permet l'introduction des préférences de l'utilisateur à la fois positives et négatives dans une requête d'une base de données"). Dans certaines techniques existantes pour l'interrogation des bases de données temporelles, les enregistrements sont soit acceptés soit rejetés en tant que résultats d'une requête. Alors que l'approche présentée ici intégrera une technique de classement bipolaire qui détermine le degré de satisfaction des enregistrements de la préférence d'une requête temporelle. Malheureusement, cette approche est incapable de modéliser les relations temporelles complexes appliquées, par exemple, dans les bases de données temporelles historiques (par exemple, l'utilisateur peut demander une période de temps, mais rejeter une partie de cette période, en spécifiant la contrainte de temps valable dans la requête). (Tudorie et al., 2012) ont proposé un modèle flou des termes vagues temporels et leurs implications dans l'évaluation des requêtes. En effet, la théorie des ensembles flous offre un cadre approprié pour modéliser et gérer des expressions temporelles vagues et pour évaluer des requêtes vagues adressées à une base de données. L'inconvénient majeur de cette approche est qu'elle ne permet pas de modéliser un grand nombre de termes temporels (tels que : *juste après* et *bien avant*).

(Galindo & J.M.Medina, 2001) ont proposé une extension pour les comparateurs temporels flous et ils ont introduit la notion des dates dans les bases de données relationnelles (*BDR*) par l'ajout de deux attributs supplémentaires dont le type de données est un des types de temps prédéfinis : *VST* (Valid Start Time) et *VET* (Valid End Time). Ces attributs permettent de vérifier la validité d'une information *T* dans le monde réel seulement au cours de la période de temps  $[T.VST ; T.VET]$ . Malheureusement, les comparateurs temporels flous présentés par Galindo et Medina pour traiter les requêtes flexibles, ne peuvent pas supporter certaines requêtes sophistiquées qui nécessitent une étape de raisonnement avant le traitement (voir dans le chapitre 3, une revue détaillée et critique de cette approche).

(RobotJazi, Reformat, Pedrycz, & Musilek, 2015) ont proposé une interface basée sur un RDF (Resource Description Framework)<sup>4</sup> nommée LORI pour interroger des données RDF qui contiennent des informations temporelles et utilisent des structures de données non triviales. Cette interface permet aux utilisateurs d'interroger des données temporelles avec des structures complexes en utilisant des prédicats de haut niveau. Ces prédicats sont construits par des données expertes basées sur des prédicats personnalisés. Par exemple, le prédicat "*approx\_at\_instant\_before*" permet de déterminer un degré de *chevauchement* entre un événement et un instant qui survient approximativement un certain nombre de temps *avant*. Le prédicat "*approx\_at\_instant*" consiste à déterminer le degré de *chevauchement* temporel entre un événement et un instant approché dans le temps.

(Billiet & Tré, 2015) ont proposé une étude détaillée d'un ensemble d'approches scientifiques existantes. Ces approches traitent l'imperfection du temps dans les systèmes d'informations (SI). La présentation de cette étude est une tentative pour identifier les défis et les opportunités ouvertes concernant la recherche sur l'imperfection de temps dans les SI. Billiet et De Tré ont classé les travaux selon trois catégories à savoir ;

- Imperfection dans les unités de temps de base (voir Table 2.1),
- Imperfection dans les relations temporelles (voir Table 2.2) et
- Imperfection dans les unités de temps de base et les relations temporelles (voir Table 2.3).

Les tableaux, ci-dessous, présentent chacun le type d'imperfection tout en précisant les unités de temps utilisées, la nature de l'incertitude (d'où vient

---

4. RDF est un modèle de graphe destiné à décrire de façon formelle les ressources Web et leurs métadonnées, de façon à permettre le traitement automatique de telles descriptions. RDF est le langage de base du Web sémantique.

l'incertitude) et la théorie utilisée ainsi que les travaux effectués.

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Intervalles	(Qiang et al., 2010), (Tré et al., 2006) et (Qiang, 2012)	Incertain due à l'imprécision et l'ambiguïté des données	Théorie des ensembles approximatifs issue de la théorie des ensembles flous
Durées	(Dubois, Fargier, & Galvagnon, 2003) et (Godo & Vila, 1995)	Incertain due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Durées	(Kopetz & Kim, 1990) et (Guil, Bosch, Bailón, & Marín, 2004)	Non spécifiée	Théorie des ensembles flous
Intervalles et instants	(Billiet, Firas, Capote, & Tré, 2013), (Tré et al., 2012), (Garrido, Marín, & Pons, 2009), (Pons et al., 2012)	Incertain due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Intervalles et instants	(Trajcevski, 2003)	N'est pas explicite	Théorie des probabilités
Intervalles et instants	(Tré, Caluwe, & Cruysen, 1997)	Incertain due au manque de connaissance	Théorie des ensembles flous

TABLE 2.1: Imperfection dans les unités de temps de base

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Intervalles	(Billiet et al., 2012), (Billiet, Firas, & Tré, 2013), (Billiet, Pons, & Tré, 2013), (Bronselaer, Pons, Tré, & Pons, 2013), (Billiet & Tré, 2014a), (Pons, Billiet, , Capote, & Tré, 2012), (Frias, Billiet, , Pons, & Tré, 1991)	Incertain due au manque de connaissance	Théorie des possibilités
Intervalles	(Aigner et al., 2005), (Asmussen et al., 2009), (Freksa, 1992), (Qiang et al., 2009)	N'est pas explicite	Théorie des ensembles approximatifs
Intervalles	(Billiet et al., 2011) et (Nagypal & Motik, 2003)	Incertain due à la nature graduelle des données ou des informations inhérentes à l'imprécision et l'ambiguïté	Théorie des ensembles flous

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Intervalles et durées	(Long, 1996), (Tré et al., 2012), (Tré et al., 1997), (Tré et al., 2006), (Dubois, Fargier, & Galvagnon, 2003), (Garrido et al., 2009), (Guil et al., 2004), (Kopetz & Kim, 1990), (Nagypal & Motik, 2003), (Qiang et al., 2010) et (Godo & Vila, 1995)	Non spécifiée	Théorie des ensembles approximatifs
Éléments du domaine temporel	(Dekhtyar, Ozcan, Ross, & Subrahmanian, 2001), (Dekhtyar, Ross, & Subrahmanian, 2001), (H. Zhang, Diao, & Immerman, 2010)	Incertainité due à la variabilité	Théorie des probabilités
Éléments du domaine temporel	(Anselma, Terenziani, & Snodgrass, 2010)	N'est pas explicite	N'est pas modélisée
Éléments du domaine temporel	(Virant & Zimic, 96) et (H. Zhang et al., 2010)	Aucune interprétation explicite	Théorie des ensembles flous

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Intervalles du domaine temporel	(Billiet & Tré, 2014b)	Incertain due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Eléments et intervalles du domaine temporel	(Chountas & Petrounias, 2000), (Dyreson & Snodgrass, 1998), (O'Connor, Tu, & Musen, 2000), (Ryabov, 2000), (Ryabov, 2001), (Tossebro & Nygard, 2002)	Incertain due à la variabilité	Théorie des probabilités
Eléments et intervalles du domaine temporel	(Caluwe, Devos, Maesfranckx, Tré, & Cruysen, 1999)	Incertain due à un niveau de précision insuffisant dans la description des données	Théorie des ensembles flous
Eléments et intervalles du domaine temporel	(Kalczynski & Chou, 2005), (Dyreson & Snodgrass, 1998), (Ryabov, 2000) et (Ryabov, 2001)	Incertain due à la source qui n'est pas clairement spécifiée	Théorie des probabilités issue de la théorie des ensembles flous

## 2.4 Conclusion

Dans ce chapitre, nous avons présenté quelques approches portant sur la gestion et le traitement des informations imparfaites dans les bases de données temporelles. Ces approches sont utilisées dans la suite de notre travail et permettent l'expression et l'évaluation des requêtes temporelles graduelles. Cette

TABLE 2.2: Imperfection dans les relations temporelles

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Intervalles	(Mitra, Gerard, Srinivasan, & Hands, 1994)	Incertain due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Intervalles	(Ryabov & Trudel, 2004)	Incertain due au manque de connaissance	Théorie des probabilités
Intervalles	(Badaloni & Giacomin, 2000) et (Badaloni & Giacomin, 2006)	Incertain due au manque de connaissance	Théorie des possibilités

étude nous a permis de voir l'importance de la structuration des requêtes graduelles pour l'interrogation des bases de données pour, d'une part, retourner les meilleurs éléments qui répondent aux critères de sélections spécifiés par l'utilisateur et, d'autre part, attribuer des degrés de satisfactions aux conditions spécifiées pour chaque réponse.

Deux types d'approches pour la gestion de l'imperfection des informations temporelles ont été abordées : (i) les approches dans le domaine d'Intelligence Artificielle et, (ii) les approches dans le contexte des bases de données. La majorité des modèles proposés dans le domaine d'Intelligence Artificielle sont des généralisations de l'approche d'Allen (Allen, 1983). Ces modèles ont la capacité de gérer non seulement des requêtes à conditions booléennes mais aussi des requêtes exprimées au moyen de prédicats flous. Quant aux approches proposés dans le domaine des bases de données, peu de travaux ont été proposés pour traiter les données imparfaites. L'objectif principal de ces approches est de déterminer le degré de satisfaction des enregistrements en utilisant les conditions imprécises. L'inconvénient majeur de ces approches est l'incapacité de gérer des requêtes sophistiquées et de modéliser un grand nombre de termes temporels. Nous avons proposé, aussi, une étude d'un ensemble d'approches scientifiques

TABLE 2.3: Imperfection dans les unités de temps de base et les relations temporelles

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Instant	(Bordogna, Bucci, Carrara, Pagani, & Rampini, 2009)	Incertain due au manque de connaissance	Théorie des possibilités
Intervalles	(Schockaert & Cock, 2008), (Schockaert, Cock, & Kerre, 2006) et (Schockaert et al., 2008)	Incertain due au manque de connaissance	Théorie des ensembles flous
Intervalles	(Bassiri, Malek, Ale-sheikh, & Amirian, 2009)	Incertain due au manque de connaissance	Théorie des ensembles approximatifs
Instants et intervalles	(Dubois et al., 1991)	Incertain due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Instants et intervalles	(Ryabov & Terziyan, 2003)	Incertain due au variabilité	Théorie des probabilités
Instants et intervalles	(Ryabov & Terziyan, 2003)	Incertain due au manque de connaissance	Pas de modèle

existantes traitant l'imperfection du temps dans les systems d'informations.

Nous remarquons, suite à cette étude de l'état de l'art, que très peu des travaux pour la gestion de l'imperfection dans les bases de données temporelles ont été proposés dans la littérature. D'où l'idée de proposer des modèles pour améliorer la qualité des résultats dans le domaine des bases de données. Cependant, nous remarquons qu'il n'existe actuellement aucun outil capable de prendre en considération à la fois, les spécifications temporelles de l'utilisateur et la gestion des données temporelles dans les bases de données pour améliorer la qualité des résultats.

Pour contribuer à résoudre ces problèmes des approches proposées dans ce chapitre, nous proposons dans la deuxième partie de ce manuscrit, la prise en considération des requêtes utilisateur temporelles graduelles dans le cadre des bases de données. Pour cela, nous proposons, dans ce qui suit, des approches pour la gestion et le traitement des informations imparfaites dans les bases de données temporelles.

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Instants et Durées	(Palma, Juarez, Campos, & Marin, 2004), (Palma, Juarez, Campos, & Marin, 2006) et (Kahn & Gorry, 1977)	Incertitude due au manque de connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Éléments du domaine temporel	(Bosch, Torres, & Marín, 2002) et (Dubois & Prade, 1989)	Incertitude due au manque connaissance	Théorie des possibilités issue de la théorie des ensembles flous
Éléments du domaine temporel	(Dubois et al., 2007)	Incertitude due au manque connaissance	Théorie des possibilités
Éléments du domaine temporel	(Bordogna, Carrara, Paganini, Pepe, & Rampini, 2009)	Incertitude due au manque connaissance	Théorie des ensembles flous
Intervalles du domaine temporel	(Galindo & J.M.Medina, 2001)	Incertitude due au manque connaissance	Théorie des ensembles flous
Éléments et des intervalles du domaine temporel	(Dubois, Hadjali, & Prade, 2003)	Incertitude due au manque connaissance	Théorie des possibilités

Unités de temps	Travaux	Nature de l'incertitude	Modèle utilisé
Eléments du domaine temporel	(Bordogna, Carrara, et al., 2009)	Incertain due au manque de connaissance	Théorie des ensembles flous
Intervalles du domaine temporel	(Galindo & J.M.Medina, 2001)	Incertain due au manque de connaissance	Théorie des ensembles flous
Eléments et des intervalles du domaine temporel	(Dubois, Hadjali, & Prade, 2003)	Incertain due au manque de connaissance	Théorie des possibilités



## **Deuxième partie**

### **Contributions**



# Requêtes en présence des relations temporelles graduelles

## 3.1 Introduction

Les bases de données temporelles permettent de gérer les aspects temporels des données qu'ils contiennent. Généralement, les informations temporelles sont exprimées d'une manière imprécise et floue. Par exemple, les périodes des révolutions sont caractérisées par des débuts et des fins naturellement graduelles et mal-définis (à savoir, "*bien après début des années 20*" ou "*à la fin des années 30*"). Dans le contexte des bases de données, peu de travaux existent pour la gestion et le traitement des informations imprécises et floues dans les bases de données temporelles.

Comme discuté dans le chapitre 2, (Tudorie et al., 2012) ont proposé un modèle flou pour les termes temporels vagues et ses implications dans l'évaluation des requêtes. Aussi (Galindo & J.M.Medina, 2001) ont proposé une extension des comparateurs temporels flous et ils ont introduit la notion des dates dans les bases de données relationnelles en ajoutant deux attributs de temps supplémentaires (*VST*, *VET*), ces deux attributs sont précis. Deux points faibles, de l'approche de Galindo et Medina, ont été identifiés (i) les comparateurs proposés ne permettent pas de traiter toutes les relations temporelles floues et (ii) le modèle proposé ne peut pas supporter les requêtes sophistiquées qui nécessitent une phase de raisonnement, par exemple, "*Trouver les découvertes*

*archéologiques qui ont eu lieu juste après la deuxième guerre mondiale sachant que cette dernière a eu lieu bien à la fin des années 30”.*

Pour remédier aux limites déjà citées dans l’interrogation des bases de données temporelles en utilisant des requêtes flexibles (qui possèdent des critères temporels flous), nous proposons dans ce chapitre une approche intégrant à la fois une extension des relations temporelles d’Allen floues et une phase de raisonnement basée sur ces relations. Cette approche utilise aussi les notions de *VST* et *VET* introduites dans (Galindo & J.M.Medina, 2001).

Pour cela, nous discutons, dans un premier temps, l’extensions floues des relations temporelles. Ensuite, nous décrivons l’approche proposée par Galindo et Medina en 2001 en tenant en compte de l’extension des comparateurs flous.

Puis, nous présentons en détail les points faibles de cette approche avec des exemples.

Dans ce qui suit, la section 4.4 décrit notre approche (Gammoudi, Hadjali, & Ben Yaghlane, 2014) qui traite les intervalles de temps flous et nous donnons un aperçu du langage *TSQLf* avec l’architecture du système développé. Enfin, la section 4.6 conclut le chapitre.

## **3.2 Extensions floues des relations temporelles**

Les relations temporelles proposées dans le formalisme qualitatif d’Allen souffrent de deux insuffisances non négligeables, à savoir :

- Manque de robustesse : une légère variation sur les bornes d’intervalles peut changer brusquement leur relation. Cette situation est clairement illustrée par les trois relations d’Allen (voir figure 3.1). D’un point de vue perception humaine, ces relations peuvent être décrites par une seule caractérisation de type *”rencontre-nuancée”*.
- Manque de gradualité : le formalisme d’Allen ne peut pas distinguer entre les deux relations (voir figure 3.2) appartenant à une même classe. Il est facile de voir que dans le cas de la première situation une importante variation sur les bornes des deux intervalles peut ne pas changer leur relation. D’un point de vue perception humaine, la première situation correspond à *”A est bien avant B”* et la seconde à *”A est juste avant B”*.

C’est pour répondre à ce type de problèmes qu’une extension floue des relations d’Allen a été proposée dans (Dubois et al., 2004) et (Dubois, Hadjali, & Prade,



FIGURE 3.1: Exemple 1 des relations d'Allen.

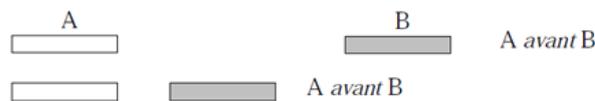


FIGURE 3.2: Exemple 2 des relations d'Allen.

2003) permettant une description plus graduelle des relations possibles entre des intervalles de temps. Plus de détail sur cette extension est discuté dans le premier chapitre (section 3).

Il convient de noter que les extensions floues des relations temporelles proposées dans la littérature ne permettent pas de prendre en compte les deux problèmes cités ci-dessus, excepté l'approche de Ohlbach (Ohlbach, 2004) et celle de Schockaert et al. (Schockaert et al., 2005), (Schockaert & Cock, 2008). Dans le cas des intervalles temporels flous, (i) le modèle de Ohlbach permet de représenter des relations temporelles imprécises de type "A plus ou moins termine B" en utilisant des mesures de chevauchement entre ensembles flous ; (ii) le modèle de Schockaert et al. permet d'exprimer des relations de type "A est juste/bien avant B" en utilisant des *ordres flous* entre instants. Ces ordres flous sont similaires aux *comparateurs flous* composant la partition floue évoquée ci-dessus.

### 3.3 Approche de Galindo et Medina

Galindo et Medina (Galindo & J.M.Medina, 2001) ont proposé une extension pour le langage *SQL* nommée *FSQL* qui permet de supporter des requêtes flexibles tout en prenant en compte les caractéristiques d'informations imprécises. Cette extension consiste à prolonger la commande de sélection " *SELECT* " pour pouvoir exprimer des requêtes flexibles .

Les principales extensions de cette commande sont :

- **Étiquettes linguistiques (*linguistic label*)** : chaque étiquette a une distribution de possibilité trapézoïdale de la forme  $(a, A, B, b)$ ...
- **Comparateurs flous** : en plus des comparateurs logiques ( $=, >, <$ ), *FSQL* inclut des comparateurs flous.

Par exemple :

— L'égalité floue :  $CDEG(A \text{ FEQ } B) = \sup_{u \in U} \{\min(A(u), B(u))\}$

— Le plus grand flou :  $CDEG(A \text{ FGT } B) =$

$$\begin{cases} 1 & \text{si } \gamma_A \geq \delta_B \\ \frac{\delta_A - \gamma_B}{(\delta_B - \gamma_B) - (\gamma_A - \delta_A)} & \text{si } \gamma_A < \delta_B \text{ y } \delta_A > \gamma_B \\ 0 & \text{sinon } (\delta_A \leq \gamma_B) \end{cases}$$

avec  $U$  désigne le domaine sous-jacent ou l'univers,  $A = [\alpha_A, \beta_A, \gamma_A, \delta_A]$ ,  $B = [\alpha_B, \beta_B, \gamma_B, \delta_B]$  et *CDEG* est une fonction utilisée dans la liste de sélection, qui affiche une colonne avec un degré de satisfaction d'une condition floue pour un attribut spécifique.

