



HAL
open science

Traitement de l'imperfection des données temporelles saisies par l'utilisateur - Application aux logiciels destinés à des patients malades d'Alzheimer

Nassira Achich

► **To cite this version:**

Nassira Achich. Traitement de l'imperfection des données temporelles saisies par l'utilisateur - Application aux logiciels destinés à des patients malades d'Alzheimer. Informatique [cs]. HESAM Université; L'Université de Sfax, 2021. Français. NNT: . tel-03675221v1

HAL Id: tel-03675221

<https://theses.hal.science/tel-03675221v1>

Submitted on 29 Dec 2021 (v1), last revised 23 May 2022 (v2)

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.

HESAM
UNIVERSITÉ



le cnam



ÉCOLE DOCTORALE SCIENCES DES METIERS DE L'INGENIEUR
[Centre d'Etude et De Recherche en Informatique et Communications]
ÉCOLE DOCTORALE EN SCIENCES ECONOMIQUES, GESTION ET
INFORMATIQUE
[Multimedia, Information Systems and Advanced Computing Laboratory]

THÈSE

présentée par : **Nassira ACHICH**
soutenue le : **09 Décembre 2021**

pour obtenir le grade de : **Docteur d'HESAM Université et de L'Université de Sfax**

préparée au : **Conservatoire national des arts et métiers et Faculté des sciences économiques et de gestion de Sfax**
Discipline/S spécialité : **Informatique**

Traitement de l'imperfection des données temporelles saisies par l'utilisateur - Application aux logiciels destinés à des patients malades d'Alzheimer

THÈSE dirigée par :
Pr. GARGOURI Faiez
Pr. METAIS Elisabeth

et co-encadrée par :
Dr. GHORBEL Fatma
Dr. HAMDI Fayçal

Jury

M. HADJ KACEM Ahmed	Professeur, Université de Sfax, Tunisie	Rapporteur
M. MEZIANE Farid	Professeur, Université de Derby, Royaume Uni	Rapporteur
M. DU MOUZA Cedric	Maître de Conférences - HDR, CNAM, France	Président
Mme KEDAD Zoubida	Maître de Conférences - HDR, UFR des sciences, Université de Versailles, France	Examinatrice
M. SILBERZTEIN Max	Professeur, Université de Franche Comté, France	Examineur
Mme ZANNI-MERK Cecilia	Professeur, INSA Rouen Normandie, France	Examinatrice

**T
H
È
S
E**

Affidavit

Je soussignée, Nassira Achich, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Pr. Elisabeth Métais et Pr. Faiez Gargouri, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect de la charte nationale de déontologie des métiers de la recherche.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Paris, le 20 Octobre 2021

Signature



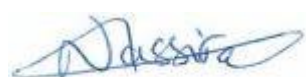
Affidavit

I, undersigned, Nassira Achich, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Pr. Elisabeth Métais and Pr. Faiez Gargouri, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with the French charter for Research Integrity.

This work has not been submitted previously either in France or abroad in the same or in a similar version to any other examination body.

Paris, October 20th, 2021

Signature



Dédicace

Aujourd'hui, je serai Docteur. Et ce n'est qu'un premier pas.
C'est le fruit d'un long labeur, de grands efforts et de grands sacrifices

Maman **Wafa**, Papa **Slim**

Ma sœur, ma petite, **Zahra** et son Mari **Amine**

Mon ange, ma nièce, **Kmar**

Mon chéri **Mohamed**

Vous êtes ma vie, toute ma vie.

A toute ma famille et tous ceux qui m'aiment et se soucient de ma réussite

Je dédie ce travail.

Remerciements

Je tiens à exprimer ma profonde gratitude avec un immense respect à toutes les personnes qui, sans leur soutien et générosité, je n'aurais jamais pu réaliser cette thèse.

Je remercie, du fond du cœur, ma directrice de thèse, **Pr. Elisabeth Métais**, pour la confiance qu'elle m'a accordée en acceptant de diriger ma thèse. Elle m'avait guidée dans mes travaux, avec une très grande expertise. J'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de mon travail de doctorat. Elle était toujours présente malgré la distance. Je lui suis reconnaissante pour toute ma vie.

Je remercie chaleureusement mon directeur de thèse, **Pr. Faiez Gargouri**, pour toutes ses qualités scientifiques et humaines. J'avais l'honneur et le privilège de travailler sous sa direction et d'apprendre et de bénéficier de sa grande compétence, son aide inestimable, sa générosité, la bonne humeur et l'intérêt manifesté à l'égard de ma recherche.

J'aime bien remercier, et témoigner ma grande reconnaissance, à **Dr. Fatma Ghorbel**. Elle est mon professeur, un modèle pour moi et un exemple de sérieux à suivre. J'aimerais lui dire Merci pour l'intérêt manifesté à l'égard de mon travail, l'effort et l'expertise que vous m'avez versée et toutes les heures consacrées pour m'encadrer tout au long ma thèse. J'apprécie vivement sa grande disponibilité et son accueil même durant la crise sanitaire du covid. Elle m'a submergé de son humanité infinie. Elle m'était une sœur et une amie et le sera toujours. 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 remercie vivement **Dr. Fayçal Hamdi**, de m'avoir encadrée. Je lui suis reconnaissante pour ses conseils précieux, pour tous ses efforts fournis et toutes ses qualités scientifiques et humaines. Merci pour la confiance que vous m'avez accordée. J'ai pris un grand plaisir à travailler avec vous et j'ai trop appris de votre expertise et votre compétence.

Qu'il me soit permis d'exprimer mes sincères remerciements à **Pr. HADJ KACEM Ahmed** et **Pr. MEZIANE Farid** pour accepter d'être mes rapporteurs de thèse. J'exprime également toute ma gratitude à **Dr. DU MOUZA Cedric**, **Dr. KEDAD Zoubida**, **Pr. SILBERZTEIN Max** et **Pr. ZANNI-MERK Cecilia**, qui m'ont fait l'honneur d'avoir accepté d'examiner ce travail.

Je remercie **l'université de Sfax, Tunisie**, le **Conservatoire Nationale des Arts et Métiers de Paris, France (CNAM)** et les laboratoires **CEDRIC** et **MIRACL** de leur soutien pour que cette thèse soit en cotutelle.

Je remercie aussi les personnels de la **maison de la retraite de Sfax en Tunisie** et les **malades d'Alzheimer** qui ont participé aux évaluations effectuées dans le cadre de cette thèse.

Résumé

Le besoin de traiter l'imperfection des données temporelles dans le Web sémantique augmente et, par conséquent, appelle à une manière standard de représenter et de raisonner sur telles données. Cette thèse présente une nouvelle vision visant à fournir des solutions pour différents types d'imperfections affectant les données temporelles.

Notre thèse comprend plusieurs contributions. La première porte sur **la définition d'une typologie de l'imperfection des données temporelles** dans laquelle nous listons les différentes imperfections qui peuvent affecter les données temporelles. Nous les classifions en imperfections directes et imperfections indirectes. Par la suite, nous **traitons les différentes imperfections définies dans la typologie**. Nous proposons une approche qui **traite l'imprécision des données temporelles** en représentant et raisonnant sur les dates et les horloges du temps dans l'ontologie avec **une vue précise**. Pour cela, nous **étendons l'approche du 4D-Fluents et l'algèbre d'intervalles d'Allen**. Nous proposons **une ontologie** basée sur nos extensions nommée « *CrispImpTimeOnto* ». Nous traitons **l'incertitude des données temporelles** dans l'ontologie, tout d'abord, en utilisant **les réseaux bayésiens pour calculer les mesures d'incertitude** et la deuxième en utilisant **la théorie des possibilités pour calculer les mesures de nécessité**. Nous étendons l'approche du 4D-Fluents pour la représentation et l'Algèbre d'intervalle d'Allen pour le raisonnement pour les deux approches. Nous proposons **deux ontologies nommées** « *UncertTimeOnto* » et « *UncertPossibTimeOnto* ». Nous conduisons une **comparaison entre les deux approches proposées**. Nous passons par la suite **aux types d'imperfection composite**. Nous présentons une contribution qui consiste à traiter, **à la fois, l'incertitude et l'imprécision des données temporelles** dans l'ontologie en utilisant **la théorie des possibilités**. Nous proposons **une ontologie**, nommée « *UncertPossibTimeOnto* ». Nous traitons **le conflit des données temporelles** dans l'ontologie en utilisant **la théorie de l'évidence** pour calculer **les masses de croyances des données**. Nous proposons une représentation sémantique en nous basant sur les mesures. Nous étendons l'algèbre d'Allen pour raisonner sur les données temporelles conflictuelles. Nous proposons une ontologie, nommée « *BeliefTimeOnto* » basée sur nos extensions. Sur la base de nos ontologies proposées, nous proposons **une ontologie de l'imperfection des données temporelles nommée** « *TemporalOntoImperfection* ».

Nous intégrons notre ontologie temporelle « *TemporalOntoImperfection* » dans la prothèse de mémoire **CAPTAIN MEMO** et nous évaluons nos approches traitant les différents

types d'imperfection. Cette évaluation fait intervenir des personnes atteintes de la maladie d'Alzheimer.

Mots-clés : Ontologie, Imperfection des données temporelles, théorie de possibilité, théorie de l'évidence, réseaux bayésiens, l'approche 4D-Fluents, l'algèbre d'intervalles d'Allen, Alzheimer.

Résumé en anglais

The need to deal with temporal data imperfection in the Semantic Web is increasing and, therefore, calls for a standard way of representing and reasoning about such data. This thesis presents a new vision aiming to provide solutions for different types of imperfections affecting temporal data, which we first define in a typology that we propose.

Our thesis includes several contributions. The first one consists in defining a typology of temporal data imperfection in which we elaborate the various imperfections that can affect temporal data. We classify them into direct imperfections and others that are indirect. Then, we move to dealing with these imperfections in ontology. We propose an approach to deal with imprecision by representing and reasoning about dates and time clocks and associated qualitative relations in ontology with a crisp view. Therefore, we extend the 4D-Fluents approach for the representation and the Allen's interval algebra for the reasoning. Then, we create an ontology based on our extensions called « *CrispImpTimeOnto* ». When dealing with uncertain temporal data in ontology and associated qualitative relations, we first propose an approach in a certain environment. We calculate certainty degrees using Bayesian networks. We extend the 4D-Fluents approach to represent the handled data and extending Allen's interval algebra for the reasoning part. Second, we propose an approach base on possibility theory to calculate the necessity measures of the handled data and by extending the same approaches for representation and reasoning. Then, we conduct a small comparison between the two proposed approaches. We offer an ontology based on our extensions called « *UncertTimeOnto* ». Temporal data may, also, be affected by multiple types of imperfections at the same time. Thus, we present a contribution which consists in dealing with simultaneously uncertain and imprecise temporal data in ontology using possibility theory. We propose an ontology, named « *UncertPossibTimeOnto* », based on our extensions. We deal with the conflict of temporal data in ontology by using the evidential theory to calculate the belief masses of the handled data. We propose a semantic representation based on the obtained measures. We also extend Allen's algebra to reason about conflicting temporal data. We propose an ontology, named « *BeliefTimeOnto* » based on our extensions. Based on the proposed ontologies in the previous chapters, we propose an ontology of temporal data imperfection called « *TemporalOntoImperfection* » that encompasses all the treated imperfections.

Our contributions are implemented and evaluated within CAPTAIN MEMO, the memory prosthesis based on the PersonLink ontology, dedicated to Alzheimer's patients. We integrate

our temporal ontology « *TemporalOntoImperfection* » in the prosthesis and we evaluate our approaches dealing with the different types of imperfection. This assessment involves people with Alzheimer's disease.

Keywords: Ontology, Temporal Data Imperfection, Possibility Theory, Evidential Theory, Bayesian Networks, 4D-Fluents Approach, Allen's Interval Algebra, Alzheimer

Table des matières

CHAPITRE 1 : Introduction générale.....	21
1.1. Contexte du travail	22
1.2. Problématiques	23
1.3. Contributions.....	25
1.4. Liste des publications	30
1.5. Organisation du mémoire.....	31
CHAPITRE 2 : Notions théoriques et état de l’art	33
2.1. Introduction.....	34
2.2. Notion d’ontologie.....	34
2.2.1 Web sémantique	34
2.2.2 Définition	35
2.2.3 Composants	37
2.2.4 Types d’ontologie.....	37
2.2.4.1 Ontologie classique (parfaite).....	38
2.2.4.2 Ontologie imparfaite	38
2.2.4.3 Ontologies temporelles.....	40
2.3. L’imperfection des données	42
2.3.1 Typologies de l’imperfection des données	43
2.3.1.1 Typologies génériques de l’imperfection des données	43
2.3.1.2 Typologies spécifiques de l’imperfection des données dans un domaine donné.....	44
2.3.1.3 Discussion	50
2.3.2 Théories de l’imperfection des données.....	50
2.3.2.1 Théorie de probabilité.....	51
2.3.2.2 Théorie de possibilité.....	53
2.3.2.3 Théorie de probabilité et théorie de possibilité : comparaison	54
2.3.2.4 Théorie de l’évidence.....	56

2.4.	Les données temporelles	57
2.4.1	Éléments, relations et ontologies temporelles	57
2.4.1.1	Éléments temporels	57
2.4.1.2	Relations temporelles	58
2.4.2	Représentation et raisonnement sur les données temporelles dans le web sémantique	59
2.4.2.1	Représentation des données temporelles dans le Web Sémantique	60
2.4.2.2	Raisonnement sur des données temporelles dans le Web Sémantique	65
2.4.2.3	Résumé et critiques	66
2.5.	Conclusion	72
CHAPITRE 3 : Typologie de l'imperfection des données temporelles		73
3.1.	Introduction	74
3.2.	Échantillon de données saisies dans le cadre de CAPTAIN MEMO	74
3.2.1	Présentation d'un exemple illustratif	74
3.2.2	Analyse : dégager les différents types d'imperfection	76
3.3.	Notre typologie de l'imperfection des données temporelles	77
3.3.1	Les imperfections directes	77
3.3.2	Les imperfections indirectes	78
3.3.3	Dépendance du contexte général des données temporelles	79
3.4.	Conclusion	80
CHAPITRE 4 : Traitement des données temporelles précises et imprécises dans OWL 2 : dates et horloges de temps		83
4.1.	Introduction	84
4.2.	Représentation des données temporelles précises et imprécises en ontologie	85
4.2.1	Données temporelles quantitatives	85
4.2.1.1	Points de temps précis et imprécis	85
4.2.1.2	Intervalles de temps imprécis	87
4.2.2	Données temporelles qualitatives	87
4.3.	Raisonnement sur les données temporelles précises et imprécises	88
4.3.1	Relations temporelles qualitatives	88

4.3.1.1	Relations entre Intervalles de temps	89
4.3.1.2	Relations entre un intervalle de temps et un point de temps	90
4.3.1.3	Relations entre les points de temps	91
4.3.2	Table de transitivité	92
4.4.	« CrispTimeOnto » : notre ontologie des données temporelles précises et imprécises	95
4.5.	Expérimentations.....	96
4.5.1	Prototype implémenté.....	96
4.5.2	Application dans CAPTAIN MEMO	97
4.6.	Conclusion	97
CHAPITRE 5 : Traitement des données temporelles certaines et incertaines dans OWL 2 : approche basée sur la théorie de probabilité.....		99
5.1.	Introduction.....	100
5.2.	Représentation des données temporelles certaines et incertaines dans OWL 2	101
5.2.1	Données temporelles quantitatives	101
5.2.1.1	Intervalles de temps incertains	101
5.2.1.2	Points de temps certains et incertains.....	102
5.2.2	Données temporelles qualitatives	102
5.3.	Raisonnement sur les données temporelles certaines et incertaines dans OWL 2	104
5.3.1	Relations temporelles qualitatives	104
5.3.1.1	Relations entre les intervalles de temps	104
5.3.1.2	Relations entre un intervalle de temps et un point de temps	108
5.3.1.3	Relations entre les points de temps	109
5.3.2	Tables de transitivité.....	110
5.4.	« UncertTimeOnto » : notre ontologie des données temporelles incertaines en OWL 2.....	112
5.5.	Expérimentation	114
5.5.1	Prototype implémenté.....	115
5.5.2	Application dans CAPTAIN MEMO	116
5.6.	Conclusion	116
CHAPITRE 6 : Traitement des données temporelles incertaines dans OWL 2 : Approche basée sur la théorie de possibilité		119

6.1. Introduction.....	120
6.2. Représentation des données temporelles incertaines dans OWL 2 : extension de l’approche 4D-Fluents.....	120
6.2.1 Données temporelles quantitatives	121
6.2.1.1 Intervalles de temps	121
6.2.1.2 Les points de temps.....	121
6.2.2 Représentation qualitative des données temporelles	121
6.2.3 Degré de nécessité	122
6.3. Raisonnement sur les données temporelles incertaines : extension de l’algèbre d’intervalle d’Allen.....	123
6.3.1 Relations temporelles qualitatives entre les intervalles de temps.....	123
6.3.2 Relations temporelles qualitatives entre un intervalle de temps et un point de temps	124
6.3.3 Relations temporelles qualitatives entre deux points de temps	125
6.3.4 Tables de transitivité.....	125
6.4. «UncertPossibTimeOnto»: notre ontologie temporelle pour les données temporelles incertaines	129
6.5. Expérimentations.....	130
6.5.1 Prototype implémenté.....	130
6.5.2 Étude de cas : application à l’ontologie PersonLink.....	131
6.5.3 Comparaison entre l’approche possibiliste et l’approche probabiliste	133
6.6. Conclusion	134
CHAPITRE 7 : Traitement des données temporelles à la fois incertaines et imprécises dans OWL2.....	137
7.1. Introduction.....	138
7.2. Mesures de possibilité et de nécessité	139
7.2.1 Mesures de possibilité et de nécessité des intervalles de temps	139
7.2.2 Mesures de possibilité et de nécessité des points de temps.....	142
7.3. Représentation des données temporelles incertaines et imprécises dans OWL 2	144
7.3.1 Données temporelles quantitatives	144
7.3.1.1 Intervalles de temps	144

7.3.1.2	Points de temps	145
7.3.2	Données temporelles qualitatives	146
7.4.	Raisonnement sur des données temporelles incertaines et imprécises dans OWL 2	147
7.4.1	Relations temporelles qualitatives	147
7.4.1.1	Relations entre deux intervalles de temps	147
7.4.1.2	Relations entre un intervalle de temps et un point de temps	149
7.4.1.3	Relations entre deux points de temps	149
7.4.2	Table de transitivité	150
7.5.	« PossibilisticTemporalOnto » : notre ontologie temporelle possibiliste pour les données temporelles à la fois incertaines et imprécises	153
7.6.	Expérimentations	154
7.6.1	Prototype implémenté	154
7.6.2	Étude de cas : Application aux patients atteints d'Alzheimer	155
7.7.	Conclusion	158
CHAPITRE 8 : Traitement des données temporelles conflictuelles dans OWL 2		161
8.1.	Introduction	162
8.2.	Masses de croyance pour les données temporelles quantitatives	163
8.2.1	Intervalles de temps	163
8.2.2	Points de temps	165
8.3.	Représentation des données temporelles conflictuelles dans OWL 2	166
8.3.1	Données temporelles quantitatives	166
8.3.2	Données temporelles qualitatives	167
8.4.	Raisonnement sur des données temporelles conflictuelles dans OWL 2	167
8.4.1	Relations temporelles qualitatives	168
8.4.1.1	Relations entre intervalles de temps	168
8.4.1.2	Relations entre un intervalle de temps et un point de temps	169
8.4.1.3	Relations entre points de temps	170
8.4.2	Tables de transitivité	170
8.4.	« BeliefTimeOnto » : notre ontologie temporelle évidentielle pour les données temporelles conflictuelles	172

8.5. Expérimentation	173
8.5.1 Prototype implémenté.....	173
8.5.2 Etude de cas : application dans CAPTAIN MEMO	174
8.6. Conclusion	178
CHAPITRE 9 : Ontologie de l'imperfection des données temporelles et évaluation dans CAPTAIN MEMO.....	179
9.1. Introduction.....	180
9.2. Ontologie de l'imperfection des données temporelles.....	180
9.2.1 Eléments de « TimeOntoImperfection »	181
9.2.2 Prototype basé sur notre ontologie « TimeOntoImperfection »	184
9.3. Evaluation dans le cadre de CAPTAIN MEMO	188
9.4. Conclusion	190
CHAPITRE 10 : Conclusion générale et Perspectives.....	191
10.1. Nos contributions	192
10.2. Perspectives.....	195
REFERENCES BIBLIOGRAPHIQUES	197

Liste des tableaux

Tableau 1. Tableau récapitulatif des typologies de l'imperfection des données existantes	47
Tableau 2. Théories et type d'imperfection(s) traité(s)	51
Tableau 3. Théorie de possibilité VS théorie de probabilité.....	55
Tableau 4. Relations d'Allen entre deux intervalles de temps précis.	59
Tableau 5. Algèbre des points de Vilain et Kautz.....	59
Tableau 6. Modélisation de la dimension temporelle dans l'ontologie.....	68
Tableau 7. Résumé des recherches qui représentent et raisonnent sur des données temporelles dans l'ontologie	70
Tableau 8. Relations temporelles entre les intervalles de temps imprécis A et B	89
Tableau 9. Relations temporelles entre un intervalle de temps A et un point P	90
Tableau 10. Relations temporelles entre les points temporels N et R.....	91
Tableau 11. Table de transitivité pour raisonner $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$	93
Tableau 12. Table de transitivité pour raisonner sur $R_3(I, J)$ à partir de $R_1(I, P)$ et $R_2(P, J)$..	93
Tableau 13. Table de transitivité pour raisonner $R_3(P, M)$ à partir de $R_1(P, L)$ et $R_2(L, M)$	95
Tableau 14. Relations temporelles entre deux intervalles de temps incertains A et B.....	105
Tableau 15. Relations temporelles entre un intervalle de temps A et un point de temps P...108	
Tableau 16. Relations temporelles entre deux points de temps P et Q	109
Tableau 17. Table de transitivité qui permet de déduire $R_3(P, Q)$ de $R_1(P, A)$ et $R_2(A, Q)$	110
Tableau 18. Table de transitivité qui permet de déduire $R_3(A, B)$ de $R_1(A, P)$ et $R_2(P, B)$	111
Tableau 19. Table de transitivité qui permet de déduire $R_3(P, M)$ à partir de $R_1(P, Q)$ et $R_2(Q, M)$	112
Tableau 20. Relations temporelles entre deux intervalles de temps incertains A et B.....	124

Tableau 21. Relation temporelle entre un intervalle de temps incertain A et un point de temps incertain P	125
Tableau 22. Relations temporelles entre les points de temps incertains P_{Np} and Q_{Nq}	125
Tableau 23. Table de transitivité qui permet de déduire $R3(P, Q)$ de $R1(P, A)$ et $R2(A, Q)$	126
Tableau 24. Table de transitivité qui permet de déduire $R3(A, B)$ de $R1(A, P)$ et $R2(P, B)$	127
Tableau 25. Table de transitivité qui permet de déduire $R3(P, M)$ à partir de $R1(P, Q)$ et $R2(Q, M)$	128
Tableau 26. Degrés de certitude obtenus à partir des théories de possibilité et de probabilité	134
Tableau 27. Relations temporelles entre les intervalles de temps à la fois incertains et imprécis I et J	148
Tableau 28. Relations temporelles entre un intervalle de temps I conjointement incertain et imprécis et un point de temps P conjointement incertain et imprécis	149
Tableau 29. Relations temporelles entre les points temporels conjointement incertains et imprécis P et Q.....	150
Tableau 30. Table de transitivité permettant de raisonner $TR3(P, L)$ à partir de $TR1(P, I)$ et $TR2(I, L)$	151
Tableau 31. Table de transitivité permettant de raisonner $TR3(P, L)$ à partir de $TR1(P, I)$ et $TR2(I, L)$	151
Tableau 32. Table de transitivité qui permet de raisonner $TR3(P, M)$ à partir de $TR1(P, L)$ et $TR2(L, M)$	152
Tableau 33. Les différents Intervalles I_i possibles	155
Tableau 34. Les différents Intervalles J_j possibles	155
Tableau 35. Mesures de possibilité et de nécessité de l'intervalle de temps conjointement incertain et imprécis I	156
Tableau 36. Mesures de possibilité et de nécessité de l'intervalle de temps conjointement incertain et imprécis J	157
Tableau 37. Extension de l'algèbre des intervalles d'Allen pour les relations temporelles qualitatives	169
Tableau 38. Table de transitivité qui permet de raisonner $R3(P, L)$ à partir de $R1(P, I)$ et $R2(I, L)$	170

Tableau 39. Table de transitivité qui permet de raisonner $R3(P, L)$ à partir de $R1(P, I)$ et $R2(I, L)$.	171
Tableau 40. Table de transitivité qui permet de raisonner $R3(P, M)$ à partir de $R1(P, L)$ et $R2(L, M)$.	172
Tableau 41. Réponses du patient Alzheimer	175
Tableau 42. Mesures de crédibilité des points de temps basées sur DBE_ALZ	175
Tableau 43. Les masses de croyance des points de temps	176
Tableau 44. Fonction de crédibilité des intervalles de temps I_i	176
Tableau 45. Masses de croyance des intervalles de temps I_i	177
Tableau 46. Précision des résultats de l'évaluation de notre approche	189

Liste des figures

Figure 1. Représentation pyramidale du web sémantique (Berners-Lee et al., 2001).....	35
Figure 2. Exemple de l'approche classique du 4D-Fluents	67
Figure 3. Échantillon fictif des réponses entrées par un malade d'Alzheimer.....	75
Figure 4. Notre typologie des de l'imperfection des données temporelles	81
Figure 5. Notre extension de l'approche 4D-Fluents pour représenter des dates et horloges de temps précis et imprécis.....	88
Figure 6. Prototype proposé basé sur notre ontologie	96
Figure 7. Notre extension de l'approche 4D-Fluents pour représenter des données temporelles certaines et incertaines.....	103
Figure 8. Réseau bayésien associé à la relation «Before _c » entre deux intervalles	105
Figure 9. Réseau bayésien associé à la relation «Overlaps _c » entre deux intervalles de temps incertains.....	106
Figure 10. Réseaux bayésiens associés à l'Exemple	107
Figure 11. Notre ontologie des données temporelles incertaines "UncertTimeOnto"	113
Figure 12. Exemple d'une règle SWRL.....	113
Figure 13. Notre ontologie crée par l'éditeur Protégé.....	Erreur ! Signet non défini.
Figure 14. L'interface utilisateur du prototype implémenté	115
Figure 15. Un exemple d'ontologie PersonLink basée sur notre approche	116
Figure 16. Notre extension du 4D-fluents pour représenter des données temporelles incertaines avec la théorie de possibilité	122
Figure 17. Règles SWRL de "UncertPossibTimeOnto"	129
Figure 18. Prototype basé sur « UncertPossibTimeOnto ».....	131
Figure 19. Distributions de possibilité de chaque borne de l'intervalle de temps incertain A	132
Figure 20. Distributions de possibilité de chaque borne de l'intervalle de temps incertain B	132

Figure 21. Un exemple d'ontologie PersonLink basée sur notre approche	133
Figure 22. Représentation sémantique des données à la fois certaines et imprécises	146
Figure 23. Notre ontologie des données temporelles à la fois incertaines et imprécises "PossibilisticTemporalOntol"	153
Figure 24. Prototype implémenté basé sur notre ontologie	154
Figure 25. Mesures de possibilité et de nécessité de I_1 basées sur les règles définies	156
Figure 26. Représentation sémantique de l'exemple dans le contexte de l'ontologie PersonLink	158
Figure 27. Représentation sémantique des données temporelles conflictuelles	167
Figure 28. Règles SWRL de notre ontologie	173
Figure 29. Prototype implémenté basée sur "BeliefTimeOnto".....	174
Figure 30. Représentation sémantique de l'exemple proposé basée sur notre approche.....	177
Figure 31. Composants de l'ontologie de l'imperfection des données temporelles.....	180
Figure 32. Les classes de l'ontologie "TimeOntoImperfection"	181
Figure 33. Propriétés d'objet de l'ontologie "TimeOntoImperfection"	182
Figure 34. Propriétés de type de données de "TimeOntoImperfection"	183
Figure 35. Exemple d'une règle SWRL.....	184
Figure 36. Prototype basé sur "TimeOntoImperfection"	185
Figure 37. Composant des relations qualitatives du prototype	186
Figure 38. Composant permettant de chercher dans les données temporelles enregistrées ..	186
Figure 39. L'ontologie de l'imperfection des données temporel	187
Figure 40. Résultats de l'évaluation	190
Figure 41. Les différentes contributions proposée dans notre thèse	194

CHAPITRE 1 : Introduction générale

Dans cette introduction générale, nous présentons le contexte de nos travaux de thèse, les problématiques associées, nos contributions, nos publications ainsi que l'organisation de ce manuscrit.

1.1. Contexte du travail

Aborder **le sujet de l'imperfection des données** relève de la nécessité d'avoir « le droit à l'erreur », de l'inexactitude, du doute, de l'envie de dire « peut-être » aux choses dont on n'est pas sûr à cent pour cent, de s'exprimer de façon imprécise, etc.

Dans le cadre **d'une application destinée à des personnes souffrant de la maladie d'Alzheimer**, des spécificités rajoutent de la complexité au traitement des données collectées. En effet, les patients atteints de la maladie d'Alzheimer souffrent de discordance de la mémoire, de changements d'attention et de concentration et d'une diminution de leurs capacités (Ghorbel, 2018). Suite à ces changements et incapacités, les données fournies par les patients atteints de la maladie d'Alzheimer sont pour la plupart **sujettes à différents types d'imperfections**. D'où, le niveau de fiabilité d'un utilisateur peut être très fluctuant.

Toutes les données peuvent être affectées par l'imperfection. **Les données temporelles** ne font pas exception. Cependant, la relation d'ordre établie sur ces **données** peut aider à détecter les erreurs. En effet, les données temporelles peuvent être représentées en utilisant **des points de temps ou des intervalles**. Elles peuvent également être conceptualisées comme **quantitatives** (définies avec précision, par exemple à l'aide des dates) ou **qualitatives** (désigne les relations temporelles qualitatives comme « avant »).

Notre travail a été effectué dans le contexte de **CAPTAIN MEMO** (Métais, et al., 2012), qui est une prothèse de mémoire « intelligente » destinée aux malades d'Alzheimer au stade précoce de la maladie pour pallier leurs problèmes mnésiques, et qui a été développée en collaboration entre le laboratoire Cedric (Cnam, France) et le laboratoire Miracl (Université de Sfax, Tunisie), en particulier le cadre de la thèse de (Ghorbel, 2018). Cette prothèse est basée sur une ontologie OWL 2 temporelle floue, appelée **MemoFuzzyOnto** (Ghorbel 2018), et une ontologie multilingue et multiculturelle pour représenter les relations familiales, appelée **PersonLink** (Herradi, 2018), qui permet de modéliser et de raisonner sur les relations interpersonnelles (par exemple, mère, voisine). **CAPTAIN MEMO** vise à offrir des interfaces utilisateurs « intelligentes » basées sur un traitement sémantique des données, et s'adaptant au patient. Elle offre plusieurs services notamment les interfaces utilisateurs du service Autobiographical Training

qui sont adaptatives, les services MEMO_Calendring et E-Diary génèrent automatiquement des données supplémentaires et le service MEMO_Calendring qui offre la possibilité de saisir des données en langage naturel.

1.2. Problématiques

Les données saisies par le patient d'Alzheimer, vivant dans l'incertitude, sont sujettes à différentes formes d'imperfection. Ceci est dû aux troubles de mémoires causés par la maladie d'Alzheimer et par le processus normal du vieillissement. Par exemple, le malade peut saisir des données contradictoires ou redondantes à cause de sa discordance mnésique. Il peut aussi confondre les mots (c.-à-d., dire un mot au lieu d'un autre).

Traiter l'imperfection des données est une tâche difficile. Cette tâche se complique quand il s'agit de données temporelles qui jouent un rôle important dans la représentation des connaissances (par exemple sous forme d'ontologies). Ces données peuvent être exposées à divers types d'imperfections. Par exemple, on trouve l'imprécision dans « J'ai vécu à Berlin dans les années 70 » et « Ma sœur s'est mariée avant de s'installer à Paris ».

La problématique de cette thèse consiste à traiter les imperfections des données saisies par les utilisateurs. Plus particulièrement **les données temporelles imparfaites représentées sous forme d'ontologie**. Ces données sont saisies par **les malades d'Alzheimer** à l'aide de la prothèse « **CAPTAIN MEMO** ».

Les défis scientifiques à relever sont multiples :

A. Les différentes formes d'imperfection des données temporelles

Comme déjà mentionné, plusieurs imperfections comme l'imprécision, l'incertitude, etc. peuvent affecter les données temporelles notamment celles saisies par des patients d'Alzheimer.

Pour la prise en compte de l'imperfection des données saisies, comprendre ses différentes formes est fondamental. Dans la littérature, différentes typologies de l'imperfection des données dans plusieurs domaines ont été proposées. Cependant, à notre connaissance, **il n'existe pas une typologie fine de l'imperfection des données temporelles**.

B. L'imprécision des données temporelles de granularité fine dans l'ontologie

Selon la littérature, **l'imprécision** est un type d'imperfection qui peut affecter les données temporelles. Ceci est plus fréquent chez les patients atteints de la maladie d'Alzheimer, chose

que nous avons remarquée lors de nos évaluations dans le cadre de CAPTAIN MEMO. Par exemple, des patients peuvent saisir des phrases comme « *vers les années 80* » ou « *vers la fin Juin* ».

Dans l'état de l'art, **plusieurs travaux ont été proposés pour représenter et raisonner sur les données temporelles précises (Herradi, et al., 2017)**, mais très peu sur **les données temporelles imprécises (Ghrobel, et al., 2018)**. Par ailleurs, à notre connaissance, aucun de ces travaux ne considère **les granularités fines des données temporelles imprécises comme les dates et les horloges de temps dans OWL 2**.

C. L'incertitude des données temporelles dans l'ontologie

L'incertitude est un autre type d'imperfection qui peut toucher les données temporelles. Il représente le type d'imperfection le plus fréquent des données saisies CAPTAIN MEMO. Par exemple une phrase comme « *Je pense qu'elle s'est mariée avant l'obtention du diplôme universitaire* ». Ceci est très justifié pour des malades d'Alzheimer vivant dans l'incertitude.

Dans le domaine du web sémantique, plusieurs approches ont été proposées pour représenter et raisonner sur les données temporelles certaines. Cependant, à notre connaissance, il n'y a pas d'approches qui **représentent et raisonnent sur les données temporelles incertaines en OWL 2**.