De la même façon, comme dans une commande *SQL*, les comparateurs flous peuvent comparer un attribut avec une constante ou deux attributs du même type.

- **Degré de satisfaction ( $\gamma$ )** : Pour chaque condition, un degré de satisfaction peut être établi avec le format  $\langle \text{condition} \rangle \text{ THOLD } \gamma$ , indiquant que la condition doit être satisfaite avec un degré minimum  $\gamma$ , avec  $\gamma \in [0, 1]$ .

Le mot réservé *THOLD* est optionnel et peut être substitué par un comparateur classique précis ( $=, \leq, \text{etc}$ ) en modifiant le sens de la requête.

Galindo et Medina ont également développé un outil nommé "*FSQL Server*" pour obtenir des réponses aux requêtes *FSQL*. Il est programmé en langage *PL/SQL* sous le SGBD Oracle. *FSQL Server* est programmé pour travailler sur les bases de données précises ou floues.

Deux types d'informations sont stockées dans la base de données :

- Base de données traditionnelles** : Elle se compose de toutes les données stockées dans les relations avec un format spécial pour représenter les valeurs des *attributs flous*. Les *attributs flous* sont classés par le système en 3 types :

- Type 1 : Attributs totalement précis (traditionnel), mais ils ont quelques étiquettes trapézoïdales linguistiques définies sur eux. Avec ces attributs

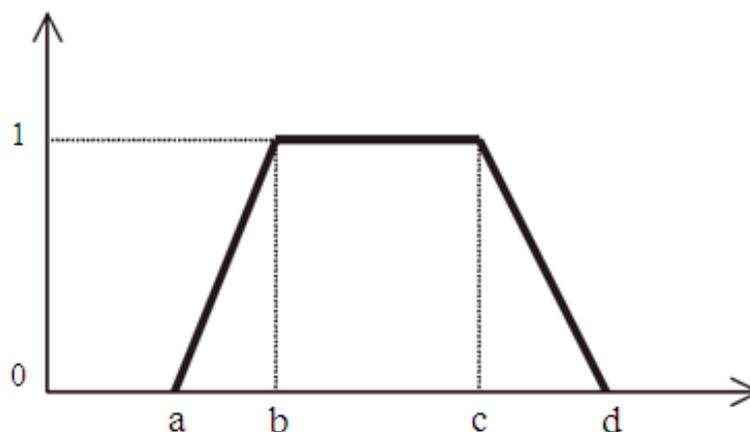


FIGURE 3.3: Trapèze flou

flous, nous pouvons utiliser toutes les constantes (tableau 3.1) dans la condition de la requête.

- Type 2 : Attributs qui admettent des données précises, ainsi que des distributions de possibilité sur un domaine sous-jacent ordonné ou universel « $U$ ». Avec ces attributs, nous pouvons stocker et utiliser toutes les constantes indiquées dans le tableau 3.1.
- Type 3 : Attributs qui sont définis sur un domaine sous-jacent non ordonné, par exemple la couleur des cheveux. Sur ces attributs certaines étiquêtes sont définies et sur ces étiquêtes une relation de similarité doit encore être définie. Avec ces attributs, nous pouvons utiliser uniquement le comparateur flou *FEQ*, comme ils n'ont aucune relation d'ordre.

**b) Méta-connaissances floues (FMB) :** elle stocke les données sur le *FRDB* (Base de données relationnelles floues) : attributs qui supportent les traitements flous, les étiquêtes linguistiques et la marge  $M$  (voir tableau 3.1 et 3.2), la relation de similarité (type 3)...

Pour intégrer les dates temporelles dans les *BDR* (Base de Données Relationnelles), Galindo et Medina ont ajouté deux attributs supplémentaires dont le type de données est un des types de temps prédéfinis : *VST* (*Valid Start Time*) et *VET* (*Valid End Time*). Ces attributs, pour un tuple  $t$ , expriment que ses informations sont valides dans le monde réel uniquement pendant la période de temps  $[t.VST, t.VET]$ . Donc, pour sélectionner toutes les versions de tuple  $t$  qui étaient valides sur certains points de temps  $T$  ou étaient valides pendant une certaine période de temps  $[T1, T2]$  (avec  $T$  et  $[T1, T2]$  peuvent être des valeurs

TABLE 3.1: Les valeurs floues qui peuvent être utilisées dans la base de données floues ou dans des requêtes *FSQL*.

Constante Floue	Signification
<i>UNKNOWN</i>	Valeur inconnue, mais l'attribut est applicable.
<i>UNDEFINED</i>	L'attribut n'est pas applicable ou il n'a pas de sens.
<i>NULL</i>	Ignorance totale : Nous ne savons rien sur l'attribut.
$[a, b, c, d]$	Trapèze flou $a \leq b \leq c \leq d$ (voir figure 3.3)
Étiquette	Étiquette linguistique (un trapèze ou un scalaire défini dans <i>FMB</i> ).
$[n, m]$	Intervalle "entre $n$ et $m$ " ( $a = b = n$ et $c = d = m$ ).
$n$	"Approximativement $n$ " (Triangle : $b = c = n$ et $-a = d - n = \text{marge}$ )

floues ou précises), il faut comparer chaque version de tuple  $t : [t.VST, t.VET]$  avec  $T$  ou  $[T1, T2]$ . Le tableau 3.3 présente l'extension des comparateurs flous.

Prenons l'exemple de (Galindo & J.M.Medina, 2001), "Récupérer toutes les tuples qui étaient valides (en tout point) pendant l'année 1994 (avec un degré minimum 0,6)". Galindo et Medina transforment cette requête à une autre

requête *FSQL* par deux étapes ;

- i) Transformer la description floue (pendant l'année 1994) en un intervalle bien précis ((01, 01, 1994), (31, 12, 1994)),
- ii) Choisir un comparateur flou parmi ceux définis dans la table 3.3 qui répond à la demande.

⇒ La condition de la requête sera alors :

$[t.VST, t.VET]$  F-chevauche [(01, 01, 1994), (31, 12, 1994)] **THOLD** 0.6

Rappelons que les comparateurs temporels flous présentés par Galindo et Medina ont des limites, et par conséquent ne peuvent pas traiter un certain type de

TABLE 3.2: Extension des comparateurs du TSQL2

Signification	Comparateurs de possibilité	Comparateurs de nécessité
Égalité floue	$CDEG(A FEQ B) = \sup_{u \in U} \{\min(A(u), B(u))\}$	$CDEG(A NFEQ B) = \inf_{u \in U} \max(1 - A(u), B(u))$
Plus grand flou (Fuzzy After)	$CDEG(A FGT B) =$ $\begin{cases} 1 & \text{si } \gamma_A \geq \delta_B \\ \frac{\delta_A - \gamma_B}{(\delta_B - \gamma_B) - (\gamma_A - \delta_A)} & \text{si } \gamma_A < \delta_B \text{ y } \delta_A > \gamma_B \\ 0 & \text{sinon } (\delta_A \leq \gamma_B) \end{cases}$	$CDEG(A NFGT B) =$ $\begin{cases} 1 & \text{si } \alpha_A \geq \delta_B \\ \frac{\beta_A - \gamma_B}{(\delta_B - \gamma_B) - (\alpha_A - \beta_A)} & \text{si } \alpha_A < \delta_B \text{ y } \beta_A > \gamma_B \\ 0 & \text{sinon } (\beta_A \leq \gamma_B) \end{cases}$
Plus grand ou égalité floue (Fuzzy After or at a Time)	$CDEG(A FGEQ B) =$ $\begin{cases} 1 & \text{si } \gamma_A \geq \beta_B \\ \frac{\delta_A - \alpha_B}{(\beta_B - \alpha_B) - (\gamma_A - \delta_A)} & \text{si } \gamma_A < \beta_B \text{ y } \delta_A > \alpha_B \\ 0 & \text{sinon } (\delta_A \leq \alpha_B) \end{cases}$	$CDEG(A NFGEQ B) =$ $\begin{cases} 1 & \text{si } \alpha_A \geq \beta_B \\ \frac{\beta_A - \alpha_B}{(\beta_B - \alpha_B) - (\alpha_A - \beta_A)} & \text{si } \alpha_A < \beta_B \text{ y } \beta_A > \alpha_B \\ 0 & \text{sinon } (\beta_A \leq \alpha_B) \end{cases}$
Plus petit flou (Fuzzy Before)	$CDEG(A FLT B) =$ $\begin{cases} 1 & \text{si } \beta_A \leq \alpha_B \\ \frac{\alpha_A - \beta_B}{(\alpha_B - \beta_B) - (\beta_A - \alpha_A)} & \text{si } \beta_A > \alpha_B \text{ y } \alpha_A < \beta_B \\ 0 & \text{sinon } (\alpha_A \geq \beta_B) \end{cases}$	$CDEG(A NFLT B) =$ $\begin{cases} 1 & \text{si } \delta_A \leq \alpha_B \\ \frac{\gamma_A - \beta_B}{(\alpha_B - \beta_B) - (\delta_A - \gamma_A)} & \text{si } \delta_A > \alpha_B \text{ y } \gamma_A < \beta_B \\ 0 & \text{sinon } (\gamma_A \geq \beta_B) \end{cases}$
Plus petit ou égalité floue (Fuzzy Before or at a Time)	$CDEG(A FLEQ B) =$ $\begin{cases} 1 & \text{si } \beta_A \leq \gamma_B \\ \frac{\delta_B - \alpha_A}{(\beta_A - \alpha_A) - (\gamma_B - \delta_B)} & \text{si } \beta_A > \gamma_B \text{ y } \alpha_A < \delta_B \\ 0 & \text{sinon } (\alpha_A \geq \delta_B) \end{cases}$	$CDEG(A NFLEQ B) =$ $\begin{cases} 1 & \text{si } \delta_A \leq \gamma_B \\ \frac{\gamma_A - \delta_B}{(\gamma_B - \delta_B) - (\delta_A - \gamma_A)} & \text{si } \delta_A > \gamma_B \text{ y } \gamma_A < \delta_B \\ 0 & \text{sinon } (\gamma_A \geq \delta_B) \end{cases}$
Beaucoup plus grand que (Much After)	$CDEG(A MGT B) =$ $\begin{cases} 1 & \text{si } \gamma_A \geq \delta_B + M \\ \frac{\gamma_B + M - \delta_A}{(\gamma_A - \delta_A) - (\delta_B - \gamma_B)} & \text{si } \gamma_A < \delta_B + M \text{ y } \delta_A > \gamma_B + M \\ 0 & \text{sinon } (\delta_A \leq \gamma_B + M) \end{cases}$	$CDEG(A NMGT B) =$ $\begin{cases} 1 & \text{si } \alpha_A \geq \delta_B + M \\ \frac{\gamma_B + M - \beta_A}{(\alpha_A - \beta_A) - (\delta_B - \gamma_B)} & \text{si } \alpha_A < \delta_B + M \text{ y } \beta_A > \gamma_B + M \\ 0 & \text{sinon } (\beta_A \leq \gamma_B + M) \end{cases}$
Beaucoup plus petit que (Much Before)	$CDEG(A MLT B) =$ $\begin{cases} 1 & \text{si } \beta_A \leq \alpha_B - M \\ \frac{\beta_B - M - \alpha_A}{(\beta_A - \alpha_A) - (\alpha_B - \beta_B)} & \text{si } \beta_A > \alpha_B - M \text{ y } \alpha_A < \beta_B - M \\ 0 & \text{sinon } (\alpha_A \geq \beta_B - M) \end{cases}$	$CDEG(A NMLT B) =$ $\begin{cases} 1 & \text{si } \delta_A \leq \alpha_B - M \\ \frac{\beta_B - M - \gamma_A}{(\delta_A - \gamma_A) - (\alpha_B - \beta_B)} & \text{si } \delta_A > \alpha_B - M \text{ y } \gamma_A < \beta_B - M \\ 0 & \text{sinon } (\gamma_A \geq \beta_B - M) \end{cases}$

TABLE 3.3: Les comparateurs flous

Expression	Equivalence
$[t.VST, t.VET] F\text{-INCLUDES } [T1, T2]$	$T1 FGEQ t.VST$ et $T2 FLEQ t.VET$
$[t.VST, t.VET] F\text{-INCLUDED-IN } [T1, T2]$	$T1 FLEQ t.VST$ et $T2 FGEQ t.VST$
$[t.VST, t.VET] F\text{-OVERLAPS } [T1, T2]$	$T1 FLEQ t.VET$ et $T2 FGEQ t.VST$
$[t.VST, t.VET] F\text{-BEFORE } [T1, T2]$	$T1 FGEQ t.VET$
$[t.VST, t.VET] F\text{-AFTER } [T1, T2]$	$T2 FLEQ t.VST$
$[t.VST, t.VET] F\text{-XBEFORE } [T1, T2]$	$T1 FGT t.VET$
$[t.VST, t.VET] F\text{-XAFTER } [T1, T2]$	$T2 FLT t.VST$
$[t.VST, t.VET] F\text{-MUCH-BEFORE } [T1, T2]$	$T1 MGT t.VET$
$[t.VST, t.VET] F\text{-MUCH-AFTER } [T1, T2]$	$T2 MLT t.VST$

requêtes. Par exemple ;

- ◇ Sélectionnez les tuples valides à la fin de la période  $[T1, T2]$  sans pour autant qu'ils aient dépassé la fin de la période  $[T3, T4]$  (nécessité d'une phase de raisonnement).
- ◇ Sélectionnez les tuples valides juste au début des deux dernières années ou courant les deux premiers mois de l'année en cours.

Par conséquent, l'approche de Galindo et Médina néglige la précision de l'égalité aux bornes des intervalles à comparer.

Nous pouvons dire donc que les relations temporelles proposées par Galindo et Medina présentent deux inconvénients non négligeables, à savoir :

- ▷ le manque d'exactitude au niveau des bornes (voir figure 3.4). D'après le tableau 3.3, on a  $[t.VST, t.VET] F\text{-INCLUDES } [T1, T2]$  est équivalente à  $T1 FGEQ t.VST$  et  $T2 FLEQ t.VET$  (avec  $FGEQ$  est une fonction qui vérifie si un intervalle de temps est suit ou égale à un autre intervalle de temps et  $FLEQ$  est une fonction qui vérifie si un intervalle de temps précède ou coïncide avec un autre intervalle de temps). D'après cette explication nous pouvons conclure clairement que l'approche de Galindo et Medina (Galindo & J.M.Medina, 2001) néglige la notion d'égalité entre deux intervalles, elle ne fait pas la différence entre "l'égalité" et le concept de "pendant" ou "de l'inclusion". D'autre part, Galindo et Medina ont défini la relation de chevauchement entre deux intervalles comme suit :

$[t.VST, t.VET] F\_OVERLAPS [T1, T2]$  est équivalente à  $T1 FLEQ t.VET$  et  $T2 FGEQ t.VST$ .

Cette définition couvre les relations  $t$  chevauche  $T$ ,  $t$  inclus  $T$ ,  $t$  commencé\_par  $T$ ,  $t$  chevauché\_par  $T$  et  $t$  rencontré\_par  $T$ .

⇒ D'où, l'imprécision des relations temporelles de Galindo et Medina.

- ▷ l'incapacité de raisonner sur les relations temporelles floues pour répondre à certaines requêtes complexes.

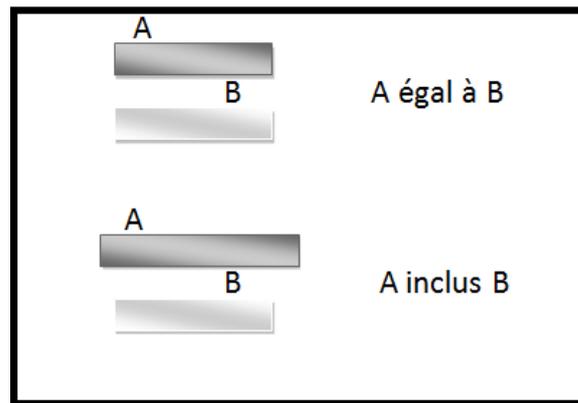


FIGURE 3.4: Inexactitude au niveau des bornes

## 3.4 Approche proposée

Dans cette section, nous présentons notre approche (Gammoudi et al., 2014) qui traite les intervalles de temps flous et nous donnons un aperçu sur le langage *TSQLf* et sur le processus de formulation et de traitement de requêtes *TSQLf*.

### 3.4.1 Modélisation

Notre approche consiste à améliorer la méthodologie de Galindo et Medina (Galindo & J.M.Medina, 2001) afin de combler les lacunes de cette dernière.

L'idée principale de notre approche vise à utiliser le principe de (Galindo & J.M.Medina, 2001) qui consiste à incorporer deux attributs supplémentaires (*VST* et *VET*) et de modéliser les relations temporelles d'Allen floues pour mieux répondre aux besoins des utilisateurs (voir tableau 3.4 avec  $t = [t.VST, t.VET]$ ,  $T = [T1, T2]$ ,  $L = (-\delta, \delta, \varepsilon, \varepsilon)$  et  $L_+^c = (\delta + \varepsilon, +\infty, \delta, +\infty)$ ).

TABLE 3.4: Relations temporelles d'Allen floues

Relation	Définition	Label
$t$ F-précède(L) T T F-suit(L) t	$T1 G(L_+^c) VET$	$fb(L)$ $fa(L)$
$t$ F-rencontre(L) T T F-rencontré_par(L) t	$VET E(L) T1$	$fm(L)$ $fmi(L)$
$t$ F-chevauche(L) T T F-chevauché_par(L) t	$T1 G(L_+^c) VST \wedge VET G(L_+^c) T1 \wedge T2 G(L_+^c) VET$	$fo(L)$ $foi(L)$
$t$ F-pendant(L) T T F-contient(L) t	$VST G(L_+^c) T1 \wedge T2 G(L_+^c) VET$	$fd(L)$ $fdi(L)$
$t$ F-commence(L) T T F-commencé_par(L) t	$VST E(L) T1 \wedge T2 G(L_+^c) VET$	$fs(L)$ $fsi(L)$
$t$ F-termine(L) T T F-terminé_par(L) t	$VET E(L) T2 \wedge VST G(L_+^c) T1$	$ff(L)$ $ffu(L)$
$t$ F-égal(L) T T F-égal(L) t	$VST E(L) T1 \wedge T2 E(L) VET$	$fe(L)$ $fe(L)$

### 3.4.2 Langage *TSQLf*

Nous proposons un langage, nommé *TSQLf*. Ce dernier est une extension du langage *SQLf* (pour plus de détail, voir le premier chapitre) en y ajoutant la dimension "temps".

L'évaluation d'une requête temporelle flexible exprimée en *TSQLf* passe par deux étapes à savoir :

- 1) La conversion de la requête temporelle graduelle à une requête *SELECT* standard basée sur des critères de recherche précis.
- 2) Le calcul du degré de satisfaction des critères pour chaque tuple candidat.

La structure des trois clauses "Select", "From" et "Where" du bloc de base *SQL* est préservée dans *TSQLf*. La clause "From" ne change pas et les différences sont principalement en deux points :

- i) Le calibrage du résultat qui présente le nombre de réponses désirées (noté par  $n$ ) ou un seuil qualitatif (noté par  $s$ ), ou les deux et,
- ii) la nature des conditions autorisées.