D. Les types d'imperfections composites : Incertitude et imprécision à la fois

Les données temporelles peuvent être sujettes à **des types composites d'imperfection**, notamment **l'incertitude et l'imprécision à la fois** comme par exemple « *Je pense que j'habitais à Paris entre les années 60 et 70* ».

A notre connaissance, l'état de l'art ne comprend pas d'approches qui **représentent et raisonnent sur des données temporelles qui sont à la fois incertaines et imprécises en OWL 2**.

E. Le conflit des données temporelles dans l'ontologie

Le conflit caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires et donc incompatibles. Dans le cas de la maladie d'Alzheimer, un patient peut, pour une même question, exprimer plusieurs réponses contradictoires. Par exemple, si on lui pose trois fois la question suivante : « *Quand avez-vous obtenu votre diplôme universitaire ?* », il peut donner trois réponses temporelles différentes comme : 1998, 1995 et 1996.

Dans la littérature, **à notre connaissance il n’y a pas d’approches qui représentent et raisonnent sur des données temporelles conflictuelles dans OWL 2.**

F. Représentation et raisonnement sur l’imperfection des données temporelles dans l’ontologie

Dans la littérature, plusieurs **ontologies de temps** ont été proposées comme OWL-TIME (Hobbs & Pan, 2004), SWRL-TEMPORAL (O’Connor & Das, 2011) et SOWL (Batsakis & Petrakis, 2011). Cependant, **elles ne considèrent pas les imperfections** qui peuvent toucher les données temporelles comme l’imprécision et l’incertitude.

A notre connaissance, **il n’y a pas une ontologie de l’imperfection des données temporelles** dans la littérature.

1.3. Contributions

Dans cette section, nous présentons nos contributions.

I. Une typologie de l’imperfection des données temporelles

La création d’une typologie est essentielle pour aider à rassembler toutes les imperfections pouvant affecter les données temporelles. Plusieurs typologies d’imperfections des données ont été proposées, mais elles ne couvrent pas certains types d’imperfection spécifiques aux données temporelles. En effet, ces dernières peuvent être soumises à plusieurs facteurs susceptibles d’interférer dans la spécification du type d’imperfection. Par ailleurs, **à notre connaissance, il n’existe pas une typologie des imperfections des données temporelles.**

Pour résoudre ce problème, nous proposons **une typologie de l’imperfection temporelle des données.** Cette typologie est divisée en **imperfections directes, imperfections indirectes** qui peuvent être déduites des imperfections directes, ainsi que **d’autres facteurs** qui peuvent interférer dans la spécification du type d’imperfection tel que le profil de la personne et le multiculturalisme.

Les imperfections directes consistent en l’incertitude, le conflit, l’erreur de saisie, l’imprécision, l’oubli, la circonlocution et l’inutilité. Les imperfections indirectes sont générées des imperfections directes comme suit : l’incertitude et l’erreur de saisie entraînant l’incohérence ; l’imprécision et l’oubli conduisent à l’incomplétude ; l’inutilité conduit à la redondance.

Ce travail a fait l’objet d’un article accepté dans KEOD 2019 : « *A Typology of Temporal Data Imperfection* » (Achich, et al., 2019, September).

II. Représentation et raisonnement sur les dates et horloges du temps précis et imprécis dans le contexte de l'ontologie avec une vue précise

Nous proposons une approche pour traiter **les données temporelles de granularité fine**, précisément **des dates et des horloges précises et imprécises** dans OWL 2 avec **une vue précise**.

Notre approche vise à **représenter et raisonner sur des données temporelles** en termes **quantitatifs** (points de temps pouvant être des dates et des horloges, et des intervalles de temps) ainsi que **des relations qualitatives** (par exemple, « *before* »). Elle traite non seulement des points de temps et des intervalles de temps précis, mais également **des points de temps et des intervalles de temps imprécis**, par exemple « *Le voyage commence au début Juin et se termine vers la mi-Juin* ». Elle ne vise pas à créer un nouveau standard, mais de réutiliser les technologies Web sémantiques existants.

Notre approche consiste à **étendre l'approche de 4D-fluents** (Welty & Fikes, 2006) avec des éléments de l'ontologie OWL-Time et **de nouveaux composants ontologiques précis que nous définissons** pour **représenter les données temporelles**. Notamment les dates et les horloges de temps, précises et imprécises. Elle consiste aussi à **étendre l'algèbre d'intervalle d'Allen** (Allen, 1983) pour **raisonner sur des intervalles de temps imprécis**. Par rapport aux travaux connexes, notre extension est **entièrement basée sur un environnement précis**. Les relations d'intervalles résultantes **préservent les propriétés de l'algèbre d'Allen standard**. Nous adaptons ces relations pour permettre de relier **un intervalle de temps et un point de temps, et deux points de temps ; où les points de temps et les intervalles peuvent être précis ou imprécis**. Toutes les relations proposées peuvent être utilisées pour le raisonnement au moyen de tables de transitivité. Nous proposons **une ontologie OWL 2 appelée « CrispTimeOnto »** basée sur nos extensions. Nous définissons également un ensemble **de règles SWRL** permettant de déduire les relations temporelles qualitatives proposées. Pour valider notre travail, nous implémentons **un prototype basé sur l'ontologie proposée** que nous intégrons dans la prothèse de mémoire **CAPTAIN MEMO**.

Ce travail a fait l'objet d'un article accepté dans DEXA 2019: « *Representing and Reasoning about Precise and Imprecise Time Points and Intervals in Semantic Web: Dealing with Dates and Time Clocks* » (Achich, et al., 2019).

III. Représentation et raisonnement sur les intervalles et points du temps certains et incertains dans le contexte de l'ontologie avec une vue certaine

Nous proposons une approche pour **représenter et raisonner sur des données temporelles certaines et incertaines** en termes de **relations qualitatives** (par exemple, « before ») et **quantitatives** (intervalles et points de temps).

Cette approche se compose de trois parties. La première partie concerne la **représentation de données temporelles certaines et incertaines dans OWL2**. Les intervalles de temps certains (c'est-à-dire parfait) peuvent être modélisés en OWL en utilisant **l'approche des 4D-fluents**. Nous l'étendons avec de nouvelles composantes ontologiques pour **représenter des données temporelles quantitatives certaines** (les points de temps) et incertaines (les points de temps et les intervalles de temps), et **des relations temporelles qualitatives** entre les intervalles de temps et les points de temps. La deuxième partie concerne **le raisonnement sur les données temporelles certaines et incertaines en étendant l'algèbre d'intervalle d'Allen**. Cette algèbre propose des relations qualitatives entre les intervalles de temps. Cependant, elle n'est pas consacrée à gérer des intervalles de temps incertains. De plus, elle n'est pas prévue de relier un intervalle de temps et un point de temps ou même deux points de temps. **Nous étendons l'algèbre d'Allen en proposant des relations temporelles qualitatives entre des intervalles de temps incertains**. Elles conservent des propriétés importantes de l'algèbre d'origine. **Nous adaptons les relations d'intervalle résultantes pour proposer des relations temporelles entre un intervalle de temps et un point de temps, et deux points de temps**. La troisième partie consiste à **proposer une ontologie OWL 2 nommée « *UncertTimeOnto* »** qui peut être intégrée dans d'autres ontologies pour gérer des données temporelles certaines et incertaines. Elle est implémentée sur la base des extensions proposées. Les inférences sont faites en utilisant des règles SWRL intégrées dans l'ontologie. Pour valider notre travail, nous avons implémenté **un prototype basé sur l'ontologie proposée** que nous avons intégré dans la prothèse de mémoire CAPTAIN MEMO.

Ce travail a fait l'objet d'un article accepté dans KES Conference 2020 : « *Approach to Reasoning about Uncertain Temporal Data in OWL 2* » (Achich, et al., 2020).

Une version étendue de ce papier est acceptée dans un IJSWIS (journal Q2) : « *Certain and Uncertain Temporal Data Representation and Reasoning in Owl 2* »

IV. Représentation et raisonnement sur les intervalles et points du temps certains et incertains dans le contexte de l'ontologie en utilisant la théorie de possibilité

Nous proposons **une approche basée sur la théorie de possibilité** pour **représenter et raisonner sur des données temporelles incertaines en termes de relations qualitatives** (par exemple, « before ») et **quantitatives** (intervalles et points de temps).

Notre première contribution consiste à **étendre l’approche 4D-fluents avec de nouveaux composants ontologiques pour représenter des intervalles de temps incertains et des points temporels incertains et associés à des relations qualitatives**. Notre deuxième contribution consiste à **étendre l’algèbre d’intervalle d’Allen pour raisonner sur les données manipulées**. Nous proposons 13 relations temporelles entre des intervalles de temps incertains. Notre extension conserve les propriétés de réflexivité / irreflexivité, de symétrie / asymétrie et de transitivité. Ces relations sont adaptées pour relier un intervalle de temps et un point temporel ou deux points temporels. Nous introduisons quatre tables de transitivité pour raisonner sur les relations temporelles résultantes. La troisième contribution consiste à **proposer une ontologie temporelle possibiliste "UncertPossibTimeOnto"** qui implémente nos extensions. Enfin, **nous comparons nos deux approches qui traitent l’incertitude**.

Ce travail a fait l’objet d’un article accepté dans la conférence FUZZ-IEEE 2021 : « *Handling Uncertain Time Intervals in Owl 2: Possibility Vs Probability Theories-based Approaches* » (Achich, et al., 2021).

V. **Représentation et raisonnement sur les données temporelles à la fois incertaines et imprécises dans le contexte de l’ontologie en utilisant la théorie de possibilité**

Nous proposons **une approche basée sur la théorie de possibilité** pour **représenter et raisonner sur des données temporelles à la fois incertaines et imprécises** en termes de **relations qualitatives** (par exemple, « before ») et **quantitatives** (intervalles et points de temps).

Notre première contribution consiste à **calculer les degrés de possibilité et de nécessité** à partir **d’un ensemble de règles que nous définissons**. Notre deuxième contribution consiste à proposer **une représentation sémantique des données temporelles manipulées et des mesures obtenues**. Notre troisième contribution consiste à **étendre l’algèbre d’intervalle d’Allen pour raisonner sur les données temporelles en questions**. Nous proposons 13 relations temporelles entre des intervalles de temps à la fois incertains et imprécis en conservant la réflexivité / irreflexivité, la symétrie / asymétrie et la transitivité. Notre quatrième contribution se concentre sur la proposition **d’une ontologie temporelle possibiliste nommée «Possibilistic-TemporalOnto»** qui implémente la représentation sémantique proposée et l’extension d’Allen. Les inférences sont faites en utilisant des règles SWRL intégrées dans l’ontologie. Pour valider

notre travail, nous avons conduit **une étude de cas dans le contexte de CAPTAIN MEMO**. Nous avons implémenté un prototype basé sur l'ontologie proposée que nous avons intégrons dans la prothèse de mémoire.

Ce travail a fait l'objet d'un article accepté dans RCIS Conference 2021 : « *Dealing with Uncertain and Imprecise Time Intervals in OWL2 : A Possibility Theory-Based Approach* » (Achich, et al., 2021).

VI. Représentation et raisonnement sur le conflit des données temporelles dans le contexte de l'ontologie en utilisant la théorie de l'évidence

Nous proposons **une approche basée sur la théorie de l'évidence pour représenter et raisonner sur des données temporelles conflictuelles en termes de relations qualitatives** (par exemple, « before ») et **quantitatives** (intervalles et points de temps).

Notre approche comporte trois contributions. La première concerne **le calcul des masses de croyance des données temporelles conflictuelles**, en passant par **les mesures de crédibilité basées sur notre approche précédente DBE_ALZ** (Ghorbel, et al., 2020) et la proposition **d'une représentation sémantique des données manipulées associées aux mesures obtenues et aux relations qualitatives associatives**. La deuxième contribution consiste à **raisonner sur des données temporelles conflictuelles en étendant l'algèbre d'intervalle d'Allen**. Nous proposons 13 relations temporelles entre des intervalles de temps conflictuels. Notre extension conserve les propriétés de réflexivité / irreflexivité, de symétrie / asymétrie et de transitivité. Ces relations sont adaptées pour relier un intervalle de temps et un point temporel ou deux points temporels. Nous introduisons quatre tables de transitivité pour raisonner sur les relations temporelles résultantes. La troisième contribution consiste à proposer **une ontologie évidentielle nommée « *BeliefTimeOnto* »** qui implémente la représentation sémantique proposée et les extensions d'algèbre d'intervalle d'Allen. Les inférences sont effectuées à l'aide des règles SWRL. Pour valider notre travail, nous avons mis en place un prototype permettant aux utilisateurs d'explorer notre approche, basée sur notre ontologie, en l'intégrant dans CAPTAIN MEMO.

VII. Ontologie de l'imperfection des données temporelles

En nous basant sur nos ontologies proposées des différentes imperfections traitées, **nous créons notre ontologie de l'imperfection des données temporelles nommée « *TimeOntoImperfection* »** qui, à notre connaissance, **n'existaient pas dans des travaux connexes**. Cette

ontologie peut être **fusionnée avec d'autres ontologies** d'application ou de domaine qui doivent être étendues **pour représenter et raisonner sur des données temporelles imparfaites**. Nous avons implémenté un prototype basé sur cette ontologie et nous l'avons intégré dans CAPTAIN MEMO pour évaluer notre travail.

1.4. Liste des publications

Revues internationales avec comité de lecture

- **Achich, N.**; Ghorbel, F. ; Hamdi, F. Metais, E. and Gargouri, F. (2021). Certain and Uncertain Temporal Data Representation and Reasoning in OWL 2. In International Journal on Semantic Web and Information Systems (IJSWIS).
- Ghorbel, F., Hamdi, F., **Achich, N.**, & Metais, E. (2020). Handling data imperfection-False data inputs in applications for Alzheimer's patients. Data & Knowledge Engineering, 130, 101864.

Conférences internationales avec comité de lecture

- **Achich, N.**; Ghorbel, F.; Hamdi, F. c.; Metais, E. and Gargouri, F. (2021). Handling Uncertain Time Intervals in Owl 2: Possibility Vs Probability Theories-based Approaches. In FUZZ-IEEE 2021, Luxemburg, Luxembourg. « Rang A ».
- **Achich N.**, Ghorbel F., Hamdi F., Metais E., Gargouri F. (2021). Dealing with Uncertain and Imprecise Time Intervals in OWL 2: A Possibility Theory-Based Approach. In: Cherfi S., Perini A., Nurcan S. (eds) Research Challenges in Information Science. RCIS 2021. Lecture Notes in Business Information Processing, vol 415. Springer, Cham. https://doi.org/10.1007/978-3-030-75018-3_35 « Rang B ».
- **Achich, N.**, Ghorbel, F., Hamdi, F., Métais, E., & Gargouri, F. (2020). Approach to Reasoning about Uncertain Temporal Data in OWL 2. In : 24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES). Procedia Computer Science, 176, 1141-1150. <https://doi.org/10.1016/j.procs.2020.09.110> « Rang B ».

- **Achich, N.** (2020, November). Dealing with Data Imperfection in OWL 2-Application to Alzheimer’s Patients Software. In International Semantic Web Conference (ISWC DC). « Rang A ».
- **Achich, N., Ghorbel, F., Hamdi, F., Metais, E., & Gargouri, F.** (2019, August). Representing and Reasoning About Precise and Imprecise Time Points and Intervals in Semantic Web: Dealing with Dates and Time Clocks. In International Conference on Database and Expert Systems Applications (pp. 198-208). Springer, Cham. https://doi.org/10.1007/978-3-030-27618-8_15 « Rang B ».
- **Achich, N., Ghorbel, F., Hamdi, F., Metais, E., & Gargouri, F.** (2019, September). A Typology of Temporal Data Imperfection. In KEOD (pp. 305-311). « Rang C ».

1.5. Organisation du mémoire

Nous présentons dans cette section l’organisation de ce manuscrit. Ce mémoire de thèse comporte **dix chapitres**.

Le **second chapitre** est dédié aux notions théoriques et l’état de l’art. Dans ce chapitre, nous présentons la notion d’ontologie qui est au centre de nos travaux sont dans le contexte de l’ontologie. Nous proposons notre définition de l’imperfection de données et les différentes formes qui peuvent y figurer. Nous présentons aussi les théories utilisées dans le traitement des imperfections des données. Puis, nous passons en revue les différentes typologies des imperfections des données ainsi que les travaux existant dans le contexte de la représentation et le raisonnement sur les données temporelles dans le web sémantique. Dans le **troisième chapitre**, nous présentons une première contribution apportée sur la définition d’une typologie de l’imperfection des données temporelles. Dans le **quatrième chapitre**, nous présentons notre approche qui sert à traiter l’imprécision des dates et les horloges du temps dans l’ontologie avec une vue précise. Dans le **cinquième et le sixième chapitre**, nous traitons l’incertitude des données temporelles dans l’ontologie en utilisant dans un premier lieu les réseaux bayésiens et en deuxième lieu la théorie des possibilités. Dans le **septième chapitre**, nous présentons notre approche qui traite l’incertitude et l’imprécision, simultanément, des données temporelles dans l’ontologie en utilisant la théorie de possibilités. Dans le **huitième chapitre**, nous traitons le

conflit des données temporelles dans l'ontologie en utilisant la théorie de l'évidence. Le **neuvième chapitre** présente notre ontologie de l'imperfection des données temporelles et les évaluations dans le cadre de CAPTAIN MEMO. Dans le **dixième chapitre**, un bilan global de la thèse est fourni en conclusion générale où nous rappelons nos différentes contributions et présentons des perspectives pour nos travaux futurs.

CHAPITRE 2 : Notions théoriques et état de l'art

2.1. Introduction

Ce chapitre présente **les notions théoriques et l'état de l'art** des travaux réalisés sur les différents sujets abordés tout au long de cette thèse. En premier lieu, nous définissons **la notion d'ontologie, ses composants et les types d'ontologie**. Puis, nous **traitons de l'imperfection des données**. Dans cette partie, nous présentons **les différentes formes d'imperfections** pouvant affecter les données, ensuite nous passons en revue **les différentes typologies de l'imperfection des données existantes**, puis nous introduisons **les théories de l'imperfection**. La dernière partie porte sur **les données temporelles**, précisément les **éléments temporels, la représentation et le raisonnement sur les données temporelles dans le web sémantique**. Nous réalisons un travail bibliographique neutre et une analyse des publications majeures en rapport avec les thèmes mentionnés.

2.2 Notion d'ontologie

La notion d'ontologie est née du besoin de manipulation de la sémantique à travers les machines. L'ontologie est une notion centrale du **web sémantique**. Pour cette raison, nous définissons le web sémantique. Puis, nous introduisons l'ontologie. Enfin, nous précisons ses différents composants.

2.2.1 Web sémantique

Le web sémantique est « une extension du web statique dans laquelle les données ont un sens bien défini, permettant aux ordinateurs et aux utilisateurs de mieux travailler en coopération (homme-homme, homme-machine et machine-machine) ». Il a été défini par Tim Berners-Lee (Berners-Lee, et al., 2001) qui est le fondateur du Consortium World Wide Web (W3C).

L'architecture du web sémantique, proposée par le W3C, s'appuie sur une représentation pyramidale de langages associés aux ontologies, comme le montre la Figure 1.

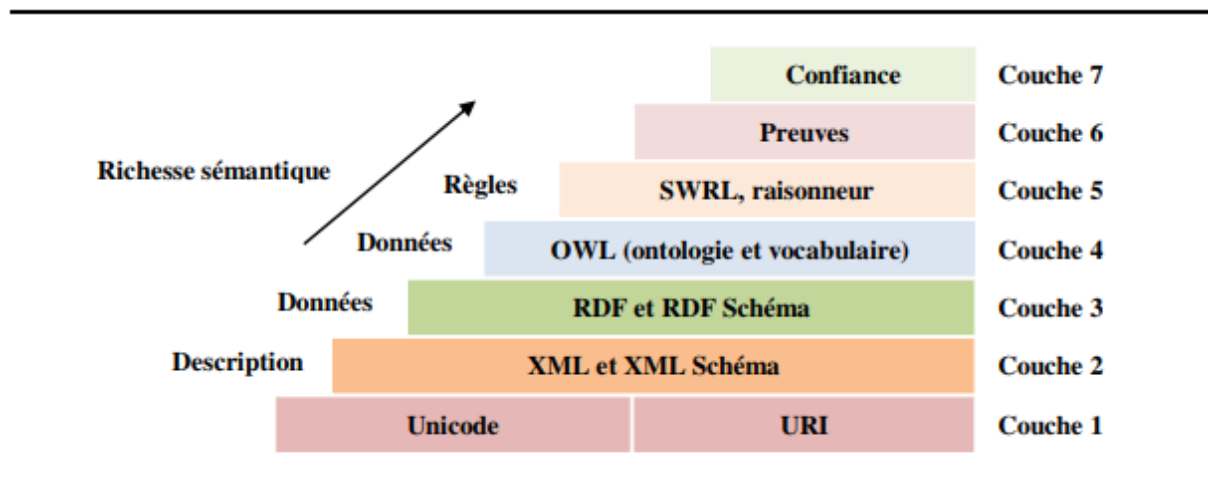


Figure 1. Représentation pyramidale du web sémantique (Berners-Lee et al., 2001).

L'interopérabilité syntaxique est assurée par la couche la plus basse de l'architecture. Les données sont codées par le standard Unicode et possèdent une adresse « **Uniforme Ressource Identifiant** » (**URI**) qui permet d'attribuer un identifiant unique et universel à un ensemble de données (Couche 1). Ces données peuvent être structurées grâce à « eXtensible Markup language » (XML). Au niveau de la couche 2, **la syntaxe XML** est considérée comme un premier niveau de la sémantique permettant aux utilisateurs de structurer les données en fonction de leur contenu. Dans la couche 3, les données sont structurées en utilisant le langage **Resource Description Framework (RDF)** (Brickley, 2004). Ceci permet d'attribuer une signification à cette structure et relier d'une façon pertinente les différents éléments. Au niveau de la couche 4, les données sont structurées avec **le langage OWL (Web Ontology Language)** (Bechhofer, et al., 2004).

2.2.2 Définition

L'ontologie est une branche de la métaphysique qui s'intéresse à l'existence, à l'être en tant qu'être et aux catégories fondamentales de l'existant (Devaux & Lamanna, 2009). En effet, ce terme est construit à partir des racines grecques « ontos », qui signifie « être » et « logos » qui signifie « l'étude », qui se traduit par « L'étude de l'être » et par extension de l'existence (Devaux & Lamanna, 2009).

En intelligence artificielle, Neeches et ses collègues ont été les premiers à proposer une définition : « Une ontologie définit les termes et relations de base du vocabulaire d'un domaine et les règles qui indiquent comment combiner termes et relations afin que le vocabulaire puisse être étendu » (Gomez-Perez et al., 2006).

En 1993, Gruber (Gruber, 1993) proposait la définition suivante : « spécification formelle d'une conceptualisation partagée ». Cette définition a été légèrement modifiée par Borst (Borst, 1997) en tant que « spécification formelle d'une conceptualisation partagée ». (Studer et al., 1998) a regroupé les deux dernières définitions ainsi : « spécification formelle et explicite d'une conceptualisation partagée ». Formelle signifie que l'ontologie doit être lisible par une machine, ce qui exclut le langage naturel. Explicite signifie que les concepts utilisés doivent être explicitement définis. La conceptualisation signifie conceptualiser le modèle abstrait d'un phénomène du monde réel en identifiant ses concepts clés (Lee, et al., 2009). Partagé signifie que l'ontologie n'est pas la propriété d'un individu, mais qu'elle représente un consensus accepté par une communauté d'utilisateurs. Pour (Giarretta & Guarino, 1995) : « une ontologie est une spécification qui rend partiellement compte d'une conceptualisation ». (Swartout et al., 1999) le définissent ainsi : « une ontologie est un ensemble de termes structurés hiérarchiquement destinés à décrire un domaine et pouvant servir de cadre à une base de connaissances ». La même notion est également développée par Gomez : « une ontologie fournit les moyens de décrire explicitement la conceptualisation des connaissances représentées dans une base de connaissances ». (Gruber, 1993) l'a définie comme suit : « Dans le contexte de l'informatique et des sciences de l'information, une ontologie définit un ensemble de primitives représentationnelles avec lesquelles modéliser un domaine de connaissance ou de discours. Les primitives de représentation sont généralement des classes (ou des ensembles), des attributs (ou propriétés) et des relations (ou des relations entre les membres de la classe) ».

Nous distinguons deux objectifs de l'ontologie :

- Le premier objectif est **descriptif**. Il consiste à former une base de connaissances en modélisant un ensemble de connaissances dans un domaine donné.
- Le deuxième objectif est **le raisonnement**. Il consiste à faire des inférences et générer automatiquement des connaissances à l'aide d'un ensemble de règles.

Le langage OWL, recommandé par le W3C en 2004, est le plus expressif des langages ontologiques (Baget, 2004). Il présente une extension du RDF et RDFS. Le but est d'ajouter

plus de formalisation dans l'expression des informations afin d'assurer leur validité. Il permet de définir des restrictions de cardinalités comme. Contrairement à RDF et RDFS, avec OWL 2, nous pouvons aussi décrire des disjonctions de classes ou des combinaisons booléennes entre les classes. OWL 2 peut aussi exprimer des caractéristiques particulières des propriétés comme inverse de (par exemple « mange » est l'inverse de « est mangé par ») et la transitivité (la propriété « plus grand que » est transitive).

2.2.3 Composants

(Charlet, et al., 2005) présentent les composants minimaux qu'une ontologie doit contenir. Le premier composant est **la classe (concept)** qui est un ensemble de ressources partageant les mêmes caractéristiques. Elle peut représenter un objet matériel, une notion, une action, un processus de raisonnement ou une idée. Chaque classe est associée à un ensemble d'individus. Une classe peut être spécifiée comme une sous-classe d'une autre classe, mettant ainsi en œuvre la relation de spécialisation. **Les propriétés** sont utilisées pour définir le contenu des classes. Les propriétés se divisent en deux catégories ; **les propriétés d'objet** (object properties) représentent des relations générales entre deux classes, tandis que **les propriétés de type de données** (datatype properties) représentent des attributs. Chaque propriété d'objet admet un « domain » (l'ensemble de ses concepts sources) et un « range » (l'ensemble de ses concepts destinations). **Les relations d'une ontologie** décrivent les différentes corrélations entre les classes. Elles incluent la relation de spécialisation (is_a) et la relation de méronymie (part_of). Les propriétés d'objets sont considérées comme des relations dans la plupart des travaux de recherche. **Les individus, appelés aussi des instances**, sont des instanciations de concepts, représentant généralement des objets du monde réel. Ils contiennent des valeurs (instances d'attributs) et sont connectés à des instances de relations. Il y a une distinction entre la base de faits qui contient les individus (ABox) et l'ontologie qui définit la terminologie (TBox).

2.2.4 Types d'ontologie

Les ontologies peuvent être classées selon leur conceptualisation. **L'ontologie supérieure (Upper Ontology)** est également appelée ontologie de niveau supérieur (top-level ontology). Elle décrit des concepts très généraux tels que l'espace, le temps, la matière, les objets, les événements, les actions, etc. Ces concepts ne dépendent pas d'un problème ou d'un domaine particulier, ils sont communs entre les utilisateurs. **Les ontologies de domaine (Domain Ontologies)** relient les concepts sémantiques dans un domaine donné. Elles fournissent les concepts et les relations pour couvrir les vocabulaires, les activités et les théories de ce domaine.

Les ontologies de tâches (Task Ontologies) fournissent un vocabulaire systématique de termes utilisés pour résoudre des problèmes liés à des tâches qui peuvent ou non être dans le même domaine. Elles fournissent un ensemble de termes par lesquels nous pouvons généralement décrire comment résoudre un type de problème. Elles comprennent des noms, des verbes et des adjectifs génériques dans les descriptions de tâches. **Les ontologies d'application (Application ontologies)** présentent le type de l'ontologie la plus spécifique, elles contiennent des concepts dépendants d'un domaine particulier et d'une tâche. Ces concepts correspondent souvent aux rôles joués par les entités du domaine lors de l'exécution d'une certaine activité.

D'autre part, les ontologies peuvent être aussi classifiées en :

2.2.4.1 Ontologie classique (parfaite)

Au cours des dernières années, les ontologies ont été largement utilisées pour représenter la connaissance de la plupart des domaines du monde réel. Elles fournissent une définition des concepts, des relations et d'autres caractéristiques liées à la modélisation des connaissances d'un domaine particulier. Grâce à ces éléments, elles sont utilisées pour modéliser la réalité (applications du monde réel).

Une **ontologie classique** est **une ontologie parfaite ne modélisant pas l'imperfection**. Cependant, ce monde comprend des inexactitudes et des imperfections qui ne peuvent être représentées par des ontologies classiques.

2.2.4.2 Ontologie imparfaite

Pour permettre aux agents de gérer l'imperfection, une extension des ontologies qui a la capacité de supporter des connaissances imparfaites, est obligatoire. Ces imperfections peuvent être incertitude, imprécision, conflit, etc.

- **Ontologie probabiliste**

Plusieurs définitions **des ontologies probabilistes** ont été proposées. (Costa, 2005) définissent l'ontologie probabiliste comme « une représentation formelle explicite des connaissances exprimant des connaissances sur un domaine d'application qui comprend : les types d'entités existant dans le domaine ; les propriétés de ces entités ; les relations entre entités ; les processus et événements qui se produisent avec ces entités ; les régularités statistiques qui caractérisent le domaine ; les connaissances non concluantes, ambiguës, incomplètes, peu fiables et dissonantes ; l'incertitude concernant toutes les formes de connaissance susmentionnées ; où

le terme entité désigne tout concept (réel, fictif, concret ou abstrait) pouvant être décrit et raisonné dans le domaine d'application ».

Selon (Costa, et al., 2008), les ontologies probabilistes fournissent une manière structurée, partageable et fondée sur des principes, de décrire de manière exhaustive les connaissances sur un domaine et l'incertitude qui l'entoure. Ils élargissent également les possibilités des ontologies standard en introduisant l'exigence d'une représentation correcte des régularités statistiques et des preuves incertaines sur les entités dans un domaine d'application. Idéalement, la représentation est dans un format qui peut être lu et traité par un ordinateur.

Selon (Carvalho, et al., 2011) les ontologies probabilistes sont utilisées pour décrire en détail la connaissance d'un domaine et l'incertitude associée à cette connaissance de manière raisonnée, structurée et partageable.

(Hlel, et al., 2015) définissent l'ontologie probabiliste simplement comme une ontologie classique contenant des connaissances incertaines. En d'autres termes, une ontologie n'est une ontologie probabiliste que si elle contient au moins une composante probabiliste (Hlel, et al., 2015).

○ **Ontologie possibiliste**

Dans la littérature, **il n'existe pas beaucoup de travaux définissant l'ontologie possibiliste. Dans le domaine temporel, il n'existe pas une ontologie temporelle possibiliste.**

(Loiseau, 2004) propose une ontologie possibiliste basée sur WordNet dans le processus de recherche d'information.

Selon (Lamine & Omri, 2012), dans une ontologie possibiliste, les relations entre les concepts sont munies chacune par un degré de possibilité et un degré de nécessité qui reflètent respectivement à quel degré cette relation est-elle possible ? Et à quel degré cette relation est-elle certaine ?

(Bourai, et al., 2014) proposent une extension possibiliste d'OWL 2, appelée Poss-OWL 2, pour traiter l'ontologie géographique incertaine. (Bal-Bourai & Mokhtari, 2016) proposent une extension possibiliste de la logique de description SROIQ(D) appelée π -SROIQ(D) en incorporant un niveau de certitude pour différents éléments de la SROIQ(D) DL pour traiter des informations géographiques incertaines

(Salem, et al., 2017) proposent une ontologie possibiliste basée sur une approche de raisonnement basé sur des cas (CBR) afin d'effectuer un algorithme de récupération sémantique possibiliste qui gère les problèmes d'ambiguïté et d'incertitude.

- **Ontologie floue**

L'ontologie floue est une extension de l'ontologie précise (Kaur & Kaur, 2012) (Zekri, et al., 2015) (Akremi, et al., 2016). Elle se base sur l'intégration de la logique floue à la définition de l'ontologie précise pour représenter et raisonner sur les données imprécises (Rodríguez, et al., 2014) . Elle est très utile dans de nombreux domaines.

D'après (Ghorbel H., et al., 2013), une ontologie floue a neuf composants : (1) les concepts précis, (2) les concepts flous, (3) les relations sémantiques précises, (4) les relations sémantiques floues, (5) les relations conceptuelles précises, (6) les relations conceptuelles floues, (7) les relations de fuzzification, (8) les axiomes et (9) les instances. Les composants précis gardent les mêmes définitions.

- **Ontologie évidentielle**

Dans la littérature, **très peu de travaux définissent une ontologie évidentielle**. (Essaid & Yaghlane, 2009) proposent « BeliefOWL » qui est une nouvelle approche pour représenter l'incertitude dans une ontologie OWL. Ils n'ont considéré que le cas de l'inclusion de l'incertitude dans les classes. Cette incertitude est modélisée via la théorie de l'évidence de Dempster-Shafer. Ils ont étendu les classes d'ontologies OWL avec des masses de croyances, puis ils ont appliqué des règles de traduction structurelle afin d'obtenir un DAG (Directed acyclic graph).

2.2.4.3 Ontologies temporelles

OWL-TIME. (Hobbs & Pan, 2004) proposent une ontologie du temps appelée « owl-time » qui décrit le contenu temporel, comprenant des instants, des intervalles, des durées et des termes de calendrier. Plusieurs chercheurs comme (Cui, et al., 2010) adoptent une approche qui réutilisent les concepts de temps définis dans OWL-Time. Cependant, cette représentation ne crée pas de modèle temporel pour décrire de manière cohérente toutes les informations temporelles d'un système. En outre, il ne spécifie pas comment ces concepts peuvent être utilisés pour représenter les propriétés des objets changeant dans le temps et ne spécifie pas non plus comment raisonner sur la relation qualitative des intervalles temporels et des instants.

Dans ce qui suit, nous listons les principaux concepts et relations de l'ontologie OWL-Time :

- L'ontologie Owl-Time est basée sur le concept «TemporalEntitiy», qui ne comprend que deux sous-classes : «Instant» et «Interval». Les intervalles sont considérés comme les choses dont l'étendue a des points intérieurs. Un instant est considéré comme un point dans le temps n'ayant pas de points intérieurs et est également considéré comme un intervalle avec une étendue nulle ayant le même début et la même fin ;
- Les prédicats « begins » et « ends » sont des relations entre des instants et des entités temporelles ;
- Il existe une relation «before» sur les entités temporelles, qui donne une directionnalité au temps. Si l'entité temporelle T1 est avant l'entité temporelle T 2, alors la fin de T1 est avant le début de T 2. Ainsi, « avant » peut être considéré comme fondamental pour les instants et dérivé pour les intervalles ;
- Les moyens pour spécifier les formats de date et d'heure dans OWL-Time incluent les propriétés suivantes d'une «DateTimeDescription»: «unitType, year, month, week, day, dayOfWeek, dayOfYear, hour, minute, second, and timeZone». La propriété «unitType» spécifie le type d'unité temporelle de la description datetime.

SWRL-Temporal (O'Connor & Das, 2011) est un modèle temporel pour représenter des informations basées sur des intervalles dans OWL. SWRL-Temporal définit la classe «Granularity», qui a des instances «Années, Mois, Jours, Heures, Minutes, Secondes et Millisecondes», pour exprimer la granularité de la représentation temporelle. Il définit également les durées en utilisant la classe «Duration», avec deux propriétés fonctionnelles: «hasCount» avec une range entière et une propriété «granularity» avec une range de granularité. L'heure valide des faits et est représentée à l'aide de la classe «ValidTime», qui à son tour a une propriété «hasGranularity» avec une range de granularité. «ValidTime» a deux sous-classes qui sont «ValidInstant» et «ValidPeriod», représentant respectivement des instants et des intervalles. Il a également une propriété de type de données «hasTime» avec une range «dateTime», tandis que ValidPeriod a des propriétés de type de données «hasStart» et «hasFinish» avec une range dateTime représentant le début et la fin de l'intervalle correspondant. De plus, SWRL-Temporal contient des règles SWRL pour le raisonnement temporel. Ces règles prédéfinies ont pour but de vérifier si les prédicats comparant une durée spécifique à celle d'un intervalle de validité (par exemple,

durationLessThan, durationEqualTo) sont vrais. Des prédicats similaires sont utilisés pour vérifier si une relation d'Allen tient ou non entre deux intervalles. Cependant, SWRL-temporal ne prend pas en charge les intervalles définis qualitatifs.

SOWL (Batsakis & Petrakis, 2011). L'ontologie tente de surmonter les problèmes de certaines approches précédentes, telles que **TOWL** (Frasincar, et al., 2010) (c'est-à-dire que TOWL n'était pas compatible avec OWL et qu'il ne prenait pas en charge la représentation d'intervalles définis qualitativement), en offrant une solution du W3C pour la représentation temporelle. Dans SOWL, les dates sont représentées en utilisant le type de données «dateTime» et les règles SWRL sont utilisées pour le raisonnement temporel. Les concepts temporels sont définis à l'aide de l'ontologie OWL-time. 4D fluents est le modèle utilisé pour représenter des objets dynamiques via l'utilisation de la classe TimeSlice et de la propriété tsTimeSlice. Une version alternative basée sur les relations n-aires est également proposée dans (Anagnostopoulos, et al., 2013). Les relations Allen présentent les propriétés d'objet entre les intervalles. En plus de cela, un mécanisme de raisonnement basé sur la cohérence de chemin a été implémenté en utilisant SWRL pour déduire les relations implicites et vérifier la cohérence des assertions temporelles. Le raisonnement est solide, complet et traitable pour l'ensemble pris en charge des relations d'intervalle (Batsakis & Petrakis, 2011). En plus du support de raisonnement mentionné ci-dessus, le support d'interrogation à SOWL en ajoutant des opérateurs temporels au-dessus de SPARQL est proposé dans (Batsakis & Petrakis, 2011). Plus précisément, (Batsakis & Petrakis, 2011) élaborent un langage de requête basé sur SPARQL qui propose des opérateurs temporels tels que « at » (time-instance) avec la représentation temporelle correspondante.

Toutes ces ontologies qui représentent l'aspect temporel ne considèrent pas les imperfections qui peuvent affecter les données temporelles.

2.3. L'imperfection des données

Les données peuvent être parfaites, c'est-à-dire des données sans erreurs. Elles peuvent aussi être sujettes à plusieurs types d'imperfections. Selon (Smets, 1997), on utilise « l'imperfection » comme étiquette la plus générale. L'imperfection peut être due à des erreurs ou à des imperfections sémantiques.

Nous considérons que le mot « imperfection » est plus général que le mot « erreur ». Une erreur constitue un type d'imperfection. (Vetrò, et al., 2016) considèrent comme « erreurs »

toute faute de frappe ou de syllabage lorsque la saisie de données est manuelle ou associe des valeurs aux mauvaises instances en raison d'une mauvaise conduite du logiciel (dans le cas de l'utilisation des applications ou logiciels). Une « erreur » représente toute erreur syntaxique ou erreur de saisies dans une application ou logiciel utilisés. Une « imperfection sémantique » désigne toute imprécision, incomplétude, confusion, incertitude, etc. qui touche les données. Nous classifions les données en données entrées par les utilisateurs et autres données issues des applications. Dans notre travail de thèse, nous s'adressons aux données entrées par les utilisateurs de logiciels pour personnes âgées et/ou malades d'Alzheimer.

L'incertitude, selon (Ammar, 2014) fait référence à la « qualité » de l'information traduisant la réalité de sa réalisation. Une information est dite incertaine s'il est impossible de juger la vérité binaire (vrai/fausse) de cette information. Les données peuvent contenir des incertitudes. Cela signifie que la valeur exacte est (partiellement) inconnue, cependant, généralement une certaine connaissance est présente de toute façon, décrivant peut-être la valeur. L'imprécision est un type d'imperfection qui fait référence au "contenu" de l'information. Ainsi, lorsque nous ne disposons pas de variables, caractéristiques, etc. pour décrire clairement le contenu de l'information, l'imperfection est appelée imprécision. L'ambiguïté d'une information est le fait que celle-ci entraîne deux ou plusieurs interprétations, car ses limites ne sont pas connues. L'incomplétude représente le manque d'informations apportées par la source. Elle peut être mesurée par la différence entre la quantité d'informations réellement fournies par la source et la quantité d'informations que la source devrait fournir. Le conflit caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires et donc incompatibles.

2.3.1 Typologies de l'imperfection des données

Nous distinguons deux types de typologies : les typologies génériques de l'imperfection des données et les typologies spécifiques de l'imperfection des données dans un domaine déterminé.

2.3.1.1 Typologies génériques de l'imperfection des données

De la littérature, nous identifions quatre typologies génériques de l'imperfection des données : la typologie de la non-précision des données selon (Niskanen, 1989), la typologie de l'imperfection des données selon (Bouchon-Meunier, 1995), la typologie de l'incertitude des données selon (Klir & Yuan, 1995) et la typologie de l'imperfection des données selon (Smets, 1997).

Typologie de la non-précision des données selon (Niskanen, 1989) : Niskanen propose une typologie de la non-précision des données. Elle fait appel à quatre concepts : l'incertitude, l'imprécision, l'ambiguïté et la généralité. Selon l'auteur, l'incertitude est un concept mesurable lié à la notion de l'erreur, tandis que l'imprécision n'est pas mesurable. Il propose trois termes relatifs à la notion de l'imprécision : l'imprécision ontologique (les objets sont naturellement imprécis), l'imprécision linguistique (les objets ne sont pas définis précisément) et l'imprécision épistémologique (les objets ne peuvent pas être perçus d'une manière précise). L'ambiguïté apparaît lorsque plusieurs points de vue existent sur le même phénomène du monde réel. La généralité est la représentation multiple de la réalité en fonction du niveau de détail.

Typologie de l'imperfection des données selon (Bouchon-Meunier, 1995) : dans un premier ouvrage (Bouchon-Meunier, 1993), l'auteur distingue deux types d'imperfection : l'incertitude et l'imprécision. Dans son deuxième ouvrage (Bouchon-Meunier, 1995), elle distingue un troisième facteur qui est l'incomplétude. Elle définit l'incertitude par la validité des données, l'imprécision des données par le caractère vague ou approximatif de la sémantique utilisée et l'incomplétude par l'absence des données.

Typologie de l'incertitude des données selon (Klir & Yuan, 1995) : Klir et Yuan proposent une typologie en deux classes : le flou et l'ambiguïté. L'ambiguïté fait référence au conflit et à la non-spécificité.

Typologie de l'imperfection des données selon (Smets, 1997) : Smets établit une classification de l'imperfection des données. Elle est éclatée en trois groupes : l'imprécision, l'incohérence et l'incertitude. L'imprécision est relative au contenu des données qui peuvent avoir ou non des erreurs, dû à une approximation. L'incohérence est relative aux données conflictuelles et contradictoires. L'incertitude est induite par des données subjectives (c'est-à-dire l'avis de l'observateur) ou objectives (c'est-à-dire les propriétés de la donnée).

2.3.1.2 Typologies spécifiques de l'imperfection des données dans un domaine donné

Nous avons identifié dans la littérature sept typologies de l'imperfection des données spécifique à un domaine donné : la typologie de l'incertitude des données géographiques (Fisher, et al., 2005), la typologie de l'imperfection d'un ensemble des données textuelles décrivant des objets ethnographiques (Olteanu, et al., 2006), la typologie de l'imperfection des données financières (Casta, 2009), la typologie de l'imperfection des données archéologiques (Desjardin, et al., 2012), la typologie de l'imperfection des données spatiales (Snoussi & Davoine, 2014)

et la typologie de l'imperfection des données du processus de récupération d'informations (Sta, 2017).

Typologie de l'incertitude des données géographiques selon (Fisher, et al., 2005) : Cette typologie est la plus utilisée dans la communauté géomatique [Zoghلامي, 2013]. (Fisher et al.) utilisent le concept de l'incertitude comme un chapeau qui englobe tous les autres concepts (autrement dit, il s'agit de l'imperfection). Une donnée est bien ou mal définie. Dans le cas où la donnée est bien définie, elle est soumise à l'incertitude. Dans les autres cas, la donnée est mal définie et l'imperfection de la donnée est due à l'imprécision, l'ambiguïté et/ou l'incomplétude. L'ambiguïté se produit lorsqu'il y a un doute sur la manière de définir un objet ou un phénomène. Deux types d'ambiguïtés sont reconnus à savoir le désaccord et le manque de spécificité.

Typologie de l'imperfection d'un ensemble des données textuelles décrivant des objets ethnographiques selon (Olteanu, et al., 2006) : Olteanu et al. se basent sur les travaux de (Fisher, 2003) et (Fisher, et al., 2005) pour classifier l'imperfection d'un ensemble de données textuelles décrivant des objets ethnographiques. Quatre types d'imperfection sont distingués : l'imprécision qui est la difficulté d'exprimer clairement et précisément les connaissances, l'incertitude qui concerne le doute sur la validité d'une information, le niveau de détail qui est relatif aux connaissances présentées à plusieurs granularités et l'incomplétude qui fait référence à l'absence de l'information.

Typologie de l'imperfection des données financières selon (Casta, 2009) : Casta établit une typologie de l'imperfection des données issues de l'activité économique qui couvre l'incertitude, l'imprécision et l'erreur. L'incertitude désigne l'état d'un agent qui n'a aucune connaissance sur l'avenir, car l'ensemble des événements possibles, ou éventualités, comprend plusieurs éléments qui n'obéissent pas à un déterminisme strict et connu. L'imprécision désigne un manque de rigueur ou d'une contrainte opérationnelle qui affecte la mesure ou, le plus souvent dans les sciences sociales, d'une ambiguïté inhérente à la formulation des concepts. L'erreur est définie par la dispersion des gains et des pertes probables autour de leur moyenne.

Typologie de l'imperfection des données archéologiques selon (Desjardin, et al., 2012) : les auteurs se fondent sur la typologie de (Fisher, et al., 2005) pour proposer une typologie de l'imperfection adaptée au contexte des données archéologiques. Ils classifient les imperfections en incertitude, imprécision, ambiguïté et incomplétude. L'incertitude est le fait qu'il y a un doute sur la validité de la connaissance. L'imprécision est la difficulté à exprimer clairement

les connaissances. Notamment lorsque ces dernières ne sont pas suffisamment définies ou manquent de précision. L'ambiguïté est le fait qu'il y a une difficulté à être d'accord. Il y a conflit si au minimum deux classifications contradictoires pour un unique objet sont possibles. Lorsqu'une définition d'une relation ou d'un objet peut amener à plusieurs sens, ou lorsque l'échelle de l'analyse est susceptible d'amener à de multiples interprétations, c'est l'ambiguïté. L'incomplétude est le fait qu'il y a des connaissances manquantes ou partielles.

Typologie de l'imperfection des données spatiales selon (Snoussi & Davoine, 2014) : Snoussi propose une classification construite spécifiquement pour caractériser les données spatiales. Cette taxonomie distingue trois types d'imperfections : une imprécision qui se produit lorsque la valeur réelle est située dans un sous-ensemble défini de valeurs, une incohérence des valeurs et une incertitude qui est la connaissance partielle de la valeur réelle de l'information.

Typologie de l'imperfection des données du processus de récupération d'informations selon (Sta, 2017) : Sta propose plusieurs types de données imparfaites au cours du processus de récupération d'informations et d'intégration des données dans les villes intelligentes. Cette imperfection peut revêtir plusieurs formes : une information incertaine qui reflète un manque de connaissances, une information imprécise qui traduit une non-spécificité, une information vague relative à une information ambiguë et une information manquante qui concerne une information incomplète ou non trouvée.

Typologie de l'imperfection des données saisies par les malades d'Alzheimer (Ghorbel, et al., 2020) : nous avons proposé plusieurs types d'imperfections dégagées des données saisies par les malades d'Alzheimer, comme l'incertitude, la confusion, la connaissance erronée, l'oubli, l'inconnaissance, l'imprécision, la circonlocution et l'inutilité qui sont des types directs, et l'incohérence qui est un type indirect.

Le Tableau 1 résume les différents types de typologies, illustrés par des schémas représentatifs pour chacune d'elle.

Tableau 1. Tableau récapitulatif des typologies de l'imperfection des données existantes

	Recherche	Type de la typologie	Concepts	Schéma
Typologie générique	[Niskanen, 1989]	Non-precision	Incertitude, imprécision, ambiguïté et généralité	<pre> graph TD NP[Non-precision] --> I[incertitude] NP --> IMP[imprécision] NP --> A[ambiguïté] NP --> G[généralité] I --> E[erreur] IMP --> O[Ontologique] IMP --> L[Linguistique] IMP --> EP[Epistémologique] </pre>
	[Bouchon-Meunier, 1995]	Imperfection	Incertitude, imprécision et incomplétude	<pre> graph TD ID[Imperfection des données] --> I[Incertitude] ID --> IMP[Imprécision] ID --> INC[Incomplétude] </pre>
	[Klir et Yuan, 1995]	Incertitude	Flou et ambiguïté	<pre> graph TD ID[Incertitude des données] --> F[Flou] ID --> A[ambiguïté] A --> C[Conflit] A --> NS[Non-spécificité] </pre>
	[Smets, 1997]	Imperfection	Incertitude, incohérence et imprécision	<pre> graph TD ID[Imperfection des données] --> I[Incertitude] ID --> INC[Incohérence] ID --> IMP[Imprécision] I --> DSE[Données sans erreurs] I --> DAE[Données avec erreurs] INC --> DAE INC --> DSE IMP --> O[Objective] IMP --> S[Subjective] DSE --> A[Absence] DSE --> F[Flou] DAE --> P[Propension] DAE --> D[Disposition] </pre>

Typologie spécifique	[Fisher et al., 2005]	Incertitude	Objet bien défini et Objet mal défini	<pre> graph TD A[Incertitude des données géographiques] --> B[Objet bien défini] A --> C[Objet mal défini] C --> D[Vague] C --> E[Ambiguïté] D --> F[Désaccord] E --> G[Manque de spécificité] </pre>
	[Olteanu et al., 2006]	Imperfection	Incertitude, imprécision, niveau de détail et incomplétude	<pre> graph TD A[Imperfection d'un ensemble de données textuelles décrivant des objets ethnographiques] --> B[Incertitude] A --> C[Imprécision] A --> D[Incomplétude] A --> E[Niveau de détail] </pre>
	[Casta, 2009]	Imperfection	Incertitude, erreur et imprécision	<pre> graph TD A[Imperfection des données financières] --> B[Incertitude] A --> C[Erreur] A --> D[Imprécision] </pre>
	[Desjardin et al., 2012]	Imperfection	Incertitude, imprécision, ambiguïté et incomplétude	<pre> graph TD A[Typologie de l'imperfection des données archéologiques] --> B[Incertitude] A --> C[Imprécision] A --> D[Ambiguïté] A --> E[Incomplétude] B --> F[Approximation] B --> G[Vague] C --> G C --> H[Conflit] D --> H D --> I[Non spécificité] E --> I E --> J[Lacune] I --> K[Absence] </pre>

	[M. Snoussi & Davoine, 2014]	Imperfection	Imprécision, incohérence et connaissance partielle	<pre> graph TD A[Imperfection des données spatiales] --> B[Imprécision] A --> C[Incohé-] A --> D[Connaissance partielle] </pre>
	[Sta, 2017]	Imperfection	Incertitude, imprécision, incomplétude et ambiguïté	<pre> graph TD A[Imperfection des données financières] --> B[Incertitude] A --> C[Incomplétude] A --> D[ambiguïté] A --> E[Imprécision] </pre>
	[Ghorbel et al, 2020]	Imperfection	<p>Types « directes » (Incertitude, confusion, erreur de saisie, connaissance erronée, oubli, inconnaisance, imprécision, circonlocution et inutilité) et</p> <p>Types « indirectes » « (incohérence, incomplétude et redondance)</p>	<pre> graph TD A[Imperfection des données saisies par des patients d'Alzheimer] --> B[Incertitude] A --> C[Erreur de saisie] A --> D[inconnaisance] A --> E[oubli] A --> F[inutilité] B --> G[confusion] C --> H[Connaissance erronée] D --> I[imprécision] E --> J[circonlocution] F --> K[Redondance] G --> H I --> L[Incomplétude] J --> L H --> M[Incohérence] </pre>

2.3.1.3 Discussion

Il existe un grand nombre de typologies d'imperfection des données. Des auteurs tels que (Niskanen, 1989) et (Bouchon-Meunier, 1995) proposent des typologies génériques tandis que d'autres, tels que (Fisher, 2005), (Desjardin et al., 2012) et (Sta, 2017), proposent des typologies pour des domaines spécifiques. Certaines de ces typologies correspondent mieux à une réalité que d'autres. En outre, il n'existe pas une définition arrêtée pour qualifier des données imparfaites, telles que l'incertitude et l'imprécision. Nous notons également que la majorité de ces typologies partagent trois concepts communs qui sont l'imprécision, l'incertitude et l'incomplétude. Les types d'imperfections sont interdépendants. Selon (Bouchon-Meunier, 1995), l'incomplétude est une source d'incertitude et l'imprécision peut aussi être associée à l'incomplétude. Selon (Smets, 1998), l'imprécision fait toujours référence à l'incomplétude. L'imprécision peut être une source d'incertitude, mais pas nécessairement (Smets, 1998).

Les données temporelles peuvent présenter plus d'imperfections que celles proposées dans les typologies génériques en raison de la complexité de ce type de données et de la spécificité qu'elles contiennent (c'est-à-dire qu'elles peuvent être numériques ou basées sur un langage naturel). Elles peuvent également dépendre du contexte général et peuvent être soumises à plusieurs facteurs susceptibles d'interférer dans la spécification du type d'imperfection. Ainsi, les typologies génériques ne peuvent pas être adaptées aux données temporelles (c'est-à-dire qu'elles sont inadéquates). Par exemple, si nous avons les informations « J'oublie la dernière fois où j'ai rendu visite à mon oncle qui vit au Japon », les données temporelles font référence à un « oubli », qui est un type d'imperfection qu'aucune des typologies génériques existantes n'inclut. Un autre exemple est « Le premier jour de la semaine, nous aurons une réunion ». Dans cet exemple, les données temporelles indiquent une « circonlocution », qui est un autre type d'imperfection qu'aucune des typologies génériques ne contient. Par ailleurs, **à notre connaissance, il n'existe pas une typologie des imperfections des données temporelles.**

2.3.2 Théories de l'imperfection des données

Plusieurs théories ont été exploitées dans plusieurs domaines pour tenir compte des imperfections de données. Parmi ces théories nous citons : **la théorie de probabilités, la théorie de possibilités et la théorie de l'évidence.** Tableau 2 présente les imperfections traitées par chaque théorie.

Tableau 2. Théories et type d'imperfection(s) traité(s)

	Théorie de probabilité	Théorie de possibilité	Théorie de l'évidence
Type d'imperfection traité	Incertitude aléatoire (par conséquent l'incohérence)	Incertitude épistémique Imprécision Incertitude+Imprécision (par conséquent l'incomplétude et l'incohérence)	Conflit (contradiction) (par conséquent l'incohérence) Incertitude Imprécision

2.3.2.1 Théorie de probabilité

Les approches probabilistes se basent sur un cadre mathématique rigoureux qui est **la théorie des probabilités**. Cette théorie est employée dans les systèmes de modélisation d'information imparfaite du type incertitude, dans la fusion de l'information et dans les systèmes d'aide à la décision.

Dans le cadre de la théorie des probabilités, l'information est modélisée à l'aide d'une distribution de probabilités $p(\cdot)$. Disposant de cette modélisation, la seule connaissance disponible sur la réalisation de divers singletons x_n , $n = 1, \dots, N$ est $p(x_n)$ qui désigne la probabilité pour que x_n se produise :

$$p : \Omega \rightarrow [0, 1]$$

$$x_n \rightarrow p(x_n)$$

Cette application doit, nécessairement, vérifier la condition de normalisation suivante :

$$\sum_{x_n \in \Omega} p(x_n) = 1.$$

La probabilité $P(A)$ d'un évènement $A \subseteq \Omega$, désigne la somme des probabilités élémentaires pour la réalisation de tous les singletons constituant l'évènement A . Elle doit vérifier les axiomes suivants :

$$P(\Omega) = 1; P(\varphi) = 0$$

$$P(A, B) = P(A) + P(B) : \text{Si } A, B \subseteq \Omega \text{ et } A \cap B = \varphi$$

$$P(A) + P(A^c) = 1, \text{ avec } A^c \text{ l'évènement complémentaire de } A$$

Réseaux bayésiens.

Le réseau bayésien, introduit par Judea Pearl vers les années 80 (Pearl, 1988), est un modèle graphique probabiliste. Il combine une représentation de la connaissance sous forme probabiliste (une incertitude sur la connaissance) et graphique (des relations de dépendance directe). Il permet d'exprimer une distribution de probabilité sur un ensemble de variables aléatoires à l'aide d'un produit de distributions locales de probabilités. Il est formé par deux composantes :

- Une première composante graphique consiste en un graphe orienté sans circuit. Ce graphe est formé par des nœuds représentant les variables pertinentes du domaine et des arcs représentant les relations de dépendance entre ces variables ;
- Une seconde composante numérique consiste en un ensemble de distributions de probabilités conditionnelles de chaque nœud dans le contexte de ses parents.

Définition. Un réseau bayésien $B = (G, P)$ est défini par :

- Un graphe dirigé sans circuit $G = (X, E)$ où X représente les nœuds et E représente les arcs ;
- Un espace probabiliste (Ω, P) ;
- Un ensemble de variables aléatoires $X = \{X_1, \dots, X_n\}$ associées aux nœuds du graphe et définies sur (Ω, P) telles que

Avec $Pa(X_i)$ est l'ensemble des parents du nœud X_i dans G .

Les réseaux bayésiens causaux sont ceux dans lesquels l'arc est orienté de la cause vers la conséquence. L'information du nœud « Cause » peut modifier les croyances que nous avons sur le nœud « Conséquence ». Dans un graphe causal, le sens de la flèche indique le sens de la causalité.

La construction d'un réseau bayésien est basée sur trois étapes. La première étape consiste en l'identification des variables. La seconde étape correspond à définir les arcs du graphe. La troisième étape correspond à spécifier numériquement les distributions de probabilités conditionnelles pour chaque variable définie.

L'inférence bayésienne consiste à calculer la probabilité a posteriori au niveau de certaines variables en prenant en compte d'autres variables (observations). Des valeurs sont affectées aux variables d'observation, puis un algorithme d'inférence est utilisé pour calculer les probabilités a posteriori des autres variables.

2.3.2.2 Théorie de possibilité

La théorie des possibilités a été introduite par Zadeh en 1978 (Zadeh, 1978), ensuite développée par Dubois et Prade en 1988 (Dubois & Prade, 1988). Cette théorie est capable, à la fois, de modéliser différents types d'informations (numérique, symbolique, qualitative, etc.) et de manipuler différentes formes d'imperfection (ambiguïté, imprécision, etc.) (Ammar, 2014). Les fonctions d'appartenance cèdent leur place aux distributions de possibilités π pour représenter à la fois l'imprécision, l'incertitude et l'ambiguïté. La théorie des possibilités fournit des modèles qui peuvent gérer, à la fois, différents types d'informations et différentes formes d'imperfections (Ammar, 2014). Contrairement à la théorie des probabilités et celle des ensembles flous. Cette théorie permet de caractériser les événements par deux mesures, la possibilité Π , et la nécessité N .

Une distribution de possibilité, notée $\pi_x(u)$, est une application de l'univers du discours à l'échelle $[0, 1]$ traduisant une connaissance partielle de l'univers (Dubois & Prade, 1998). D'après (Zadeh, 1978), la mesure de possibilité d'un événement A , notée Π , est déduite de la distribution de possibilité $\pi_x(u)$ par :

$$\Pi(A) = \sup_{u \in A} \pi_x(u)$$

$$N(A) = 1 - \Pi(A)$$

La mesure de possibilité est insuffisante pour caractériser la réalisation d'un événement, car elle n'est pas capable de discriminer deux situations extrêmes :

- L'ignorance totale : $\Pi(A) = 1$ et $N(A) = 1$
- La certitude totale : $\Pi(A) = 1$ et $N(A) = 0$ d'un événement A .

Dans les deux cas, $\Pi(A)$ vaut 1, l'évènement A est complètement possible. Il est donc nécessaire de compléter la mesure de possibilité par une mesure de nécessité (ou de certitude) que l'évènement soit réalisé.

Dans la théorie des possibilités, la certitude est modélisée par une mesure de possibilité élevée et une mesure de nécessité élevée. La certitude absolue est traduite par une mesure de possibilité $\Pi(A) = 1$ et une mesure de nécessité $N(A) = 1$.

L'impossibilité est traduite par une mesure de possibilité nulle $\Pi(A) = 0$ et une mesure de nécessité nulle $N(A) = 0$.

Un degré de possibilité est la borne supérieure d'un degré de probabilité et le degré de nécessité est sa borne inférieure étant donné qu'un évènement certain doit être probable et celui probable doit être possible

$$N(A) \leq P(A) \leq \Pi(A)$$

Deux branches de la théorie des possibilités :

- Approche ordinale : Les valeurs possibles d'une grandeur sont rangées par ordre de plausibilité. Ensuite, une échelle qualitative finie ou ordinale de degrés de possibilité est utilisée.
- Approche numérique : Les degrés de possibilités sont des valeurs réelles de $[0, 1]$ ou des entiers naturels. Il faut les articuler avec les probabilités.

2.3.2.3 Théorie de probabilité et théorie de possibilité : comparaison

La théorie des possibilités est une théorie de l'incertitude consacrée au traitement des informations incertaines, imprécises et, par conséquent, incomplètes. C'est différent de la théorie des probabilités. « *Ce qui est certain doit être probable. Ce qui est probable doit être possible* » (Zadeh, 1978). Tableau 3 résume ces différences.

- La première différence à noter est que la théorie de probabilité traite l'incertitude aléatoire (c'est-à-dire intrinsèquement, présente dans un phénomène aléatoire de sorte que l'écart de connaissances ne peut pas être rempli). C'est une théorie quantitative. Alors que la théorie de possibilité est capable de traiter l'incertitude épistémique, (c'est-à-dire due à une méconnaissance partielle, par exemple en raison de données manquantes, douteuses ou incohérentes), l'imprécision et l'incertitude et l'imprécision à la fois. En effet, la théorie des possibilités fournit des modèles capables de gérer à la fois différents types d'informations et différentes formes d'imperfection (Dubois & Prade, 1990). C'est une théorie qui peut être quantitative ou qualitative.
- De point de vue de la modélisation de l'incertitude : **la théorie des probabilités** (en supposant une distribution unique) ne peut modéliser l'ignorance (même partielle) de manière naturelle. Les distributions de probabilité uniformes sur des ensembles finis d'événements élémentaires expriment mieux le hasard que l'ignorance, à savoir l'égalité des chances d'occurrence d'événements élémentaires.

L'ignorance signifie plutôt que chaque occurrence possible est considérée comme tout aussi plausible par un agent, car il n'y a aucune preuve disponible qui étaye l'une d'entre elles, c'est-à-dire qu'il manque simplement des informations. Aucune mesure de probabilité ne peut expliquer un tel état de manque de connaissances (car même une distribution de probabilité uniforme conduit à attribuer des probabilités inégales à au moins deux événements contingents possibles en général). En revanche, **la théorie des possibilités** gère la situation d'ignorance totale d'une manière non biaisée (sans faire aucune hypothèse préalable).

- Une autre différence entre la théorie des probabilités et la théorie des possibilités est que **l'encodage probabiliste** des connaissances est numérique et repose sur une hypothèse d'additivité alors que **l'encodage possibiliste** peut être purement qualitatif, donc moins gourmand en données.
- Le **raisonnement possibiliste** est moins difficile à calculer que le **raisonnement probabiliste** (mais les résultats obtenus peuvent également être moins informatifs).

Tableau 3. Théorie de possibilité VS théorie de probabilité

Théorie de probabilité	Théorie de possibilité
Incertitude aléatoire	Incertitude épistémique Imprécision Incertitude + Imprécision
1 seule mesure : $P(A) = 1 - P(\text{non } A)$	2 mesures : $\Pi(A)$ et $N(A) = 1 - \Pi(\text{non } A)$
$P(A) = P(\text{non } A) = 1/2$ représente l'aléatoire	$\Pi(A) = 1$; $N(A) = 0$ représente l'ignorance
Très quantitatif +	Qualitatif ou quantitatif (proche du qualitatif)

Amplification erreurs	Max, Min Erreurs fixes
Résultats précis Parfois arbitraires Parfois difficiles à obtenir	Résultats prudents parfois trop peu informants mais plus sûrs. Plus facile en générale à obtenir

2.3.2.4 Théorie de l'évidence

Bien que les concepts de base de la théorie de l'évidence aient trouvé leur origine au 17^{ème} siècle dans les travaux du mathématicien Jacob Bernoulli aux années 1654-1705, et ensuite dans les travaux de Johann Heinrich Lambert aux années 1728-1777, les fondements de cette théorie n'ont été posés qu'en 1967 par Arthur Dempster (Dempster, 1967). Puis, Shafer (Shafer, 1976) a montré leur intérêt pour la modélisation des informations imprécises ou incertaines, provenant des sources redondantes, complémentaires ou éventuellement non indépendantes. Cette modélisation repose sur la définition des fonctions de masse m appelées aussi masses de croyance, de crédibilité Cr et de plausibilité Pl .

La théorie de l'évidence représente un outil puissant pour représenter des informations imparfaites permet de modéliser toutes les formes d'imperfection, y compris les conflits.

Dans le cadre de la théorie de l'évidence, l'ensemble Ω est appelé cadre de discernement. La fonction de masse $m(\cdot)$ caractérise la croyance en la réalisation d'un évènement A . Elle est définie alors par :

$$m : 2^\Omega \rightarrow [0, 1]$$

$$\sum_{A \subseteq \Omega} m(A) = 1$$

Une masse m peut être représentée de manière équivalente par deux mesures non additives : une fonction de crédibilité (Believability) $Bel: 2^\Omega \rightarrow [0, 1]$ et une fonction de plausibilité $Pl: 2^\Omega \rightarrow [0, 1]$, définie par:

$$Bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B)$$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B)$$

Où $B \subseteq A$ est équivalent à l'ensemble de toutes les disjonctions de A , c'est-à-dire $B \in 2^A$. Cette fonction de crédibilité doit vérifier la propriété suivante : $Bel(\emptyset) = 0$.

La fonction de plausibilité $Pl(.)$ permet de mesurer l'intensité avec laquelle nous ne doutons pas de A . Également, elle peut s'interpréter comme la croyance maximale en A .

La relation entre Bel et Pl est donnée par la relation :

$$Pl(A) = 1 - Bel(\bar{B})$$

2.4 Les données temporelles

Les données temporelles ont une relation d'ordre. Ceci peut aider à détecter les erreurs. Les données temporelles peuvent être représentées en utilisant des points de temps ou des intervalles de temps. Elles peuvent également être conceptualisées comme quantitatives (définies avec précision, par exemple à l'aide des dates) ou qualitatives (désigne les relations temporelles qualitatives comme « avant »).

2.4.1 Éléments, relations et ontologies temporelles

Intervalles ou points ? La question la plus fondamentale dans tout modèle temporel est le choix de la primitive temporelle de base. La littérature mentionne deux primitives usuelles (Vila, 1994) (Chomicki, 1994) : les instants de temps (ou points de temps, *chronos*) et les intervalles de temps. Un modèle temporel peut être basé sur une des deux primitives ou sur les deux en même temps. Dans la littérature, il y a un débat en cours pour savoir quelle primitive est la plus appropriée, sans gagnant clair. Alors que dans la communauté de recherche sur les bases de données temporelles, l'instant temporel est plus couramment utilisé (Etzion, 1998), dans la communauté de l'intelligence artificielle, l'intervalle de temps ou les approches mixtes sont les plus populaires (Vila, 1994). Dans (Vila, 1994), il a été démontré que le choix de la primitive de base dépend principalement des exigences de l'application. (Nagypál & Motik, 2003) pensent que les intervalles de temps sont plus proches de la perception intuitive humaine du temps.

2.4.1.1 Éléments temporels

Selon l'enquête présentée par (Ermolayev, et al., 2014) et le groupe de travail W3C qui a conçu l'ontologie TIME (<http://time.di.unimi.it/>), l'élément temporel est une entité utilisée comme élément constitutif d'une théorie du temps. Ces éléments pourraient être « Points de temps » et « Intervalles de temps ».

Points de temps. Une ligne de temps est constituée de points temporels ou d'instant. La valeur d'un instant reflète sa position sur la ligne du temps. Selon (Ermolayev, et al., 2008), si

l'instant est le présent alors la valeur est égale à zéro, si l'instant est strictement avant le présent, la valeur est négative et si l'instant est strictement après le présent alors la valeur est positive.