Par conséquent, la formulation du bloc de base d'une requête *TSQLf* est donnée par l'expression suivante :

**Select** [*distinct*] [ $n$  |  $s$  |  $n, s$ ] *attributs*.

**From** *relations*

**Where** *cond-temp-flou*

où *cond-temp-flou* peut contenir des conditions booléennes et des relations temporelles d'Allen floues (voir Table 3.4).

Dans la clause [*Where*] peut apparaître divers types des relations temporelles d'Allen floues ou d'une combinaison de ces dernières, ainsi que des conditions de jointure.

### 3.4.3 Processus de traitement de requêtes *TSQLf*

Dans cette section, nous présentons l'architecture de notre système développé pour traiter les requêtes *TSQLf*. L'architecture présentée dans la figure 3.5 décrit les différents modules nécessaires pour le traitement d'une requête *TSQLf*. Dans

le module *Interface*, l'utilisateur saisit une requête temporelle graduelle en utilisant une interface graphique, celle-ci lui donne en premier lieu la possibilité de choisir les attributs, les tables et d'introduire des conditions temporelles floues. Puis, le système sollicite l'utilisateur de déterminer l'intervalle de validité et l'intervalle de tolérance pour chaque condition temporelle floue (i.e. le paramètre flou L).

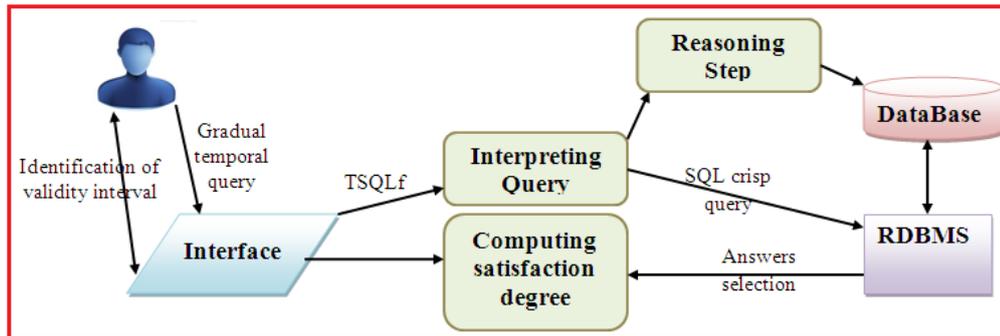
Par la suite, le système transmet la requête au module *Interpreting Query* qui la transforme en une requête non graduelle (sans conditions floues) en utilisant :

- i) les intervalles de validité déjà introduits au cours du module précédent,
- ii) les attributs *VST* et *VET* de la base de données et
- iii) une ou plusieurs relations d'Allen floues qui répondent à la requête.

Le résultat de ce module est une requête *TSQLf* prête à être envoyée à un système de gestion de base de données afin de sélectionner les attributs qui répondent aux critères de recherche temporels flous.

Le dernier module, *Computing Satisfaction Degree*, assure le calcul du degré de satisfaction d'une requête *TSQLf*.

Les résultats retournés sont attachés d'un degré de satisfaction, puis affichés sur l'interface de l'utilisateur. Dans le cas où la requête *TSQLf* nécessite une phase de raisonnement, le système appelle le module *Reasoning Step*. Ce module exploite le mécanisme d'inférence des relations d'Allen floues.

FIGURE 3.5: Processus de traitement de requêtes *TSQlf*.

### 3.5 Exemples illustratifs

Pour mieux expliquer notre proposition, nous présentons ci-dessous l'exemple du domaine d'archéologie. Considérons la table Archéologie (voir tableau 3.5) qui présente des vestiges matériels des temps préhistoriques. Le schéma de la table est *Archéologie* (*Code\_Ar*, *Nom\_Ar*, *Lieu*, *Date\_découverte*, *VST\_Dc*, *VET\_Dc*, *Date\_Daté*, *VST\_Dd*, *VET\_Dd*). Où *VST\_Dc* signifie la date de l'entrée en vigueur de la validité de *Date\_découverte*, *VET\_Dc* signifie la date de la fin de validité de la *Date\_découverte*, *VST\_Dd* signifie la date de commencement de la validité de *Date\_Daté* and *VET\_Dd* signifie la date de l'expiration de la date de validité *Daté*.

Etant donné les requêtes suivantes :

**Q1** : "Afficher les découvertes archéologiques qui ont eu lieu bien après milieu 2011 et juste à la fin Juillet 2013 (avec un degré minimum de 0,8)."

L'utilisateur dans la requête **Q1** a précisé une période [Bien après milieu 2011, Juste à la fin Juillet 2013]. La spécification temporelle de l'utilisateur est incertaine donc d'après le principe de *TSQlf*, le système demande l'utilisateur de définir une période de validité [*VST*, *VET*] et deux seuils de validité  $\alpha$  et  $\beta$ .

Par exemple,  $VST = 01/08/2011$ ,  $VET = 30/07/2013$ ,  $\alpha = 2$  mois et  $\beta = 1$  mois.

Donc on aura la forme trapézoïdale suivant [01/06/2011, 01/08/2011, 30/07/2013, 30/08/2013]. Selon **Q1** l'utilisateur demande les découvertes pendant la période [Bien après milieu 2011, Juste à la fin Juillet 2013] alors la relation d'Allen que nous allons utiliser est *F\_contains*. Pour le calcul des degrés ; Les découvertes, qui ont eu lieu pendant [01/08/2011 ; 30/07/2013], ont un degré = 1 (c'est à dire  $t.VST\_Dc > 01/08/2011$  et  $t.VET\_Dc < 30/07/2013$ ).

Les découvertes, qui ont eu lieu en dehors de la période de tolérance définie par l'utilisateur (c'est à dire  $t.VST\_Dc < 01/06/2011$  ou  $t.VET\_Dc > 30/08/2013$ ), ont

un degré = 0.

Si non, on aura un degré entre [0, 1] selon la formule des comparateurs flous définie dans le premier chapitre.

Enfin, on peut écrire la **requête TSQlf** comme suit ;

**Select** 0.8 *Nom\_Ar, Lieu*

**From** *Archéologie*

**Where** [*t.VST\_Dc, t.VET\_Dc*] *F\_Contains* [01/08/2011, 30/07/2013] ;

Le résultat de cette requête est les tuples (qui ont un degré supérieur à 0.8) *A011*, *A120* et *A075* (avec un degré = 1) et *A042* (avec un degré = 0.9).

**Q2** : "Afficher les lieux de découvertes archéologiques qui ont eu lieu bien après la fin de l'année 2008 et datées juste après 440 ans avant JC (avec un degré minimum de 0.7)."

**Requête TSQlf** :

**Select** 0.7 *Lieu*

**From** *Archéologie*

**Where** [*t.VST\_Dc, t.VET\_Dc*] *F\_After* [01/01/2009, 31/12/2009] and [*t.V-ST\_Dd, t.VET\_Dd*] *F\_Started-by* [440 BC, 450 BC].

Le résultat de cette requête est le tuple *A002* (avec un degré = 1).

**Q3** : "Afficher les noms des archéologies qui ont eu lieu juste avant ou avec la *Date\_Dc* de l'archéologie *A011* (avec un degré minimum de 0,6)" où l'intervalle de validité de l'archéologie *A011* est [10/07/2013, 31/07/2013].

**Requête TSQlf** :

**Select** 0.6 *Nom\_Ar*

**From** *Archéologie*

**Where** [*t.VST\_Dc, t.VET\_Dc*] *F\_Before* [10/07/2013, 31/07/2013] or [*t.V-ST\_Dc, t.VET\_Dc*] *F\_Overlaps* [10/07/2013, 31/07/2013].

Le résultat de cette requête est les tuples *A120* (avec un degré = 1), *A075* (avec un degré = 0.9), *A101* et *A111* (avec un degré = 0.8), *A002* (avec un degré = 0.7) *A015* et *A224* (avec un degré = 0.6).

**Q4** : "Afficher les découvertes archéologiques qui ont eu lieu juste à la fin de la *Date\_Dc* de l'archéologie *A012* (avec un degré minimum de 0.7)" où la *Date\_Dc* de *A012* contient l'intervalle de validité de *A011* qui est [10/07/2013, 31/07/2013].

Cette requête nécessite une phase de raisonnement.

TABLE 3.5: Table d'Archéologie.

Code_Ar	Name_Ar	Location	Date_Discovery	VST_Dc	VET_Dc	Date_Dated	VST_Dd	VET_Dd
A011	Pyramid of six meters in height	Lima	Recently	10/07/2013	31/07/2013	5000 years ago	2987 BC	2988 BC
A015	Cone rocks	Lake Tiberias	In 2003	05/02/2003	08/08/2003	2050 years ago	37 BC	38 BC
A120	Church	Island of the City	Little ago	20/06/2013	25/07/2013	End 158	08/09/158	18/12/158
A002	Chanel monumental	Narbanne	In 2010	05/03/2010	10/11/2010	2455 years ago	440 BC	442 BC
A020	Hunting weapon	South Africa	Early 2009	03/01/2009	05/04/2009	500000 years ago	497985 BC	497987 BC
A042	Mammoth Skeleton	Seine-et-Marne	Summer 2013	05/06/2013	01/09/2013	Before beginning 159	15/11/189	13/12/189
A075	Sort of mini-dinosaur	Africa	October 2012	02/10/2012	31/10/2012	Before the end of 260	20/09/260	15/12/260
A101	Indian prints	Brazil	Before the end of November 2011	18/11/2011	30/11/2011	More than 3000 years BC	3000 BC	3002 BC
A111	Corps soldiers Allemends	Carspach	Before the end of 2011	12/10/2011	15/12/2011	In 1918	10/07/1918	15/12/1918
A224	Statues banned by the Nazis	Berlin	Beginning in November 2010	01/11/2010	10/11/2010	During the second war mandial	01/09/1939	02/09/1945

Soit  $A$  : un intervalle de validité de  $A011$ ,  $B$  : la  $Date\_Dc$  de  $A012$  et  $C = [t.VST\_Dc, t.VET\_Dc]$  : un intervalle de validité des tuples à sélectionner. D'après la requête **Q4**, on a  $A$  Pendant  $B$  et  $B$  Rencontre  $C$ . En utilisant les règles d'inférences appliquées sur les relations d'Allen (pour plus de détails voir chapitre 1), nous pouvons conclure que  $C$  Après  $A$  c'est à dire  $[t.VST\_Dc, t.VET\_Dc] F\_After [10/07/2013, 31/07/2013]$ .

**Requête TSQlf :**

**Select** 0.7 \*

**From** Archéologie

**Where**  $[t.VST\_Dc, t.VET\_Dc] F\_After [10/07/2013, 31/07/2013]$  ;

Montrons sur quelques exemples comment les axiomes décrivant le comportement de la transitivité des relations temporelles d'Allen floues peuvent être établis :

- i) Soit la requête **Q5** "Afficher les découvertes archéologiques qui ont eu lieu avant la révolution Tunisienne (avec un degré minimum 0.8)."

**Requête TSQlf :**

**Select** 0.8 \*

**From** Archéologie

**Where**  $[t1.VST\_Dc, t1.VET\_Dc] F\_Before [17/12/2010, 15/01/2011]$  ;

Soit  $A$  : le résultat de la requête **Q5**. Supposons, maintenant, la requête **Q6**  
 ”Afficher les découvertes archéologiques qui ont eu lieu avant la révolution  
 Egyptienne (avec un degré minimum 0.9).”

On sait que la révolution Tunisienne (dénotée par  $B$ ) a eu lieu avant la révolution  
 Egyptienne (dénotée par  $C = [t2.VST\_Dc, t2.VET\_Dc]$ ). Alors  $A F\_Before (L_1) B$   
 et  $B F\_Before (L_2) C$ , la relation temporelle qui tient entre  $A$  et  $C$  peut être  
 obtenue comme suit :

$$A F\_Before(L_1) B \Leftrightarrow 17/12/2010 G(L_1)_+^c \text{ t1.VET\_Dc,}$$

$$B F\_Before (L_2) C \Leftrightarrow t2.VST\_Dc G(L_2)_+^c \text{ 15/01/2011.}$$

Maintenant, par application de la règle  $R_1$  des règles d’inférence (voir chapitre  
 1) sur  $17/12/2010 G(L_1)_+^c \text{ t1.VET\_Dc}$ , on obtient  $15/01/2011 G(L_1)_+^c \text{ t1.VET\_Dc}$   
 puisque  $15/01/2011 > 17/12/2010$ .

La règle de transitivité  $R_6$  appliquée sur  $t2.VST\_Dc G(L_2)_+^c \text{ 15/01/2011 et}$   
 $15/01/2011 G(L_1)_+^c \text{ t1.VET\_Dc}$ , implique que  $t2.VST\_Dc G((L_2)_+^c \oplus (L_1)_+^c)$   
 $t1.VET\_Dc$ . Ceci signifie que  $A F\_Before(L_2 \oplus L_1) C$  puisque  $(L_1)_+^c \oplus (L_2)_+^c =$   
 $(L_1 \oplus L_2)_+^c$ . Enfin, le résultat de la requête **Q5** satisfait amplement au critère de la  
 requête **Q6** (avec un degré = 1).

ii) **Q7** : ”Afficher les découvertes archéologiques qui ont commencé avec la  
 deuxième guerre mondiale (avec un degré minimum 0.6).”

Soit  $A = [t1.VST\_Dc, t1.VET\_Dc]$  et  $B = [03/09/1939, 10/12/1945]$ , **Requête**  
**TSQlf** :

**Select** 0.6 \*

**From** Archéologie

**Where**  $A F\_Starts(L_1) B$  ;

Supposons maintenant que nous avons le résultat (dénoté par  $C = [t2.VST\_Dc,$   
 $t2.VET\_Dc]$ ) d’une requête **Q8** sachant que  $B F\_started-by (L_2) C$ . D’après les  
 définitions de  $F\_Starts(L_1)$  et  $F\_started-by (L_2)$ , les deux équivalences suivantes  
 tiennent :

$$A F\_Starts(L_1) B \Leftrightarrow t1.VST\_Dc E(L_1) \text{ 03/09/1939 } \wedge \text{ 10/12/1945 } G((L_1)_+^c)$$

$$t1.VET\_Dc,$$

$$B F\_Started-by (L_2) C \Leftrightarrow 03/09/1939 E(L_2) \text{ t2.VST\_Dc } \wedge \text{ 10/12/1945 } G((L_2)_+^c)$$

$$t2.VET\_Dc.$$

La règle de transitivité  $R_5$ , appliquée à  $t1.VST\_Dc E(L_1) \text{ 03/09/1939 et}$   
 $03/09/1939 E(L_2) \text{ t2.VST\_Dc}$ , permet de déduire  $t1.VST\_Dc E(L_1 \oplus L_2)$

t2.VST\_Dc (qui sera dénotée par  $H_1$ ). Les relations atomiques qui pourraient donc tenir entre  $A$  et  $C$  sont :

$$fs = H_1 \wedge H_3 \text{ (} H_3 \text{ désigne } t2.VET\_Dc > t1.VET\_Dc)$$

$$fe = H_1 \wedge H'_3 \text{ (} H'_3 \text{ désigne } t2.VET\_Dc = t1.VET\_Dc)$$

$$fsi = H_1 \wedge H''_3 \text{ (} H''_3 \text{ désigne } t2.VET\_Dc < t1.VET\_Dc)$$

D'une manière similaire, comme ci-dessus, on peut aussi montrer que la disjonction  $fs \vee fe \vee fsi$  se réduit à  $H_1$ . Ce qui signifie que  $A < fs..fsi > (L_1 \oplus L_2)$

$C$ .

### 3.6 Conclusion

Dans ce chapitre, nous avons discuté les premières fondations pour la construction d'un système d'interrogation intelligent des bases de données temporelles. Le concept clé de notre approche est les relations temporelles d'Allen floues pour exprimer certains critères temporels vagues. Nous avons aussi fourni une interface conviviale pour permettre aux utilisateurs de soumettre leurs requêtes temporelles floues. Le système développé peut passer par une étape de raisonnement pour résoudre des requêtes sophistiquées exprimant les besoins complexes des utilisateurs.

Dans le chapitre suivant, nous prévoyons à border les intervalles de temps avec des bornes floues.

# Requêtes en présence des intervalles temporels flous

## 4.1 Introduction

Plusieurs application du monde réel nécessite la gestion et le traitement des informations temporelles. Les relations temporelles d'Allen sont les formalismes les plus utilisées pour gérer et modéliser les données temporelles. Comme mentionné dans le chapitre précédent, très peu de travaux sont considérés pour résoudre le problème de la modélisation des requêtes graduelles pour la gestion des bases de données temporelles. Dans le domaine des bases de données, nous pouvons citer quelques travaux qui ont abordé la question de l'imperfection des informations temporelles (à savoir, (Billiet et al., 2011), (Frias et al., 1991), (Deng, Liang, & Zhang, 2008), (Tudorie et al., 2012), (Galindo & J.M.Medina, 2001)).

Dans (Gammoudi et al., 2014), nous avons proposé une extension, nommée *TSQLf*, du langage *SQLf* (Pivert & Bosc, 2012) par l'addition du dimension temps. *TSQLf* permet d'exprimer des requêtes utilisateurs contenant des critères temporels flous.

Il est à noter que toutes les approches ci-dessus considèrent les relations temporelles floues seulement entre les intervalles de temps reguliers (leurs bornes inferieures et superieures sont des instants précis). Alors que dans les applications du monde réel, les intervalles de temps sont généralement décrits

par des bornes mal-définies.

Dans ce chapitre, nous présentons deux modèles pour résoudre de cette problématique à savoir ; premièrement, nous discutons une nouvelle idée pour introduire quelques flexibilités dans la définition des relations entre deux intervalles de temps flous. Le concept clé de ce modèle est une relation de tolérance floue qui permet de définir les relations d'Allen tolérantes en utilisant les intervalles dilatés et érodés de l'intervalle initial. Des indices de comparaison flous sont utilisés pour comparer deux intervalles de temps flous. Cette extension des relations d'Allen est intégrée dans notre système *Fuzz-TIME* déjà développé. Deuxièmement, nous introduisons une deuxième extension des relations temporelles d'Allen entre les intervalles de temps flous. Ce modèle exprime les différentes positions entre deux intervalles de temps flous suivant l'interprétation disjonctive.

Ce chapitre est structuré comme suit, dans un premier lieu, nous rappelons les opérations de dilatation et d'érosion pour passer par la suite à la discussion d'une nouvelle sémantique des relations d'Allen tolérantes. Les objectifs de la section 5.3 sont doubles, nous introduisons la modélisation des relations d'Allen tolérantes basée sur les opérations de dilatation et d'érosion. Ensuite, nous présentons les indices de comparaisons entre deux intervalles de temps flous pour estimer à quel point une relation d'Allen tolérante est satisfaite. Dans la deuxième partie de ce chapitre, nous allons présenter les relations d'Allen entre les intervalles de temps flous. Pour se faire, nous commençons par rappeler la différence entre les ensembles disjonctifs et les ensembles conjonctifs. Puis nous présentons les relations d'Allen entre les intervalles de temps flous dans le cadre d'une sémantique disjonctive. Enfin, la section 5.5 conclut le chapitre.

## 4.2 Opérations de dilatation et d'érosion

Nous discutons ici les bases d'une nouvelle sémantique, fondée sur la relation de tolérance  $E(L)$  (discutée précédemment dans le chapitre 1), des relations d'Allen.

Chaque relation d'Allen dite *tolérante* est exprimée en termes d'une classe de relations d'Allen classiques appliquées sur les versions dilatées et érodées des intervalles  $A$  et  $B$  au moyen de la relation de tolérance  $E(L)$ .

Pour mieux comprendre les opérations de dilatation et d'érosion, nous prenons un ensemble flou où  $A$  représentant un intervalle de temps, et une relation

d'égalité approchée  $E(L)$ . (Dubois, Hadjali, & Prade, 2003) ont associé à  $A$  une paire imbriquée d'ensembles flous quand la relation paramétrée  $E(L)$  est utilisée comme une relation de tolérance. En effet,

- on peut construire un ensemble flou d'instantns proches de  $A$  tel que  $A \subseteq A^L$ . Ceci est l'opération de dilatation.
- on peut construire un ensemble flou d'instantns proches de  $A$  tel que  $A_L \subseteq A$ . Ceci est l'opération de l'érosion.

### 4.2.1 Opération de dilatation

Dilater l'ensemble flou d'instantns  $A$  par  $L$  produira un ensemble flou  $A^L$  défini par :

$$\mu_{A^L}(r) = \sup_s \min(\mu_{E[L]}(s, r), \mu_A(s)) \quad (4.1)$$

$$= \sup_s \min(\mu_L(r - s), \mu_A(s)) \quad (4.2)$$

$$= \mu_{A \oplus L}(r). \quad (4.3)$$

Ce qui permet d'écrire que :

$$A^L = A \oplus L,$$

où  $\oplus$  est l'opération d'addition étendue pour les ensembles flous (Dubois & Prade, 1988).  $A^L$  rassemble les éléments de  $A$  et des éléments en dehors de  $A$  qui sont un peu proches d'un élément dans  $A$ . Voir Figure 4.1.

#### Lemme 1.

Étant donné un ensemble flou d'instantns  $A$  et une relation de tolérance  $E[L]$ , nous avons  $A \subseteq A^L$ .

*Preuve.* Évidente.

Nous pouvons facilement vérifier que l'ensemble flou d'instantns  $A^L$  est moins restrictive que  $A$ , mais sémantiquement proche de  $A$ . Ainsi,  $A^L$  peut être considéré comme une variante étendue de  $A$ . En terme de t.m.f., si  $A = (a, a', \alpha, \alpha')^1$  et  $L = (-\delta, \delta, \epsilon, \epsilon)$  alors  $A^L = (a - \delta, a' + \delta, \alpha + \epsilon, \alpha' + \epsilon)$ , voir Figure 4.1.

#### Exemple :

Si  $A = (15, 19, 2, 1)$  et  $L = (-1, 1, 0.5, 0.5)$  alors

---

1. avec  $[a, a']$  (resp.  $[a - \alpha, a' + \alpha']$ ) représente le noyau (resp. support) de  $A$ .

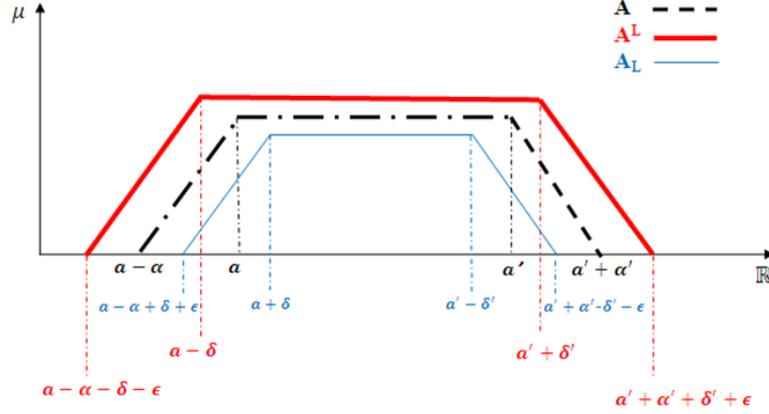


FIGURE 4.1: Dilatation et érosion d'un intervalle de temps.

$$A^L = (14, 20, 2.5, 1.5)$$

## 4.2.2 Opération d'érosion

Soit  $L \oplus X = A$  une équation où  $X$  est une variable inconnue. La résolution de cette équation est largement discutée dans (Dubois & Prade, 1983a). Il a été démontré que la plus grande solution de cette équation est donnée par,

$\bar{X} = A \ominus (-L) = A \ominus L$  (comme  $L = -L$ ) et où  $\ominus$  est la soustraction étendue de Minkowski définie par (Dubois & Prade, 1983a) :

$$\mu_{A \ominus L}(r) = \inf_s (\mu_L(r - s) \Rightarrow_{\mathbb{T}} \mu_A(s)) \quad (4.4)$$

où  $\mathbb{T}$  est une t-norme, et  $\Rightarrow_{\mathbb{T}}$  est la R-implication induite par  $\mathbb{T}$  et définie par  $\Rightarrow_{\mathbb{T}}(u, v) = \sup\{\lambda \in [0, 1], \mathbb{T}(u, \lambda) \leq v\}$ , pour  $u, v \in [0, 1]$ . Nous utilisons la même T-norme  $\mathbb{T}$  ( $= \min$ ) que dans l'opération de dilatation qui implique que  $\Rightarrow_{\mathbb{T}}$  est l'implication de Gödel.

Soit  $(E[L])_r = \{s, \mu_{E[L]}(s, r) > 0\}$  l'ensemble des éléments qui sont à proximité de  $r$  dans le sens de  $E[L]$ . L'expression ci-dessus peut être interprétée comme le degré d'inclusion de  $(E[Z])_r$  en  $A$ . Cela signifie que  $r$  appartient à  $A \ominus L$  si tous les éléments  $s$  qui sont proches de  $r$  sont dans  $A$ . Par conséquent, l'inclusion  $A \ominus L \subseteq A$  est vérifiée. Cette opération est très utile en langage naturel pour intensifier le sens des termes vagues. L'érosion de l'ensemble flou  $A$  par  $L$  donne ainsi l'ensemble flou  $A_L$  définie par,

$$A_L = A \ominus L$$

**Lemme 2.**

Étant donné un ensemble flou d'instant  $A$  et une relation de tolérance  $E[L]$ , nous avons  $A_L \subseteq A$ .

*Preuve.* Évidente.

L'ensemble flou  $A_L$  est plus précis que l'ensemble flou d'origine  $A$  mais il reste encore pas trop loin de  $A$  sémantiquement parlant. Si  $A = (a, a', \alpha, \alpha')$  et  $L = (-\delta, \delta, \epsilon, \epsilon)$  alors  $A_L = A \ominus L = (a + \delta, a' - \delta, \alpha - \epsilon, \alpha' - \epsilon)$  à condition que  $\alpha \geq \epsilon$  et  $\alpha' \geq \epsilon$ . Figure 4.1 illustre cette opération.

**Exemple :**

Si  $A = (15, 19, 2, 1)$  et  $L = (-1, 1, 0.5, 0.5)$  alors

$$A_L = (16, 18, 1.5, 0.5)$$

Dans le cas précis,  $A \ominus L = [a, a'] \ominus [-\delta, \delta] = [a + \delta, a' - \delta]$  (tandis que  $A \oplus Z = [a - \delta, a' + \delta]$ ).

On peut facilement vérifier que la proposition suivante est vraie :

**Proposition 1.** En utilisant les t.m.f de  $A^L$  et  $A_L$  données ci-dessus, nous avons :

- $(A^L)_L = (A_L)^L = A$
- $(A^L)^L = A \oplus 2L$
- $(A_L)_L = A \ominus 2L$

### 4.3 Relations d'Allen Tolérantes

Les objectifs de cette section sont doubles. Nous introduisons la modélisation des relations d'Allen tolérantes basée sur les opérations de dilatation et d'érosion. Ensuite, nous présentons les indices de comparaisons entre deux intervalles de temps flous pour démontrer jusqu'à quel point une relation d'Allen tolérante est satisfaite.

### 4.3.1 Modélisation

En utilisant les opérations de dilatation et d'érosion définies ci-dessus, nous pouvons construire des bases pour définir une extension fondée sur une relation de tolérance des relations d'Allen. Par exemple, la relation temporelle tolérante correspondant à la relation d'Allen *meets* s'écrit :

$$A \text{ toler-meets}(L) B \text{ correspondrait à } A_L \text{ before } B_L \text{ and } A^L \text{ overlaps } B^L,$$

La déclaration  $A \text{ toler-meets}(L) B$  signifie que la relation temporelle entre les deux intervalles de temps  $A$  et  $B$  est perçue comme une variante de la relation *meets* moyennant une certaine tolérance exprimée par l'indicateur  $L$ . Ceci correspond à une perception humaine qui est souvent rencontrée dans les applications du monde réel (comme dans la gestion des données historiques temporelles, la planification et l'ordonnancement, la compréhension du langage naturel, etc.).

La relation *toler-meets* regroupe une classe de relations d'Allen (c'est-à-dire, *before* et *overlaps*) appliquées sur les intervalles de temps érodés et dilatés correspondant aux intervalles de temps  $A$  et  $B$ . Cela revient à calculer les relations traditionnelles d'Allen, *before* et *overlaps*, sur des intervalles temporels flous (comme  $A^L$  et  $A_L$  sont des ensembles flous, comme indiqué dans la section précédente).

Soit  $A = (a, a', \alpha, \alpha')$  un intervalle de temps flou avec  $\tilde{a} = (a, a, \alpha, 0)$  et  $\tilde{a}' = (a', a', 0, \alpha')$  les deux bornes floues de validité de  $A$ .

On peut écrire  $A = (a, a', \tilde{a}, \tilde{a}')$  (où  $[a, a']$  est le noyau de  $A$  et  $\tilde{a}$  et  $\tilde{a}'$  sont ses bornes floues). Sous cette forme :

- $A^L$  est défini par  $(a - \delta, a' + \delta, \tilde{a}^{(L)}, \tilde{a}'^{(L)})$  où  $\tilde{a}^{(L)} = (a - \delta, a - \delta, \alpha + \epsilon, 0)$  et  $\tilde{a}'^{(L)} = (a' + \delta, a' + \delta, 0, \alpha' + \epsilon)$ .
- $A_L$  est défini par  $(a + \delta, a' - \delta, \tilde{a}_{(L)}, \tilde{a}'_{(L)})$  où  $\tilde{a}_{(L)} = (a + \delta, a + \delta, \alpha - \epsilon, 0)$  et  $\tilde{a}'_{(L)} = (a' - \delta, a' - \delta, 0, \alpha' - \epsilon)$ .

#### Exemple :

Soit  $A$  un intervalle de temps flou qui exprime la période du début des années 20 jusqu'à la fin des années 20. Alors,  $A$  est défini par  $(a, a', \tilde{a}, \tilde{a}')$  avec  $\tilde{a} = (1920, 1920, 3, 0)$  et  $\tilde{a}' = (1930, 1930, 0, 2)$ . Admettons que  $\alpha = 3$ ,  $\alpha' = 2$  et  $L = (-3, 3, 1, 1)$ . Nous pouvons facilement déduire que :

- $A^L = (1917, 1933, \tilde{a}^{(L)}, \tilde{a}'^{(L)})$  où  $\tilde{a}^{(L)} = (1917, 1917, 4, 0)$  et  $\tilde{a}'^{(L)} = (1933, 1933, 0, 3)$ .

•  $A_L = (1923, 1927, \tilde{a}_{(L)}, \tilde{a}'_{(L)})$  où  $\tilde{a}_{(L)} = (1923, 1923, 2, 0)$  et  $\tilde{a}'_{(L)} = (1927, 1927, 0, 1)$ .

D'une manière semblable, les contreparties de toutes les relations d'Allen tolérantes peuvent être établies :

- $A$  *toler-meets* (L)  $B$  correspondrait à  $A_L$  *before*  $B_L \wedge A^L$  *overlaps*  $B^L$ .
- $A$  *toler-before* (L)  $B$  correspondrait à  $A^L$  *toler-meets*  $B^L$ .
- $A$  *toler-overlaps* (L)  $B$  correspondrait à  $A_L$  *toler-meets*  $B_L$ .
- $A$  *toler-during* (L)  $B$  correspondrait à  $A^L$  *toler-equals*  $B_L$ .
- $A$  *toler-starts* (L)  $B$  correspondrait à  $A_L$  *during*  $B^L \wedge A^L$  *overlaps*  $B_L$ .
- $A$  *toler-finishes* (L)  $B$  correspondrait à  $A^L$  *overlapped by*  $B_L \wedge A_L$  *during*  $B^L$ .
- $A$  *toler-equals* (L)  $B$  correspondrait à  $A^L$  *contains*  $B_L \wedge A_L$  *during*  $B^L$ .

### 4.3.2 Indices de comparaison

Étant donné que les relations d'Allen tolérantes introduites sont de nature graduelle, leur satisfaction est alors une question de degré. Cette section vise à calculer à quel point une relation d'Allen tolérante est satisfaite. Pour ce faire, nous avons besoin d'un indice pour comparer deux intervalles de temps flous.

D'abord, nous rappelons les indices de comparaison d'intervalles flous proposés par (Dubois & Prade, 1983b).

Soit deux intervalles de temps flous  $M$  et  $N$  exprimés respectivement par deux quadruplets de la forme  $(m_1, m_2, \alpha_1, \alpha_2)$  et  $(n_1, n_2, \beta_1, \beta_2)$ . Fondamentalement, il y a quatre indices pour interpréter à quel point  $M$  est plus grand que  $N$  (voir Figure 4.2).

$$d(M > N) = \inf_{x,y} \{ \max(1 - \mu_M(x), 1 - \mu_N(y)) : x \leq y \} \quad (4.5)$$

$$d(M >^+ N) = \inf_x \sup_y \{ \max(1 - \mu_M(x), \mu_N(y)) : x \geq y \} \quad (4.6)$$

$$d(M >^- N) = \sup_x \inf_y \{ \max(\mu_M(x), 1 - \mu_N(y)) : x < y \} \quad (4.7)$$

$$d(M \geq N) = \sup_{x,y} \{ \min(\mu_M(x), \mu_N(y)) : x \geq y \} \quad (4.8)$$

L'équation (4.5) exprime la certitude que  $x$  est plus grand que  $N$ , sachant que  $x$  est  $M$ . Elle signifie que  $M$  est nécessairement plus grand que  $N$ . Elle peut être

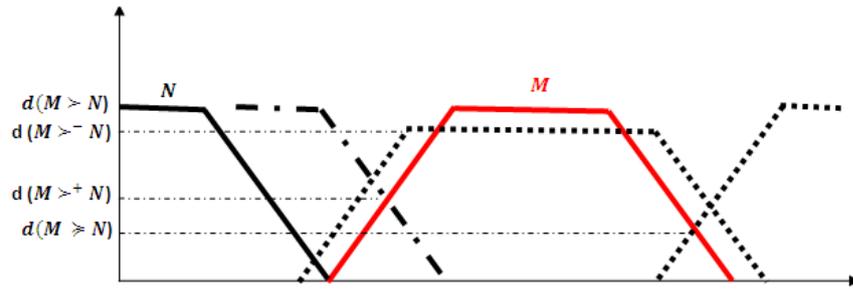


FIGURE 4.2: Indices possibilistes de comparaison des intervalles flous.

exprimée en fonction d'un degré de nécessité de la proposition que  $M$  est strictement supérieur à  $N$  comme suit ;

$$\mathfrak{N}_M(]N, +\infty)) = 1 - \sup_{x \leq y} \min(\mu_M(x), \mu_N(y)) \quad (4.9)$$

L'équation (4.9) se réfère au degré d'inclusion de l'ensemble flou  $M$  dans  $]N, +\infty)$ . Sachant que ;

$$\mathfrak{N}_M(]N, +\infty)) = \text{Ness}(x > N \mid x \text{ est } M)$$

L'équation (4.6) exprime la certitude que  $x$  est plus grand où égal à  $N$ , sachant que  $x$  est  $M$ . Elle peut être exprimée en fonction d'un degré de nécessité de la proposition que  $M$  est supérieur ou égal à  $N$  comme suit ;

$$\mathfrak{N}_M([N, +\infty)) = \inf_x \sup_{y \leq x} \max(1 - \mu_M(x), \mu_N(y)) \quad (4.10)$$

L'équation (4.10) se réfère au degré d'inclusion de l'ensemble flou  $M$  dans  $[N, +\infty)$ . Sachant que ;

$$\mathfrak{N}_M([N, +\infty)) = \text{Ness}(x \geq N \mid x \text{ est } M)$$

L'équation (4.7) exprime la possibilité que  $x$  est plus grand que  $N$ , sachant que  $x$  est  $M$ . Elle peut être exprimée en fonction d'un degré de possibilité de la proposition que  $M$  est strictement supérieur à  $N$  comme suit ;

$$\prod_M([N, +\infty)) = \sup_x \inf_{y: y \geq x} \min(\mu_M(x), 1 - \mu_N(y)) \quad (4.11)$$

L'équation (4.11) se réfère au degré de correspondance partielle de l'ensemble flou  $M \cap ]N, +\infty)$  des nombres strictement supérieur à  $N$ , étant donné qu'ils sont limités par  $M$ . Sachant que ;

$$\prod_M([N, +\infty)) = \text{Poss}(x > N \mid x \text{ est } M)$$

L'équation (4.8) exprime la possibilité que  $x$  est plus grand ou égal à  $N$ , sachant que  $x$  est  $M$ . Elle peut être exprimée en fonction d'un degré de possibilité de la proposition que  $M$  est supérieur ou égal à  $N$  comme suit ;

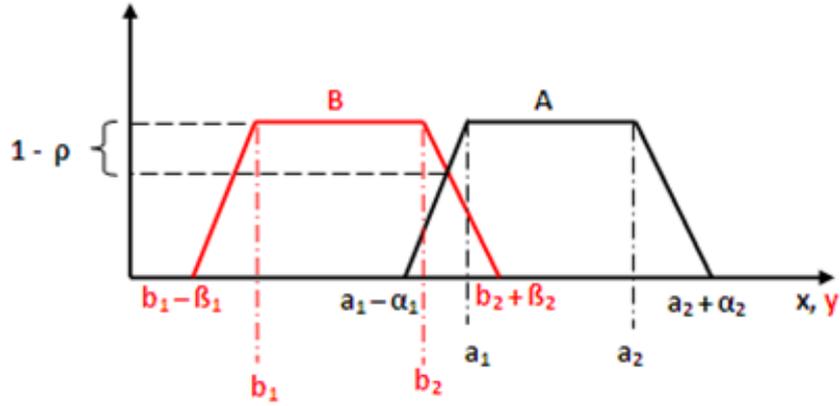
$$\prod_M([N, +\infty)) = \sup_{x, y: x \geq y} \min(\mu_M(x), \mu_N(y)) \quad (4.12)$$