Intervalles de temps. (Allen, 1983) considère les intervalles de temps comme des entités primitives. Il considère ces éléments comme ceux qui ont une durée en se référant à une relation « pendant » permettant de définir les hiérarchies d'intervalles. Pratiquement toutes les autres approches proposées ne considèrent pas les intervalles comme leurs éléments primitifs, mais comme des structures basées sur des points (Ermolayev, et al., 2014). Selon (Ermolayev, et al., 2014), les intervalles pourraient être convexes (co) ou non convexes (n), ouverts (o) ou fermés (c), bornés (b) ou illimités (u) et flous (f) ou croustillant (cr). Par exemple, Alur et al. (1996) admettent que les intervalles de temps sont ouverts, semi-ouverts ou fermés ; borné ou illimité - c'est-à-dire ayant l'une des formes suivantes : $[a, b]$, $[a, b)$, $[a, \infty)$, $(a, b]$, (a, b) , (a, ∞) , où $a \leq b$ pour $a, b \in \mathbb{R}_{\geq 0}$.

2.4.1.2 Relations temporelles

Les relations temporelles pourraient être :

- Relation basée sur des points
- Relation basée sur l'intervalle
- Relation point-intervalle
- Relation intervalle-point

Nous présentons dans ce qui suit les deux formalismes temporels qualitatifs les plus cités : l'algèbre des intervalles d'Allen (Allen, 1983) et l'algèbre des points de Vilain et Kautz (Vilain & Kautz, 1986).

2.4.1.2.1 Algèbre des intervalles d'Allen


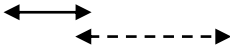

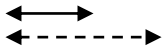
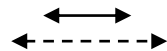
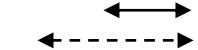
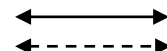
Allen a défini 13 relations de base qui décrivent toutes les manières possibles d'ordonner les extrémités de deux intervalles temporels précis. Ces relations, dites relations d'Allen, sont habituellement dénotées par : Before, After, Meets, Met-by, Overlaps, Overlapped-by, During, Contains, Starts, Started-by, Finishes, Finished-by, et Equal.

Soit $I = [I^-, I^+]$ et $J = [J^-, J^+]$ deux intervalles temporels précis non réduits à un point. Les relations d'Allen 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 I Starts J correspond à :

$$(I^- = J^-) \wedge (I^- < J^+) \wedge (I^+ > J^-) \wedge (I^+ < J^+)$$

Le Tableau 4 résume les 13 relations d'Allen entre les deux intervalles temporels précis I et J.

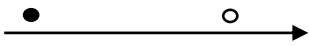
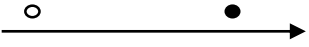
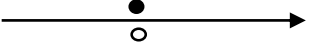
Tableau 4. Relations d'Allen entre deux intervalles de temps précis.

Relation (I, J)	Relations entre les bornes	Illustration	Inverse
Before	$I^+ < J^-$		After
Meets	$I^+ = J^-$		Met-by
Overlaps	$(I^- < J^-) \wedge (I^+ > J^-) \wedge (I^+ < J^+)$		Overlapped-by
Starts	$(I^- = J^-) \wedge (I^+ < J^+)$		Started-by
During	$(I^- > J^-) \wedge (I^+ < J^+)$		Contains
Ends	$(I^- > J^-) \wedge (I^+ = J^+)$		Ended-by
Equals	$(I^- = J^-) \wedge (I^+ = J^+)$		Equals

2.4.1.2.2 Vilain et Kautz

Le second modèle temporel qualitatif assez utilisé est celui introduit par Vilain et Kautz (Vilain & Kautz, 1986). Dans ce modèle, trois relations temporelles qualitatives permettent de comparer deux instants T1 et T2 : T1 Précède (Précède) T2, T1 Follow (Suit) T2 et T1 est Same (Égale à) T2, comme le montre le Tableau 5.

Tableau 5. Algèbre des points de Vilain et Kautz

Relation	Symbole	Illustration	Inverse
T1 Précède T2	<		T2 Suit T1
T1 Suit T2	>		T2 Précède T1
T1 Égale à T2	=		T2 Égale à T1

2.4.2 Représentation et raisonnement sur les données temporelles dans le web sémantique

La représentation des données temporelles dans l'ontologie est un besoin crucial. Cependant, les langages d'ontologie tels que OWL fournissent un support minimal, car ils sont tous

basés sur des relations binaires qui relient simplement deux instances. Cela explique l'émergence de nombreuses approches pour représenter et raisonner sur les données temporelles en ontologie.

2.4.2.1 Représentation des données temporelles dans le Web Sémantique

À partir de la littérature, nous identifions plusieurs approches permettant la modélisation de la dimension temporelle dans l'ontologie : les logiques de description temporelles (Artale & Franconi, 2000), versioning (Klein & Fensel, 2001), réification (Noy, et al., 2006), relations naires (Noy, et al., 2006) et 4D-fluents (Welty & Fikes, 2006). Les Tableaux 5 et 6 résument ces approches ainsi que les travaux utilisant et étendant ces approches.

Nous classons les approches en deux catégories : (i) les approches qui étendent la syntaxe OWL ou RDF en définissant de nouveaux opérateurs pour incorporer les données temporelles, et (ii) les approches qui sont implémentées directement en utilisant OWL ou RDF pour représenter les données temporelles sans étendre leurs syntaxes. La première catégorie comprend la logique de description temporelle, les « concrete domains » et le RDF temporel (Gutierrez, et al., 2005)

La logique de description temporelle étend les logiques de description standard avec de nouvelles sémantiques temporelles telles que « until ». Cette approche conserve la décidabilité et ne provoque pas la redondance des données. Les principales composantes du formalisme de la logique de Description (DL) sont les classes, leurs propriétés et les individus. En raison de l'importance des aspects temporels, les logiques de description temporelle (TDL) (Artale & Franconi, 2000), (Lutz, et al., 2008) étendent la logique de description standard en ajoutant des capacités expressives au DL non temporelles, avec des constructions temporelles supplémentaires telles que « toujours dans le passé », « parfois dans le futur ». Cependant, elles nécessitent d'étendre la syntaxe et la sémantique OWL, qui est une tâche fastidieuse.

Concrete Domains nécessite l'introduction de types de données et d'opérateurs supplémentaires dans OWL. Plusieurs extensions basées sur cette approche ont été proposées. OWL-MeT (Keberle, et al., 2007) étend le langage OWL tout en fournissant des mécanismes de raisonnement. TL-OWL (Kim, et al., 2008) combine les constructions OWL avec les relations d'Allen pour implémenter des classes temporalisées. (Milea, et al., 2012) proposent une nouvelle approche appelée OWL temporel (TOWL) qui est une extension de l'Ontology Web Language Description Logics (OWL-DL). Les auteurs introduisent trois extensions: (1) Concrete

Domains, permettant la représentation des restrictions en utilisant des prédicats binaires de domaine concret, (2) Temporal Representation, introduisant des points de temps, des relations entre les points de temps, les intervalles et les relations entre les 13 intervalles d'Allen dans le langage, et (3) TimeSlices / Fluents, mettant en œuvre une vision perdurantiste des individus et permettant la représentation d'aspects temporels complexes.

RDF temporel étend RDF en annotant les propriétés avec les informations sur l'intervalle de temps qu'elles conservent. Cela nécessite également d'étendre la syntaxe de RDF. Cependant, Temporal RDF ne permet pas d'exprimer de relations qualitatives, car elle est limitée à l'utilisation de triplets RDF uniquement. Des exemples de représentation temporelle basée sur Temporal RDF sont présentés dans (Hurtado & Vaisman, 2006). Les auteurs améliorent Temporal RDF avec la prise en compte d'intervalles non définis. (Gutierrez, et al., 2006) présentent un cadre complet pour incorporer le raisonnement temporel dans RDF, produisant des graphes RDF temporels. Ils définissent une notion syntaxique de graphes RDF temporels. stRDF (Koubarakis & Kyzirakos, 2010) est un modèle de données de contraintes qui étend RDF et qui permet de représenter des données spatiales et temporelles. Il n'exprime pas de données temporelles qualitatives et n'a aucun support de raisonnement.

La deuxième catégorie comprend :

Le **Versionning** est décrit comme la capacité de gérer les changements dans l'ontologie en créant différentes versions de l'ontologie (Klein & Fensel, 2001). Lorsqu'une ontologie est modifiée, une nouvelle version est créée pour représenter l'évolution temporelle de l'ontologie. Il souffre de la redondance des données en raison du grand nombre de versions créées chaque fois qu'un changement unique se produit dans la version d'ontologie précédente. Toutes ces versions créées sont considérées comme indépendantes les unes des autres du fait de l'absence de relations entre les concepts et les propriétés (qui évoluent) des différentes versions. Les langages d'ontologie, tels que OWL, sont basés sur des relations binaires et ne prennent pas en compte la dimension temporelle entre les différentes versions d'une ontologie. La recherche de n'importe quel élément dans l'ontologie nécessite des recherches exhaustives dans toutes les versions créées (Batsakis, et al., 2017). La plupart des implémentations adoptent une variété de stratégies d'optimisation pour garantir que des copies entières de l'ontologie ne sont pas générées pour chaque nouvelle version. (Völkel & Groza, 2006) introduisent une approche de gestion des versions centrée sur RDF et une implémentation nommée SemVersion. Cette approche fournit un versionnage structurel et sémantique pour les modèles RDF et les langages d'ontologie basés sur RDF comme RDFS, en considérant l'enrichissement des nœuds blancs comme

une technique pour identifier les nœuds vides. (Bedi & Marwaha, 2007) introduisent une approche qui combine les concepts de versionnage de trame temporelle et de slot avec l'ontologie pour créer des ontologies temporelles étiquetées avec versionnage embarqué. Les auteurs proposent la création d'ontologies OWL à balises temporelles, avec deux nouvelles balises, «rdf:Validity» et «rdf:Timestamp», et un schéma est présenté pour la valeur des «rdf:ID» et «rdf:Resource» Des balises pour rendre les ontologies étiquetées temporelles cohérentes avec les ontologies non temporelles. (Grandi, 2013) propose un schéma de stockage basé sur une relation temporelle qui peut être utilisée pour représenter et gérer la structure de classe d'une ontologie multi-version (c'est-à-dire incarnant une hiérarchie de classes en forme d'arbre). (Grandi, 2016) étend (Grandi, 2013) en considérant les ontologies avec une hiérarchie de classes structurée comme un graphe dirigé général, qui prend également en charge les classes d'héritage et d'intersection multiples, et en montrant comment les ontologies multi-versions doivent être traitées, pour le traitement des requêtes de personnalisation basées sur l'ontologie. (Zekri, et al., 2017) proposent τ OWL, un framework de type τ XSchema, qui permet de créer une ontologie temporelle OWL 2 à partir d'une ontologie OWL 2 conventionnelle et d'un ensemble d'annotations logiques et physiques.

La **réification** est une technique pour représenter les relations n-aires en utilisant un langage qui n'autorise que des relations binaires (relations reliant deux instances) comme OWL (Batsakis & Petrakis, 2011). Elle propose de représenter une relation n-aire comme une relation R entre deux objets A et B, à un instant t. D'où l'expression R (A, B, t). Cependant, cette approche souffre de la redondance des données à chaque création d'un nouvel objet quand une relation temporelle doit être représentée (Batsakis & Petrakis, 2011). De plus, elle offre des capacités de raisonnement et d'inférence OWL limitées (Harbelot, et al., 2013). Des exemples de représentation temporelle basée sur la réification sont présentés dans (Cui, et al., 2010). Les auteurs combinent la représentation temporelle avec l'application de règles SWRL spécifiques pour représenter les récits cliniques. (O'Connor & Das, 2011) proposent une approche pour représenter et interroger des données temporelles dans OWL. Leur approche utilise un modèle temporel léger pour encoder la dimension temporelle des données. Il utilise également le langage SWRL (Semantic Web Rule Language) basé sur OWL et le langage de requête OWL basé sur SWRL (SQWRL) pour raisonner et interroger les informations temporelles représentées à l'aide du modèle proposé.

L'approche des **relations N-aires** propose de représenter une relation N-aire comme deux propriétés liées chacune à un nouvel objet (plutôt que comme l'objet d'une propriété, comme

le fait la réification). Elle ne nécessite qu'un seul objet supplémentaire pour chaque intervalle de temps et conserve la sémantique des propriétés. Cependant, il souffre de la redondance des données dans le cas des propriétés inverses et symétriques (Noy, et al., 2006). Plusieurs outils ont été proposés à partir des relations N-aires. CNTRO (Cui, et al., 2010) adoptent l'approche N-aires combinée aux concepts de temps définis dans OWL-Time. (Batsakis & Petrakis, 2011) enrichissent l'approche des relations N-aires avec des relations temporelles qualitatives entre intervalles de temps. (Santos, et al., 2011) présentent un plug-in pour l'éditeur Protégé nommé « FONTE » prenant en charge l'édition d'ontologies temporelles basées sur les relations N-aires. (O'Connor & Das, 2011) combinent l'approche des relations N-aires avec les règles SWRL pour représenter des données quantitativement temporelles. (Preventis, et al., 2012) présentent un outil similaire pour les relations 4D-fluents et N-aires. (Motik, 2012) propose une extension de RDF et OWL supportant les relations N-aires avec représentation temporelle. Un support de requête est également proposé et implémenté. (Anagnostopoulos, et al., 2013) proposent CHORONOS, un Plugin Tab pour l'éditeur Protégé. L'ontologie temporelle est ajoutée à l'ontologie statique basée sur l'approche des relations n-aires. Les auteurs choisissent de représenter à la fois des informations qualitatives temporelles et quantitatives.

L'approche des **graphes nommés** « **Named Graph** » (Tappolet & Bernstein, 2009) représentent le contexte temporel d'une propriété par inclusion d'un triplet représentant la propriété dans un graphe nommé. Les graphes nommés ne font pas partie de la spécification OWL (c'est-à-dire qu'il n'y a pas de constructions OWL traduites en graphes nommés) (Motik, 2010) et ils ne sont pas supportés par les raisonneurs OWL (Batsakis & Petrakis, 2011). (Tappolet & Bernstein, 2009) introduisent un langage de requête temporelle basé sur SPARQL, combiné à un mécanisme d'indexation qui ne s'applique qu'aux intervalles temporels définis quantitativement.

L'approche **4D-Fluents** représente en OWL les intervalles de temps et leur évolution. Elle maintient l'expressivité de OWL et supporte le raisonnement. Elle définit une tranche de temps pour encoder la dimension temporelle occupée par un concept variant dans le temps. Elle est représentée sous la forme d'objets à 4 dimensions, le quatrième étant la tranche de temps. Les changements sont limités, car ils se produisent uniquement sur les parties temporelles de l'ontologie tout en gardant la partie statique inchangée. Par conséquent, cela réduit le problème de la redondance. Cependant, elle introduit deux objets supplémentaires pour chaque relation temporelle au lieu d'un comme c'est le cas pour la réification et les relations n-aires. Les instances de temps et les intervalles de temps sont représentés comme des instances d'une classe

« TimeInterval », qui à son tour est liée à des concepts variant dans le temps. Selon (Harbelot, et al., 2013), elle a une structure simple permettant de transformer facilement une ontologie statique en une ontologie dynamique. L'approche 4D-fluents a inspiré plusieurs implémentations. (Krieger, 2010) qui a proposé le système MUSING utilise à la fois une approche basée sur les fluents 4D et une approche alternative basée sur l'extension de RDF avec des annotations temporelles en conjonction avec OWL-Time. Cependant, MUSING ne fournit aucun support de raisonnement qualitatif. SOWL (Batsakis & Petrakis, 2011), une ontologie spatio-temporelle OWL, utilise l'approche 4D-fluents comme composante temporelle. Dans SOWL, l'approche 4D-fluents, est étendue avec des relations temporelles qualitatives entre des intervalles de temps dont les extrémités droite et gauche ne sont pas spécifiées. Ceci est mis en œuvre en introduisant des relations temporelles sous forme de relations d'objet entre des intervalles de temps. Cela peut être l'une des relations d'Allen. Les relations temporelles des relations qualitatives sont prises en charge via les règles SWRL. (Batsakis & Petrakis, 2011) a travaillé sur l'augmentation du pouvoir expressif de la représentation en améliorant le 4D-fluent avec des relations temporelles qualitatives entre des intervalles de temps dont les points de départ et d'arrivée ne sont pas spécifiés. Comme les 4D-fluents nécessitent des intervalles temporels fermés pour la représentation des informations temporelles, ils ont représenté les points d'extrémité inconnus avec des relations d'Allen. Par exemple, si l'intervalle t_1 est connu et t_2 est inconnu, mais que t_2 commence lorsque t_1 se termine, alors il est possible d'affirmer que t_1 « rencontre » t_2 . PROTON (Papadakis, et al., 2011) est basé sur la 4D-Fluents. Cette ontologie est combinée avec un raisonnement basé sur le calcul de situation implémenté dans Prolog. (Anagnostopoulos, et al., 2013) proposent un mécanisme de raisonnement temporel basé sur le modèle SOWL. (Harbelot, et al., 2013) et (Herradi, et al., 2017) proposent un raisonnement sur les relations temporelles au moyen de règles SWRL intégrées dans l'ontologie et représentées sur la base de l'extension de l'approche 4D-fluents proposée par (Batsakis & Petrakis, 2011). La relation temporelle peut être l'une des relations d'Allen. (Batsakis, et al., 2017) étendent l'approche 4D-fluents pour représenter des instants de temps. (Herradi, et al., 2017) montrent la relation entre la sémantique OWL 2 et la représentation des intervalles de temps en introduisant une représentation qualitative de l'information temporelle basée sur un ensemble de règles SWRL, qui permet un mécanisme de raisonnement solide et complet. (Ghorbel, et al., 2018) proposent une approche pour représenter et raisonner sur des données temporelles quantitatives et qualitatives imprécises en OWL 2. Les auteurs étendent l'approche 4D-fluents pour représenter des intervalles et des points de temps imprécis. Cette extension traite également des relations qualitatives entre ces

derniers. Ils étendent également l’algèbre d’intervalle d’Allen. Ils proposent des relations d’intervalle temporelles précises et imprécises. L’inférence est effectuée à l’aide d’un ensemble de règles SWRL et Mamdani IF-THEN.

2.4.2.2 Raisonnement sur des données temporelles dans le Web Sémantique

La plupart des approches proposées étendent l’algèbre d’intervalle d’Allen. Certaines étendent cette algèbre pour proposer des relations temporelles entre des intervalles de temps parfaits (Freksa, 1992), (Madkour, et al., 2016), (Herradi, 2018) (Li, et al., 2020) et (Nys, et al., 2018). Dans (Dubois & Prade, 1989), les auteurs représentent un intervalle de temps précis comme une paire de distributions de possibilité qui définissent les valeurs possibles des bornes d’un intervalle. Plusieurs approches étendent l’algèbre des intervalles d’Allen pour proposer des relations temporelles entre des intervalles de temps imparfaits. (Guesgen, et al., 1994) proposent des relations temporelles floues vues comme des ensembles flous de relations d’Allen comme la relation « Fuzz-Meets » qui recouvre la relation « Meets » d’Allen. (Ryabov & Trudel, 2004) proposent une approche basée sur la théorie des possibilités pour modéliser des relations d’intervalle temporel incertaines en attribuant un degré de préférence à chaque relation d’Allen de base. Cette approche est proposée dans le cadre d’un réseau d’intervalles temporels probabilistes ; où les nœuds sont des intervalles de temps et les arêtes sont des relations d’intervalle incertaines. (Badaloni & Giacomini, 2006) proposent une extension floue de l’algèbre d’Allen. Ils associent un degré de préférence à chaque relation temporelle. Certaines approches proposent des relations temporelles entre des intervalles de temps imprécis tels que (Nagypál & Motik, 2003), (Sadeghi & Goertzel, 2014) et (Gammoudi, et al., 2017). Cependant, ces extensions sont basées sur des théories liées aux données imparfaites et ne peuvent pas être supportées dans le cadre de certaines ontologies. Par exemple, les approches proposées dans (Noy, et al., 2006) et (Gammoudi, et al., 2017) sont basées sur la théorie des ensembles flous. De plus, la plupart de ces extensions ne conservent pas toutes les propriétés de l’algèbre d’Allen originale. Par exemple, dans (Noy, et al., 2006), la relation « Equals » n’est pas réflexive. Par ailleurs, les compositions des différentes relations ne sont pas étudiées par les auteurs. (Gammoudi, et al., 2017) généralisent les relations d’Allen pour les rendre applicables à des intervalles de temps imprécis de manière conjonctive et disjonctive. Cependant, les auteurs ne proposent pas le tableau de composition des relations proposées.

2.4.2.3 Résumé et critiques

Le tableau 6 présente une comparaison des approches de représentation temporelle en ontologie. Ces approches sont comparées par rapport à leur conformité avec les standards existants du Web sémantique, le langage supporté (RDF / S, OWL) et la compatibilité avec le support de raisonnement.

Le tableau 7 résume les travaux de recherches selon les approches de représentation temporelle sur lesquelles elles se basent, les langages de raisonnement utilisés, les informations temporelles représentées (qualitatives et / ou quantitatives) et l'utilisation ou l'extension les relations d'Allen. Le constat est le suivant :

- La plupart des recherches étendent l'approche 4D-fluents.
- Presque toutes les recherches utilisent les relations d'Allen (sans les étendre) dans le cadre parfait (précis) même quand les théories des ensembles flous, de probabilités et de possibilités sont utilisées, sauf pour (Ghorbel et al, 2018) qui étendent l'algèbre Allen et l'approche de Vilain et Kautz dans un contexte imprécis.
- La plupart des travaux ne représentent que des données temporelles quantitatives et ne traitent pas les données qualitatives.
- La plupart des recherches qui prennent en charge les capacités de raisonnement utilisent le langage de règles SWRL. Cependant, dans (Ghorbel et al, 2018) l'inférence est faite en utilisant un ensemble de règles SWRL et Mamdani IF-THEN.
- La plupart des recherches mentionnées ne traitent que des données temporelles parfaites, précisément les données précises comme dans (Herradi, 2018). (Ghorbel et al., 2018) représentent des données temporelles imprécises quantitatives et qualitatives en OWL 2.

Plusieurs approches de représentation et de raisonnement sur des données temporelles dans des domaines spécifiques (autres que le web sémantique) tels que les bases de données, les images de télédétection et la récupération d'informations ont été proposées (Pons, et al., 2014) et (Messaoudi, et al., 2019). La plupart des ces approches ne traitent que les données temporelles parfaites et négligent celles qui sont imparfaites. Elles ne sont pas destinées à gérer les points de temps et les relations temporelles qualitatives entre un intervalle de temps et un ou deux points de temps. Notre approche devrait s'appuyer sur les constructions OWL existantes. Ainsi, nous écartons les approches « Temporal Description logics », « Concrete Domains

» et « Temporal RDF ». Nous écartons également l'approche « Named Graphs », car elle ne prend pas en charge OWL et il n'est pas une solution conforme au W3C.

Pour nos travaux, nous choisissons d'étendre l'approche des 4D-Fluents pour représenter des données temporelles quantitatives imprécises et les relations temporelles qualitatives associées (les dates et les horloges de temps) et pour représenter des données temporelles quantitatives incertaines et les relations temporelles qualitatives associées.

Par rapport à la réification, aux relations N-aires et au versioning, l'approche 4D-fluents minimise la redondance des données, car les changements se produisent uniquement sur les parties temporelles et gardent la partie statique inchangée. La Figure 2 montre une application de l'approche classique du 4D-Fluents sur l'expression suivante : « *Myriam a vécu en Espagne de 2001 au 2010* ».

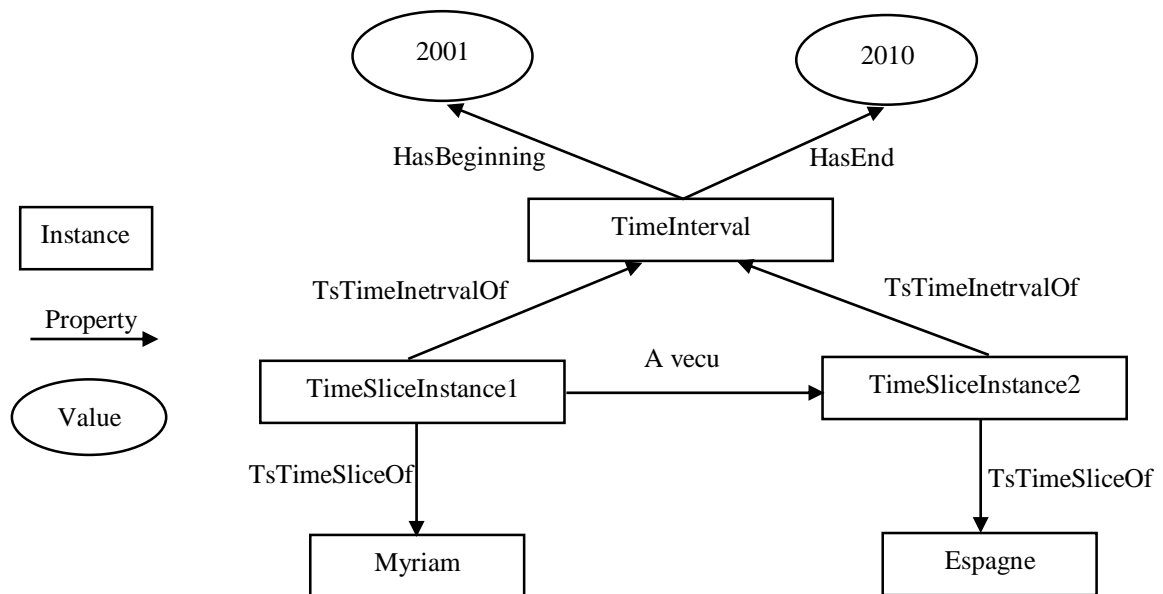


Figure 2. Exemple de l'approche classique du 4D-Fluents

Tableau 6. Modélisation de la dimension temporelle dans l'ontologie

	OWL/ RDF		Raisonnement		Interrogation		W3C		Avantages et Inconvénients
	Etendre OWL/ RDF syntax	Implémenté sans étendre OWL/RDF	Supporte le raisonnement	Ne supporte pas le raisonnement	Supporte l'interrogation	Ne supporte pas l'interrogation	W3C compliance	Not W3C compliance	
Temporal description logics	*		*		*		*		(-)Nécessite l'extension de OWL ou RDF (S), ce qui est une tâche fastidieuse. (+)Ne souffre pas de la redondance des données. (+)Conserve la décidabilité.
Concrete Domains	*			*		*		*	(-)Nécessite l'extension de OWL ou RDF (S), ce qui est une tâche fastidieuse.
Temporal RDF	*			*	*			*	(+)Pas de classes / objets supplémentaires Propriétés. (+)Utilisation de rdfs: label pour stocker l'heure valide. (-)Nécessite l'extension de RDF (S), ce qui est une tâche fastidieuse.

									(-)N'a pas toute l'expressivité de OWL.
Reification		*	*		*		*		(-)Souffre de la redondance des données. (-)Offre des capacités limitées de raisonnement et d'inférence OWL.
Versioning		*		*			*		(-)Souffre de la redondance des données. (-)La recherche nécessite des recherches exhaustives dans toutes les versions créées.
N-ary Relations		*	*		*		*		(+)La relation n-aire devient un propriété d'objet (binaire), d'où la propriété d'objet OWL les axiomes sont applicables. (-)Souffre de la redondance des données (un nouveau RDF Nœud noir par relation N-aire).
4D-Fluents		*	*				*		(+)Les changements se produisent seulement sur les parties temporelles ; (+) Elle permet le raisonnement . (+) Elle maintient l'expressivité d'OWL.
Named graphs			*		*	*			(+) Pas de classes / propriétés d'objet supplémentaires. (-) Solution non conforme au W3C. (-) Ne proposez pas de support de raisonnement

Tableau 7. Résumé des recherches qui représentent et raisonnent sur des données temporelles dans l'ontologie

Recherche	Description	Approche	Raisonnement	Allen		Représentation des données			
				Utiliser	Etendre	Qualitative	Quantitative	Parfaites	Imparfaites
TL-OWL [Kim et al. (2008)]	Combine les constructions OWL avec les relations d'Allen pour implémenter des classes temporalisées	Concrete Domains		*				*	
[Tappolet and Bernstein, 2009]	Introduire un langage de requête temporelle basé sur SPARQL	Temporal RDF						*	
stRDF [Koubarakis and Kyzirakos, 2010]	Modèle de données de contraintes qui étend RDF avec la capacité de représenter des données spatiales et temporelles	Temporal RDF						*	
SOWL [Batsakis and Petrakis, 2010]	Ontologie pour représenter et raisonner sur des informations spatio-temporelles dans OWL	4D-Fluents	SWRL	*				*	
TOWL [Frasinkar et al., 2010]	Etend la syntaxe OWL avec des concepts temporels et prend en charge les intervalles de temps quantitatifs.	Concrete domain + 4D-Fluents		*			*	*	
[Connor and Das, 2010]	Approche de représentation et d'interrogation des données temporelles dans OWL	Reification	SWRL					*	

MUSING [Krieger, 2010]	Utilise à la fois une approche basée sur 4D fluents et une approche alternative basée sur l'extension de RDF avec annotation temporelle en conjonction avec OWL-Time.	4D-Fluents						*	
PROTON [Papadakis et al., 2011]	Ontologie combinée à un raisonnement basé sur le calcul de situation implémenté dans Prolog	4D-Fluents		*				*	
[O'Connor and Das, 2011]	Représenter des données quantitativement temporelles	4D-Fluents	SWRL				*	*	
CHRONOS [Anagnostopoulos et al. (2013)]	Moteur de raisonnement pour les informations temporelles qualitatives	N-ary relations + 4D-Fluents		*		*		*	
[Harbelot et al., 2013] and [Herradi et al., 2017]	Une approche basée sur de l'extension des 4D-fluents	4D-Fluents	SWRL	*				*	
[Zekri et al., 2017]	Une approche générique pour la gestion des versions de schéma dans OWL 2	Versionning						*	
[Ghorbel et al, (2018)]	Une approche pour représenter et raisonner sur des données temporelles quantitatives et qualitatives imprécises en ontologie	4D-Fluents	SWRL + Mamdani IF-THEN rules		*	*	*	*	*
[Herradi, 2018]	Introduire une représentation qualitative des informations temporelles précise basée sur un ensemble de règles SWRL	4D-Fluents	SWRL	*		*		*	

2.5 Conclusion

L'objectif du présent chapitre est l'introduction des principaux concepts liés à notre travail de recherche, à savoir la notion de l'ontologie. Nous avons mis l'accent sur l'imperfection des données et les typologies de l'imperfection des données. Nous avons étudié des typologies génériques et d'autres spécifiques aux domaines. Nous avons ensuite passé en revue les différentes représentations des données temporelles en Web sémantique, qui sont la base de cette thèse, et nous avons étudié et comparé les travaux existants qui se basent sur ces représentations.

Dans le chapitre suivant, nous proposons notre typologie de l'imperfection des données temporelles. Nous présentons un échantillon d'un malade d'Alzheimer à partir duquel nous dégageons tous les types d'imperfections pouvant affecter les données temporelles.

CHAPITRE 3 : Typologie de l'imperfection des données temporelles

3.1. Introduction

Dans cette section, nous introduisons **notre typologie des imperfections de données temporelles**. Elle est basée sur des typologies existantes et rassemble des exemples réels. Nous divisons notre typologie en **imperfections directes**, **imperfections indirectes** et **les facteurs** qui peuvent intervenir dans la spécification du type d'imperfection. Les imperfections directes sont celles qui peuvent être déduites directement des données fournies : incertitude, erreur de saisie, imprécision, conflit, oubli et inutilité. Les imperfections indirectes sont celles que l'on peut déduire des imperfections directes. Nous distinguons trois types d'imperfections indirectes : *(i)* L'incohérence générée à partir de l'incertitude et de l'erreur de frappe, *(ii)* L'incomplétude générée à partir de l'imprécision et l'oubli, et *(iii)* La redondance générée à partir de l'inutilité. La granularité est le contexte général des données temporelles qui permet de déterminer le type d'imperfection.

3.2. Échantillon de données saisies dans le cadre de CAPTAIN MEMO

Dans cette section, nous présentons un échantillon réel illustré par la Figure 3, d'informations contenant des données temporelles saisies par un malade d'Alzheimer et son entourage dans le cadre de CAPTAIN MEMO vu qu'elle présente le domaine d'application de cette thèse. Puis, nous dégageons les éventuels types d'imperfection liés à ces données.

3.2.1 Présentation d'un exemple illustratif

Dans cet exemple établi à partir d'échantillons issus des différents tests d'utilisation de notre prototype, nous supposons qu'une malade d'Alzheimer qui s'appelle « *Mariam* », de premier stade et ayant 72 ans, réponds à 09 questions relatives à sa vie personnelle avec seulement un attribut : donnée temporelle. La Figure 3 montre cet échantillon de données.