L'équation (4.12) se réfère au degré de correspondance partielle de l'ensemble flou  $M \cap [N, +\infty)$  des nombres supérieur ou égal à  $N$ , étant donné qu'ils sont limités par  $M$ . Sachant que ;

$$\prod_M([N, +\infty)) = \text{Poss}(x \geq N \mid x \text{ est } M)$$

Etant donné ces degrés de comparaison entre deux intervalles de temps flous, nous pouvons utiliser l'équation (4.5) redéfinie en terme de degré de nécessité par l'équation (4.9) afin d'évaluer la mesure avec laquelle un intervalle de temps flou  $A$  est supérieur à un autre intervalle de temps flou  $B$  (voir Figure 4.3, noté par  $d(A > B)$  (avec  $A = (a_1, a_2, \alpha_1, \alpha_2)$  et  $B = (b_1, b_2, \beta_1, \beta_2)$ ) comme suit ;

$$d(A > B) = 1 - \sup_{x \leq y} \min(\mu_A(x), \mu_B(y)) \quad (4.13)$$

FIGURE 4.3: Indice de comparaison  $d(A > B)$ .

$$= \begin{cases} 1 & \text{si } a_1 - \alpha_1 \geq b_2 + \beta_2 \\ 1 - \rho & \text{si } a_1 - \alpha_1 < b_2 + \beta_2 \text{ et } a_1 > b_2 \\ 0 & \text{sinon} \end{cases}$$

$$\text{avec } \rho = \frac{w - (a_1 - \alpha_1)}{\alpha_1} \text{ et } w = \frac{b_2 \alpha_1 + a_1}{1 + \alpha_1}$$

Pour illustrer la formule (4.13), considérons l'exemple suivant (où  $A = [25/10/2015, 28/10/2015, 1, 1]$  qui exprime le temps environ 25 Octobre et 28 Octobre) :

**Cas 1 :**  $B = [19/10/2015, 20/10/2015, 2, 2]$

$$a_1 - \alpha_1 > b_2 + \beta_2 \Rightarrow 24/07/2015 > 20/07/2015 \text{ alors}$$

$$d(A > B) = 1.$$

**Cas 2 :**  $B = [19/10/2015, 24/10/2015, 2, 2]$

$$a_1 - \alpha_1 < b_2 + \beta_2 \Rightarrow 24/07/2015 < 26/07/2015 \text{ et}$$

$$a_1 > b_2 \Rightarrow 25/07/2015 > 24/07/2015 \text{ alors}$$

$$d(A > B) = 1 - \rho = 1 - \frac{b_2 - a_1 + 1 + \alpha_1}{1 + \alpha_1} = 0.5$$

**Cas 3 :**  $B = [19/10/2015, 26/10/2015, 2, 2]$

$$a_1 - \alpha_1 < b_2 + \beta_2 \Rightarrow 24/07/2015 < 27/07/2015 \text{ et}$$

$$a_1 < b_2 \Rightarrow 25/10/2015 < 26/10/2015 \text{ alors}$$

$$d(A > B) = 0$$

Maintenant, sur la base de la formule (4.13), nous fournissons dans le tableau 4.1 (colonne 3) un moyen pour calculer les degrés de satisfaction des relations d'Allen tolérantes. Par exemple, le degré de satisfaction  $A \text{ toler-meets } (L) B$  s'écrit :  $A \text{ toler-meets } (L) B = \min(d(\tilde{b}_L > \tilde{a}'_L), d(\tilde{b}^L > \tilde{a}^L), d(\tilde{a}'^L > \tilde{b}^L), d(\tilde{b}'^L > \tilde{a}^L))$ .

TABLE 4.1: Relations d'Allen tolérantes (où  $d(\tilde{b}_L > \tilde{a}'_L) = 1 - \sup_{x \leq y} \min(\mu_{\tilde{b}_L}(x), \mu_{\tilde{a}'_L}(y))$ )

Relation d'Allen tolérante	Interprétation	Définition
$A$ toler-meets(L) $B$	$A_L$ before $B_L \wedge A^L$ overlaps $B^L$	$\min(d(\tilde{b}_L > \tilde{a}'_L), d(\tilde{b}^L > \tilde{a}^L), d(\tilde{a}'^L > \tilde{b}^L), d(\tilde{b}'^L > \tilde{a}^L))$
$A$ toler-before(L) $B$	$A^L$ toler-meets $B^L$	$\min(d(\tilde{b} > \tilde{a}'), d(\tilde{b}^{2L} > \tilde{a}^{2L}), d(\tilde{b}'^{2L} > \tilde{a}'^{2L}), d(\tilde{a}'^{2L} > \tilde{b}^{2L}))$
$A$ toler-overlaps(L) $B$	$A_L$ toler-meets $B_L$	$\min(d(\tilde{b}_{2L} > \tilde{a}'_{2L}), d(\tilde{b} > \tilde{a}), d(\tilde{a}' > \tilde{b}), d(\tilde{b}' > \tilde{a}'))$
$A$ toler-during(L) $B$	$A^L$ toler-equals $B_L$	$\min(d(\tilde{b}_{2L} > \tilde{a}^{2L}), d(\tilde{a}'^{2L} > \tilde{b}'_{2L}), d(\tilde{a} > \tilde{b}), d(\tilde{b}' > \tilde{a}'))$
$A$ toler-starts(L) $B$	$A_L$ during $B^L \wedge A^L$ overlaps $B_L$	$\min(d(\tilde{a}_L > \tilde{b}^L), d(\tilde{b}'^L > \tilde{a}'_L), d(\tilde{b}_L > \tilde{a}^L), d(\tilde{a}'^L > \tilde{b}_L), d((\tilde{b}'_L > \tilde{a}^L))$
$A$ toler-finishes(L) $B$	$A^L$ overlapped by $B_L \wedge A_L$ during $B^L$	$\min(d(\tilde{a}^L > \tilde{b}_L), d(\tilde{b}'_L > \tilde{a}^L), d(\tilde{a}'^L > \tilde{b}'_L), d(\tilde{a}_L > \tilde{b}^L), d(\tilde{b}'^L > \tilde{a}'_L))$
$A$ toler-equals(L) $B$	$A^L$ contains $B_L \wedge A_L$ during $B^L$	$\min(d(\tilde{b}_L > \tilde{a}^L), d(\tilde{a}'^L > \tilde{b}'_L), d(\tilde{a}_L > \tilde{b}^L), d(\tilde{b}'^L > \tilde{a}'_L))$

## 4.4 Relations d'Allen entre les intervalles de temps flous

Généralement, les relations d'Allen temporelles et leurs contreparties floues sont définies sur des intervalles de temps réguliers/flous interprétés comme des ensembles conjonctifs de dates. Néanmoins, dans le monde réel, les bornes des intervalles de temps sont considérées comme des ensembles disjonctifs pour représenter une connaissance incomplète (sur les valeurs de ces bornes). Autrement, une seule valeur est possible pour chaque borne. Dans cette section, nous commençons par rappeler la différence entre les ensembles disjonctifs et les ensembles conjonctifs. Puis nous présentons les relations d'Allen entre les intervalles de temps flous.

#### 4.4.1 Ensembles disjonctifs vs ensembles conjonctifs

Une information incomplète est souvent représentée d'une manière imprécise. L'imprécision est toujours représentée par une disjonction d'éléments mutuellement exclusifs ; l'un d'eux est la vraie valeur. Par exemple, pour dire que l'âge de Paul est entre 20 et 25 ans, i.e,

$$v = \text{Age}(\text{Paul}) \in \{20, 21, 22, 23, 24, 25\},$$

cela veut dire que  $v = 20$  ou  $v = 21$  ou  $v = 22$  ou  $v = 23$  ou  $v = 24$  ou  $v = 25$ .

L'ensemble utilisé pour représenter une information incomplète est appelé un *ensemble disjonctif*. Elle contraste avec la vue conjonctive de l'ensemble considéré comme une collection d'éléments.

Un *ensemble conjonctif* représente une information précise. Par exemple, considérons la quantité  $v = \text{soeurs}(\text{Jack})$  dont le domaine est l'ensemble des sous-ensembles de noms possibles des soeurs de Jack. L'information  $v = \{\text{Lina}, \text{Acil}\}$  est précise et signifie que les soeurs de *Jack* sont *Lina* et *Acil*.

#### 4.4.2 Relations temporelles : Vue disjonctive

Ici, nous discutons une autre idée pour définir les relations d'Allen temporelles entre les intervalles de temps flous dans le cadre de la sémantique disjonctive. Soit  $\tilde{A} = [\tilde{a}, \tilde{a}']$  et  $\tilde{B} = [\tilde{b}, \tilde{b}']$  deux intervalles de temps flous où les bornes  $\tilde{a}$  et  $\tilde{a}'$  (resp.  $\tilde{b}$  et  $\tilde{b}'$ ) sont définies d'une manière floue incomplète. Considérons maintenant les relations d'Allen temporelles pour définir la position de  $\tilde{A}$  par rapport à  $\tilde{B}$ .

Par exemple,

$\tilde{A}$  *before*  $\tilde{B}$  correspondrait à

$$\exists v_1 \in \tilde{a}', \exists v_2 \in \tilde{b} / v_1 < v_2.$$

Cela veut dire qu'il suffit qu'ils existent  $v_1$  appartient à  $\tilde{a}'$  et  $v_2$  appartient à  $\tilde{b}$  tel que  $v_1 < v_2$ , pour dire que la relation  $\tilde{A}$  *before*  $\tilde{B}$  est vérifié à une certaine mesure. On peut facilement mesurer le degré de satisfaction de  $\tilde{A}$  *before*  $\tilde{B}$  comme suit :

$$d(\tilde{A} \text{ before } \tilde{B}) = \min(\max \mu_{\tilde{a}'}(v_1), \max \mu_{\tilde{b}}(v_2), \mu_{<}(v_1, v_2))$$

$$\text{où } \mu_{<}(v_1, v_2) = \begin{cases} 1 & \text{si } v_1 < v_2 \\ 0 & \text{sinon} \end{cases}$$

De la même façon, on peut calculer  $\tilde{A} \text{ meets } \tilde{B}$  qui correspondrait à

$$\exists v_1 \in \tilde{a}', \exists v_2 \in \tilde{b} / v_1 = v_2$$

et  $d(\tilde{A} \text{ meets } \tilde{B})$  qui s'écrit

$$d(\tilde{A} \text{ meets } \tilde{B}) = \min(\max \mu_{\tilde{a}'}(v_1), \max \mu_{\tilde{b}}(v_2), \mu_{=}(v_1, v_2))$$

$$\text{où } \mu_{=}(v_1, v_2) = \begin{cases} 1 & \text{si } v_1 = v_2 \\ 0 & \text{sinon} \end{cases}$$

Le Tableau 4.2 résume toutes les relations d'Allen entre deux intervalles de temps flous  $\tilde{A}$  et  $\tilde{B}$ .

**Exemple :**

Soit  $\tilde{A} = [\tilde{a}, \tilde{a}']$  et  $\tilde{B} = [\tilde{b}, \tilde{b}']$  deux intervalles de temps flous.  $\tilde{A}$  représente la période allant du début des années 30 jusqu'à la fin des années 30 tandis que  $\tilde{B}$  représente la période allant du début des années 40 jusqu'à la fin des années 40.

Supposons que  $\tilde{a} = (1930, 1930, 2, 0)$ ,  $\tilde{a}' = (1939, 1939, 0, 2)$ ,  $\tilde{b} = (1940, 1940, 2, 0)$  et  $\tilde{b}' = (1949, 1949, 0, 2)$ . Le problème est de vérifier si la relation *meets* est vérifiée entre  $\tilde{A}$  et  $\tilde{B}$ .

Il est facile de vérifier que  $\tilde{A} \text{ meets } \tilde{B}$  avec un degré égal à 0.65. En effet,

(i)  $\exists v_1 = 39.7$  tel que  $\mu_{\tilde{a}'}(v_1) = 0.65$

(ii)  $\exists v_2 = 39.7$  tel que  $\mu_{\tilde{b}}(v_2) = 0.65$

(iii)  $v_1 = v_2$

Alors,  $d(\tilde{A} \text{ meets } \tilde{B}) = \min(0.65, 0.65, 1) = 0.65$ .

TABLE 4.2: Relations d'Allen entre deux intervalles de temps flous.

Relation d'Allen	Interprétation	Définition
$\tilde{A}$ Before $\tilde{B}$	$\exists v_1 \in \tilde{a}', \exists v_2 \in \tilde{b} / v_1 < v_2$	$\text{Min} (\text{Max } \mu_{\tilde{a}'}(v_1), \text{Max } \mu_{\tilde{b}}(v_2) : v_1 < v_2)$
$\tilde{A}$ Meets $\tilde{B}$	$\exists v_1 \in \tilde{a}', \exists v_2 \in \tilde{b} / v_1 = v_2$	$\text{Min}(\text{Max } \mu_{\tilde{a}'}(v_1), \text{Max } \mu_{\tilde{b}}(v_2) : v_1 = v_2)$
$\tilde{A}$ Overlaps $\tilde{B}$	$\exists v_1 \in \tilde{a}, \exists v_2 \in \tilde{a}', \exists v_3 \in \tilde{b}, \exists v_4 \in \tilde{b}' / v_1 < v_3 \wedge v_3 < v_2 \wedge v_2 < v_4$	$\text{Min}\{(\text{Max } \mu_{\tilde{a}}(v_1), \text{Max } \mu_{\tilde{b}}(v_3) : v_1 < v_3), (\text{Max } \mu_{\tilde{b}}(v_3), \text{Max } \mu_{\tilde{a}'}(v_2) : v_3 < v_2), (\text{Max } \mu_{\tilde{a}'}(v_2), \text{Max } \mu_{\tilde{b}'}(v_4) : v_2 < v_4)\}$
$\tilde{A}$ During $\tilde{B}$	$\exists v_1 \in \tilde{a}, \exists v_2 \in \tilde{a}', \exists v_3 \in \tilde{b}, \exists v_4 \in \tilde{b}' / v_3 < v_1 \wedge v_2 < v_4$	$\text{Min}\{(\text{Max } \mu_{\tilde{b}}(v_3), \text{Max } \mu_{\tilde{a}}(v_1) : v_3 < v_1), (\text{Max } \mu_{\tilde{a}'}(v_2), \text{Max } \mu_{\tilde{b}'}(v_4) : v_2 < v_4)\}$
$\tilde{A}$ Starts $\tilde{B}$	$\exists v_1 \in \tilde{a}, \exists v_2 \in \tilde{a}', \exists v_3 \in \tilde{b}, \exists v_4 \in \tilde{b}' / v_1 = v_3 \wedge v_2 < v_4$	$\text{Min}\{(\text{Max } \mu_{\tilde{a}}(v_1), \text{Max } \mu_{\tilde{b}}(v_3) : v_1 = v_3), (\text{Max } \mu_{\tilde{a}'}(v_2), \text{Max } \mu_{\tilde{b}'}(v_4) : v_2 < v_4)\}$
$\tilde{A}$ Finishes $\tilde{B}$	$\exists v_1 \in \tilde{a}, \exists v_2 \in \tilde{a}', \exists v_3 \in \tilde{b}, \exists v_4 \in \tilde{b}' / v_1 < v_3 \wedge v_2 = v_4$	$\text{Min}\{(\text{Max } \mu_{\tilde{a}}(v_1), \text{Max } \mu_{\tilde{b}}(v_3) : v_1 < v_3), (\text{Max } \mu_{\tilde{a}'}(v_2), \text{Max } \mu_{\tilde{b}'}(v_4) : v_2 = v_4)\}$
$\tilde{A}$ Equals $\tilde{B}$	$\exists v_1 \in \tilde{a}, \exists v_2 \in \tilde{a}', \exists v_3 \in \tilde{b}, \exists v_4 \in \tilde{b}' / v_1 = v_3 \wedge v_2 = v_4$	$\text{Min}\{(\text{Max } \mu_{\tilde{a}}(v_1), \text{Max } \mu_{\tilde{b}}(v_3) : v_1 = v_3), (\text{Max } \mu_{\tilde{a}'}(v_2), \text{Max } \mu_{\tilde{b}'}(v_4) : v_2 = v_4)\}$

## 4.5 Conclusion

Dans ce chapitre, nous avons présenté deux approches pour gérer les intervalles de temps flous non réguliers dont les bornes des intervalles sont définies d'une manière imprécise/floue. Le concept clé de notre premier modèle est une relation de tolérance floue qui permet de définir des relations d'Allen tolérantes en utilisant les intervalles dilatés et érodés de l'intervalle initial. Pour comparer ces intervalles de temps, nous avons proposé des indices de comparaison flous. Dans le deuxième modèle, nous avons introduit une extension des relations temporelles d'Allen entre les intervalles de temps flous. Ce modèle exprime les différentes positions entre deux intervalles de temps flous selon deux interprétations (conjonctive et disjonctive). Nous avons intégré ces deux approches dans notre système *Fuzz-TIME* déjà développé.

Nous présentons dans le chapitre suivant le système développé pour implémenter les modèles proposés. Ensuite, nous montrons les résultats des différentes expérimentations que nous avons menées et qui nous ont permis d'évaluer et de valider les approches proposées.

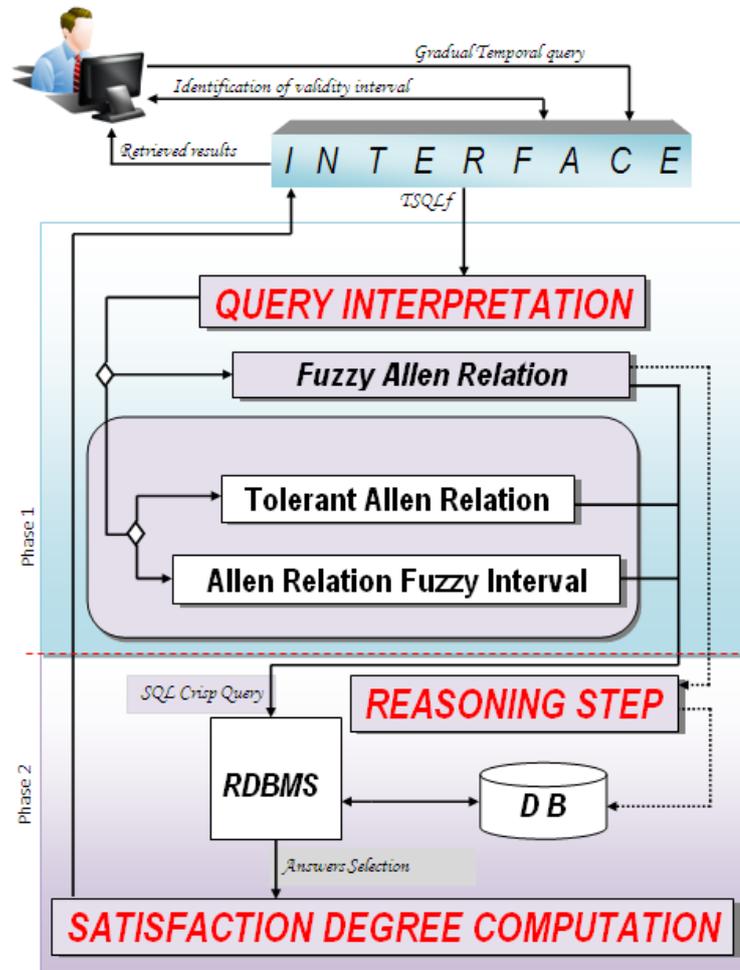


# Implémentation et Expérimentation

## 5.1 Introduction

Dans ce chapitre, nous présentons le système *Fuzz-TIME* développé pour la gestion de l'imperfection dans les bases de données temporeles et les expérimentations conduites pour valider les approches proposées et évaluer la pertinence de ses résultats. Ce système représente donc une mise en oeuvre des méthodes présentées dans les chapitres 4 et 5. Cette mise en oeuvre est rendue possible grâce au système *Fuzz-TIME* qui permet de gérer les requêtes temporelles flexibles proposées par les utilisateurs et de traiter les spécifications imprécises et floues dans les bases de données temporelles. Notre apport dans ce cadre a porté sur la définition des conditions temporelles floues et la gestion des intervalles de temps flous.

En section 6.2, nous décrivons le système *Fuzz-TIME*. Nous présentons dans cette section, en premier lieu, l'architecture du système *Fuzz-TIME*. Ensuite, nous décrivons comment nous avons l'implémenté. La section 6.3, décrit un exemple illustratif pour mieux expliquer notre proposition, nous présentons dans cette section un exemple du domaine d'archéologie avec un ensemble de requêtes temporelles graduelles. Enfin, la section 6.3 conclut le chapitre.

FIGURE 5.1: Architecture du système *Fuzz-TIME*.

## 5.2 Le système *Fuzz-TIME*

Afin d'évaluer les approches déjà décrites dans les deux chapitres précédents, nous avons développé un système nommé *Fuzz-TIME* (pour *Fuzzy Temporal Information Management and Exploitation*). Nous présentons dans cette section, en premier lieu, l'architecture du système *Fuzz-TIME*. Ensuite, nous décrivons comment nous avons implémenté les différents modules du système.

### 5.2.1 Architecture du système

Le système *Fuzz-TIME* comprend en réalité deux étapes principales et chaque étape contient un ensemble de modules.

Étape 1 : Elle commence par une requête temporelle graduelle (avec des spécifications temporelles floues) proposée par l'utilisateur à travers une interface graphique du système *Fuzz-TIME*. Ce dernier demande à l'utilisateur de définir un intervalle de validité pour chaque spécification temporelle floue. Ensuite, une requête *TSQLf* est générée et envoyée à un module d'interprétation. Ce module est composé d'un ensemble de sous-modules ; chacun présente une alternative pour la gestion des requêtes *TSQLf*. Le résultat de ce module correspond à une requête *SQL* classique.

Étape 2 : Cette étape débute généralement, si la requête nécessite une phase de raisonnement, par un appel à un module de raisonnement. Ce module utilise le mécanisme d'inférence basé sur des règles de transitivité des relations d'Allen floues. La requête passe, par la suite, par le système de gestion de base de données pour sélectionner les tuples qui répondent aux critères temporels en question. Puis, les tuples sélectionnées passent à travers le dernier module qui calcule le degré de satisfaction de chaque tuple avec les critères de sélection. Enfin, le résultat (les tuples sélectionnées avec leurs degrés) apparaît à l'utilisateur.

Comme indiqué auparavant, nous avons développé le système *Fuzz-TIME* pour le traitement des requêtes *TSQLf*. La figure 5.1 donne un aperçu sur l'architecture du système *Fuzz-TIME* et détaille ses différents modules.

Les fonctionnalités de chaque module sont résumées ci-dessous :

**Interface** : l'utilisateur saisit une requête temporelle graduelle en utilisant une interface graphique, cette dernière lui donne, tout d'abord, la possibilité de choisir les attributs, les tables et la construction de (des) condition(s) temporelle(s) floue(s). Deuxièmement, le système demande à l'utilisateur d'identifier l'intervalle de validité et l'intervalle de tolérance pour chaque condition temporelle floue (à savoir le paramètre flou  $L$ ).

**Fuzzy Allen Relation (FAR)** : si l'utilisateur définit un intervalle de validité avec des bornes précises, alors le module *FAR* est déclenché. Plus de détails sur ce module peuvent être trouvés dans le troisième chapitre. Notant que dans le cas où la requête *TSQLf* nécessite une phase de raisonnement, le système appelle le module *Reasoning Step*.

**Reasoning Step** : ce module exploite les mécanismes d'inférence basés sur des règles de transitivité des relations d'Allen floues. Dont l'objectif est de calculer la fermeture transitive des relations temporelles floues.

**Tolerant Allen Relation (TAR)** : si l'utilisateur définit un intervalle de validité avec des bornes floues, le système peut gérer la requête utilisateur à l'aide du module *TAR*, le système génère automatiquement les intervalles érodé et dilaté. Ensuite, le système transforme la requête en une requête *SQL* standard en utilisant le principe des opérations de dilatation et d'érosion. Plus de détails sur ce module peuvent être trouvés dans le quatrième chapitre. Le résultat de ce module est une requête prête à être envoyée au système de gestion de base de données afin de sélectionner les attributs qui répondent aux critères de recherche temporels flous.

**Allen Relations between Fuzzy Time Intervals (ARFI)** : si l'utilisateur définit un intervalle de validité avec des bornes floues, le système peut également gérer la requête utilisateur en utilisant le module *ARFI*. Ce module génère pour chaque borne un ensemble flou exprimé en terme de t.m.f. Ensuite, le système compare chacune des deux bornes floues en utilisant l'interprétation des relations d'Allen définies dans le chapitre précédent.

**Satisfaction Degree Computation** : ce module assure le calcul du degré de satisfaction de la requête *TSQLf* en question. Les résultats retournés sont attachés avec un degré de satisfaction et affichés sur l'interface utilisateur.

## 5.2.2 Implémentation et interfaces

Nous présentons dans cette section l'environnement de développement ainsi que certaines interfaces conviviales. L'outil, que nous avons développé, agit comme une interface *JAVA* avec le SGBD *Oracle* et génère des requêtes *TSQLf* directement exécutables par des appels à des fonctions et des bloc *PL/SQL* stockés. Dans ce qui suit nous donnons un aperçu, dans un premier lieu, sur le développement des relations d'Allen floues (*FAR*). Puis, sur l'implémentation d'une extension basée sur les relations d'Allen tolérantes. Enfin, nous présentons la réalisation des relations d'Allen entre les intervalles flous.

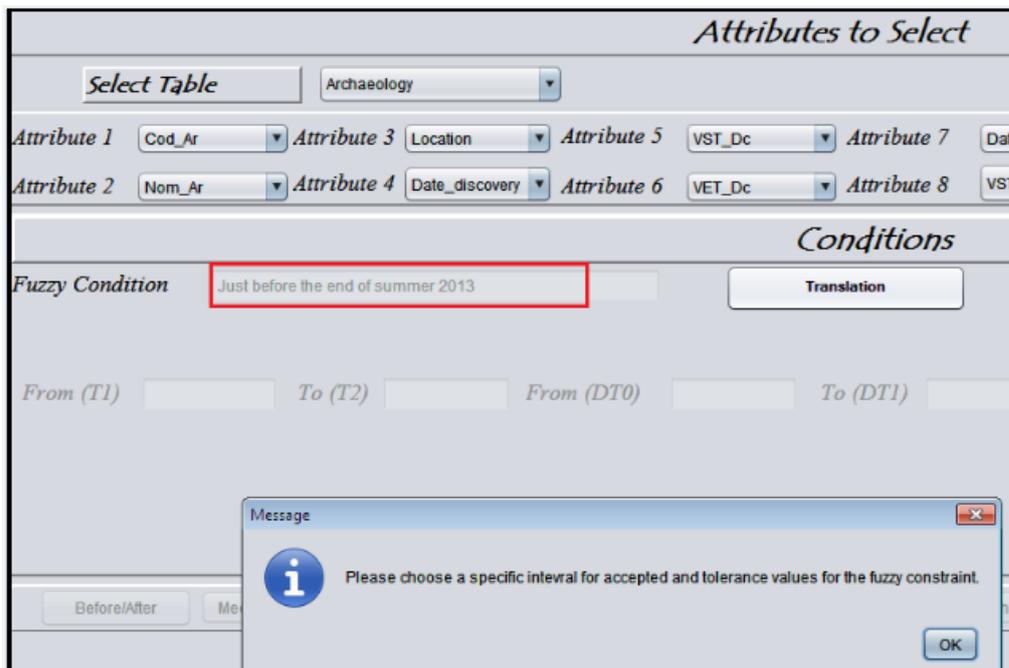


FIGURE 5.2: Définition d'une condition temporelle floue.

### A. Fuzzy Allen Relation (FAR)

L'interface du système *Fuzz-TIME* est connectée à une base de données pour stocker des tables d'un domaine qui intègre des aspects temporels flous. De cette façon, il est possible d'ajouter pour chaque table, qui contient un *attribut temporel flou*, deux attributs temporels spécifiques (*VST*, *VET*). Ces deux attributs expriment les bornes de l'intervalle validité de l'attribut temporel flou. Notre approche basée sur les relations temporelles d'Allen floues pour comparer deux intervalles de temps précis  $[T1, T2]$  et  $[t.VST, t.VET]$  (présentent, respectivement, les deux intervalles de validité de la condition floue de la requête et de l'attribut temporel flou de la base de données).

*Fuzz-TIME* offre à l'utilisateur une interface graphique pour l'identification des attributs à sélectionner, les tables à utiliser et la définition des conditions de filtrage. Selon le principe de la requête *TSQLf*, ces conditions peuvent être floues et bien évidemment booléennes. La figure 5.2 montre un exemple d'une requête là où l'utilisateur a saisi une (des) condition(s) temporelle(s) floue(s).

Suite à l'introduction d'une (des) condition(s) temporelle(s) floue(s), le système demande l'utilisateur pour définir un intervalle de validité et un intervalle de

FIGURE 5.3: Définition des intervalles de validité et de tolérance.

tolérance pour chaque condition temporelle floue (voir figure 5.3).

*Fuzz-TIME* génère, par la suite, une requête *TSQLf* et la soumet à la base de données pour sélectionner les tuples qui correspondent aux critères temporels flous. Finalement, un degré de satisfaction est calculé et attribué à chaque tuple sélectionnée (voir figure 5.4).

### B. Tolerant Allen Relation (TAR)

Les relations temporelles d'Allen floues modélisées par la requête *TSQLf* et implémentées dans le système *Fuzz-TIME*, traitent l'aspect temporelle flou dans la clause "where". En se basant sur le principe de la requête *TSQLf*, l'utilisateur doit définir les bornes de l'intervalle de validité de chaque condition floue par deux dates précises (*VST*, *VET*). En effet, dans les applications du monde réel, les bornes sont généralement définies d'une manière incertaine/graduelle. Par conséquent, un utilisateur se trouve parfois dans des cas où il ne peut pas définir les bornes d'un intervalle de validité par des dates précises. D'où l'idée d'étendre le système par des relations d'Allen tolérantes (*TAR*) pour répondre à cette question.

COD_AR	NOM_AR	LOCATION	DATE_DISCOV...	VST_DC	VET_DC	DATE_DATED	VST_DD	VET_DD	DEGREE
A011	Pyramid of six ...	Lima	Recently	2013-07-10 00:...	2013-07-31 00:...	5000 years ago	2987 BC	2988 BC	0
A015	Cone rocks	Lake Tiberias	In 2003	2003-02-05 00:...	2003-08-08 00:...	2050 years ago	37 BC	38 BC	1
A120	Church	Island of the City	Little ago	2013-06-20 00:...	2013-07-25 00:...	End 158	08/09/158	18/12/158	1
A002	Chanel monum...	Narbonne	In 2010	2010-03-05 00:...	2010-11-10 00:...	2455 years ago	440 BC	442 BC	1
A020	Hunting weapon	South Africa	Early 2009	2006-01-03 00:...	2009-04-05 00:...	500000 years a...	497985 BC	497987 BC	1
A042	Mammoth Skel...	Seine-et-Marne	Summer 2013	2013-06-05 00:...	2013-09-01 00:...	Before Beadinni...	15/11/158	13/12/158	0.52459016393...

FIGURE 5.4: Résultat d'une requête *TSQLf*.

Après la modélisation des relations d'Allen tolérantes (voir quatrième chapitre), nous avons amélioré le système *Fuzz-TIME* tout en gardant la possibilité où l'utilisateur peut introduire des bornes précises en appliquant l'approche des relations d'Allen floues (*FAR*) (voir figure 5.5).

L'extension (*TAR*) intégrée dans le système *Fuzz-TIME* offre, en premier lieu, la possibilité de définir une (des) condition(s) temporelle(s) floue(s). Puis, la définition des intervalles de validité et de tolérance pour chaque condition temporelle floue. Enfin, la possibilité offerte à l'utilisateur d'introduire les bornes floues pour chaque intervalle de validité (voir figure 5.6).

Après la définition des intervalles de validité et de tolérance, le système informe l'utilisateur que les bornes introduites sont floues. Par conséquent, le système sollicite l'utilisateur de définir un ensemble flou  $A$  qui correspond à la condition temporelle floue et bien évidemment d'introduire l'ensemble flou  $L$ . L'objectif principale de la définition de ces ensembles flous ( $A$  et  $L$ ) est de générer automatiquement une paire d'intervalles (ceux dilaté et érodé). Ensuite, une requête *TSQLf* est générée et soumise à la base de données pour sélectionner les tuples qui correspondent aux critères temporels flous. Finalement, un degré de satisfaction est calculé et attribué à chaque ligne sélectionnée (voir figure 5.7 et figure 5.8).

The screenshot shows the 'Fuzz-Time: Fuzzy Temporal Information Managing and Exploitation' web application. The main header features the 'Fuzz-Time' logo and navigation buttons like 'Return', 'Refresh', and 'Logout'. Below the header, there's a section for 'Management archaeologies' with a list of attributes to display (code, Name, location, date discovery, VST\_DC, VET\_DC, Date dated, VST\_DD, VET\_DD) and a 'translate' button. The main configuration area is titled 'I want to display all selected attributes their date:' and includes a dropdown menu set to 'are during this interval'. A text input field contains 'The years 2010 and 2011', which is highlighted with a green diamond. Below this, a red warning message states: 'Please specify only two dates if the condition will be considered fuzzy'. Two date input fields are present: 'I0' with '01-02-2010' and 'I1' with '31-10-2011', both highlighted with yellow rectangles. Further down, there are tolerance settings for 'TO' (10) and 'T1' (15) with the text 'according to their dates disco...'. At the bottom, there's a table with columns 'Title 1', 'Title 2', 'Title 3', and 'Title 4'.

FIGURE 5.5: Définition des conditions temporelles floues avec des bornes précises.

This screenshot shows the same web application interface as Figure 5.5, but with different inputs. The dropdown menu is set to 'are before this interval'. The text input field contains 'Six months ago', highlighted with a red diamond. The date input fields are 'I0' with 'Middle of March' and 'I1' with 'End of March', both highlighted with red rectangles. A red warning message is visible: 'Please specify only two dates if the condition will be considered fuzzy'. A modal message box is displayed in the center, containing an information icon and the text: 'You have entered a bound(s) with a fuzzy value. You will pass to an improved TSQL query.' with an 'OK' button. The rest of the interface, including the attribute list, tolerance settings, and table, remains the same as in Figure 5.5.

FIGURE 5.6: Définition des bornes temporelles floues.

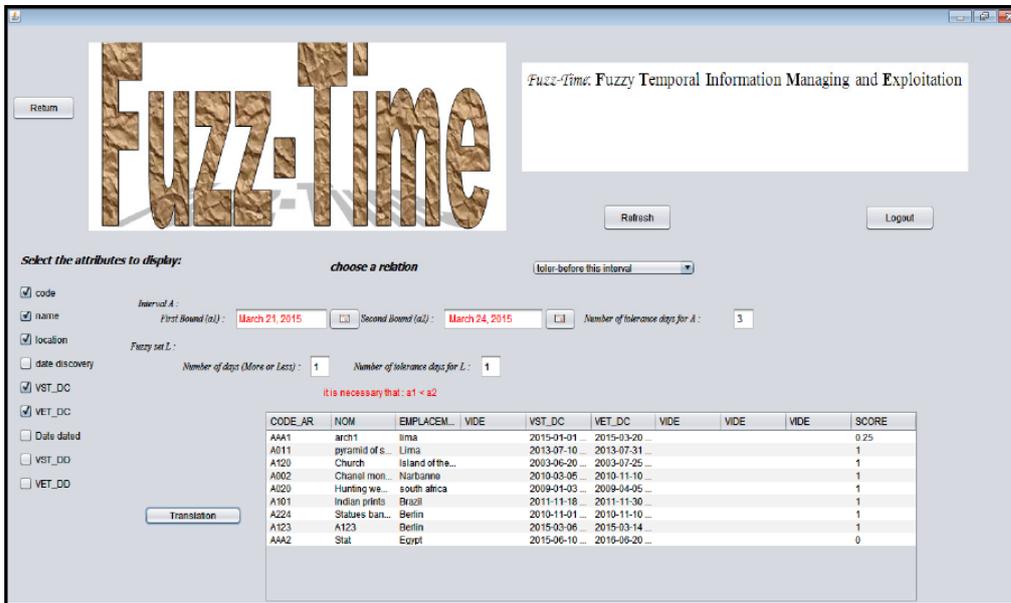


FIGURE 5.7: Résultat d'une requête *TSQLf* 1.

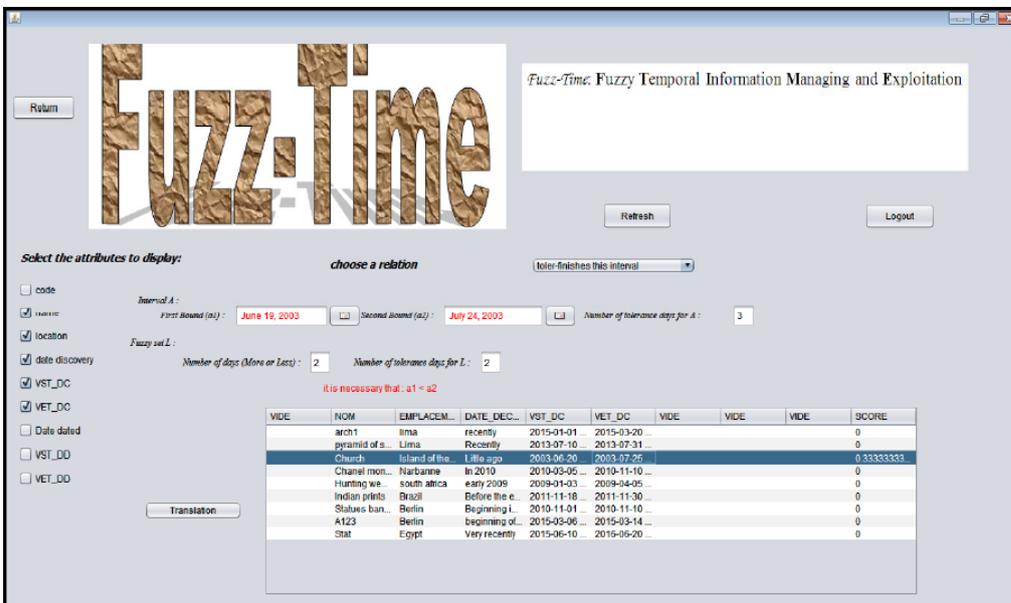


FIGURE 5.8: Résultat d'une requête *TSQLf* 2.