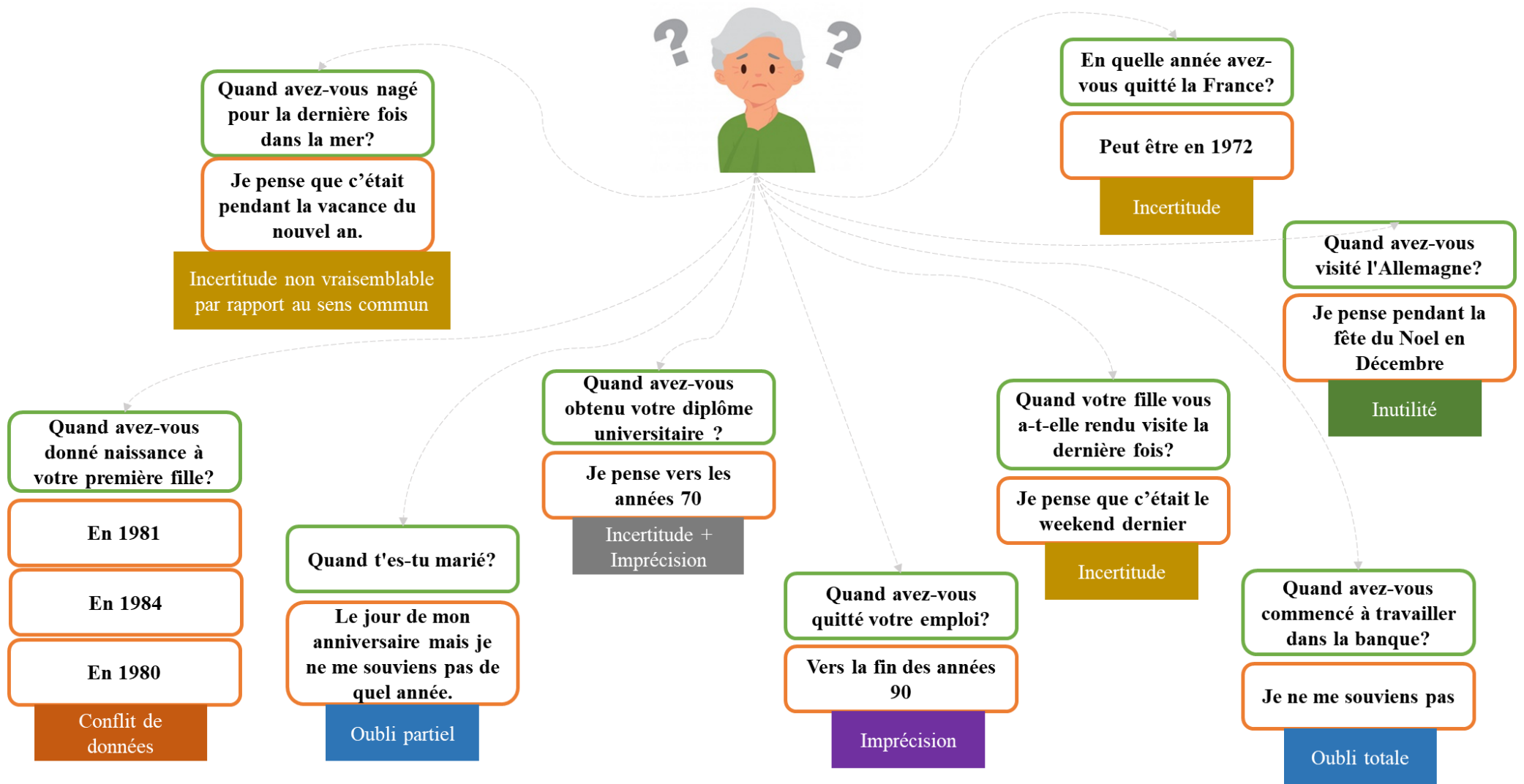


Figure 3. Échantillon fictif des réponses entrées par un malade d'Alzheimer

3.2.2 Analyse : dégager les différents types d'imperfection

En analysant les données saisies par « *Mariam* », nous avons distingué plusieurs types d'imperfection. Nous les détaillons dans ce qui suit.

Des données imprécises. Nous remarquons l'existence des données temporelles imprécises. Par exemple, nous avons demandé à la patiente de nous citer l'année pendant laquelle elle avait quitté son emploi. Elle a répondu d'une façon imprécise (« *vers la fin des années 90* »). Par la suite, l'année de sa retraite est, avec une probabilité croissante, en 1995, ou 1996, ou 1997, ou 1998, ou 1999.

Des données incertaines fiables par rapport au sens commun. Ce sont des données que *Mariam* doute de leur véracité. Par exemple, la patiente doute de l'année pendant laquelle elle a quitté le pays (« *Peut-être en 1972* »).

Des données non vraisemblables par rapport au sens commun. Nous notons l'existence de données fausses qui ne sont pas vraisemblables par rapport au sens commun.

Des données incomplètes dues à un oubli partiel. Nous remarquons l'existence des données partiellement incomplètes. La patiente se souvient partiellement de la date de son mariage, elle se rappelle uniquement du jour de son anniversaire, mais elle oublie l'année.

Des données totalement oubliées. Nous remarquons dans certains cas que la patiente oublie totalement les données. Par exemple, elle a oublié quand elle a commencé à travailler dans la banque.

Des données contradictoires. Nous remarquons un conflit des données. Nous avons posé la même question pendant 3 jours successifs et elle a répondu de 3 manières différentes. Nous avons demandé à la patiente l'année de naissance de sa fille et on a eu 3 dates différentes.

Des données incertaines et imprécises à la fois. Certaines données exprimées par la patiente englobent deux types d'imperfections, qui sont l'incertitude et l'imprécision à la fois. Par exemple, « *Je pense vers les années 70* ». L'incertitude est exprimée par « *Je pense* » et l'imprécision est exprimée par « *vers les années 70* ».

Des données inutiles. Nous remarquons que certaines données sont inutiles. Par exemple, « *Pendant Noël en Décembre* ». C'est inutile de dire « *en Décembre* », car il est évident que Noël se fête en Décembre.

3.3 Notre typologie de l'imperfection des données temporelles

3.3.1 Les imperfections directes

Dans notre typologie, illustrée par la Figure 4., nous distinguons sept types d'imperfections directes : l'incertitude, l'erreur de saisies, le conflit de données, l'imprécision, l'oubli, la cir-conlocution et l'inutilité qui seront détaillées une à une et illustrées par des exemples.

L'incertitude. Selon (Smets, 1999), chaque affirmation est vraie ou fausse. Cependant, la connaissance de la déclaration, si elle est incertaine, ne permet pas de décider si elle est vraie ou fausse. L'incertitude est une connaissance partielle de la valeur réelle des données. C'est le manque d'information. Dans notre typologie, l'incertitude comprend deux catégories. La première catégorie inclut des données incertaines qui respectent les règles du bon sens. Cela peut être dû à un doute. Par exemple, « *je ne suis pas sûr que ce soit l'été dernier ou celui d'avant quand ma tante est venue de Paris pour passer les vacances avec nous* ». Dans cet exemple, la personne doute des données temporelles, mais son affirmation respecte les règles du sens commun. La deuxième catégorie comprend des données incertaines non fiables en rapport avec les règles générales du sens commun. Par exemple, « *Nous avons nagé dans la mer pendant les vacances du Nouvel An* ». Ces données ne respectent pas les règles du bon sens puisque les vacances du Nouvel An sont en hiver et non en été.

Erreur de saisie. Dans certains cas, comme l'utilisation d'un outil, les données saisies par l'utilisateur peuvent être sujettes à des erreurs. Par exemple, si nous avons les informations suivantes : « *L'employé de la municipalité a mal saisi ma date de naissance* », la date de naissance insérée est une erreur de frappe.

Le conflit de données. Le conflit caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires et donc incompatibles. Il peut figurer dans de nombreux contextes, comme dans la maladie d'Alzheimer. En effet, un patient Alzheimer peut répondre différemment plusieurs fois à une même question. Par exemple, lorsqu'il lui demande « *Quand avez-vous obtenu votre diplôme universitaire ?* », il peut donner trois réponses temporelles différentes comme dans : 1998, 1995 et 1996.

L'imprécision. La plupart des données temporelles sont imprécises. L'imprécision mène à l'ambiguïté des informations. Les données temporelles imprécises ont deux catégories. Le premier inclut ceux qui sont traduisibles. Par exemple, « *Nous avons déménagé vers les années 90* ». « *Vers les années 90* » est une donnée temporelle numérique qui peut être traduite en un

intervalle de temps comportant deux bornes [1990 .. 1999]. Un autre exemple, si nous avons l'information suivante : « *Pendant les vacances d'été, ils se sont rendus en Espagne* », « *Pendant les vacances d'été* » représente des données temporelles basées sur un langage naturel pouvant être traduites en un intervalle de temps [01/06 .. 31/08] ayant deux limites. La deuxième catégorie inclut ceux qui ne sont pas traduisibles. Plusieurs exemples peuvent être introduits, tels que « *J'ai changé d'emploi après la naissance de ma fille* » où « *Après la naissance de ma fille* » est imprécis, car la naissance de la fille est inconnue. Un autre exemple « *Il était malade avant sa mort* » où « *Avant sa mort* » est également imprécis, car nous ne savons pas combien de temps avant sa mort il était malade. Un autre type de données temporelles basées sur le langage naturel et non traduisibles sont les adverbes du temps, tel que « *De nos jours* », « *Actuellement* », « *Maintenant* » et « *plus tard* ».

L'oubli. Les données temporelles peuvent être manquantes à cause d'un oubli. L'oubli peut être partiel ou total. Par exemple, si nous avons l'information « *C'était en mars, mais j'ai oublié l'année* ». Dans cet exemple, l'oubli est partiel. La personne se souvient du mois, mais elle a oublié l'année. Un autre exemple « *J'ai oublié la dernière fois que j'ai rendu visite à ma tante qui vit à Paris* ». Dans cet exemple, l'oubli est total. La personne a oublié totalement les données temporelles.

Circonlocution. Les données temporelles peuvent faire l'objet d'une circonlocution. Par exemple, « *Le premier jour de la semaine, nous aurons une réunion de travail* ». La personne a utilisé la déclaration « *le premier jour de la semaine* », mais il veut dire « *lundi* ». Un autre exemple, « *Le mariage a lieu deux jours après le week-end* ». La personne remplace « *mardi* » par la déclaration « *deux jours après le week-end* ».

L'inutilité. Cette imperfection signifie que les informations fournies sont inutiles et qu'elles n'apportent aucune valeur ajoutée.

3.3.2 Les imperfections indirectes

Nous distinguons trois types d'imperfections indirectes : incohérence, incomplétude et redondance. Ces imperfections sont déduites des types d'imperfections directes.

L'incohérence. Pour définir l'incohérence, nous nous sommes basés sur (Gavignet, et al., 2016). Selon elle, l'incohérence fait référence à « l'existence de données contradictoires sur le même objet ». Ainsi, l'incertitude, l'erreur de saisie et le conflit pouvant exister dans une donnée temporelle conduisent à l'incohérence.

L'incomplétude. Des données temporelles manquantes et imprécises peuvent conduire à une incomplétude. Par exemple, si nous avons l'information « *Mon oncle viendra le week-end prochain* »; « *Week-end prochain* » est imprécis et, donc incomplet. Ce pourrait être « samedi » ou « dimanche ».

La redondance. Elle peut être conclue lorsque l'inutilité interfère dans les données temporelles données. Par exemple, si nous avons l'information « *Ça fait longtemps que mes grands-parents sont morts. Ça fait plusieurs années* », nous pouvons déduire deux données temporelles, « *Il y a longtemps* » et « *il y a plusieurs années* », qui donnent le même sens, ce qui est inutile et redondant.

3.3.3 Dépendance du contexte général des données temporelles

La granularité représente l'échelle ou le niveau de détail dans un ensemble de données. Le contexte d'une donnée temporelle détermine le type d'imperfection. Par exemple, si nous avons « *Nous pouvons nous rencontrer le week-end prochain* », « *Week-end prochain* » représente une donnée temporelle imprécise. En effet, ce pourrait être « *samedi* » ou « *dimanche* » ou les deux. Cependant, si nous avons « *Nous passerons le week-end prochain à Berlin* », il ressort clairement du contexte que le week-end prochain correspond à tout le week-end (c'est-à-dire samedi et dimanche). En outre, plusieurs facteurs peuvent interférer dans la détermination du type d'imperfection, tels que le profil de la personne et le multiculturalisme.

Le profil de la personne. De la personne elle-même, nous pouvons décider le type d'imperfection. Par exemple, les données temporelles fournies par un patient atteint d'Alzheimer sont probablement incertaines. Par exemple, « *je ne suis pas sûr s'il faisait jour ou nuit* ». Un autre exemple concerne les enfants qui se perdent souvent dans les mots « *hier* », « *demain* » et « *aujourd'hui* ».

Multiculturalisme. L'aspect multiculturalisme reflète les différences culturelles où les idées sont centrées sur les croyances des différentes sociétés. A titre indicatif, en prenant l'exemple précédent du « *weekend* », « *le weekend* » en Europe est « *samedi et dimanche* ». Il diffère de celui du monde arabe qui est « *vendredi et samedi* ».

3.4. Conclusion

Dans ce chapitre, nous avons présenté notre propre typologie de l'imperfection des données temporelles. Cette typologie inclut onze imperfections ; huit directes et trois indirectes. Les imperfections indirectes sont déduites des imperfections directes qui sont elles déduites à partir des données saisies. La typologie inclut également des facteurs qui peuvent être utilisés pour déterminer le type d'imperfection comme le profil de la personne (des patients Alzheimer), et le multiculturalisme.

Dans le prochain chapitre, nous présentons notre approche pour représenter et raisonner sur les données temporelles imprécises de granularités fines, précisément les dates et les horloges de temps imprécises, avec une vue précise.

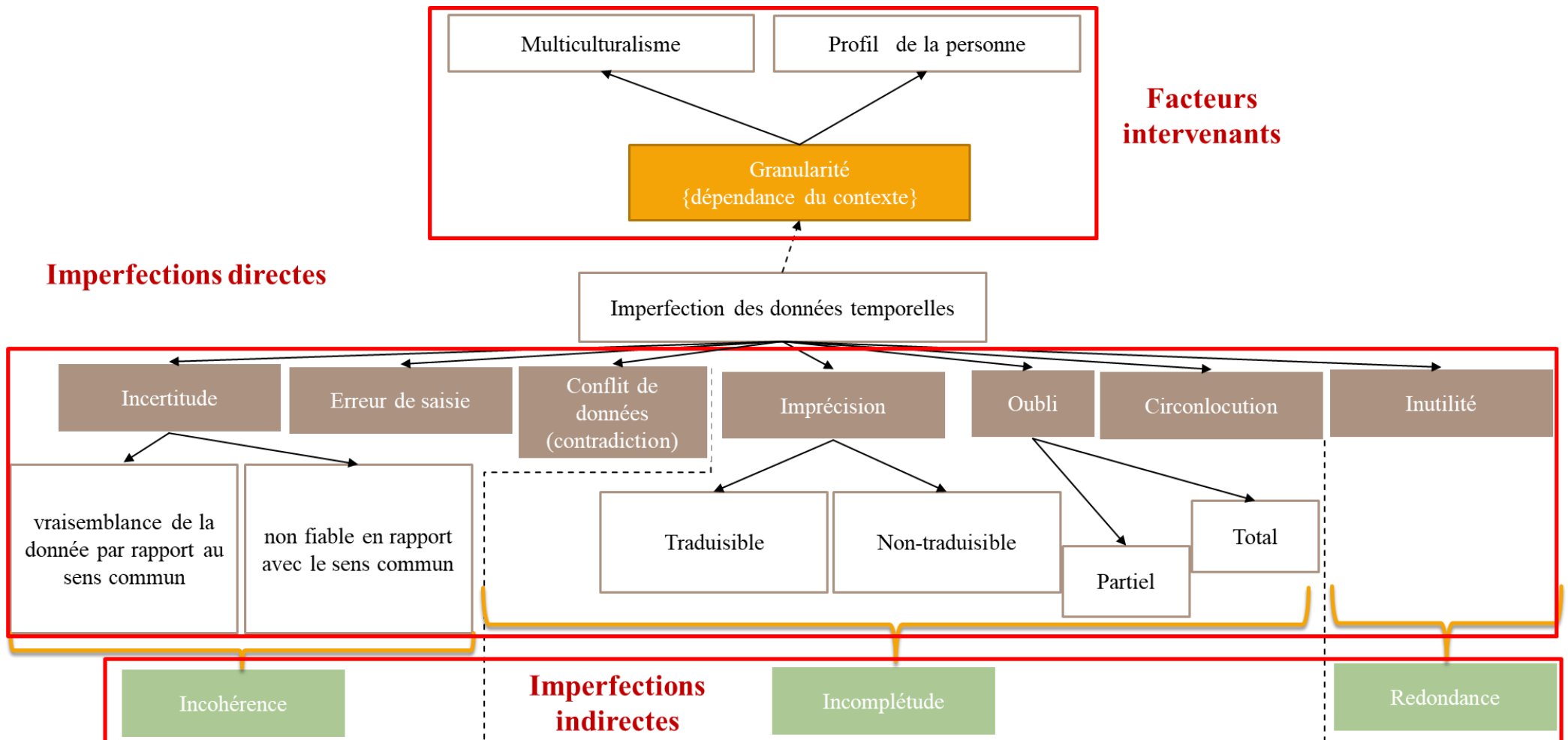


Figure 4. Notre typologie des de l'imperfection des données temporelles

CHAPITRE 4 : Traitement des données temporelles précises et imprécises dans OWL 2 : dates et horloges de temps

4.1. Introduction

Comme nous avons mentionné dans l'état de l'art, **les données temporelles fournies par les personnes sont souvent imprécises**. Par exemple, s'ils donnent l'information « **Les vacances commencent vers la fin du mois de mai** », une mesure imprécise est introduite. En effet, « **vers la fin du mois de mai** » pourrait être le 28, 29, 30 ou 31 Mai.

Dans le domaine du Web sémantique, de nombreuses approches ont été proposées pour représenter et raisonner sur des données temporelles précises. Cependant, la plupart d'entre elles **ne traitent que des intervalles de temps et les relations qualitatives associées, c'est-à-dire qu'elles ne sont pas destinées à traiter des points de temps et relations qualitatives entre un intervalle de temps et un point de temps ou deux points de temps**. Par ailleurs, à **notre connaissance, aucune approche ne traite les données temporelles imprécises de granularité fine** dans l'ontologie.

Dans ce chapitre, nous proposons **une approche pour représenter et raisonner sur des concepts évoluant dans le temps dans l'ontologie, avec une vue précise**. Les données temporelles **quantitatives**, c'est-à-dire **les intervalles de temps et les points de temps** (qui peuvent être des dates ou des horloges de temps), ainsi que les **données qualitatives** (par exemple, « avant ») sont prises en compte. Notre approche prend en compte non seulement des points de temps et des intervalles précis, mais également **des points de temps et des intervalles de temps imprécis**, tel que « *Le voyage commence le 5 juin et se termine vers la fin juillet* ». Nous n'adoptons que les standards existants du Web sémantique. Notre approche repose sur trois phases :

- 1) **Représenter les dates et les horloges du temps précises et imprécises dans une ontologie en étendant l'approche de 4D-fluents avec de nouveaux composants ontologiques précis.**
- 2) **Raisonner sur des données temporelles précises et imprécises en étendant l'algèbre d'intervalle d'Allen.** Ce dernier propose 13 relations temporelles entre des intervalles de temps précis. Cependant, cette algèbre n'est pas conçue pour gérer des intervalles de temps imprécis. De plus, il n'est pas prévu de relier un intervalle de temps et un point de temps, ni même deux points de temps. **Nous étendons cette algèbre pour proposer des relations temporelles qualitatives entre des intervalles de temps imprécis.** Les propriétés de réflexivité / irréflexivité, de symétrie / asymétrie et de transitivité sont

préservées. Nous adaptons les relations d'intervalle résultantes **pour proposer des relations temporelles entre un intervalle de temps et un point de temps, et deux points de temps pouvant être précis ou imprécis**. Toutes les relations temporelles que nous proposons peuvent être utilisées pour le raisonnement temporel au moyen **de tables de transitivité**.

- 3) Proposer **une ontologie** OWL 2 nommée « *CrispTimeOnto* » basée sur nos extensions. Elle infère les relations temporelles qualitatives proposées via un ensemble de **règles SWRL**.

La suite du chapitre est structurée comme suit. Dans la section 4.2, nous présentons notre représentation des données temporelles quantitatives et qualitatives, précises et imprécises dans OWL 2 en étendant l'approche 4D-fluents. Dans la section 4.3 nous présentons notre extension de l'algèbre d'Allen pour raisonner sur les données temporelles. Dans la section 4.4, nous proposons notre ontologie « *CrispTimeOnto* » basée sur nos extensions. Dans la section 4.5, nous présentons l'expérimentation de notre approche. Enfin, dans la dernière section, nous terminons par une conclusion.

4.2. Représentation des données temporelles précises et imprécises en ontologie

Nous étendons l'approche de 4D-fluents pour représenter, dans une ontologie, des données temporelles quantitatives précises et imprécises, ainsi que des relations temporelles qualitatives associées.

4.2.1 Données temporelles quantitatives

Nous étendons l'approche de 4D-fluents avec de nouvelles classes et relations pour représenter : (i) des points de temps précis et imprécis et (ii) des intervalles de temps imprécis. Certaines de ces classes et relations sont importées à partir de l'ontologie OWL-Time.

4.2.1.1 Points de temps précis et imprécis

Nous introduisons une classe « *TimePoint* » et une propriété d'objet « *TsTimePoint* ». Ce dernier relie une instance de « *TimeSlice* » et une instance de « *TimePoint* ». Pour exprimer les dates et les horloges de temps, nous utilisons la classe nommée « *time:DateTimeDescription* » définie dans l'ontologie OWL-Time (« *time* » étant le préfixe de l'ontologie OWL-Time).

Nous représentons des points de temps précis : dates et horloges. Pour les dates D, Mo et Y représentent respectivement le jour, le mois et l'année. Nous utilisons trois propriétés (*Datatype Property*) de OWL-Time nommées « *time:day* », « *time:month* » et « *time:year* » pour associer, respectivement, la classe « *time:DateTimeDescription* » à « D », « Mo » et « Y ». Par exemple, si nous avons « *Le voyage commence le 05 juin 2019* », le « *05 juin 2019* » est représenté comme une date précise. « *time:day* » a comme co-domaine (ou *range*) « 05 », « *time:month* » a comme co-domaine « Juin » et « *time:year* » a comme co-domaine « 2019 ». De même, pour les horloges de temps, prenons « S », « Mi » et « H », respectivement, secondes précises, minutes et heures. Nous utilisons trois propriétés (*Datatype Property*) de OWL-Time nommées « *time:second* », « *time:minute* » et « *time:hour* », pour connecter respectivement la classe « *time:DateTimeDescription* » avec « S », « Mi » et « H ». Par exemple, si nous avons « *Le petit-déjeuner à l'hôtel commence à 7h30* », « *07h30* » est représenté par une horloge de temps précise. « *time:second* » a comme co-domaine « 00 », « *time:minute* » a comme co-domaine « 30 » et « *time:hour* » a comme co-domaine « 07 ».

Nous présentons des points de temps imprécis (dates et horloges). Pour les dates, soit D, Mo et Y, respectivement, jour, mois et année imprécis. Nous les représentons par des ensembles ascendants disjonctifs $\{D^{(1)} \dots D^{(d)}\}$, $\{Mo^{(1)} \dots Mo^{(mo)}\}$ et $\{Y^{(1)} \dots Y^{(y)}\}$. À titre d'exemple, « *le prix des billets de train était beaucoup moins cher pendant les années 70* », « *pendant les années 70* » est représenté par l'ensemble ascendant disjonctif $\{1970 \dots 1979\}$. Nous définissons respectivement pour D, Mo et Y les propriétés de type de données (*Datatype Property*) suivantes : « *HasDayFrom* » et « *HasDayTo* », « *HasMonthFrom* » et « *HasMonthTo* », « *HasYearFrom* » et « *HasYearTo* ». Elles sont toutes connectées à la classe « *time:DateTimeDescription* ». Par exemple, « *Le voyage commence début juin 2019* », « *Début Juin 2019* » est représenté comme une date imprécise, car la partie « jour » est imprécise. « *HasDayFrom* » a comme co-domaine « 01st » et « *HasDayTo* » a comme co-domaine « 07th ». De même, pour les horloges, prenons « S », « Mi » et « H », respectivement, secondes, minutes et heures imprécises, représentées par des ensembles ascendants disjonctifs $\{S^{(1)} \dots S^{(s)}\}$, $\{Mi^{(1)} \dots Mi^{(mi)}\}$ et $\{H^{(1)} \dots H^{(h)}\}$. Par exemple, « *Nous devrions terminer le petit-déjeuner au plus à 10 heures* », où « *au plus à 10 heures* » est représenté par le jeu ascendant disjonctif $\{07 \dots 10\}$. Nous définissons pour « S », « Mi » et « H », respectivement, deux propriétés de type de données : « *HasSecondsFrom* » et « *HasSecondsTo* », « *HasMinutesFrom* » et « *HasMinutesTo* », « *HasHoursFrom* » et « *HasHoursTo* ». Ils sont tous connectés à la classe « *time:DateTimeDescription* ». Par exemple, si nous avons « *Nous quittons l'hôtel après le déjeuner avant 17h. L'heure du*

déjeuner est comprise entre 12h et 14h », « *Après le déjeuner avant 17h* » est représentée par une horloge imprécise. « *HasHoursFrom* » et « *HasHoursTo* » ont les ranges « 14h » et « 17h ».

4.2.1.2 Intervalles de temps imprécis

Un intervalle de temps imprécis a des limites de début et de fin. Nous les représentons à l'aide des instances de « *TimePoint* ». Dans notre extension, « *time:hasBeginning* » et « *time:hasEnd* » sont des propriétés d'objet définissant les limites de début et de fin de l'intervalle. Ils associent une instance de « *TimeInterval* » (domaine) et une instance de « *TimePoint* » (co-domaine). Par exemple : « *Nous visiterons le parc national à 14h30 et nous partirons le soir. Le parc national ferme à 22h00* ». Nous représentons l'heure de la fermeture du parc « 22h00 » comme une horloge précise. Nous représentons la durée de la visite comme un intervalle de temps imprécis. La borne supérieure est représentée par une horloge précise. La propriété « *time:hasBeginning* » a comme co-domaine « 14:30 ». La borne inférieure est représentée par une horloge de temps imprécise. Cela pourrait être entre 19h00 et 22h00. La propriété « *time:hasEnd* » a comme co-domaine « *jusqu'au soir* ». Les propriétés « *hasHoursFrom* » et « *hasHoursTo* » ont respectivement comme co-domaines « 19 » et « 22 ».

4.2.2 Données temporelles qualitatives

Quatre relations temporelles peuvent exister entre des points de temps et des intervalles de temps : relations point-point, point-intervalle, point-intervalle et intervalle-intervalle. Par conséquent, nous définissons quatre propriétés d'objet.

- 1) La propriété « *RelationPoints* » relie deux instances de la classe « *TimePoint* » pour représenter des relations point à point.
- 2) La propriété « *RelationIntervalPoint* » relie une instance de la classe « *TimeInterval* » (domaine) et une instance de la classe « *TimePoint* » (range) afin de représenter la relation Interval-Point.
- 3) La propriété « *RelationPointInterval* » relie une instance de la classe « *TimePoint* » (domaine) et une instance de la classe « *TimeInterval* » (range) pour représenter la relation point-intervalle.
- 4) « *RelationIntervals* » relie deux instances de la classe « *TimeInterval* » pour représenter les relations intervalle-intervalle.

La Figure 5 représente notre extension de l'approche 4D-Fluents.

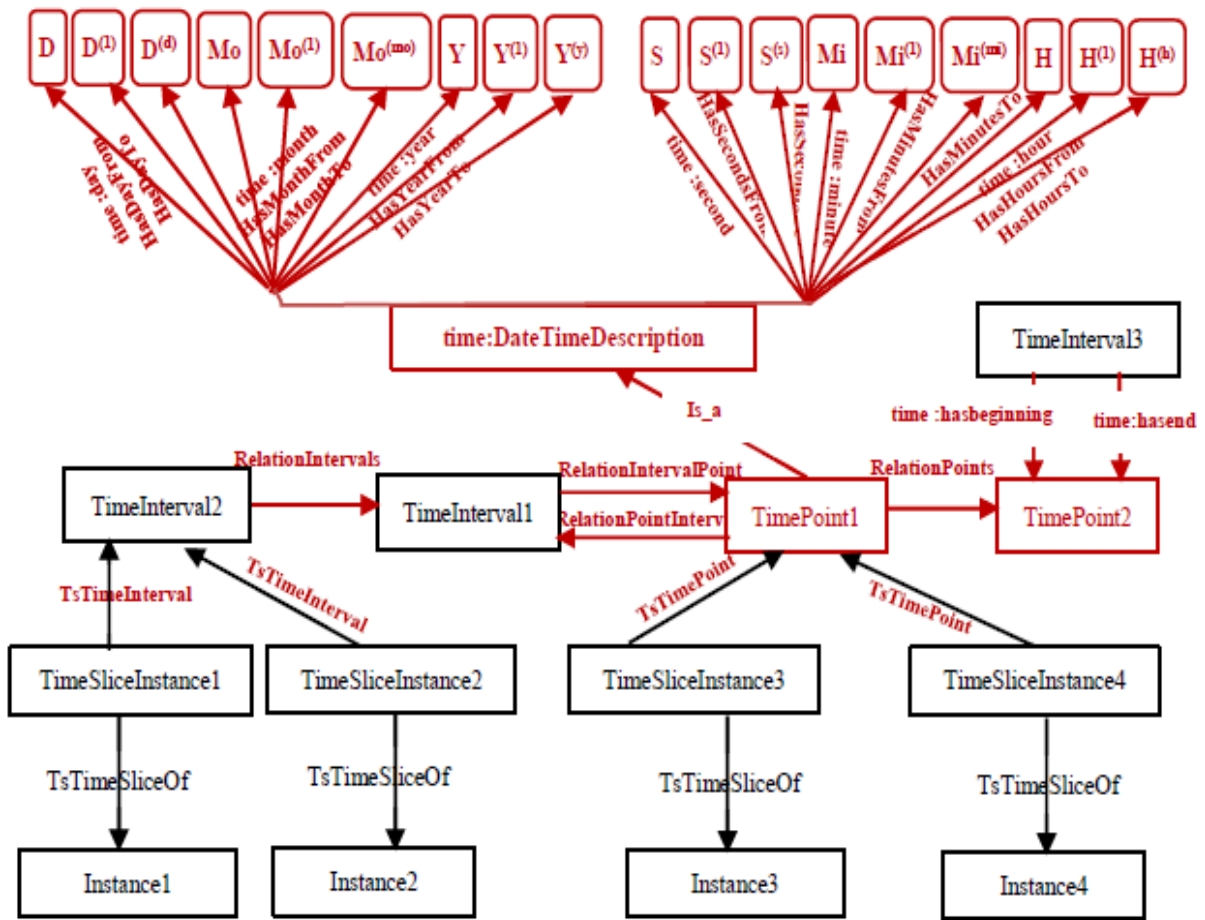


Figure 5. Notre extension de l’approche 4D-Fluents pour représenter des dates et horloges de temps précis et imprécis

4.3. Raisonnement sur les données temporelles précises et imprécises

Notre approche étend l’algèbre d’intervalle d’Allen pour : (i) raisonner sur des données temporelles quantitatives précises et imprécises pour déduire des relations temporelles qualitatives et (ii) raisonner sur des relations temporelles qualitatives pour en inférer de nouvelles.

4.3.1 Relations temporelles qualitatives

Nous proposons des relations temporelles qualitatives entre des intervalles de temps imprécis. Ces relations sont, par la suite, utilisées pour relier un intervalle de temps et un point de temps ou deux points de temps.

4.3.1.1 Relations entre Intervalles de temps

Les relations temporelles proposées sont basées sur une relation d'ordre entre les points de temps contenus dans les intervalles. Elles peuvent être exprimées à l'aide de comparateurs de temps, tels que ceux proposés dans Vilain et Algèbre de Kautz. Lorsque nous considérons des intervalles de temps précis, notre approche se réduit à l'algèbre d'intervalle d'Allen. Nous redéfinissons les 13 relations d'Allen pour proposer des relations temporelles entre des intervalles de temps imprécis.

Par exemple, nous redéfinissons la relation « *Before* » entre les deux intervalles de temps imprécis $A = [A^-, A^+]$ et $B = [B^-, B^+]$ comme :

$$\forall A^{+(a)} \in A^+, \forall B^{-(b)} \in B^- / \text{Precedes}(A^{+(a)}, B^{-(b)})$$

Cela signifie que le point de temps le plus récent de A^+ ($A^{+(E)}$) doit précéder le point de temps le plus ancien de B^- ($B^{-(1)}$) :

$$\text{Precedes}(A^{+(E)}, B^{-(1)})$$

De même, nous redéfinissons les autres relations d'Allen, comme le montre le Tableau 8.

Tableau 8. Relations temporelles entre les intervalles de temps imprécis A et B

Relation (A,B)	Définition	Inverse (B, A)
Before(A, B)	$\text{Precedes}(A^{+(E)}, B^{-(1)})$	After(B, A)
Meets(A, B)	$\text{Min}(\text{Same}(A^{+(1)}, B^{-(1)}) \wedge \text{Same}(A^{+(E)}, B^{-(B)}))$	Met-by(B, A)
Overlaps(A,B)	$\text{Min}(\text{Precedes}(A^{-(B)}, B^{-(1)}) \wedge \text{Precedes}(B^{-(B)}, A^{+(1)}) \wedge \text{Precedes}(A^{+(E)}, B^{+(1)}))$	Overlapped-by(B, A)
Starts(A, B)	$\text{Min}(\text{Same}(A^{-(1)}, B^{-(1)}) \wedge \text{Same}(A^{-(B)}, B^{-(B)}) \wedge \text{Precedes}(A^{+(E)}, B^{+(1)}))$	Started-by(B, A)
During(A, B)	$\text{Min}(\text{Precedes}(B^{-(B)}, A^{-(1)}) \wedge \text{Precedes}(A^{+(E)}, B^{+(1)}))$	Contains(B, A)
Ends(A, B)	$\text{Min}(\text{Precedes}(B^{-(B)}, A^{-(1)}) \wedge \text{Same}(A^{+(1)}, B^{+(1)}) \wedge \text{Same}(A^{+(E)}, B^{+(E)}))$	Ended-by(B, A)

Equals(A, B)	$\text{Min} (\text{Same} (A^{-{(1)}, B^{-{(1)}}}) \wedge \text{Same} (A^{-{(B)}, B^{-{(B)}}}) \wedge \text{Same} (A^{+{(1)}, B^{+{(1)}}}) \wedge \text{Same} (A^{+{(E)}, B^{+{(E)}}}))$	Equals(B, A)
--------------	---	--------------

Par exemple, soit $A = [A^-, A^+]$ et $B = [B^-, B^+]$ des intervalles de temps imprécis ; où $A^- = \{01/01/2001 \dots \text{mi-avril } 2001\}$; $A^+ = \{\text{fin avril } 2001 \dots 30/06/2001\}$; $B^- = \{20/11/1997 \dots \text{au milieu de l'année prochaine}\}$ et $B^+ = \{10/05/1998 \dots \text{au début de } 1999\}$ Nous concluons que : $\text{Before}(B, A) = 1$.

Les relations temporelles proposées entre des intervalles de temps imprécis préservent de nombreuses propriétés de l'algèbre d'Allen. Nous obtenons des généralisations des propriétés de réflexivité / irréflectivité, de symétrie / asymétrie et de transitivité.

4.3.1.2 Relations entre un intervalle de temps et un point de temps

Nous adaptons les relations temporelles qualitatives entre les intervalles de temps afin de proposer des relations entre un intervalle de temps et un point de temps, comme indiqué dans le Tableau 9. Les relations entre un intervalle de temps précis et un point de temps précis sont basées sur la définition des relations temporelles standard d'Allen. Les relations entre un intervalle de temps imprécis et un point de temps imprécis sont basées sur notre définition de relation temporelle entre des intervalles de temps imprécis.