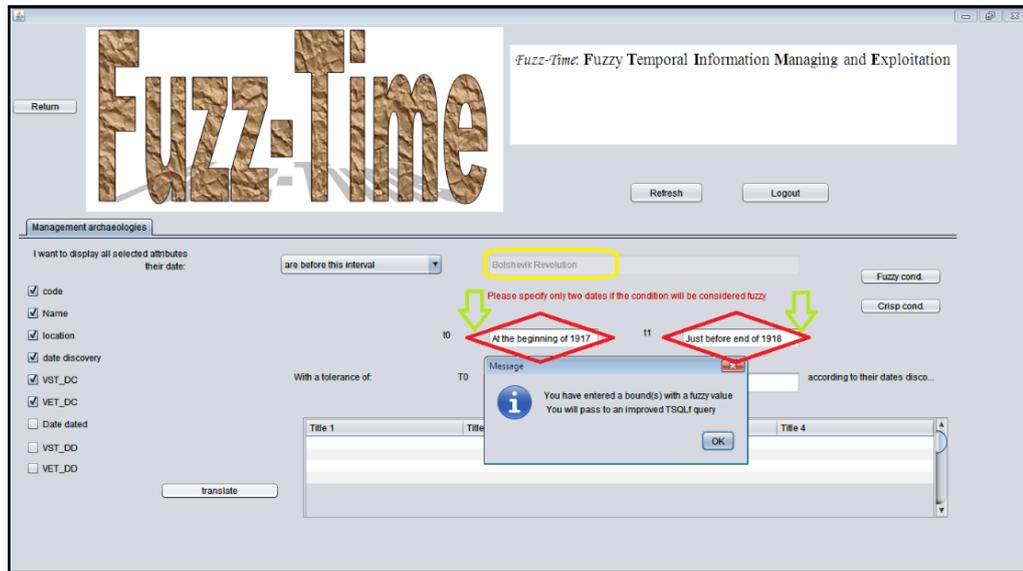


FIGURE 5.9: Définition des conditions temporelles floues.

### C. Allen Relation between Fuzzy Time Interval (ARFI)

Le système *Fuzz-TIME*, avec l'implémentation des relations temporelles d'Allen floues et les relations d'Allen tolérantes, est en mesure de répondre aux requêtes temporelles graduelles. En effet, si un utilisateur définit un intervalle de validité avec des bornes bien précises alors le système applique l'approche des relations temporelles d'Allen floues. Si la définition des bornes est effectuée d'une manière imprécise, alors le système applique l'approche des relations d'Allen tolérantes. Nous avons proposé dans le quatrième chapitre la modélisation des relations d'Allen entre les intervalles de temps flous. Cette proposition présente une autre alternative pour répondre aux requêtes des utilisateurs dont les bornes des intervalles de validité sont définies d'une façon imprécise. Ensuite, une requête *TSQlf* est générée et soumise à la base de données pour sélectionner les tuples qui correspondent aux critères temporels flous. Finalement, un degré de satisfaction est calculé et attribué à chaque tuple sélectionné (voir figure 5.10). L'outil *Fuzz-TIME* offre toujours à l'utilisateur la possibilité de définir une (des) condition(s) temporelle(s) floue(s) (voir figure 5.9). Puis, la définition des intervalles de validité et de tolérance pour chaque condition temporelle floue. Une fois les bornes des intervalles sont définies d'une manière imprécise, le système informe l'utilisateur qu'il est nécessaire d'introduire l'ensemble flou  $A$  qui correspond à la condition temporelle floue et l'ensemble flou  $L$ .

The screenshot shows the Fuzz-Time web application interface. At the top, there is a logo for 'Fuzz-Time' and a title 'Fuzz-Time: Fuzzy Temporal Information Managing and Exploitation'. Below the logo, there are search filters and a table of results.

**Select the attributes to display:**

- code
- name
- location
- date discovery
- VST\_DC
- VET\_DC
- Date dated
- VST\_DD
- VET\_DD

**Interval A:** First Bound (a1): June 19, 2003; Second Bound (a2): July 24, 2003; Number of tolerance days for A: 3

**Fuzzy set L:** Number of days (More or Less): 2; Number of tolerance days for L: 2

It is necessary that: a1 < a2

VIDE	NOM	EMPLACEMENT	DATE_DEC.	VST_DC	VET_DC	VIDE	VIDE	VIDE	SCORE
	arch1	lima	recently	2015-01-01 ...	2015-03-20				0
	pyramid of s.	Lima	Recently	2013-07-10 ...	2013-07-31				0
	Church	Island of the	Little ago	2003-06-20 ...	2003-07-25				0.33333333
	Chanel mon...	Narbanne	In 2010	2010-03-05 ...	2010-11-10				0
	Hurling wa...	south africa	early 2009	2009-01-03 ...	2009-04-05				0
	Indian prints	Brazil	Before the e...	2011-11-18 ...	2011-11-30				0
	Status ban...	Berlin	Beginning L...	2010-11-01 ...	2010-11-10				0
	A123	Berlin	beginning of...	2015-03-06 ...	2015-03-14				0
	Stat	Egypt	Very recently	2015-06-10 ...	2016-06-20				0

FIGURE 5.10: Résultat d'une requête *TSQLf*.

### 5.3 Exemple Illustratif

Pour mieux expliquer notre proposition, nous présentons ci-dessous un exemple du domaine d'archéologie. Considérons la table *Archéologie* (voir tableau 5.1) qui présente des vestiges matériels préhistoriques. Le schéma de la table est défini comme suit ;

*Archéologie* (*Code\_Ar*, *Nom\_Ar*, *Lieu*, *Date\_découverte*, *VST\_Dc*, *VET\_Dc*, *Date\_Daté*, *VST\_Dd*, *VET\_Dd*). Où *VST\_Dc* signifie la date de l'entrée en vigueur de la validité de *Date\_découverte*, *VET\_Dc* signifie la date de la fin de validité de la *Date\_découverte*, *VST\_Dd* signifie la date de commencement de la validité de *Date\_Daté* and *VET\_Dd* signifie la date de l'expiration de la date de validité *Daté*.

Nous montrons ci-après le fonctionnement du système *Fuzz-TIME* en proposant certaines requêtes particulières adressées à la base de données du domaine archéologique du tableau 5.1.

#### Q1 : Afficher les découvertes archéologiques qui ont eu lieu au cours de la révolution Bolchevique.

L'utilisateur a introduit dans cette requête une contrainte temporelle floue (à savoir, *révolution Bolchevique*), nous devons donc procéder à la génération d'une requête *TSQLf*. D'après le principe de requêtes *TSQLf*, le système doit demander à l'utilisateur de spécifier un intervalle de validité pour la contrainte temporelle floue proposée par l'utilisateur. L'utilisateur peut

TABLE 5.1: Table d'Archéologie.

Code_Ar	Name_Ar	Location	Date_Discovery	VST_Dc	VET_Dc	Date_Dated	VST_Dd	VET_Dd
A011	Pyramid of six meters in height	Lima	Recently	10/07/2013	31/07/2013	5000 years ago	2987 BC	2988 BC
A015	Cone rocks	Lake Tiberias	In 2003	05/02/2003	08/08/2003	2050 years ago	37 BC	38 BC
A120	Church	Island of the City	Little ago	20/06/2013	25/07/2013	End 158	08/09/158	18/12/158
A002	Chanel monumental	Narbanne	In 2010	05/03/2010	10/11/2010	2455 years ago	440 BC	442 BC
A020	Hunting weapon	South Africa	Early 2009	03/01/2009	05/04/2009	500000 years ago	497985 BC	497987 BC
A042	Mammoth Skeleton	Seine-et-Marne	Summer 2013	05/06/2013	01/09/2013	Before beginning 159	15/11/189	13/12/189
A075	Sort of mini-dinosaur	Africa	October 2012	02/10/2012	31/10/2012	Before the end of 260	20/09/260	15/12/260
A101	Indian prints	Brazil	Before the end of November 2011	18/11/2011	30/11/2011	More than 3000 years BC	3000 BC	3002 BC
A111	Corps soldiers Allemends	Carspach	Before the end of 2011	12/10/2011	15/12/2011	In 1918	10/07/1918	15/12/1918
A224	Statues banned by the Nazis	Berlin	Beginning in November 2010	01/11/2010	10/11/2010	During the second war mandial	01/09/1939	02/09/1945

définir un intervalle de validité avec deux bornes précises [*VST*, *VET*]. Par exemple, [*VST*, *VET*] = [02/01/1917, 20/01/1918]. Alors, la requête *TSQLf* équivalente à *Q1* s'écrit comme suit :

**Select \***

**From** *Archéologie*

**Where** *A.VST* > 01/02/1917 and *A.VET* < 02/01/1918 ;

**Q2 : Afficher les découvertes archéologiques qui ont eu lieu juste après la Seconde Guerre mondiale.**

Encore une fois, l'utilisateur a introduit dans *Q2* une contrainte temporelle floue (à savoir, *Seconde Guerre mondiale*). D'une manière similaire à celle ci-dessus, le système *Fuzz-TIME* génère la requête *TSQLf* qui correspond à *Q2*. Deux cas sont possibles :

(i) l'utilisateur définit un intervalle de validité avec deux bornes précises [*VST*, *VET*]. Par exemple [01/09/1939, 02/09/1945]. Donc, la requête *TSQLf* s'écrit comme suit :

**Select \***

**From** *Archéologie*

**Where** *A.VET* > 02/09/1949 ;

(ii) l'utilisateur définit un intervalle de validité avec deux bornes floues  $[VST, VET]$ . Par exemple,  $[Avant\ la\ fin\ de\ 1939, Juste\ avant\ fin\ de\ 1945]$ .

Les bornes de l'intervalle de validité sont alors floues. Par conséquent, la requête *TSQLf* générée devrait suivre une des deux alternatives (à savoir, le module *TAR* ou le module *ARFI*).

† **Cas du module *TAR*** : le système sollicite l'utilisateur de définir un ensemble flou  $A = (a, a', \alpha, \alpha')$  pour l'intervalle de validité flou et un paramètre de tolérance floue  $L = (-\delta, \delta, \epsilon, \epsilon)$ . Par exemple,  $A = (20/09/1939, 10/09/1945, 7, 5)$  et  $L = (-3, 3, 2, 1)$ . Ensuite, le système génère automatiquement deux intervalles (l'intervalle dilaté et l'intervalle érodé). Il génère par la suite une requête *TSQLf* en utilisant les relations d'Allen tolérantes proposées dans (Gammoudi, Hadjali, & Ben Yaghlane, 2016a).

‡ **Cas du module *ARFI*** : le système demande l'utilisateur de définir un ensemble flou  $\tilde{A} = [\tilde{a}, \tilde{a}']$  pour l'intervalle de validité flou. Par exemple :  $\tilde{a} = (1939, 1939, 3, 0)$  et  $\tilde{a}' = (1945, 1945, 0, 4)$ . Par la suite le système applique le principe de l'approche des relations d'Allen entre les intervalles temporelles floues proposées dans (Gammoudi, Hadjali, & Ben Yaghlane, 2016b). Dans ce cas, la relation temporelle d'Allen "After" définie entre deux intervalles temporelles flous sera utilisée.

**Q3 : Afficher les découvertes archéologiques qui ont eu lieu juste au début de l'indépendance française et pendant les années 40.**

L'utilisateur a introduit dans cette requête deux contraintes temporelles floues (*début de l'indépendance française* et *pendant les années 40*). Nous devons donc procéder à la génération d'une requête *TSQLf*. Dans ce cas l'utilisateur doit spécifier un intervalle de validité pour chaque spécification temporelle floue proposée.

Nous supposons que l'utilisateur définit les deux intervalles, tel que,  $[10 / 03/1946, 30/10/1950]$  pour la *période de l'indépendance française*, et  $[25/01/1940, 28/11/1949]$  pour *les années 40*. Il est à noter que nous pouvons raisonner sur les deux périodes en utilisant les deux relations temporelles floues suivantes :

$[A.VST, A.VET]$  *Fuzz-Stats*  $[10/03/1946, 30/10/1950]$  et  $[A.VST, A.VET]$  *Fuzz-During*  $[25/01/1940, 28/11/1949]$ .

A cette fin, **Q3** passe par une phase de raisonnement. Par conséquent, nous obtenons la requête *TSQLf* suivante :

```
Select *  
From Archéologie  
Where A.VST = 10/03/1946 and A.VET < 30/10/1950 ;
```

## 5.4 Conclusion

Nous avons présenté dans ce chapitre le système *Fuzz-TIME* développé pour la gestion des bases de données temporelles avec des spécifications incomplètes prenant en considération les critères de sélection flous de l'utilisateur pour chaque requête. Ce système a servi comme support pour valider les approches proposées et évaluer la pertinence de ses résultats. Il permet à l'utilisateur de spécifier sa requête et l'ensemble des critères (précis/flous) ainsi que les paramètres de sélection (le nombre de réponses désirées ou un seuil qualitatif ou les deux). Pour cela, nous avons développé une interface conviviale qui permet à l'utilisateur de définir sa requête.

Nous avons donné un aperçu sur l'architecture du système *Fuzz-TIME* qui est composé de plusieurs modules. Chaque module présente une fonctionnalité et les sous modules présentent les alternatives qu'une requête *TSQLf* devrait suivre.

Nous avons également présenté les expérimentations qui ont été réalisées et décrites pour évaluer l'efficacité et la pertinence des approches proposées. À travers ces expérimentations, nous avons montré que le système *Fuzz-TIME* peut suivre une phase de raisonnement pour gérer les requêtes sophistiquées. Le système développé est très pertinent, car il permet à l'utilisateur de définir en premier lieu les spécifications temporelles précises et/ou floues puis lui donner la main pour identifier les intervalles de validité de chaque spécification floue.

Enfin, nous avons conclut le chapitre par un exemple illustratif pour mieux expliquer notre proposition.

# Conclusion générale

Pour conclure ce manuscrit, nous présentons dans ce qui suit un résumé des travaux que nous avons réalisés au cours de cette thèse et qui portent, d'une manière générale, sur l'amélioration des modèles de gestion des informations imprécises dans les bases de données temporelles et donc de la pertinence des résultats retournés. Les perspectives de recherche qui pourraient être envisagées suite à ces travaux sont également discutées.

## Synthèse

Plus précisément, ces travaux contribuent à la modélisation et l'interrogation des bases de données temporelles contenant des informations imparfaites. Très peu de travaux, pour la gestion des données temporelles incomplètes dans le contexte des bases de données, existent dans la littérature. L'ensemble de ces travaux se fondent sous la théorie des ensembles flous.

Dans un premier temps, nous nous sommes intéressés à la problématique d'exploitation et la manipulation des informations temporelles mal-définies. Nous avons proposé une approche sur la définition d'un nouveau modèle de représentation de données temporelles entachées d'imprécision. Ce modèle a pris comme point de départ les travaux de (Dubois, Hadjali, & Prade, 2003) sur l'extension floue des relations temporelles d'Allen. Ces premiers travaux ont permis de développer un premier prototype nommé *Fuzz-TIME*. Nous avons pris en compte, dans cette proposition, le modèle proposé par (Galindo & J.M.Medina, 2001) en définissant pour chaque spécification temporelle imprécise, un intervalle de validité. Enfin, les expérimentations réalisées pour valider et évaluer l'approche proposée ont montré que le système *Fuzz-TIME* retourne des résultats pertinents à condition

que l'utilisateur définit des intervalles de validités précis. Alors que dans les applications du monde réel, les intervalles de validités sont généralement définis par des bornes graduelles.

Dans un second temps, nous avons proposé une deuxième approche composé (i) premièrement d'une nouvelle idée pour introduire quelques flexibilités dans la définition des relations entre deux intervalles de temps flous. Le concept clé de ce modèle est une relation de tolérance floue qui permet de définir les relations d'Allen tolérantes en utilisant les intervalles dilaté et érodé de l'intervalle initial. Des indices de comparaison flous sont utilisés pour identifier la position d'un intervalle de temps flou par rapport à un autre. Cette extension des relations d'Allen est intégrée dans notre système *Fuzz-TIME* déjà développé (ii) deuxièmement, d'une extension des relations temporelles d'Allen entre les intervalles de temps flous. Ce modèle exprime les différentes positions entre deux intervalles de temps flous suivant deux interprétations (conjonctive et disjonctive).

Nous nous sommes intéressés par la suite aux expérimentations qui ont été réalisées et décrites pour évaluer l'efficacité et la pertinence des approches proposées. À travers ces expérimentations, nous avons montré que le système *Fuzz-TIME* peut suivre une phase de raisonnement pour gérer les requêtes sophistiquées. Le système développé est très pertinent, car il permet à l'utilisateur de définir en premier lieu les spécifications temporelles précises et/ou floues puis lui donner la main pour identifier les intervalles de validité de chaque spécification floue.

### Perspectives

Les perspectives suscitées par ces travaux sont multiples. Une partie d'entre elles concernent l'établissement de l'ensemble complet des règles de compositions des relations temporelles d'Allen entre les intervalles de temps flous et la seconde vise à incorporer des phases de raisonnement dans notre système.

- **Définition des phases de raisonnement.** À travers le système *Fuzz-TIME*, nous avons montré la faisabilité et la pertinence de notre approche, en ajoutant des règles d'inférence. On peut raisonner sur la base des relations d'Allen floues établies par le calcul de la fermeture transitive des intervalles temporelles floues,
- **Enrichissement des expérimentations.** Un des aspects qui nous semble important à améliorer dans les expérimentations menées est d'étendre le

système *Fuzz-TIME* par des phases de raisonnement dans le cas des relations d'Allen tolérantes et aussi dans le cas de la sémantique disjonctive. Ceci permet, d'une part, de rendre notre système plus intelligent pour répondre aux requêtes des utilisateurs, d'autre part, de construire un benchmark plus complet et reflétant plus la réalité.

- **Prise en compte de l'incertitude temporelle.** Nous comptons aborder la question des relations temporelles incertaines dans l'esprit de notre récent travail (voir (Salem, Hadjali, Yaghlane, & Gammoudi, 2015)). L'incertitude est modélisée dans le cadre des fonctions de croyance.



# Références bibliographiques

- Aigner, W., Miksch, S., Thurnher, B., & Biffl, S. (2005). Planninglines : Novel glyphs for representing temporal uncertainties and their evaluation. *In Proc. of the 9th International Conference on Information Visualisation*, 457-463.
- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Comm. of the ACM*, 26, 832-843.
- Anselma, L., Terenziani, P., & Snodgrass, R. (2010). Valid-time indeterminacy in temporal relational databases : A family of data models. *In Proc. of 17th International Symposium on Temporal Representation and Reasoning*, 139-145.
- Asmussen, K., Qiang, Y., DeMaeyer, P., & DeWeghe, N. V. (2009). Triangular models for studying and memorising temporal knowledge. *In Proc. of the International Conference on Education, Research and Innovation IA-TED'09*, 1849-1859.
- Badaloni, S., Fade, M., & Giacomini, M. (2004). Integrating quantitative and qualitative fuzzy temporal constraints. *Journal of AI Communications*, 17, 187-200.
- Badaloni, S., & Giacomini, M. (2000). A fuzzy extension of allen's interval algebra. *In Proc. of LNAI, Springer*, 1792, 155-165.
- Badaloni, S., & Giacomini, M. (2002). Fuzzy extension of interval-based temporal sub-algebras. *In Proc. of the 9th Inter. Conf. on Information Processing and Management of Uncertainty in Knowledgebased Systems (IPMU'02)*, 1119-1126.
- Badaloni, S., & Giacomini, M. (2006). The algebra  $IA^{fuz}$  : a framework for qualitative fuzzy temporal. *Journal of Artificial Intelligence*, 170 (10), 872-902.
- Barro, S., Marin, R., Mira, J., & Paton, A. (1994). A model and a language for

- the fuzzy representation and handling of time. *Fuzzy Sets and Systems*, 17, 141-164.
- Bassiri, A., Malek, M., Alesheikh, A., & Amirian, P. (2009). Temporal relationships between rough time intervals. In : *Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y., Gavrilova, M.L. (eds.) ICCSA'09*, 5592, 543-552.
- Billiet, C., Firas, J. P., Capote, O. P., & Tré, G. D. (2013). A comparison of approaches to model uncertainty in time intervals. In *Advances in Intelligent Systems Research*, 626-633.
- Billiet, C., Firas, J. P., & Tré, G. D. (2013). Bipolarity in the querying of temporal databases. In *Proc. of SRI PAS/IBS PAN, new trends edn.*, 21-37.
- Billiet, C., Pons, J., Capote, O. P., & Tré, G. D. (2012). Evaluating possibilistic valid-time queries. In *Proc. of the 13th International Conference ON Information in Processing And Managment of Uncertainty in Knowledge-Based Systems IPMU 2012, Part I. CCIS*, 297, 410-419.
- Billiet, C., Pons, J., & Tré, G. D. (2013). Bipolar querying of valid-time intervals subject to uncertainty. In *Proc. of the 10th International Conference on Flexible Query Answering Systems (FQAS'2013)*, 8132, 401-412.
- Billiet, C., Pons, J. E., Matthé, T., Tré, G. D., & Capote, O. P. (2011). Bipolar fuzzy querying of temporal databases. In *Proc. of the 9th International Conference on Flexible Query Answering Systems*, 7022, 60-71.
- Billiet, C., & Tré, G. D. (2014a). Combining uncertainty and vagueness in time intervals. In *Angelov, P. (ed.) Intelligent Systems 2014. AISC*, 322, 353-364.
- Billiet, C., & Tré, G. D. (2014b). Twodimensional visualization of discrete time domain intervals subject to uncertainty. In *Proc. of the 6th International Conference on Fuzzy Computation Theory and Applications*, 137-145.
- Billiet, C., & Tré, G. D. (2015). The role of computational intelligence in temporal information retrieval : A survey of imperfect time in information systems. In *Proc. of the 13th International Conference on Flexible Query Answering Systems (FQAS'2015)*, 400, 41-756.
- Bordogna, G., Bucci, F., Carrara, P., Pagani, M., & Rampini, A. (2009). Extending inspiremetadata to imperfect temporal descriptions. *International Journal of Spatial Data Infrastructures Research, GSDI-11*, 250-264.
- Bordogna, G., Carrara, P., Pagani, M., Pepe, M., & Rampini, A. (2009). Managing imperfect temporal metadata in the catalog services of spatial data infrastructures compliant with inspire. In *Proc. of the Joint 2009 International Fuzzy Systems Associations World Congress and 2009 European Society of Fuzzy Logic and Technology Conference*, 915-920.
- Bosc, P., Galibourg, M., & Hamon, G. (1988). Fuzzy quering with sql : extensions

- and implementation aspects. *Journal of Fuzzy Sets and Systems*, 28, 333-349.
- Bosc, P., & Pivert, O. (1995). Sqlf : A relational database language for fuzzy querying. *IEEE Transactions on Fuzzy Systems*, 3, 1-17.
- Bosch, A., Torres, M., & Marín, R. (2002). Reasoning with disjunctive fuzzy temporal constraint networks. *In Proc. of the 9th International Symposium on Temporal Representation and Reasoning*, 36-43.
- Bron-selaer, A., Pons, J., Tré, G. D., & Pons, O. (2013). Possibilistic evaluation of sets. *In Proc. of International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 21(3), 325-346.
- Caluwe, R. D., Devos, F., Maesfranckx, P., Tré, G. D., & Cruyssen, B. V. D. (1999). The semantics and modelling of flexible time indications. *Studies in Fuzziness and Soft Computing*, chap. 11, 229-256.
- Chittaro, L., & Montanari, A. (2000). Temporal representation and reasoning in artificial intelligence. *Issues and approaches, Annals of Mathematics and Artificial Intelligence*, 28, 47-106.
- Chountas, P., & Petrounias, I. (2000). Modelling and representation of uncertain temporal information. *In Proc. of Requirements Engineering*, 5(3), 144-156.
- Cárdenas, M., Marin, R., & Navarette, I. (2001). Fuzzy temporal constraint logic : a valid resolution principle. *Journal of Fuzzy sets and Systems*, 117, 231-250.
- Dekhtyar, A., Ozcan, F., Ross, R., & Subrahmanian, V. (2001). Probabilistic temporal databases ii. *Calculus and query processing*, 255-263.
- Dekhtyar, A., Ross, R., & Subrahmanian, V. (2001). Probabilistic temporal databases i. *Algebra. ACM Transactions on Database Systems*, 26(1), 41-95.
- Deng, L., Liang, Z., & Zhang, Y. (2008). A fuzzy temporal model and query language for fter databases. *Eighth International Conference on Intelligent Systems Design and Applications*, 3, 77-82.
- Drakengren, T., & Jonsson, P. (1997). Towards a complete classification of tractability in allen's algebra. *In Proc. of IJCAI*, 1466-1471.
- Dubois, D., Fargier, H., & Galvagnon, V. (2003). On latest starting times and floats in activity networks with ill-known durations. *European Journal Of Operational Research*, 147, 266-280.
- Dubois, D., Fargier, H., & Prade, H. (1996). Possibility theory in constraint satisfaction problems : Handling priority, preference and uncertainty. *Journal of Applied Intelligence*, 6, 287-309.
- Dubois, D., Hadjali, A., & Prade, H. (2003). Fuzziness and uncertainty in tem-

- poral reasoning. *Journal of Universal Computer Science, Special Issue on Spatial and Temporal Reasoning* (H. Guesgen, F. Anger, G. Ligozat, R. Rodriguez, Edts.), 9(9), 1168-1194.
- Dubois, D., Hadjali, A., & Prade, H. (2004). Traitement de l'imprécision et de l'incertitude d'information temporelle. *Actes de la 14ème Conférence de Reconnaissance des Formes et Intelligence Artificielle (RFIA'2004)*, Toulouse, 103-112.
- Dubois, D., Hadjali, A., & Prade, H. (2007). A possibility theory-based approach to the handling of uncertain relations between temporal points. *International Journal of Intelligent Systems*, 22(2), 157-179.
- Dubois, D., Lang, J., & Prade, H. (1991). Timed possibilistic logic. *Fundamentae Informaticae*, 15, 211-234.
- Dubois, D., & Prade, H. (1983a). Inverse operations for fuzzy numbers. *proc. IFAC Symp. on Fuzzy Info., Knowledge representation and Decision Analysis*, 391-395.
- Dubois, D., & Prade, H. (1983b). Ranking fuzzy numbers in the setting of possibility theory. *Journal of Informations Sciences*, 30, 183-224.
- Dubois, D., & Prade, H. (1988). Possibility theory. *Plenum Press*, 326-345.
- Dubois, D., & Prade, H. (1989). Processing fuzzy temporal knowledge. *IEEE Trans. on Systems, Man and Cybernetics*, 19, 729-744.
- Dubois, D., Prade, H. M., & Prade, H. (2000). Fundamentals of fuzzy sets. *Springer*, 7, 601-621.
- Dyreson, C., & Snodgrass, R. (1998). Supporting valid-time indeterminacy. *ACM Transactions on Database Systems*, 23(1), 1-57.
- Freksa, C. (1992). Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54, 199-227.
- Frias, J. P., Billiet, C., Pons, O., & Tré, G. D. (1991). Aspects of dealing with imperfect data in temporal databases. In : Pivert, O., Zadrozny, S. (eds.) *Flexible Approaches in Data, Information and Knowledge Management. ACI*, 497, 189-220.
- Galindo, J., & J.M.Medina. (2001). Ftsql2 : Fuzzy time in relational databases. In *Proc. of the 2nd International Conference in Fuzzy Logic and Technology*, 5-7.
- Gammoudi, A., Hadjali, A., & Ben Yaghlane, B. (2014). An intelligent flexible querying approach for temporal databases. In *Proc. of the 7th international Conference on Intelligent Systems (IS'14)*, Springer, 322, 523-534.
- Gammoudi, A., Hadjali, A., & Ben Yaghlane, B. (2016a). An advanced tool for managing fuzzy complex temporal information. In *Proc. of the 6th In-*

- ternational conference on Computer Science and Information Technology (CCSIT 2016)*, 6 (1), 97-116.
- Gammoudi, A., Hadjali, A., & Ben Yaghlane, B. (2016b). Modelling temporal relations between fuzzy time intervals : A disjunctive view. *In Proc of the 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2016)*, 798, 50-57.
- Garrido, C., Marín, N., & Pons, O. (2009). Fuzzy intervals to represent fuzzy valid time in a temporal relational database. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 17(Suppl. 1), 173-192.
- Godo, L., & Vila, L. (1995). Possibilistic temporal reasoning based on fuzzy temporal constraints. *Proc. of the 14th Inter. Joint Conf. on Artificial Intelligence (IJCAI'95), Montréal*, 1916-1922.
- Guesgen, H. W., Hertzberg, J., & Philpott, A. (1994). Towards implementing fuzzy allen relations. *Proc. ECAI-94 Workshop on Spatial and Temporal Reasoning, Amsterdam, The Netherlands*, 49-55.
- Guil, F., Bosch, A., Bailón, A., & Marín, R. (2004). A fuzzy approach for mining generalized frequent temporal patterns. *In Proc. of the 4th IEEE International Conference on Data Mining, Workshop on Alternative Techniques for Data Mining and Knowledge Discovery*, 6.
- Kacprzyk, J., & Ziolkowski, A. (1986). Database queries with fuzzy linguistic quantifiers. *Journal of IEEE Transactions on Systems, Man and Cybernetics*, 16, 474-478.
- Kahn, K., & Gorry, G. (1977). Mechanizing temporal knowledge. *Artificial Intelligence*, 9(1), 87-108.
- Kalczynski, P., & Chou, A. (2005). Temporal document retrieval model for business news archives. *Information Processing Management*, 41, 635-650.
- Khatib, L. (1994). Reasoning with non-convex time intervals. *PhD thesis, Florida Institute of Technology, USA*, 255-271.
- Kopetz, H., & Kim, K. (1990). Temporal uncertainties in interactions among real-time objects. *In Proc. of the 9th Symposium on Reliable Distributed Systems*, 165-174.
- Ligozat, G. (1991). On generalized interval calculi. *In Proc. of AAI*, 234-240.
- Long, W. (1996). Temporal reasoning for diagnosis in a causal probabilistic knowledge base. *Artificial Intelligence in Medicine*, 8(3), 193-215.
- Meunier, B. B., & Yao, J. (1992). Linguistic modifiers and imprecise categories. *International Journal of Intelligent Systems*, 7, 25-36.
- Mitra, D., Gerard, M., Srinivasan, P., & Hands, A. (1994). A possibilistic interval constraint problem : Fuzzy temporal reasoning. *In Proc. of the 3rd IEEE*

- Conference on Fuzzy Systems*, 1434-1439.
- Nagypal, G., & Motik, B. (2003). A fuzzy model for representing uncertain, subjective, and vague temporal knowledge in ontologies. *In Proc. of OTM Confederated International Conferences On The Move to Meaningful Internet Systems*, 2888, 906-923.
- O'Connor, M., Tu, S., & Musen, M. (2000). Representation of temporal indeterminacy in clinical databases. *In Proc. of the AMIA Symposium*, 615-619.
- Ohlbach, H. (2004). Relations between fuzzy time intervals. *Proceedings of the 11th International Symposium on Temporal Representation and Reasoning*(44-51).
- Ortega-Binderberger, M., Chakrabarti, K., & Merhotra, S. (2002). An approach to integrating query refinement in sql. *In Proc. Int. Conf. on Extending Data Base Technology*, 15-33.
- Palma, J., Juarez, J., Campos, M., & Marin, R. (2004). A fuzzy approach to temporal modelbased diagnosis for intensive care units. *In Proc. of the 16th European Conference on Artificial Intelligence*, 868-872.
- Palma, J., Juarez, J., Campos, M., & Marin, R. (2006). Fuzzy theory approach for temporal model-based diagnosis : An application to medical domains. *Artificial Intelligence in Medicine*, 268-278.
- Petridis, S., Paliouras, G., & Perantonis, S. (2010). Allen's hourglass : probabilistic treatment of interval relations. *In Proc. of the 17th International Symposium on Temporal Representation and Reasoning (TIME'10)*, 589-602.
- Pivert, O., & Bosc, P. (2012). Fuzzy preference queries to relational databases. *Imperial College Press*, 330.
- Pons, J., Billiet, C., , Capote, O. P., & Tré, G. D. (2012). A possibilistic valid-time model. *In Proc. of the 13th International Conference on Information in Processing and Managment of Uncertainty in Knowledge-Based Systems IPMU 2012, Part I. CCIS*, 297, 420-429.
- Qian, D., & Lu, Y. (1994). A strategy of problem solving in a fuzzy reasoning network. *Fuzzy Sets and Systems*, 33, 137-154.
- Qiang, Y. (2012). Modelling temporal information in a two-dimensional space -avisualization perspective. *Phd thesis, Ghent University*, 546-561.
- Qiang, Y., Asmussen, K., Delafontaine, M., Stichelbaut, B., Tré, G. D., Maeyer, P. D., & Weghe, N. V. D. (2009). Visualising rough time intervals in a two-dimensional space. *In Proc. of IFSA World Congress/EUSFLAT Conference*, 1480-1485.
- Qiang, Y., Delafontaine, M., Asmussen, K., Delafontaine, M., Stichelbaut, B.,

- Maeyer, P. D., & Weghe, N. V. D. (2010). Modelling imperfect time intervals in a two-dimensional space\*. *In Proc. of Control And Cybernetics*, 39 (4), 983-1010.
- RobotJazi, M., Reformat, M. Z., Pedrycz, W., & Musilek, P. (2015). Lori : Linguistically oriented rdf interface for querying fuzzy temporal data. *In Proc. of the International Conference on Flexible Query Answering Systems, FQAS'15*, 400, 337-352.
- Ryabov, V. (2000). Uncertain relations between indeterminate temporal intervals. *In Proc. of the 10th International Conference on Management of Data*, 87-95.
- Ryabov, V. (2001). Probabilistic estimation of uncertain temporal relations. *Revista Colombiana de Computacion*, 2(2).
- Ryabov, V., & Puuronen, S. (2001). Probabilistic reasoning about uncertain relations between temporal points. *In Proc. of the 8th International Symposium on Temporal Representation and Reasoning (TIME'2001)*, IEEE Computer Society, 1530-1511.
- Ryabov, V., & Terziyan, V. (2003). Industrial diagnostics using algebra of uncertain temporal relations. *In Proc. of the 21st International Multi-Conference on Applied Informatics*, 351-356.
- Ryabov, V., & Trudel, A. (2004). Probabilistic temporal intervals networks. *In Proc. of the 11th International Symposium on Temporal Representation and Reasoning (TIME'2004)*, IEEE Computer Society, 64-67.
- Salem, N. E. H., Hadjali, A., Yaghlane, B. B., & Gammoudi, A. (2015). Une approche crédibiliste pour la gestion de l'incertitude temporelle. In (p. 243-249).
- Santos, J. E., & Young, J. (1999). Probabilistic temporal networks : A unified framework for reasoning with time and uncertainty. *Journal of Approximate Reasoning*, 20 (3), 263-291.
- Schockaert, S., & Cock, M. D. (2008). Temporal reasoning about fuzzy intervals. *Artificial Intelligence*, 172, 1158-1193.
- Schockaert, S., & Cock, M. D. (2009). Efficient algorithms for fuzzy qualitative temporal reasoning. *IEEE Transactions on Fuzzy Systems*, 17(4), 794-808.
- Schockaert, S., Cock, M. D., & Kerre, E. (2005). Imprecise temporal interval relations. *In Proc. Work. On Fuzzy Logic and applications (WILF)*, 3849, 108-113.
- Schockaert, S., Cock, M. D., & Kerre, E. (2006). Imprecise temporal interval relations. *In : Bloch, I., Petrosino, A., Tettamanzi, A.G.B. (eds.) WILF'05*, 3849, 108-113.

- Schockaert, S., Cock, M. D., & Kerre, E. (2008). Fuzzifying allen's temporal interval relations. *IEEE Transactions on Fuzzy Systems*, 16(2), 517-533.
- Tahani, V. (1977). A conceptuel framework for fuzzy query processing ; a step toward very intelligent database systems. *Journal of Information Processing and Management*, 13, 289-303.
- Tossebro, E., & Nygard, M. (2002). Uncertainty in spatiotemporal databases. In : *Yakhno, T. (ed.) ADVIS 2002. AIS*, 2457, 43-53.
- Tré, G. D., Bronselaer, A., Billiet, C., Qiang, Y., de Weghe, N. V., Maeyer, P. D., ... Pons, O. (2012). Visualising and handling uncertain time intervals in a twodimensional triangular space. In *Proc. of the Second World Conference on Soft Computing*, 585-592.
- Tré, G. D., Caluwe, R. D., & Cruyssen, B. V. D. (1997). Dealing with time in fuzzy and uncertain object-oriented database models. In *Proc of the EUFIT 97 Conference*, 1157-1161.
- Tré, G. D., de Weghe, N. V., de Caluwe, R., & Maeyer, P. D. (2006). Towards a flexible visualization tool for dealing with temporal data. In *Proc. of the 3rd International Conference on Flexible Query Answering Systems (FQAS'2006)*, 4027, 109-120.
- Trajcevski, G. (2003). Probabilistic range queries in moving objects databases with uncertainty. In *Proc of the 3rd ACM International Workshop on Data Engineering for Wireless and Mobile Access*, 39-45.
- Tudorie, C., Vlase, M., Nica, C., & Muntranu, D. (2012). Modeling fuzzy temporal criteria in database querying. In *Proc. of the 16th International Conference on System Theory, Control and Computing (ICSTCC'12)*, 112, 1-6.
- Vila, L. (1994). A survey on temporal reasoning in artificial intelligence. *AI Communications*, 7(1), 4-28.
- Vila, L., & Godo, L. (1994). On fuzzy temporal constraint networks. *Mathware Soft Computing*, 1, 315-334.
- Vilain, N., & Kautz, H. (1986). Constraint propagation algorithms for temporal reasoning. In *Proc.5th Nat. Conf. of the American Association for AI (AAAI'86)*, Morgan Kaufmann, 377-382.
- Virant, J., & Zimic, N. (96). Attention to time in fuzzy logic. *Fuzzy Sets and Systems*, 82 (1), 39-49.
- Wainer, J., & Sandri, S. (1998). Fuzzy temporal categorical and intensity information in diagnosis. *Lectures Notes in AI*, 1515, 181-190.
- Zhang, H., Diao, Y., & Immerman, N. (2010). Recognizing patterns in streams with imprecise timestamps. In *Proc. of the VLDB Endowment*, 3(1-2), 244-255.

- Zhang, K., & Trudel, A. (2006). Efficient heuristics for solving probabilistic interval algebra networks. *In Proc. of the 13th International Symposium on Temporal Representation and Reasoning (TIME'06)*, 111-120.