Tableau 9. Relations temporelles entre un intervalle de temps A et un point P

Relation(P, A)	Définition	Inverse(A, P)
Relations temporelles entre un intervalle de temps précis A et un instant précis P		
Before(P, A)	$\text{Precedes}(P, A^-)$	After(A, P)
After(P, A)	$\text{Precedes}(A^+, P)$	Before(A, P)
Starts(P, A)	$\text{Same}(P, A^-)$	Started-by(A, P)
During(P, A)	$\text{Precedes}(A^-, P) \wedge \text{Precedes}(P, A^+)$	Contains(A, P)
Ends(P, A)	$\text{Same}(P, A^+)$	Ended-by(A, P)
Relations temporelles entre un intervalle de temps A imprécis et un instant P imprécis		

Before(P, A)	$\text{Precedes}(P^{(P)}, A^{-(1)})$	After(A, P)
After(P, A)	$\text{Precedes}(A^{+(E)}, P^{(1)})$	Before(A, P)
Starts(P, A)	$\text{Same}(P^{(1)}, A^{-(1)}) \wedge \text{Same}(P^{(P)}, A^{-(B)})$	Started-by(A, P)
During(P, A)	$\text{Precedes}(A^{-(E)}, P^{(1)}) \wedge \text{Precedes}(P^{(P)}, A^{+(1)})$	Contains(A, P)
Ends(P, A)	$\text{Same}(P^{(1)}, A^{+(1)}) \wedge \text{Same}(P^{(P)}, A^{+(E)})$	Ended-by(A, P)

Par exemple, soit P_1 un point de temps imprécis représenté à l'aide de l'ensemble ascendant disjonctif {Tôt le matin ... heure du déjeuner} et $A_1 = [A_1^-, A_1^+]$ l'intervalle de temps imprécis, où $A_1^- = \{\text{Après l'heure du déjeuner... l'après-midi}\}$; $A_1^+ = \{\text{Soir... à minuit}\}$. Nous concluons que : $\text{Before}(P_1, A_1) = 1$.

4.3.1.3 Relations entre les points de temps

Nous adaptons les relations temporelles qualitatives entre les intervalles de temps pour proposer des relations entre les points de temps, comme indiqué dans le Tableau 10. Par exemple, supposons que N et R soient deux points de temps imprécis ; représentés respectivement en utilisant les ensembles ascendants disjonctifs {*Juste avant le déjeuner ... Une heure avant le déjeuner*} et {*Juste après le déjeuner ... Juste avant le dîner*}. Nous concluons que : $\text{Après}(R, N) = 1$.

Tableau 10. Relations temporelles entre les points temporels N et R

Relation(N, R)	Définition	Inverse(R, N)
Relations temporelles entre des instants précis N et R		
Before(N, R)	$\text{Precedes}(N, R)$	After(R, N)

Equals(N, R)	Same(N, R)	Equals(R, N)
Relations temporelles entre les points de temps imprécis N et R		
Before(N, R)	Precedes(N ^(P) , R ⁽¹⁾)	After(R, N)
Equals(N, R)	Same(N ⁽¹⁾ , R ⁽¹⁾) \wedge Same(N ^(P) , R ^(L))	Equals(R, N)

Par exemple, soit N et R deux points de temps imprécis ; où N est représenté en utilisant l'ensemble ascendant disjonctif { *Juste avant les vacances ... Un jour avant le voyage* } et R est représenté en utilisant l'ensemble ascendant disjonctif { *Premier jour du voyage ... Un jour avant la fin du voyage* }. Nous concluons que : $After(R, N) = 1$.

4.3.2 Table de transitivité

La table de transitivité d'Allen nous permet d'obtenir de $R_1(A, B)$ et $R_2(B, C)$ que $R_3(A, C)$ existe, où $A = [A^-, A^+]$, $B = [B^-, B^+]$ et $C = [C^-, C^+]$ sont des intervalles de temps précis et R_1, R_2 et R_3 sont des relations d'Allen. En se basant sur le Tableau 6, on peut déduire de "During(A,B)" et "Meet(B,C)" que "Before(A,C)" existe. "During(A,B)" signifie " $Precedes(B^-, A^{-(1)}) \wedge Precedes(A^{+(E)}, B^{+(1)})$ ", et "Meet(B,C)" signifie " $Same(B^{+(1)}, C^{-(1)}) \wedge Same(B^{+(E)}, C^{-(B)})$ ". Considérant des relations précises, notre table de transitivité coïncide avec celle d'Allen.

Nous introduisons trois tables de transitivité pour raisonner sur les relations temporelles qualitatives entre un intervalle de temps et un point de temps et les relations temporelles qualitatives entre deux points de temps.

Relations temporelles qualitatives entre un intervalle de temps et un point de temps. Nous introduisons deux tables de transitivité pour raisonner sur les relations temporelles qualitatives entre un intervalle de temps et un point de temps. Soit I et J des intervalles de temps et P et L des points de temps. Ils peuvent être tous précis ou tous imprécis. Le Tableau 11 montre la table de transitivité qui permet de raisonner $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$. R_1 est une

relation point-intervalle ; R_2 est une relation intervalle-point et R_3 est une relation point-point. Certaines relations de composition sont « indécidables ». Par exemple, les relations « After(P, I) » et « Before(I, L) » donnent trois relations possibles: « Before(P, L) », « Equals(P, L) » et « After(P, L) ».

Tableau 11. Table de transitivité pour raisonner $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$

$R_1(P, I)$ $R_2(I, L)$	Before(P, I)	After(P, I)	Starts(P, I)	During(P, I)	Ends(P, I)
Before(I, L)	Before(P, L)	L'une des relations de point dans le temps	Before(P, L)	Before(P, L)	Before(P, L)
After(I, L)	L'une des relations de point dans le temps	After(P, L)	After(P, L)	After(P, L)	After(P, L)
Started-by(I, L)	Before(P, L)	After(P, L)	Equals(P, L)	After(P, L)	After(P, L)
Contains(I, L)	Before(P, L)	After(P, L)	Before(P, L)	L'une des relations de point dans le temps	After(P, L)
Ended-by(I, L)	Before(P, L)	After(P, L)	Before(P, L)	Before(P, L)	Equals(P, L)

Le Tableau 12 permet de raisonner $R_3(I, J)$ à partir de $R_1(I, P)$ et $R_2(P, J)$. R_1 est une relation intervalle-point; R_2 est une relation point-intervalle et R_3 est une relation intervalle-intervalle.

Tableau 12. Table de transitivité pour raisonner sur $R_3(I, J)$ à partir de $R_1(I, P)$ et $R_2(P, J)$

$R_1(I, P)$ $R_2(P, J)$	Before(I, P)	After(I, P)	Started-by(I, P)	Contains(I, P)	Ended-by(I, P)
Before(P, J)	Before(I, J)		Contains(I, J) ∨ Ended-by(I, J) ∨	Contains(I, J) ∨ Ended-by(I, J) ∨	Before(I, J)

		L'une des relations d'intervalle de temps	Met-by(I, J) ∨ Before(I, J) ∨ Overlaps(I, J)	Overlapped-by(I, J) ∨ Meets(I, J) ∨ Before(I, J)	
After(P, J)	L'une des relations d'intervalle de temps	After(I, J)	After(I, J)	After(I, J) ∨ Met-by(I, J) ∨ During(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J)	During(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J) ∨ Met-by(I, J) ∨ After(I, J)
Starts(P, J)	Before(I, J)	After(I, J) ∨ Met-by(I, J) ∨ Overlapped-by(I, J) ∨ Ended-by(I, J) ∨ Contains(I, J)	Started-by(I, J) ∨ Equals(I, J) ∨ Starts(I, J)	Contains(I, J) ∨ Ended-by(I, J) ∨ Overlaps(I, J)	Meets(I, J)
During(P, J)	During(I, J) ∨ Starts(I, J) ∨ Overlaps(I, J) ∨ Meets(I, J) ∨ Before(I, J)	After(I, J) ∨ Met- by(I, J) ∨ Over- lapped-by(I, J) ∨ Ended-by(I, J) ∨ Contains(I, J)	Contains(I, J) ∨ Ends(I, J) ∨ Overlapped- by(I, J)	Overlaps(I, J) ∨ Overlapped-by(I, J) ∨ Starts(I, J) ∨ Started-by(I, J) ∨ During(I, J) ∨ Con- tains(I, J) ∨ Ends(I, J) ∨ Ended-by(I, J) ∨ Equals(I, J)	Overlapped- by(I, J) ∨ Starts(I, J) ∨ During(I, J)
Ends(P, J)	During(I, J) ∨ Starts(I, J) ∨ Overlaps(I, J) ∨ Meets(I, J) ∨ Before(I, J)	After(I, J)	Met-by(I, J)	Contains(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J)	Equals(I, J) ∨ Ended-by (I, J) ∨ During(I, J)

Relations temporelles qualitatives entre les points de temps. Nous introduisons une table de transitivité pour raisonner sur les relations temporelles qualitatives qui sont établies entre les points temporels. Le Tableau 13 permet de déduire $R_3(P, M)$ à partir de $R_1(P, L)$ et $R_2(L, M)$, où P, L et M sont des points temporels précis ou imprécis. R_1 , R_2 et R_3 sont des relations point-point.

Tableau 13. Table de transitivité pour raisonner $R_3(P, M)$ à partir de $R_1(P, L)$ et $R_2(L, M)$

$R_2(L, M) \backslash R_1(P, L)$	Before(P, L)	Equals(P, L)	After(P, L)
Before(L, M)	Before(P, M)	Before(P, M)	L'une des relations de point dans le temps
Equals(L, M)	Before(P, M)	Equals(P, M)	After(P, M)
After(L, M)	L'une des relations de point dans le temps	After(P, M)	After(P, M)

4.4. « CrispTimeOnto » : notre ontologie des données temporelles précises et imprécises

Sur la base de nos extensions de l'approche 4D-fluents et de l'algèbre d'intervalle d'Allen, nous mettons en œuvre notre ontologie temporelle OWL 2. Nous utilisons l'éditeur d'ontologie PROTÉGÉ¹.

Nousinstancions les propriétés d'objets {« RelationIntervals », « RelationIntervalPoint », « RelationPointInterval » et « Relation-Points »} en fonction de l'extension d'Allen. Par exemple, « RelationIntervals » peut être l'une des relations d'Allen pour relier deux instances de la classe « TimeInterval ». Cela signifie que 13 propriétés d'objet sont associées : {« BeforeIntervals », « MeetsIntervals », « OverlapsIntervals », « StartsIntervals », « PendantIntervals », « EndsIntervals », « AfterIntervals », « Met-byIntervals », « Overlapped-byIntervals », « Started-byIntervals », « ContainsIntervals », « Ended-byIntervals » et « EqualsIntervals »}.

Notre ontologie inclut un ensemble de règles SWRL pour déduire des relations temporelles qualitatives manquantes. Pour chaque relation temporelle, nous associons une règle SWRL pour la déduire des données temporelles quantitatives données par l'utilisateur. Sur la base des tables de transitivité, nous associons une règle SWRL pour chaque relation de transitivité.

Soit P, L et S des points temporels imprécis. Par exemple, nous associons deux règles SWRL pour déduire les propriétés d'objet « BeforePoints » comme suit :

$$\text{TimePoint}(?p) \wedge \text{TimePoint}(?l) \wedge \text{TimePoint}(?s) \wedge \text{BeforePoints}(?p, ?l) \wedge \text{EqualsPoints}(?l, ?s) \rightarrow \text{BeforePoints}(?p, ?s)$$

¹ <https://protege.stanford.edu/>

4.5. Expérimentations

Pour valider notre approche, nous implémentons un prototype basé sur l'approche proposée. Ce prototype est basé sur l'ontologie proposée.

4.5.1 Prototype implémenté

Notre prototype basé sur notre ontologie « *CrispTimeOnto* », offre des interfaces utilisateur pour explorer notre approche. Il est implémenté à base du langage JAVA. Il utilise JENA API² pour manipuler l'ontologie précise.

L'architecture de ce prototype comprend trois composants principaux. Tout d'abord, l'utilisateur instancie notre ontologie via l'interface utilisateur illustrée à la Figure 6. Cette interface permet à l'utilisateur de saisir des données temporelles.

Après chaque nouvelle entrée de données temporelles par un utilisateur, le composant « Inférence de données temporelles qualitatives » est automatiquement exécuté pour déduire les données manquantes. Ce composant est basé sur l'ensemble proposé de règles SWRL. Le troisième composant permet aux utilisateurs d'interroger notre ontologie via des requêtes SPARQL.

The screenshot displays a web-based user interface for managing temporal data, divided into four main sections:

- Add Time Interval:** Contains a text input for 'Time Interval's Name'. Below it are two columns for 'Imprecise Beginning Bound' and 'Imprecise Ending Bound', each with 'From' and 'To' date pickers. At the bottom are 'Precise Beginning Bound' and 'Precise Ending Bound' with 'From' and 'To' date pickers, and a 'Save' button.
- Add Time Point:** Contains a text input for 'Time Point's Name' (pre-filled with 'Holidays'). Below it are 'Imprecise Time Point' and 'Precise Time Point' sections, each with 'From' and 'To' date pickers. The 'Precise Time Point' section has pre-filled values: 'Precise Beginning Bound' (empty), 'Precise Ending Bound' (14 | 09 | 2019). A 'Save' button is at the bottom.
- Add Qualitative Temporal Relation:** Features two dropdown menus for 'Time Interval/Time' (both set to 'I1') and a dropdown for 'Qualitative Temporal Relation' (set to 'Before'). A 'Save' button is at the bottom.
- Saved Temporal Data:** Shows a 'Search Criterion' dropdown (set to 'All Time Intervals') and a large empty list area with scrollbars.

Figure 6. Prototype proposé basé sur notre ontologie

² <https://jena.apache.org/>

4.5.2 Application dans CAPTAIN MEMO

Nous appliquons notre approche à l'ontologie PersonLink qui fait partie de la prothèse mémoire CAPTAIN MEMO. « PersonLink » est une ontologie multiculturelle et multilingue pour le stockage, la modélisation et le raisonnement sur les relations familiales. Les relations interpersonnelles évoluent avec le temps. Cependant, des entrées de données temporelles imprécises sont particulièrement nombreuses lorsqu'elles sont introduites par un patient Alzheimer. Nous avons intégré notre approche de représentation des données temporelles dans l'ontologie PersonLink pour gérer des intervalles de temps imprécis. Le but étant de l'étendre pour représenter et raisonner sur des données temporelles précises et imprécises (nous fusionnons CrispTimeOnto et PersonLink).

Par exemple : « *John était marié à Béatrice depuis environ 10 ans. John était marié à Maria juste après avoir obtenu son doctorat et cela a duré 15 ans. John a obtenu son doctorat en Octobre 1980* ». Soit $A = [A^-, A^+]$ et $B = [B^-, B^+]$ deux intervalles de temps imprécis représentant, respectivement, les durées du mariage de John avec Béatrice et Maria. Supposons que $A^- = \{2007 \dots 2009\}$, $A^+ = 2018$, $B^- = \{\text{Octobre } 1980 \dots 1983\}$ et $B^+ = \{1995 \dots 1998\}$. Les relations temporelles qualitatives associées « BeforIntervals » et « AfterIntervals » sont inférées de A et B sur la base du composant « Inférence de données temporelles qualitatives ».

4.6. Conclusion

Dans ce chapitre, nous avons introduit notre approche pour représenter et raisonner sur des données temporelles en termes de relations quantitatives et aussi qualitatives. Elle permet la manipulation des points de temps précis et imprécis et des intervalles de temps imprécis. Elle est basée uniquement sur les normes existantes du Web sémantique.

Notre approche comporte trois contributions. La première consiste à étendre l'approche 4D-fluents avec des composants précis pour représenter des points de temps précis et imprécis, des dates et des horloges précises, et des intervalles de temps imprécis. La deuxième contribution consiste à étendre l'algèbre des intervalles d'Allen pour raisonner sur des données temporelles précises et imprécises. Cette extension garde les propriétés d'Allen qui sont réflexivité/ir-réflexivité, la symétrie/asymétrie et la transitivité. Quatre tables de transitivité pour raisonner sur les relations temporelles résultantes ont été proposées. La troisième contribution concerne la création d'une ontologie basée sur la première et la deuxième contribution. Pour valider notre travail, nous avons développé un prototype, basée sur cette ontologie, qui permet aux utilisateurs d'explorer notre approche.

L'extension de l'algèbre des intervalles d'Allen que nous avons proposée dans ce travail ne se limite pas au domaine du Web sémantique. Elle peut également être appliquée à d'autres domaines de recherche tels que les bases de données. Dans ce cas, des inférences peuvent être faites à l'aide de langages d'interrogation de base de données tels que SQL. L'intérêt principal de notre approche est qu'elle peut être implémentée avec des standards classiques et que les chercheurs ne sont pas obligés d'apprendre les technologies liées la logique floue. Nous considérons que les normes de représentation des données précises sont plus matures et prennent mieux en charge la mise à l'échelle que les outils flous.

Le prochain chapitre présente notre approche basée sur les réseaux bayésiens pour représenter et raisonner sur les données temporelles quantitatives et qualitatives incertaines, avec une vue certaine.

**CHAPITRE 5 : Traitement des données
temporelles certaines et incertaines dans
OWL 2 : approche basée sur la théorie
de probabilité**

5.1. Introduction

Comme mentionné dans notre typologie proposée dans le chapitre 3, les données temporelles sont sujettes à l'incertitude. En effet, ce genre d'imperfection pourrait être fréquent surtout dans le cadre d'une application comme CAPTAIN MEMO. Par exemple, un patient d'Alzheimer peut saisir « *Je pense que c'était de 2013 à 2020* », « *Elle a quitté le pays peut-être à 11 heures* » et « *Je pense qu'elle s'est mariée avant l'obtention du diplôme universitaire* ».

Dans le domaine du Web sémantique, plusieurs approches ont été proposées pour représenter et raisonner sur des données temporelles certaines. Cependant, la plupart d'entre elles ne traitent que des intervalles de temps et les relations qualitatives associées, c'est-à-dire qu'ils ne sont pas destinés à gérer des points de temps et des relations qualitatives entre un intervalle de temps et un point de temps ou deux points de temps. Par ailleurs, à notre connaissance, il n'existe pas d'approches dédiées au traitement des données temporelles incertaines dans OWL 2.

Nous proposons une approche, basée sur OWL 2, pour **représenter et raisonner sur des données temporelles certaines et incertaines en termes de relations qualitatives et quantitatives, avec une vue certaine**. Elle comporte trois contributions :

- 1) La représentation des données temporelles « certaines » et « incertaines » dans OWL 2. **L'approche 4D-fluents ne couvre que les intervalles de temps certains dans OWL. Nous l'étendons avec de nouvelles composantes ontologiques « certaines » pour représenter des données temporelles quantitatives certaines et incertaines** (c'est-à-dire des intervalles de temps incertains et des points temporels certains et incertains), et des relations temporelles qualitatives entre les intervalles de temps et les points. **Nous utilisons les réseaux bayésiens pour calculer les mesures de certitudes.**
- 2) **Le raisonnement sur les données temporelles certaines et incertaines en étendant l'algèbre d'intervalle d'Allen.** Cette algèbre ne prend en compte que les relations temporelles entre intervalles de temps certains. Nous l'étendons pour gérer les relations temporelles entre des intervalles de temps incertains et des points temporels certains et incertains. Notre extension préserve des propriétés importantes de l'algèbre d'origine. Nous adaptons les relations entre intervalles pour proposer des relations temporelles entre un intervalle de temps et un point de

temps, et deux points de temps. Toutes ces relations sont utilisées par la suite pour le raisonnement temporel via **des tables de transitivité**.

- 3) **La proposition une ontologie OWL 2 appelée « *UncertTimeOnto* »** qui peut être intégrée dans d'autres ontologies pour traiter des données temporelles certaines et incertaines. Elle est implémentée sur la base des extensions proposées. Les inférences sont effectuées à l'aide de **règles SWRL** intégrées dans l'ontologie.

Le reste de ce chapitre est organisé comme suit. La section 5.2 présente notre extension de l'approche 4D-fluents pour représenter les données temporelles incertaines dans OWL 2. La section 5.3 présente notre extension de l'algèbre d'Allen pour raisonner sur les données temporelles incertaines. La section 5.4 présente notre ontologie « *UncertTimeOnto* » basée sur nos extensions. La section 5.5 introduit les expérimentations de notre approche dans le cadre de CAPTAIN MEMO. Dans la dernière section, nous terminons par une conclusion.

5.2. Représentation des données temporelles certaines et incertaines dans OWL 2

Nous étendons l'approche 4D-fluents, avec des composantes ontologiques « certaines », pour représenter les données temporelles quantitatives « certaines » et « incertaines » ainsi que les relations temporelles qualitatives associées dans OWL 2.

5.2.1 Données temporelles quantitatives

Nous étendons l'approche 4D-fluents pour représenter (i) des intervalles de temps incertains et (ii) des points de temps certains et incertains. La Figure 7 illustre notre extension.

5.2.1.1 Intervalles de temps incertains

Soit $A = [A^-_{a^-}, A^+_{a^+}]$ un intervalle de temps incertain. A^- et A^+ sont respectivement les bornes supérieures et inférieures ; et a^- et a^+ sont, respectivement, les degrés de certitude associés. Nous étendons l'approche 4D-fluents pour représenter des intervalles de temps incertains en introduisant deux propriétés de type de données (*Datatype Property*) certaines nommées « *HasBeginnigCertainty* » et « *HasEndCertainty* ». Ces deux propriétés sont associées à la classe prédéfinie « *TimeInterval* ». « *HasBeginnigCertainty* » et « *HasEndCertainty* » représentent, respectivement, les degrés de certitude associés à la borne supérieure a^- et à la borne inférieure a^+ . Par exemple, si nous avons les informations suivantes données par un patient atteint de la maladie d'Alzheimer : « *Je pense que Marie a enseigné à l'Université de Paris*

Sorbonne de 2001 à 2010 ». Dans ce cas « $[2001_{0.2}, 2010_{0.6}]$ » est un intervalle de temps incertain, avec « $a^- = 0.2$ » et « $a^+ = 0.6$ » sont les degrés de certitude associés aux deux bornes. Les deux degrés de certitude sont représentés dans l'ontologie en utilisant « *HasBeginnigCertainty* » et « *HasEndCertainty* ». La première propriété a comme co-domaine (*range*) « 0.2 » et la seconde a comme co-domaine « 0.6 ».

5.2.1.2 Points de temps certains et incertains

Nous étendons l'approche 4D-fluents pour représenter des points de temps certains et incertains. Ainsi, une classe certaine « *TimePoint* » est introduite.

Soit P un point de temps certain. Nous introduisons une propriété de type de données (*Datatype Property*) nommée « *HasTimePoint* » qui relie la classe « *TimePoint* » à P. Par exemple, pour l'expression « *Nicolas a quitté le pays en 2000* », « 2000 » est un point de temps certain lié à la classe « *TimePoint* » en utilisant la propriété de type de données « *HasTimePoint* ».

Soit Q_q un point de temps incertain, où « q » est le degré de certitude associé au point de temps Q. Nous utilisons la propriété de type de données « *HasTimePoint* » pour relier la classe « *TimePoint* » à Q. Pour représenter le degré de certitude associé à Q, nous proposons une propriété de type de données certaine nommée « *PointCertainty* » qui est associée à la classe « *TimePoint* ». Prenons par exemple l'information suivante donnée par une patiente atteinte de la maladie d'Alzheimer : « *Janette a donné naissance à son premier enfant peut-être en 2004* », « $2004_{0.9}$ » est un point de temps incertain avec un degré de certitude associé « $q = 0.9$ ». « 2004 » est lié à la classe « *TimePoint* » en utilisant la propriété de type de données « *HasTimePoint* » et « 0.9 » est le degré de certitude associé représenté dans l'ontologie à l'aide de la propriété de type de données « *PointCertainty* ».

5.2.2 Données temporelles qualitatives

Pour représenter des relations temporelles certaines et incertaines entre des intervalles de temps et des points de temps, nous proposons quatre propriétés d'objet. La propriété « *RelationIntervals* » relie deux instances de la classe « *TimeInterval* » pour représenter les relations Intervalle-Intervalle. La propriété « *RelationIntervalPoint* » relie une instance de la classe « *TimeInterval* » (domaine) et une instance de la classe « *TimePoint* » (*range*) pour représenter les relations Intervalle-Point. La propriété « *RelationPointInterval* » relie une instance de la classe

« TimePoint » (domaine) et une instance de la classe « TimeInterval » (*range*) pour représenter les relations point-intervalle. « RelationPoints » relie deux instances de la classe « TimePoint » pour représenter les relations Point-Point.

Pour les relations temporelles qualitatives incertaines, quatre propriétés d'objet sont également proposées pour représenter le degré de certitude associé à une relation donnée :

- « RelationIntervalsCertainty » qui relie deux instances de la classe « TimeInterval » pour représenter un degré de certitude associé à une relation intervalle-intervalle.
- « RelationIntervalPointCertainty » qui relie une instance de la classe « TimeInterval » et une instance de la classe « TimePoint » pour représenter un degré de certitude associé à une relation Interval Point.
- « RelationPointIntervalCertainty » qui relie une instance de la classe « TimePoint » et une instance de la classe « TimeInterval » pour représenter un degré de certitude associé à une relation point-intervalle.
- « RelationPointsCertainty » qui relie deux instances de la classe « TimePoint » pour représenter un degré de certitude associé à une relation point-point.

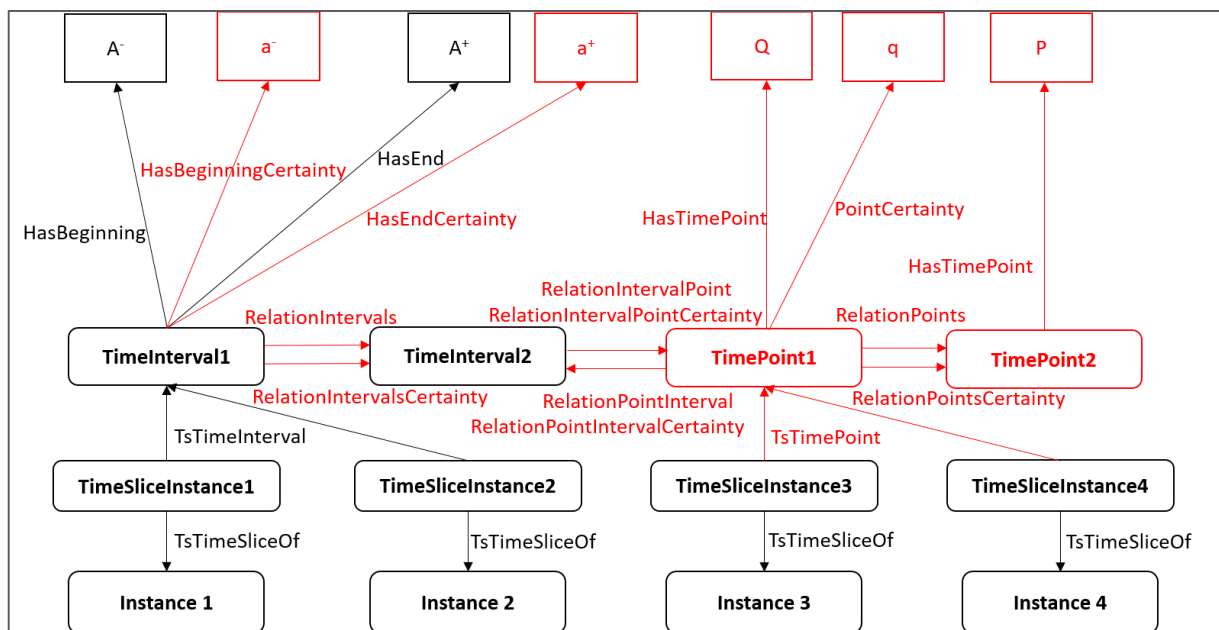


Figure 7. Notre extension de l'approche 4D-Fluents pour représenter des données temporelles certaines et incertaines

5.3. Raisonnement sur les données temporelles certaines et incertaines dans OWL 2

Notre approche consiste à étendre l’algèbre des intervalles d’Allen pour : (i) raisonner sur des données temporelles quantitatives certaines et incertaines pour inférer des relations temporelles qualitatives et (ii) raisonner sur les relations temporelles qualitatives pour en inférer de nouvelles relations.

5.3.1 Relations temporelles qualitatives

Nous proposons dans cette section des relations temporelles qualitatives entre des intervalles de temps incertains. Nous prouvons que les relations qui en résultent préservent de nombreuses propriétés algébriques d’Allen. Ensuite, nous adaptons ces relations pour relier un intervalle de temps et un point temporel ou deux points temporels.

5.3.1.1 Relations entre les intervalles de temps

En considérant les intervalles de temps certains, notre approche se réduit à l’algèbre d’intervalles d’Allen. Nous redéfinissons les 13 relations d’Allen pour proposer des relations temporelles entre des intervalles de temps incertains. Soit $A = [A^-_{a-}, A^+_{a+}]$ et $B = [B^-_{b-}, B^+_{b+}]$ deux intervalles de temps incertains. Par exemple, nous redéfinissons la relation « *Before*(A, B) » comme: « *Before_c*(A, B) »; où « c » est le degré de certitude associé à la relation *Before* entre A et B . Cela signifie que la borne inférieure incertaine de l’intervalle A est inférieure à la borne supérieure incertaine de B .

$$Before_c(A, B) \longrightarrow A^+_{a+} < B^-_{b-}$$

Le degré de certitude « c » est déduit des degrés de certitude a^+ et b^- en utilisant un réseau bayésien. Ce dernier représente, selon la littérature, une technique adaptée pour travailler avec l’incertitude typique des applications réelles (Lamma, et al., 2006), (Hlel, et al., 2018) et (Fareh, 2019). Selon (Ding & Peng, 2004), les réseaux bayésiens ont un pouvoir expressif et une capacité de raisonnement probabiliste rigoureuse et efficace. En outre, il représente un outil graphique puissant pour représenter, apprendre et calculer des distributions de probabilités (Gu, et al., 2004). Les réseaux bayésiens peuvent être analysés par des non-experts et permettent une mise en correspondance de l’ontologie avec le modèle probabiliste grâce à son modèle graphique convivial (Njah, et al., 2016). Il fournit un mode arqué nous permettant de représenter des ontologies (Hlel, et al., 2018).

Par exemple, si nous avons le degré de certitude $a^+ = 0.4$ et le degré de certitude $b^- = 0.5$ alors $c = 0.43$, comme le montre la Figure 8.

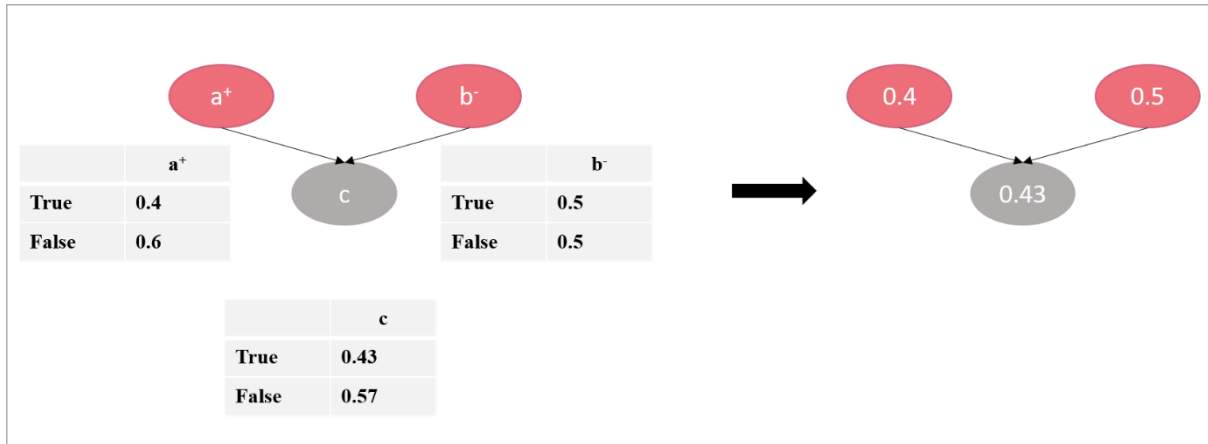


Figure 8. Réseau bayésien associé à la relation « $Before_c$ » entre deux intervalles

Nous classons les relations proposées en deux catégories. La première couvre les relations qui consistent en une formule unique sans aucune conjonction (c'est-à-dire « $Before_c$ », « $after_c$ », « $Meets_c$ » et « $Met-by_c$ »). La seconde couvre le reste des relations. Celles-ci sont définies sur la base d'un ensemble de conjonctions (par exemple, « $Overlaps_c$ »). Ainsi, comme pour la relation « $Before_c$ », nous redéfinissons la relation « $Meets_c(A, B)$ »; comme indiqué dans le Tableau 14.

Nous redéfinissons la relation « $Overlaps(A, B)$ » comme : « $Overlaps_c(A, B)$ »; où « c » est le degré de certitude associé à la relation « $Overlaps$ » entre A et B. Cela signifie que la borne supérieure incertaine de A est inférieure à la borne supérieure incertaine de B; la borne inférieure incertaine de A est supérieure à la borne supérieure incertaine de B; et la borne inférieure incertaine de l'intervalle A est inférieure à la borne inférieure incertaine de B.

$$Overlaps_c(A, B) \longrightarrow (A^-_{a^-} < B^-_{b^-}) \wedge (A^+_{a^+} > B^-_{b^-}) \wedge (A^+_{a^+} < B^+_{b^+})$$

Tableau 14. Relations temporelles entre deux intervalles de temps incertains A et B

Relation(A, B)	Définition	Inverse(B, A)
$Before_c(A, B)$	$A^+_{a^+} < B^-_{b^-}$	$After_c(B, A)$
$Meets_c(A, B)$	$A^+_{a^+} = B^-_{b^-}$	$Met-by_c(B, A)$
$Overlaps_c(A, B)$	$(A^-_{a^-} < B^-_{b^-}) \wedge (A^+_{a^+} > B^-_{b^-}) \wedge (A^+_{a^+} < B^+_{b^+})$	$Overlapped-by_c(B, A)$
$Starts_c(A, B)$	$(A^-_{a^-} = B^-_{b^-}) \wedge (A^+_{a^+} < B^+_{b^+})$	$Started-by_c(B, A)$
$During_c(A, B)$	$(A^-_{a^-} > B^-_{b^-}) \wedge (A^+_{a^+} < B^+_{b^+})$	$Contains_c(B, A)$

Ends _c (A, B)	$(A^-_{a-} > B^-_{b-}) \wedge (A^+_{a+} = B^+_{b+})$	Ended-by _c (B, A)
Equals _c (A, B)	$(A^-_{a-} = B^-_{b-}) \wedge (A^+_{a+} = B^+_{b+})$	Equals _c (B, A)

Le degré de certitude « c » est déduit des degrés de certitude « c₁ », « c₂ » et « c₃ » associés respectivement aux conjonctions (A⁻_{a-} > B⁻_{b-}), (A⁺_{a+} > B⁻_{b-}) et (A⁺_{a+} < B⁺_{b+}), en utilisant quatre réseaux bayésiens. Le premier consiste à définir le degré de certitude « c₁ ». Le second à définir le degré de certitude « c₂ ». Le troisième à définir le degré de certitude « c₃ ». Et le dernier à définir le degré de certitude « c » de la relation entre « c₁ », « c₂ » et « c₃ », comme le montre la Figure 9 .

Nous redéfinissons les autres relations d'Allen de la même manière que la relation « Overlaps_c(A, B) », comme indiqué dans le Tableau 14.

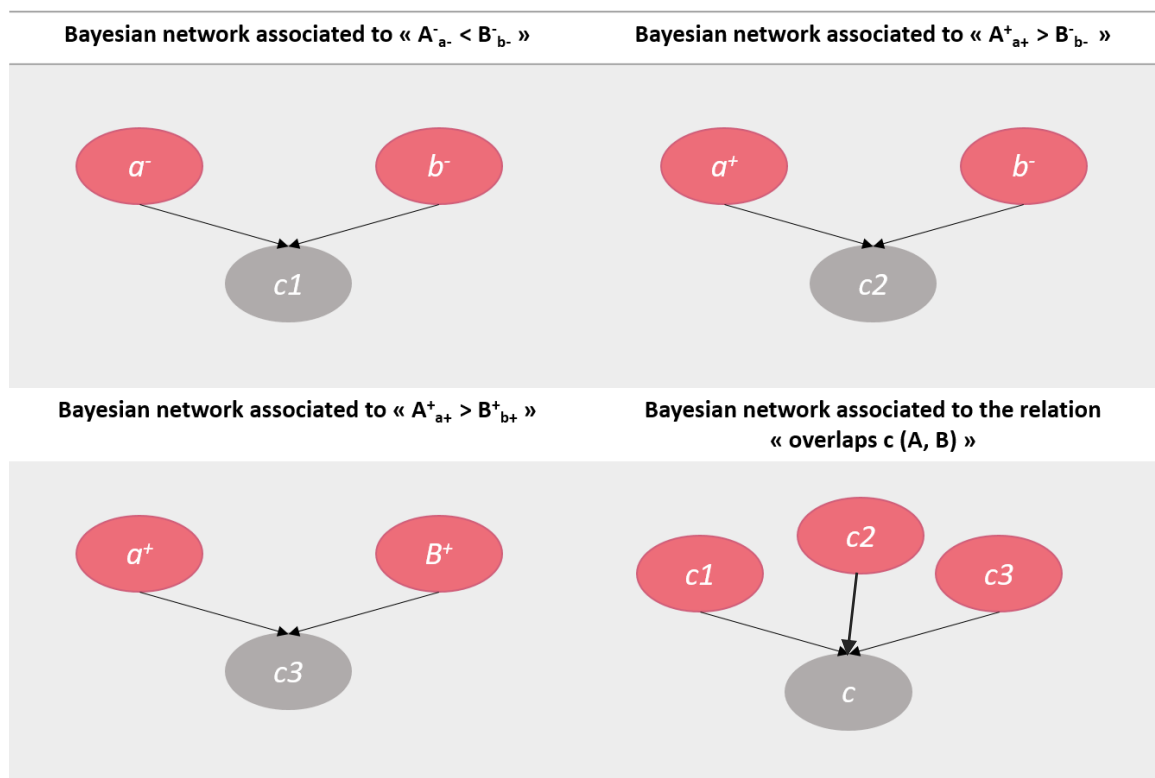


Figure 9. Réseau bayésien associé à la relation « Overlaps_c » entre deux intervalles de temps incertains.

Exemple 1: Marina, une patiente d’Alzheimer à un stade précoce, a déclaré : “*Je ne suis pas certaine si mon mari est décédé entre 2000 et 2005. Peut-être qu’il est décédé entre 1999 et 2005*”.

Soit $A = [A^-_{a-}, A^+_{a+}]$ et $B = [A^-_{b-}, B^+_{b+}]$ sont des intervalles de temps incertains; où $A = [2000, 2005]$, $B = [1999, 2005]$; et $a^- = 0,7$, $a^+ = 0,9$, $b^- = 0,2$ et $b^+ = 0,9$.

Il tient que : « Ends_c(A, B) = 1; Où $(A^-_{a-} > B^-_{b-}) \wedge (A^+_{a+} = B^+_{b+}) = 1$ »; avec un degré de certitude « c » égal à « 0.66 ». Ce dernier est déduit des degrés de certitude « c₁ » et « c₂ » de $(A^-_{a-} > B^-_{b-})$ et $(A^+_{a+} = B^+_{b+})$ en utilisant trois réseaux bayésiens, comme le montre la Figure 10.

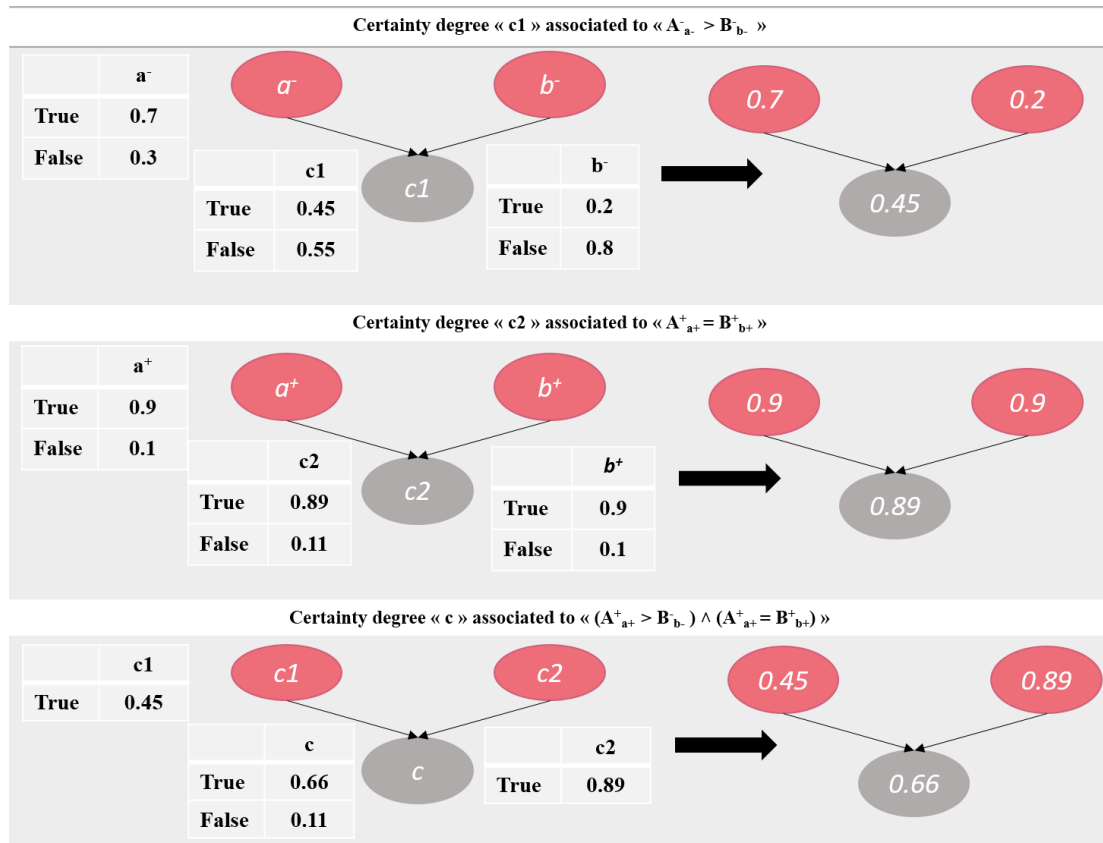


Figure 10. Réseaux bayésiens associés à l'Exemple

Les relations proposées entre des intervalles de temps incertains préservent de nombreuses propriétés de l'algèbre d'Allen.

Réflexivité / Irréflexivité : les relations { « Before_c », « After_c », « Meets_c », « Met-by_c », « Overlaps_c », « Overlapped-by_c », « Starts_c », « Staretd-by_c », « During_c », « Contains_c », « Ends_c » et « Ended-by_c » } entre des intervalles de temps incertains sont irréflexives. Soit R une de ces relations. Il tient que : $R(A, A) = 0$. Par exemple :

$$Before_c(A, A) = (A^+_{a+} < A^-_{a-}) = 0 \text{ comme } A^+_{a+} > A^-_{a-}$$

« Equals_c » est réflexif. Il tient que :

$$Equal_c(A, A) = (A^-_{a-} = A^-_{a-}) \wedge (A^+_{a+} = A^+_{a+}) = 1$$

Symétrie / Asymétrie: les relations {« Before_c », « After_c », « Meets_c », « Met-by_c », « Overlaps_c », « Overlapped-by_c », « Starts_c », « Staretd-by_c », « During_c », « Contains_c », « Ends_c » et « Ended-by_c »} entre des intervalles de temps incertains sont asymétriques. Soit R une de ces relations. Il tient que : $R(A, B) \text{ and } R(B, A) \rightarrow A = B$

Par exemple, on peut déduire de « Before_c(A, B) » et « Before_c(B, A) » que « A = B » est vrai. En effet, $(A^+_{a+} < B^-_{b-})$ » est déduit de « Before_c(A, B) », et « $(B^+_{b+} < A^-_{a-})$ » est déduit de « Before_c(B, A) ». De « $(A^+_{a+} < B^-_{b-})$ » et « $(B^+_{b+} < A^-_{a-})$ », nous concluons que « A = B ».

« Equals_c » est symétrique. Il tient que :

$$Equals_c(A, B) = Equals_c(B, A)$$

Transitivité : Les relations {« Before_c », « After_c », « Overlaps_c », « Overlapped-by_c », « Starts_c », « Staretd-by_c », « During_c », « Contains_c », « Equals_c »} entre des intervalles de temps incertains sont transitives. Soit R une de ces relations. Soit S = [S⁻_{s-}, S⁻_{s+}] un intervalle de temps incertain. Il tient que:

$$R(I, J) \text{ and } R(J, S) \Rightarrow R(I, S)$$

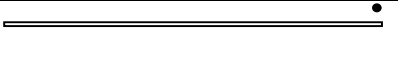
Par exemple, nous pouvons déduire de « Before_c(A, B) » et « Before_c(B, S) » que « Before_c(A, S) » est valide. En effet, « $(A^+_{a+} < B^-_{b-})$ » est déduit de « Before_c(A, B) » et « $(B^+_{b+} < S^-_{s-})$ » est déduit de « Before_c(B, S) », par conséquent « $(A^+_{a+} < S^-_{s-})$ » est valide. Cela signifie que « Before_c(A, S) » est valide.

5.3.1.2 Relations entre un intervalle de temps et un point de temps

Nous avons adapté les relations temporelles qualitatives entre les intervalles de temps pour proposer des relations entre un intervalle de temps et un point temporel, c'est-à-dire des relations intervalle-point et point-intervalle, comme indiqué dans le Tableau 15.

Tableau 15. Relations temporelles entre un intervalle de temps A et un point de temps P

Relation(P, A)	Définition	Illustration	Inverse(A, P)
Relation temporelle entre un intervalle de temps certain A = [A⁻, A⁺] et un point de temps certain P			
Before(P, A)	$P < A^-$	• <u> </u>	After(A, P)
After(P, A)	$A^+ < P$	<u> </u> •	Before(A, P)
Starts(P, A)	$P = A^-$	• <u> </u>	Started-by(A, P)
During(P, A)	$(A^- < P) \wedge (P < A^+)$	<u> </u> •	Contains(A, P)

Ends(P, A)	$P = A^+$		Ended-by(A, P)
Relation temporelle entre un intervalle de temps incertain $A = [A^-_{a-}, A^+_{a+}]$ et un point de temps incertain P_p			
Before _c (P _p , A)	$P_p < A^-_{a-}$		After _c (A, P _p)
After _c (P _p , A)	$A^+_{a+} < P_p$		Before _c (A, P _p)
Starts _c (P _p , A)	$P_p = A^-_{a-}$		Started-by _c (A, P _p)
During _c (P _p , A)	$(A^-_{a-} < P_p) \wedge (P_p < A^+_{a+})$		Contains _c (A, P _p)
Ends _c (P _p , A)	$P_p = A^+_{a+}$		Ended-by _c (A, P _p)

Exemple 2 : Soit P et Q deux points de temps certains ; où P = 1960 et Q = 1962 ; et A = [A⁻, A⁺] un intervalle de temps certain où A = [1960, 1963]. Il tient que :

$$Starts(P, A) = 1 ; During(Q, A) = 1$$

5.3.1.3 Relations entre les points de temps

Les relations temporelles qualitatives entre les intervalles de temps sont adaptées pour proposer des relations entre les points de temps, comme le montre le Tableau 16.

Tableau 16. Relations temporelles entre deux points de temps P et Q

Relation(P, Q)	Définition	Inverse(Q, P)
Relations temporelles entre les points de temps certains P et Q		
Before(P, Q)	$P < Q$	After(Q, P)
Equals(P, Q)	$P = Q$	Equals(Q, P)
Relations temporelles entre les points de temps incertains P _p and Q _q		
Before _c (P, Q)	$P_p < Q_q$	After _c (Q, P)
Equals _c (P, Q)	$P_p = Q_q$	Equals _c (Q, P)

Exemple 3 : Soit P, Q et L trois points de temps certains ; où P = 1960, Q = 1970 et L = 1940. On considère que :

$$Before(P, Q) = 1 ; After(P, L) = 1$$

5.3.2 Tables de transitivité

Nous proposons des tables de transitivité pour dériver de nouvelles données temporelles à partir des relations temporelles qualitatives entre les intervalles de temps et les points de temps.

Relations temporelles qualitatives entre les intervalles de temps. La table de transitivité d'Allen nous permet de déduire de $R_1(A, B)$ et $R_2(B, C)$ que $R_3(A, C)$ tient ; où A, B et C sont des intervalles de temps certains et R_1, R_2 et R_3 sont les relations d'Allen. Nous généralisons ces déductions en utilisant les trois intervalles de temps incertains $A = [A^-, A^+]$, $B = [B^-, B^+]$ et $C = [C^-, C^+]$.

Sur la base du tableau 14, « $\text{Before}_c(A, K)$ » peut être déduit de « $\text{During}_c(A, B)$ » et « $\text{Meet}_c(B, K)$ ». En effet, « $(B^-_{b-} < A^-_{a-}) \wedge (A^+_{a+} < B^+_{b+})$ » est déduit de « $\text{During}_c(A, B)$ » et « $(B^+_{b+} = K^-_{k-})$ » est déduit de « $\text{Meet}_c(B, K)$ ». Avec « $(A^+_{a+} < B^+_{b+})$ » et « $(B^+_{b+} = K^-_{k-})$ » nous concluons que « $\text{Before}_c(A, K)$ » est vrai. La table de transitivité proposée coïncide avec celle d'Allen.

Relations temporelles qualitatives entre un intervalle de temps et un point dans le temps.

Deux tables de transitivité sont introduites pour raisonner sur les relations temporelles entre un intervalle de temps et un point de temps. Soit A et B des intervalles de temps et P et Q des points de temps. Ils peuvent être tous certains ou incertains. Le Tableau 17 montre la table de transitivité qui permet de déduire $R_3(P, Q)$ de $R_1(P, A)$ et $R_2(A, Q)$. R_1 est une relation point-intervalle, R_2 est une relation intervalle-point et R_3 est une relation point-point. Parmi ces relations de composition, certaines peuvent donner lieu à plusieurs décisions (donc indécidable).

Par exemple, les relations « $\text{After}_c(P, A)$ » et « $\text{Before}_c(A, Q)$ » donnent trois relations possibles : « $\text{Before}_c(P, Q)$ », « $\text{Equals}_c(P, Q)$ » et « $\text{After}_c(P, Q)$ ».

Tableau 17. Table de transitivité qui permet de déduire $R_3(P, Q)$ de $R_1(P, A)$ et $R_2(A, Q)$

$R_1(P, A)$ / $R_2(A, Q)$	$\text{Before}_c(P, A)$	$\text{After}_c(P, A)$	$\text{Starts}_c(P, A)$	$\text{During}_c(P, A)$	$\text{Ends}_c(P, A)$
$\text{Before}_c(A, Q)$	$\text{Before}_c(P, Q)$	L'une des relations des points de temps	$\text{Before}_c(P, Q)$	$\text{Before}_c(P, Q)$	$\text{Before}_c(P, Q)$

After_c(A, Q)	L'une des relations des points de temps	After _c (P, Q)	After _c (P, Q)	After _c (P, Q)	After _c (P, Q)
Started-by_c(A, Q)	Before _c (P, Q)	After _c (P, Q)	Equals _c (P, Q)	After _c (P, Q)	After _c (P, Q)
Contains_c(A, Q)	Before _c (P, Q)	After _c (P, Q)	Before _c (P, Q)	L'une des relations des points de temps	After _c (P, Q)
Ended-by_c(A, Q)	Before _c (P, Q)	After _c (P, Q)	Before _c (P, Q)	Before _c (P, Q)	Equals _c (P, Q)

Le Tableau 18 liste les différentes relations $R_3(A, B)$ inférées à partir de $R_1(A, P)$ et $R_2(P, B)$ où R_1 est une relation intervalle-point, R_2 est une relation point-intervalle et R_3 est une relation intervalle-intervalle.

Tableau 18. Table de transitivité qui permet de déduire $R_3(A, B)$ de $R_1(A, P)$ et $R_2(P, B)$

R₁(A, P) \ R₂(P, B)	Before_c(A, P)	After_c(A, P)	Started-by_c(A, P)	Contains_c(A, P)	Ended-by_c(A, P)
Before_c(P, B)	Before _c (A, B)	L'une des relations d'intervalle de temps	Contains _c (A, B) ∨ Ended-by _c (A, B) ∨ Met-by _c (A, B) ∨ Before _c (A, B) ∨ Overlaps _c (A, B)	Contains _c (A, B) ∨ Ended-by _c (A, B) ∨ Overlapped-by _c (A, B) ∨ Meets _c (A, B) ∨ Before _c (A, B)	Before _c (A, B)
After_c(P, B)	L'une des relations d'intervalle de temps	After _c (A, B)	After _c (A, B)	After _c (A, B) ∨ Met-by _c (A, B) ∨ During _c (A, B) ∨ Started-by _c (A, B) ∨ Overlapped-by _c (A, B)	During _c (A, B) ∨ Started-by _c (A, B) ∨ Overlapped-by _c (A, B) ∨ Met-by _c (A, B) ∨ After _c (A, B)
Starts_c(P, B)	Before _c (A, B)	After _c (A, B) ∨ Met-by _c (A, B) ∨ Overlapped-by _c (A, B) ∨ Ended-by _c (A, B)	Started-by _c (A, B) ∨ Equals _c (A, B) ∨ Starts _c (A, B)	Contains _c (A, B) ∨ Ended-by _c (A, B) ∨ Overlaps _c (A, B)	Meets _c (A, B)

		B) \vee Contains _c (A, B)			
During_c(P, B)	During _c (A, B) \vee Starts _c (A, B) \vee Overlaps _c (A, B) \vee Meets _c (A, B) \vee Before _c (A, B)	After _c (A, B) \vee Met-by _c (A, B) \vee Overlapped-by _c (A, B) \vee Ended-by _c (A, B) \vee Contains _c (A, B)	Contains _c (A, B) \vee Ends _c (A, B) \vee Overlapped-by _c (A, B)	Overlaps _c (A, B) \vee Overlapped-by _c (A, B) \vee Starts _c (A, B) \vee Started-by _c (A, B) \vee During _c (A, B) \vee Contains _c (A, B) \vee Ends _c (A, B) \vee Equals _c (A, B) \vee Ended-by _c (A, B)	Overlapped-by _c (A, B) \vee Starts _c (A, B) \vee During _c (A, B)
Ends_c(P, B)	During _c (A, B) \vee Starts _c (A, B) \vee Overlaps _c (A, B) \vee Meets _c (A, A) \vee Before _c (A, B)	After _c (A, B)	Met-by _c (A, B)	Contains _c (A, B) \vee Started-by _c (A, B) \vee Overlapped-by _c (A, B)	Equals _c (A, B) \vee Ended-by _c (A, B) \vee During _c (A, B)

Relations temporelles qualitatives entre les points de temps. Nous introduisons une table de transitivité pour raisonner sur les relations temporelles qualitatives entre les points de temps. Elle permet d'inférer des relations $R_3(P, M)$ à partir de $R_1(P, Q)$ et $R_2(Q, M)$; où P, Q et M sont des points de temps certains ou incertains. R_1 , R_2 et R_3 sont des relations point-point.

Tableau 19. Table de transitivité qui permet de déduire $R_3(P, M)$ à partir de $R_1(P, Q)$ et $R_2(Q, M)$

$R_1(P, Q)$	Before _c (P, Q)	Equals _c (P, Q)	After _c (P, Q)
Before _c (Q, M)	Before _c (P, M)	Before _c (P, M)	L'une des relations de points de temps
Equals _c (Q, M)	Before _c (P, M)	Equals _c (P, M)	After _c (P, M)
After _c (Q, M)	L'une des relations de points de temps	After _c (P, M)	After _c (P, M)

5.4. « UncertTimeOnto » : notre ontologie des données temporelles incertaines en OWL 2

Cette section présente notre ontologie OWL 2, appelée « UncertTimeOnto »³ (Figure 11). Elle implémente notre extension de l'approche 4D-fluents et de l'algèbre d'Allen. Nous instancions, sur la base de notre extension de l'algèbre d'Allen, les propriétés d'objet suivantes :

³ <https://cedric.cnam.fr/isid/ontologies/UncertTimeOnto.owl#>

« RelationIntervals », « RelationIntervalsCertainty », « RelationIntervalPoint », « RelationIntervalPointCertainty », « RelationPointInterval », « RelationPointIntervalCertainty », « RelationPoints » et « RelationPointsCertainty ». Par exemple, « RelationIntervals » peut être l'une des relations d'Allen. En d'autres termes, 13 propriétés d'objet existent : « BeforeIntervals », « MeetsIntervals », « OverlapsIntervals », « StartsIntervals », « DuringIntervals », « EndsIntervals », « AfterIntervals », « Met-byIntervals », « Overlapped-byIntervals », « Started-byIntervals », « ContainsIntervals », « EndedbyIntervals » et « EqualsIntervals ».

Notre ontologie comporte également un ensemble de règles SWRL pour déduire de nouvelles relations temporelles. Une règle SWRL pour la déduire des données temporelles quantitatives fournies par l'utilisateur. Sur la base des tables de transitivité, nous associons une règle SWRL pour chaque relation de transitivité. La Figure 12 montre un exemple d'une règle SWRL.

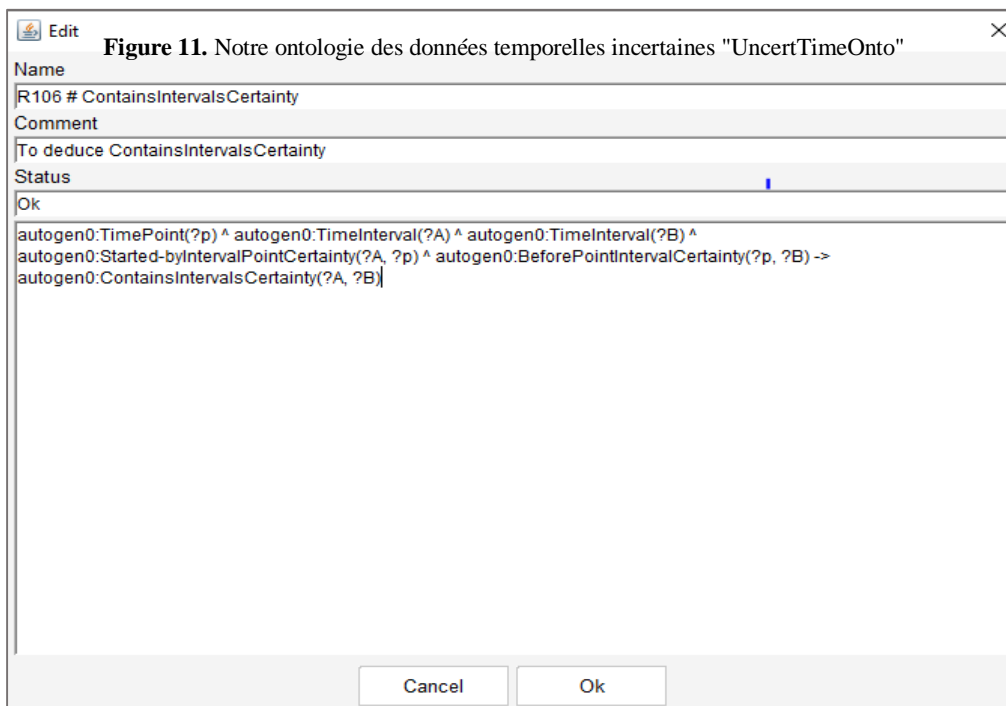


Figure 12. Exemple d'une règle SWRL

La règle SWRL suivante permet de déduire la propriété d'objet « EqualsPoints » entre deux points temporels M et P :

$$TimePoint(?m) \wedge TimePoint(?p) \wedge HasTimePoint(?m, ?mt) \wedge HasTimePoint(?p, ?pt) \wedge swrlb : equal(?pt, ?mt) \rightarrow EqualsPoints(?p, ?m)$$

« UncertTimeOnto » inclut 2 classes, 6 propriétés de type de données, 64 propriétés d'objet et 180 règles SWRL. Elle peut être intégrée dans d'autres ontologies certaines ou probabilistes pour traiter des données temporelles certaines et incertaines. Nous utilisons pour l'implémenter l'éditeur d'ontologie Protégé, comme le montre la Figure 13.

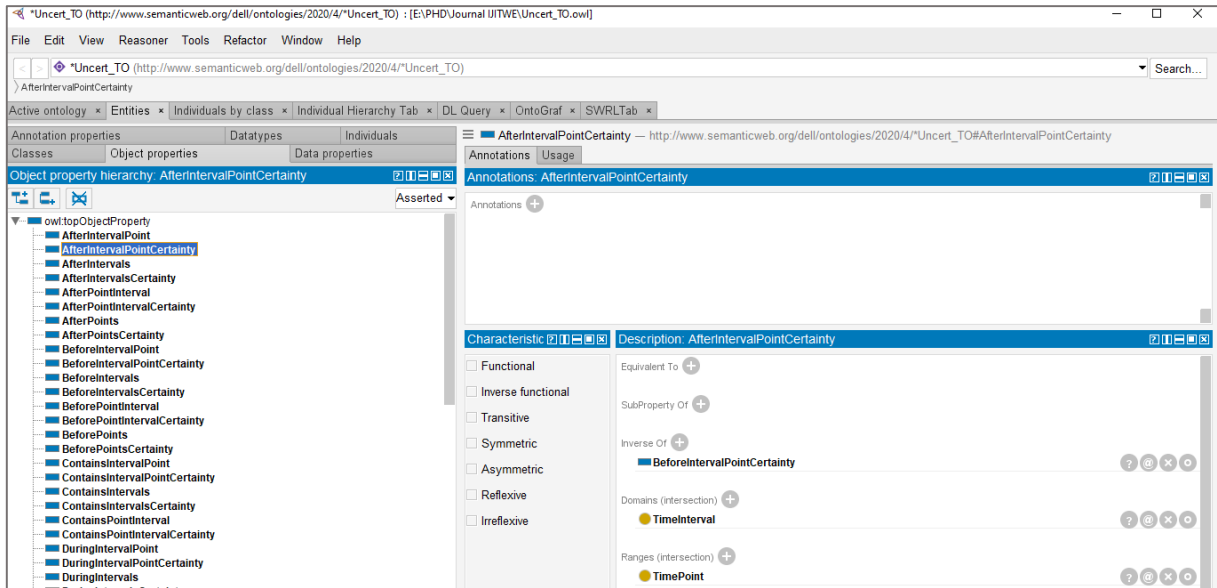


Figure 13. Notre ontologie créée par l'éditeur Protégé

5.5. Expérimentation

Pour valider notre approche, nous menons une étude de cas dont le but est de montrer l'intérêt de notre approche dans le cadre de la prothèse de mémoire CAPTAIN MEMO.

Nous implémentons un prototype basé sur l'ontologie proposée. Ensuite, nous l'intégrons dans CAPTAIN MEMO pour gérer la dimension temporelle des données décrites par l'ontologie PersonLink.

5.5.1 Prototype implémenté

Pour valider notre approche, nous avons implémenté avec le langage Java, un prototype basé sur notre ontologie « *UncertTimeOnto* ». Nous avons utilisé Netica-J⁴ pour implémenter les réseaux bayésiens et JENA API⁵ pour manipuler les ontologies.

Ce prototype comporte trois composants : « *UncertTimeOnto Population* », « *Qualitative Temporal Data Inference* » et « *Querying* ». Tout d’abord, l’utilisateur instancie « *UncertTimeOnto* » via l’interface utilisateur illustrée dans la Figure 14. Cette interface permet à l’utilisateur de saisir des données temporelles certaines/incertaines. Après chaque nouvelle saisie de données temporelles, le composant « *Inférence de données temporelles qualitatives* » est automatiquement exécuté pour déduire de nouvelles données et les degrés de certitude associés. Ce composant est basé sur les règles SWRL. Le troisième composant permet aux utilisateurs d’interroger l’ontologie via des requêtes SPARQL.

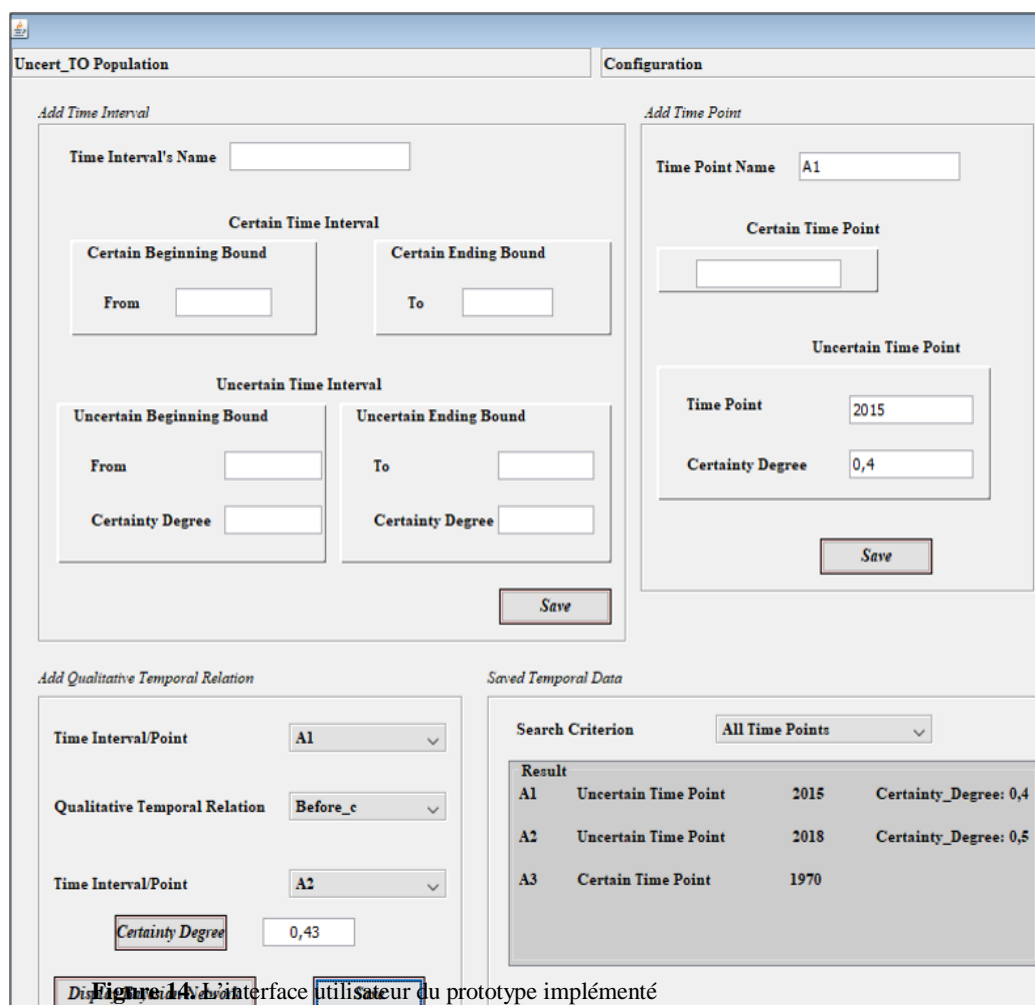


Figure 14. L'interface utilisateur du prototype implémenté

⁴ <https://www.norsys.com/netica-j.html#download>

⁵ <https://jena.apache.org/>

5.5.2 Application dans CAPTAIN MEMO

Nous intégrons notre prototype implémenté dans la prothèse de mémoire CAPTAIN MEMO pour gérer les données certaines/incertaines décrites par l'ontologie « PersonLink ».

Par exemple, prenons l'information suivante donnée par une patiente d'Alzheimer : « *Je pense qu'Helena a vécu en Chine de 1987 à 1990. Ensuite, elle a quitté la Chine pendant quelques années puis elle est revenue peut-être en 1996. Enfin, elle s'est installée à Hong-Kong en 2000* ». Soit $A = [1987_{0.6}, 1990_{0.2}]$ un intervalle de temps incertain représentant la durée de vie d'Helena en Chine. Soit $Q = 1996_{0.8}$ un point de temps incertain représentant le retour d'Helena à la Chine. Soit « $P = 2000$ » le point de temps certain représentant l'année de la vie d'Helena à Hong-Kong. La Figure 15 illustre une partie de l'ontologie PersonLink qui représente l'intervalle de temps incertain A, l'intervalle de temps incertain Q et le moment certain P.

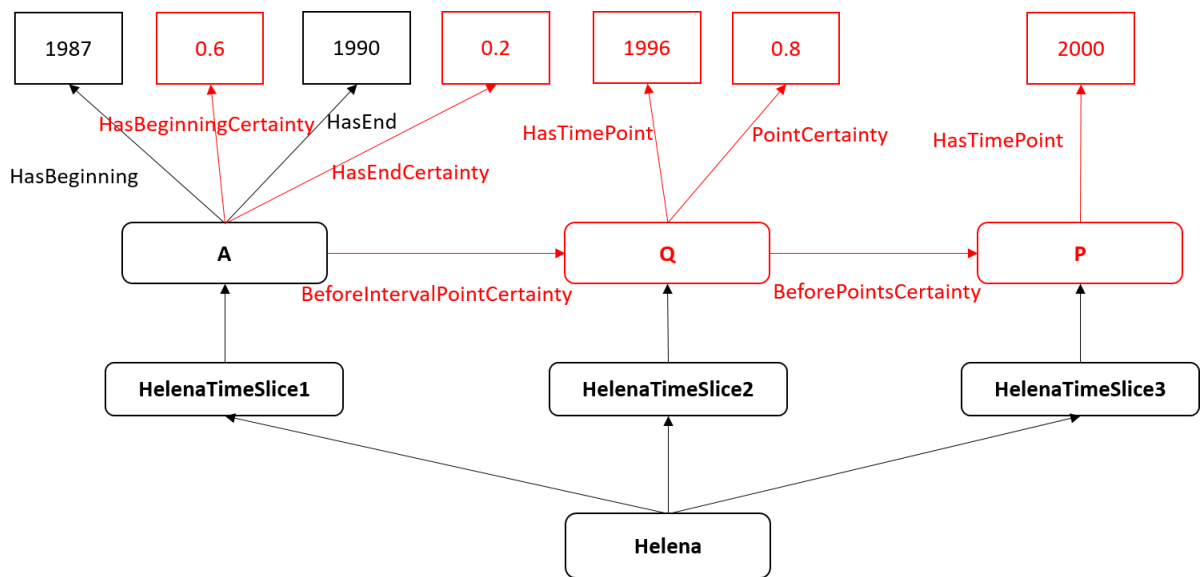


Figure 15. Un exemple d'ontologie PersonLink basée sur notre approche

5.6. Conclusion

Dans ce chapitre, nous avons proposé une approche de représentation et de raisonnement sur des données temporelles certaines et incertaines en OWL 2. Cette approche comporte trois contributions. La première consiste à étendre 4D-fluents avec de nouveaux composants ontologiques « certains » pour représenter des intervalles de temps incertains et des points de temps certains/incertains ainsi que des relations temporelles qualitatives entre les intervalles de temps

et les points. La seconde consiste à raisonner sur des données temporelles certaines et incertaines en étendant l’algèbre des intervalles d’Allen. Nous avons proposé 13 relations temporelles entre intervalles de temps incertains. Notre extension a l’avantage de garder les propriétés de réflexivité/irréflexivité, de symétrie/asymétrie et de transitivité. Ces relations permettent de relier un intervalle de temps et un instant ou deux instants ; où les intervalles de temps et les points peuvent être certains ou incertains. Nous avons introduit quatre tables de transitivité pour inférer les différentes relations temporelles. La troisième contribution consiste à proposer l’ontologie « UncertTimeOnto » qui implémente à la fois l’approche 4D-fluents et les extensions de l’algèbre d’Allen. Les inférences sont effectuées à l’aide de règles SWRL. Pour valider notre travail, nous avons implémenté un prototype permettant aux utilisateurs d’explorer notre approche, basée sur notre ontologie.

L’intérêt principal de notre approche, par rapport à d’autres travaux, est qu’elle peut être implémentée avec les standards classiques du Web sémantique et ne nécessite pas d’acquérir des connaissances liées aux ontologies probabilistes. Nous considérons que les standards classiques sont plus matures et supportent mieux la mise à l’échelle que les outils probabilistes.

Dans le chapitre qui suit, nous présentons notre approche qui traite les données temporelles incertaines en utilisant la théorie de possibilité. Nous comparons par la suite les approches probabilistes et possibilistes proposées.

CHAPITRE 6 : Traitement des données
temporelles incertaines dans OWL 2 :
Approche basée sur la théorie de possibi-
lité

6.1. Introduction

Comme mentionné dans les précédents chapitres, l'incertitude est souvent présente dans des données saisies manuellement. C'est le cas des données temporelles où il est possible d'avoir une incertitude dans la description de type « *Je ne suis pas certain si Mariana était ma voisine en 2002* ». Traiter ce type de données est ce que nous abordons spécifiquement dans ce chapitre.

Dans le domaine du web sémantique, un nombre considérable d'approches ont été proposées pour traiter des données temporelles parfaites. Dans le chapitre précédent, notre approche proposée qui traite l'incertitude est basée sur la théorie des probabilités.

Dans ce chapitre, nous proposons une autre approche qui **traite l'incertitude**, mais qui est basée sur **la théorie des possibilités**. Elle est dédiée à représenter et raisonner sur des intervalles de temps et des points temporels incertains, et les relations temporelles qualitatives associées. Elle consiste à **étendre l'approche 4D-fluents pour représenter des données temporelles en OWL2**. Nous utilisons **la théorie de possibilité pour calculer les degrés de nécessité** (certitude) et nous **étendons l'algèbre des intervalles d'Allen pour raisonner sur des relations temporelles incertaines**. Les relations d'intervalle résultantes préservent les propriétés de l'algèbre d'origine. Nous adaptons ces relations pour relier un intervalle de temps et un point de temps, et deux points de temps. Nous les utilisons par la suite pour le raisonnement temporel via des tables de transitivité. Nous proposons **une ontologie temporelle possibiliste OWL 2 basée sur nos extensions nommée « *UncertPossibTimeOnto* »**. Les relations temporelles qualitatives proposées sont inférées via des règles SWRL. Nous implémentons un prototype basé sur cette ontologie et nous l'appliquons à l'ontologie PersonLink. **Nous comparons nos deux approches qui traitent l'incertitude**.

Dans la suite du chapitre, nous présentons nos contributions de représentation et de raisonnement dans les sections 6.2 et 6.3. Ensuite, nous présentons notre ontologie dans la section 6.4. La section 6.5 introduit l'expérimentation de notre approche dans le cadre de CAPTAIN MEMO ainsi que la comparaison entre les deux approches probabilistes et possibilistes. Nous terminons par une conclusion.

6.2. Représentation des données temporelles incertaines dans OWL 2 : extension de l'approche 4D-Fluents

Nous étendons l'approche 4D-fluents avec des composants ontologiques « certains » pour représenter des données temporelles quantitatives « incertaines » et les relations temporelles qualitatives associées dans OWL 2. La Figure 16 représente notre extension.

6.2.1 Données temporelles quantitatives

Nous étendons l'approche 4D-fluents pour représenter : (i) des intervalles de temps incertains et (ii) des points de temps incertains.

6.2.1.1 Intervalles de temps

Soit $A = [A^-_{Nba}, A^+_{Nea}]$ un intervalle de temps incertain. A^- et A^+ sont respectivement les bornes de début et de fin. Chaque borne a un degré de nécessité respectivement Nba et Nea . Nous étendons 4D-fluents pour représenter des intervalles de temps incertains. Nous introduisons deux propriétés de type de données certaines nommées « HasBeginNecessity » et « HasEndNecessity ». Ces deux propriétés sont associées à la classe prédéfinie dans l'approche 4D-fluents « TimeInterval ». « HasBeginNecessity » et « HasEndNecessity » représentent, respectivement, les degrés de nécessité associés à Nba et Nea .

6.2.1.2 Les points de temps

Soit Q_{Nq} un point de temps incertain, où Nq est le degré de nécessité. Nous introduisons une propriété de type de données « HasTimePoint » pour relier la classe « TimePoint » et Q . Pour représenter le degré de nécessité associé à Q , nous proposons une propriété de type de données certaine nommée « PointNecessity » que nous associons à la classe « TimePoint ».

6.2.2 Représentation qualitative des données temporelles

Quatre relations temporelles peuvent être établies entre les points de temps et les intervalles de temps : les relations Point-Point, Intervalle-Point, Point-Intervalle et Intervalle-Intervalle.

Pour représenter les degrés de nécessité des relations temporelles incertaines entre :

- des intervalles de temps : nous créons une propriété d'objet certaine nommée « RelationIntervalsNecessity » qui relie deux instances de la classe « TimeInterval » ;
- un intervalle de temps et un point : nous créons une propriété d'objet certaine nommée « RelationIntervalPointNecessity » qui relie une instance de la classe « TimeInterval » et une instance de la classe « TimePoint » ;
- un point temporel et un intervalle de temps : nous créons une propriété d'objet certaine nommée « RelationPointIntervalNecessity » qui relie une instance de la classe « TimePoint » et une instance de la classe « TimeInterval » ;

- des points temporels : nous créons une propriété d'objet certaine nommée « RelationPointsNecessity » qui relie deux instances de la classe « TimePoint ».

6.2.3 Degré de nécessité

Nous proposons une approche pour traiter les intervalles de temps et les points de temps incertains dans OWL 2 en utilisant la théorie des possibilités. Cette théorie caractérise les événements par deux mesures qui sont la possibilité Π et la nécessité (ou certitude) N .

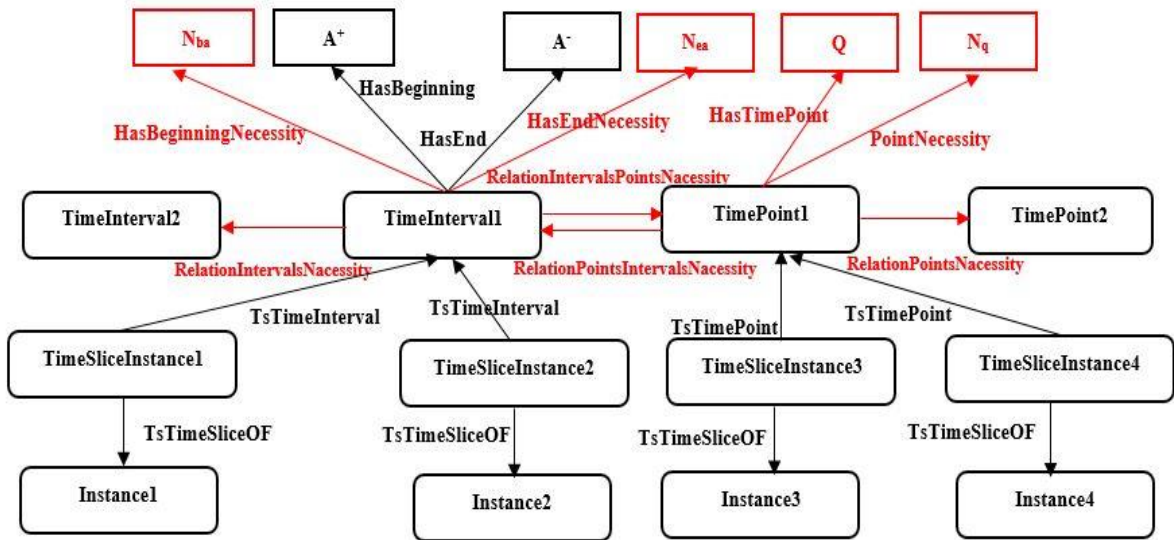


Figure 16. Notre extension du 4D-fluents pour représenter des données temporelles incertaines avec la théorie de possibilité

Rappelons que la théorie des possibilités est basée sur des distributions de possibilités. Une distribution de possibilité, notée $\pi_x(u)$, est une application de l'univers du discours à l'échelle $[0, 1]$ traduisant une connaissance partielle de l'univers (Dubois & Prade, 1998).

Notre objectif est d'estimer le degré de nécessité des données temporelles incertaines. En effet, selon (Zadeh, 1978), « un événement probable (ou incertain) doit être possible ». Selon (Dubois & Prade, 1990), un degré de possibilité est la limite supérieure d'un degré de probabilité et le degré de nécessité est sa limite inférieure puisqu'un événement incertain doit être probable et le probable doit être possible (c.-à-d. $N(A) \leq P(A) \leq \Pi(A)$). En effet, la nécessité N représente le degré de certitude de l'information. Selon (Ammar, 2014), Si « $N = 1$ », alors l'information est certaine (certitude absolue). Dans notre cas, l'information temporelle est incertaine, donc N doit être strictement inférieur à 1 ($N < 1$).

Exemple : Soit $A = [A^-_{Nba}, A^+_{Nea}]$ et $B = [B^-_{Nbb}, B^+_{Neb}]$ sont des intervalles de temps incertains où $A = [1960, 1963]$, $B = [1940, 1963]$; et « $Nba = 0.73$ », « $Nea = 0.61$ », « $Nbb =$

0.57 » et « Neb = 0.85 ». Il tient que « Ends_N(A, B) = 1 ; Où (A⁻_{Nba} > B⁻_{Nbb}) ∧ (A⁺_{Nea} = B⁺_{Neb}) = 1 » ; avec un degré de nécessité « N » égal à « 0.57 ». Ce dernier est déduit des degrés de nécessité « N₁ » et « N₂ » de la conjonction de « (A⁻_{Nba} > B⁻_{Nbb}) » et « (A⁺_{Nea} = B⁺_{Neb}) ».

6.3. Raisonnement sur les données temporelles incertaines : extension de l’algèbre d’intervalle d’Allen

Nous étendons l’algèbre d’Allen pour : (i) raisonner sur des données temporelles quantitatives incertaines pour inférer des relations temporelles qualitatives et (ii) raisonner sur des relations temporelles qualitatives pour en inférer de nouvelles.

6.3.1 Relations temporelles qualitatives entre les intervalles de temps

Nous étendons l’algèbre des intervalles d’Allen pour raisonner sur des intervalles de temps incertains. Pour les intervalles de temps certains, notre approche se réduit à l’algèbre d’intervalle d’Allen. Nous redéfinissons les 13 relations d’Allen pour proposer des relations temporelles entre des intervalles de temps incertains. Soit « A = [A⁻_{Nba}, A⁺_{Nea}] » et « B = [B⁻_{Nbb}, B⁺_{Neb}] » deux intervalles de temps incertains. Par exemple, nous redéfinissons la relation « Before(A, B) » comme : « Before_N(A, B) » ; où « N » est le degré de nécessité associé à la relation « Before » entre A et B. Cela signifie que la borne inférieure incertaine de l’intervalle A est inférieure à la borne supérieure incertaine de B.

$$Before_N(A, B) \rightarrow A^+_{Nea} < B^-_{Nbb}$$

Le degré de nécessité N est déduit des degrés de nécessité de conjonction N_{ba} et N_{eb} .

$$(N = N_{ba} \wedge N_{eb})$$

La conjonction fournit un résultat plus précis que si chaque source est considérée séparément (Dubois & Prade, 1998). Les degrés de possibilité et de nécessité dans ce cas peuvent être calculés en utilisant la fonction « min » (Ammar, 2014) :

$$N = \min(N_{ba}, N_{eb})$$

Par exemple, si nous avons le degré de nécessité N_{ba} égal à « 0,4 » et le degré de nécessité N_{eb} égal à « 0,5 » alors « N » est égal à « 0,4 » ;

Nous classons les relations proposées en deux catégories. La première couvre les relations qui consistent en une formule sans aucune conjonction (c’est-à-dire « Before_N », « After_N », « Meet_N » et « Met-by_N »). La seconde couvre le reste des relations. Celles-ci sont définies sur la base d’un ensemble de conjonctions (par exemple, « Overlaps_N »). Ainsi, comme pour la

relation « Before_N », nous redéfinissons la relation « Meets_N(A, B) » ; comme illustré dans le Tableau 18.

Nous redéfinissons la relation « Overlaps(A, B) » comme suit : « Overlaps_N(A, B) » ; où « N » est le degré de nécessité associé à cette relation. Cela signifie que la borne supérieure incertaine de A est inférieure à la borne supérieure incertaine de B, que la borne inférieure incertaine de A est supérieure à la borne supérieure incertaine de B et que la borne inférieure incertaine de l'intervalle A est inférieure à la borne inférieure incertaine de B.

$$\text{Overlaps}_N(A, B) \rightarrow (A^-_{Nba} < B^-_{Nbb}) \wedge (A^+_{Nea} > B^-_{Nbb}) \wedge (A^+_{Nea} < B^+_{Neb})$$

Le degré de nécessité « N » est déduit des degrés de nécessité « N₁ », « N₂ » et « N₃ » associés respectivement aux conjonctions $(A^-_{Nba} < B^-_{Nbb})$, $(A^+_{Nea} > B^-_{Nbb})$ et $(A^+_{Nea} < B^+_{Neb})$. La première consiste à définir le degré de nécessité « N₁ », la seconde celui de « N₂ » et la troisième celui de « N₃ ». Enfin, le degré de nécessité « N » est défini à partir de « N₁ », « N₂ » et « N₃ ». Nous redéfinissons les autres relations d'Allen de la même manière que la relation « Overlaps_N(A, B) », comme indiqué dans le Tableau 20.

Tableau 20. Relations temporelles entre deux intervalles de temps incertains A et B

Relation(A, B)	Définition	Inverse(B, A)
Before _N (A, B)	$A^+_{Nea} < B^-_{Nbb}$	After _N (B, A)
Meets _N (A, B)	$A^+_{Nea} = B^-_{Nbb}$	Met-by _N (B, A)
Overlaps _N (A, B)	$(A^-_{Nba} < B^-_{Nbb}) \wedge (A^+_{Nea} > B^-_{Nbb}) \wedge (A^+_{Nea} < B^+_{Neb})$	Overlapped-by _N (B, A)
Starts _N (A, B)	$(A^-_{Nba} = B^-_{Nbb}) \wedge (A^+_{Nea} < B^+_{Neb})$	Started-by _N (B, A)
During _N (A, B)	$(A^-_{Nba} > B^-_{Nbb}) \wedge (A^+_{Nea} < B^+_{Neb})$	Contains _N (B, A)
Ends _N (A, B)	$(A^-_{Nba} > B^-_{Nbb}) \wedge (A^+_{Nea} = B^+_{Neb})$	Ended-by _N (B, A)
Equals _N (A, B)	$(A^-_{Nba} = B^-_{Nbb}) \wedge (A^+_{Nea} = B^+_{Neb})$	Equals _N (B, A)

6.3.2 Relations temporelles qualitatives entre un intervalle de temps et un point de temps

Nous adaptons les relations temporelles qualitatives entre les intervalles de temps pour proposer des relations entre un intervalle de temps et un point temporel, c'est-à-dire des relations intervalle-point et point-intervalle ; comme illustré dans le Tableau 21.

Tableau 21. Relation temporelle entre un intervalle de temps incertain A et un point de temps incertain P

Relation(P, A)	Définition	Illustration	Inverse(A, P)
Relation temporelle entre un intervalle de temps incertain A = [A^{-Nba-}, A^{+Nea}] et un point de temps incertain P_{Np}			
Before _N (P _{Np} , A)		$P_{Np} < A^{-Nba-}$	After _N (A, P _p)
After _N (P _{Np} , A)		$A^{+Nea} < P_{Np}$	Before _N (A, P _p)
Starts _N (P _{Np} , A)		$P_{Np} = A^{-Nba-}$	Started-by _N (A, P _p)
During _N (P _{Np} , A)		$(A^{-Nba-} < P_{Np}) \wedge (P_{Np} < A^{+Nea})$	Contains _N (A, P _p)
Ends _N (P _{Np} , A)		$P_p = A^{+Nea}$	Ended-by _N (A, P _p)

6.3.3 Relations temporelles qualitatives entre deux points de temps

Les relations temporelles qualitatives entre intervalles de temps sont adaptées pour représenter des relations entre points de temps, comme le montre le Tableau 22.

Tableau 22. Relations temporelles entre les points de temps incertains P_{Np} et Q_{Nq}

Relation(P, Q)	Définition	Inverse(Q, P)
Relations temporelles entre les points de temps incertains P _{Np} et Q _{Nq}		
Before _N (P, Q)	$P_{Np} < Q_{Nq}$	After _N (Q, P)
Equals _N (P, Q)	$P_{Np} = Q_{Nq}$	Equals _N (Q, P)

Exemple : soit P et Q deux points de temps incertains ; où P_{Np} = 09h45 est associé à un degré de nécessité « Np = 0,5 » et Q_{Nq} = 12h00 est associé à un degré de nécessité « Nq = 0,7 ». Nous concluons que After_N(Q_{Nq}, P_{Np}) = 1. En effet, After_N(Q_{Nq}, P_{Np}) = Q_{Nq} > P_{Np} = 1.

6.3.4 Tables de transitivité

Nous proposons des tables de transitivité pour dériver de nouvelles données temporelles à partir des relations temporelles qualitatives entre les intervalles de temps et les points de temps.

Relations temporelles qualitatives entre les intervalles de temps. La table de transitivité d'Allen nous permet de déduire de R₁(A, B) et R₂(B, C) que R₃(A, C) est valide ; où A, B et C

sont des intervalles de temps certains et R_1 , R_2 et R_3 sont des relations d'Allen. Nous généralisons ces déductions en utilisant les trois intervalles de temps incertains $A = [A^-, A^+]$, $B = [B^-, B^+]$ et $C = [C^-, C^+]$.

Sur la base du Tableau 18, « $\text{Before}_N(A, K)$ » peut être déduit de « $\text{During}_N(A, B)$ » et « $\text{Meet}_N(B, K)$ ». En effet, « $(B^-_{Nbb^-} < A^-_{Nba^-}) \wedge (A^+_{Nea^+} < B^+_{Neb^+})$ » est déduit de « $\text{During}_N(A, B)$ » et « $(B^+_{Neb^+} = K^-_{Nbk^-})$ » est déduit de « $\text{Meet}_N(B, K)$ ». Avec « $(A^+_{Nea^+} < B^+_{Neb^+})$ » et « $(B^+_{Neb^+} = K^-_{Nbk^-})$ » nous concluons que « $\text{Before}_N(A, K)$ » est valide. La table de transitivité proposée coïncide avec celle d'Allen.

Relations temporelles qualitatives entre un intervalle de temps et un point dans le temps.

Deux tables de transitivité sont créées pour raisonner sur les relations temporelles entre un intervalle de temps et un point de temps. Soit A et B des intervalles de temps et P et Q des points de temps. Ils peuvent être tous certains ou incertains. Le Tableau 23 représente la table de transitivité qui permet de déduire $R_3(P, Q)$ de $R_1(P, A)$ et $R_2(A, Q)$. R_1 est une relation point-intervalle ; R_2 est une relation intervalle-point et R_3 est une relation point-point. Certaines relations de composition peuvent donner lieu à plusieurs décisions (donc indécidable).

Par exemple, les relations « $\text{After}_N(P, A)$ » et « $\text{Before}_N(A, Q)$ » donnent trois relations possibles : « $\text{Before}_N(P, Q)$ », « $\text{Equals}_N(P, Q)$ » et « $\text{After}_N(P, Q)$ ».

Tableau 23. Table de transitivité qui permet de déduire $R_3(P, Q)$ de $R_1(P, A)$ et $R_2(A, Q)$

$R_1(P, A)$ $R_2(A, Q)$	$\text{Before}_N(P, A)$	$\text{After}_N(P, A)$	$\text{Starts}_N(P, A)$	$\text{During}_N(P, A)$	$\text{Ends}_N(P, A)$
$\text{Before}_N(A, Q)$	$\text{Before}_N(P, Q)$	L'une des relations des points de temps	$\text{Before}_N(P, Q)$	$\text{Before}_N(P, Q)$	$\text{Before}_N(P, Q)$
$\text{After}_N(A, Q)$	L'une des relations des points de temps	$\text{After}_N(P, Q)$	$\text{After}_N(P, Q)$	$\text{After}_N(P, Q)$	$\text{After}_N(P, Q)$
$\text{Started-by}_N(A, Q)$	$\text{Before}_N(P, Q)$	$\text{After}_N(P, Q)$	$\text{Equals}_N(P, Q)$	$\text{After}_N(P, Q)$	$\text{After}_N(P, Q)$

Contains_N(A, Q)	Before _N (P, Q)	After _N (P, Q)	Before _N (P, Q)	L'une des relations des points de temps	After _N (P, Q)
Ended-by_N(A, Q)	Before _N (P, Q)	After _N (P, Q)	Before _N (P, Q)	Before _N (P, Q)	Equals _N (P, Q)

Tableau 24 représente la table de transitivité qui permet de déduire $R_3(A, B)$ à partir de $R_1(A, P)$ et $R_2(P, B)$ où R_1 est une relation intervalle-point, R_2 est une relation point-intervalle et R_3 est une relation intervalle-intervalle.

Tableau 24. Table de transitivité qui permet de déduire $R_3(A, B)$ de $R_1(A, P)$ et $R_2(P, B)$

R₁(A, P) \ R₂(P, B)	Before_N(A, P)	After_N(A, P)	Started-by_N(A, P)	Contains_N(A, P)	Ended-by_N(A, P)
Before_N(P, B)	Before _N (A, B)	L'une des relations d'intervalle de temps	Contains _N (A, B) ∨ Ended-by _N (A, B) ∨ Met-by _N (A, B) ∨ Before _N (A, B) ∨ Overlaps _N (A, B)	Contains _N (A, B) ∨ Ended-by _N (A, B) ∨ Overlapped-by _N (A, B) ∨ Meets _N (A, B) ∨ Before _N (A, B)	Before _N (A, B)
After_N(P, B)	L'une des relations d'intervalle de temps	After _N (A, B)	After _N (A, B)	After _N (A, B) ∨ Met-by _N (A, B) ∨ During _N (A, B) ∨ Started-by _N (A, B) ∨ Overlapped-by _N (A, B)	During _N (A, B) ∨ Started-by _N (A, B) ∨ Overlapped-by _N (A, B) ∨ Met-by _N (A, B) ∨ After _N (A, B)
Starts_N(P, B)	Before _N (A, B)	After _N (A, B) ∨ Met-by _N (A, B) ∨ Overlapped-by _N (A, B) ∨ Ended-by _N (A, B) ∨ Contains _N (A, B)	Started-by _N (A, B) ∨ Equals _N (A, B) ∨ Starts _N (A, B)	Contains _N (A, B) ∨ Ended-by _N (A, B) ∨ Overlaps _N (A, B)	Meets _N (A, B)
During_N(P, B)	During _N (A, B) ∨ Starts _N (A, B) ∨ Overlaps _N (A, B) ∨ Meets _N (A, B)	After _N (A, B) ∨ Met-by _N (A, B)	Contains _N (A, B) ∨ Ends _N (A, B) ∨ Overlapped-by _N (A, B)	Overlaps _N (A, B) ∨ Overlapped-by _N (A, B) ∨ Starts _N (A, B) ∨	Overlapped-by _N (A, B) ∨

	$B) \vee \text{Before}_N(A, B)$	$B) \vee \text{Overlapped-by}_N(A, B) \vee \text{Ended-by}_N(A, B) \vee \text{Contains}_N(A, B)$		$\text{Started-by}_N(A, B) \vee \text{During}_N(A, B) \vee \text{Contains}_N(A, B) \vee \text{Ends}_N(A, B) \vee \text{Equals}_N(A, B) \vee \text{Ended-by}_N(A, B)$	$\text{Starts}_N(A, B) \vee \text{During}_N(A, B)$
Ends_N(P, B)	$\text{During}_N(A, B) \vee \text{Starts}_N(A, B) \vee \text{Overlaps}_N(A, B) \vee \text{Meets}_N(A, A) \vee \text{Before}_N(A, B)$	$\text{After}_N(A, B)$	$\text{Met-by}_N(A, B)$	$\text{Contains}_N(A, B) \vee \text{Started-by}_N(A, B) \vee \text{Overlapped-by}_N(A, B)$	$\text{Equals}_N(A, B) \vee \text{Ended-by}_N(A, B) \vee \text{During}_N(A, B)$

Relations temporelles qualitatives entre les points de temps. Nous créons une table de transitivité pour raisonner sur les relations temporelles qualitatives qui peuvent être établies entre les points temporels. Cette table permet de déduire $R_3(P, M)$ à partir de $R_1(P, Q)$ et $R_2(Q, M)$ où P, Q et M sont des points de temps certains ou incertains. R_1, R_2 et R_3 sont des relations point-point.

Tableau 25. Table de transitivité qui permet de déduire $R_3(P, M)$ à partir de $R_1(P, Q)$ et $R_2(Q, M)$

$R_1(P, Q)$ / $R_2(Q, M)$	Before_N(P, Q)	Equals_N(P, Q)	After_N(P, Q)
Before_N(Q, M)	$\text{Before}_N(P, M)$	$\text{Before}_N(P, M)$	L'une des relations de points de temps
Equals_N(Q, M)	$\text{Before}_N(P, M)$	$\text{Equals}_N(P, M)$	$\text{After}_N(P, M)$
After_N(Q, M)	L'une des relations de points de temps	$\text{After}_N(P, M)$	$\text{After}_N(P, M)$

Exemple : Soit P_{Np}, Q_{Nq} et L_{Nl} trois points temporels incertains. De « $\text{Before}_N(P, Q) = 1$ », nous déduisons « $\text{After}_N(Q, P) = 1$ ». En nous basant sur la relation de transitivité, nous déduisons « $\text{After}_N(Q, L)$ » de « $\text{After}_N(Q, P)$ » et « $\text{After}_N(P, L)$ ». Alors : $\text{After}_N(Q, L) = 1$.

6.4. «UncertPossibTimeOnto»: notre ontologie temporelle pour les données temporelles incertaines

Nous proposons une ontologie temporelle possibiliste appelée «UncertPossibTimeOnto»⁶ qui implémente notre extension de l'approche 4D-fluents et notre extension de relations d'Allen.

Name	Rule	Comment
R1#BeforePointIntervalNecessity	autogen0:TimeInterval(?A) ^ autogen0:HasEnd(?A, ?Ae) ^ autogen0:TimePoint(?p) ^ swrlb:lessThan(?pt, ?Ab) ^ autogen0:HasTimePoint(?p, ?pt) ^ autogen0:HasB...	To deduce BeforePointInt
R10#AfterPointsNecessity	autogen0:TimePoint(?m) ^ autogen0:TimePoint(?p) ^ autogen0:HasTimePoint(?m, ?mt) ^ autogen0:HasTimePoint(?p, ?pt) ^ autogen0:PointNecessity(?m, ?mn) ^ ...	To deduce AfterPointsNec
R100#BeforePointsNecessity	autogen0:TimePoint(?m) ^ autogen0:AfterPointsNecessity(?q, ?m) ^ autogen0:TimePoint(?p) ^ autogen0:BeforePointsNecessity(?p, ?q) ^ autogen0:TimePoint(?q)...	To deduce BeforePointsN
R101#AfterPointsNecessity	autogen0:TimePoint(?m) ^ autogen0:AfterPointsNecessity(?q, ?m) ^ autogen0:TimePoint(?p) ^ autogen0:EqualsPointsNecessity(?p, ?q) ^ autogen0:TimePoint(?q)...	To deduce AfterPointsNec
R102#BeforePointsNecessity	autogen0:TimePoint(?m) ^ autogen0:AfterPointsNecessity(?q, ?m) ^ autogen0:TimePoint(?p) ^ autogen0:AfterPointsNecessity(?p, ?q) ^ autogen0:TimePoint(?q)...	To deduce AfterPointsNec
R103#BeforeIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:BeforeIntervalPoint...	To deduce BeforeIntervals
R104#BeforeIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:AfterIntervalPointNe...	To deduce BeforeIntervals
R105#AfterIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:AfterIntervalPointNe...	To deduce AfterIntervalsN
R106#ContainsIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:Started-byIntervalP...	To deduce ContainsInterv
R107#Ended-byIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:Started-byIntervalP...	To deduce Ended-byInterv
R108#Met-byIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:Started-byIntervalP...	To deduce Met-byIntervald
R109#BeforeIntervalsNecessity	autogen0:TimeInterval(?A) ^ autogen0:TimePoint(?p) ^ autogen0:TimeInterval(?B) ^ autogen0:BeforePointIntervalNecessity(?p, ?B) ^ autogen0:Started-byIntervalP...	To deduce BeforeIntervald
R11#AfterIntervalPoint	autogen0:TimeInterval(?A) ^ autogen0:HasBeginning(?A, ?Ab) ^ autogen0:HasEnd(?A, ?Ae) ^ autogen0:TimePoint(?p) ^ autogen0:HasTimePoint(?p, ?pt) ^ swrlb:d...	To deduce AfterIntervalPo

Control Rules Asserted Axioms Inferred Axioms OWL 2 RL

OWL axioms successfully transferred to rule engine.
 Number of SWRL rules exported to rule engine: 180
 Number of OWL class declarations exported to rule engine: 2
 Number of OWL individual declarations exported to rule engine: 0
 Number of OWL object property declarations exported to rule engine: 64
 Number of OWL data property declarations exported to rule engine: 6
 Total number of OWL axioms exported to rule engine: 455
 The transfer took 4390 milliseconds(s).
 Press the 'Run Drools' button to run the rule engine.
 Successful execution of rule engine.
 Number of inferred axioms: 299
 The process took 225 milliseconds(s).
 Look at the 'Inferred Axioms' tab to see the inferred axioms.
 Press the 'Drools->OWL' button to translate the inferred axioms to OWL knowledge.

Figure 17. Règles SWRL de "UncertPossibTimeOnto"

Notre ontologie temporelle possibiliste est une ontologie classique contenant de nouveaux composants représentant des mesures de nécessité. Nous utilisons l'éditeur d'ontologie Protégé et le raisonneur Pellet. Nous déduisons via un ensemble de règles SWRL notre extension de l'algèbre d'Allen comme le montre la Figure 17. Notre ontologie contient 2 classes, 64 propriétés d'objet, 6 propriétés de données et 180 règles swrl.

⁶ https://cnam-my.sharepoint.com/:u:/g/personal/nassira_achich_auditeur_lecnam_net/EfwSQrjXNjRJCZhLNaghWEB_BCfxP5frSdyxGaYxAlz0Q?e=sCP

Par exemple, soit P et Q deux points temporels incertains. Nous proposons une règle SWRL pour déduire la propriété d'objet « *EqualsPointsNecessity* » entre deux points de temps incertains P et Q :

$$\begin{aligned} & \textit{TimePoint}(?p) \wedge \textit{TimePoint}(?q) \wedge \textit{HasTimePoint}(?p; ?pt) \\ & \wedge \textit{HasTimePoint}(?q; ?qt) \wedge \textit{PointNecessity}(?p, ?Np) \wedge \textit{Point-} \\ & \textit{Necessity}(?q, ?Nq) \wedge \textit{swrlb} : \textit{equal}(?pt, ?qt) \\ & \Rightarrow \textit{EqualsPointsNecessity}(?p; ?q) \end{aligned}$$

6.5. Expérimentations

Dans cette section, nous présentons le prototype implémenté basé sur notre ontologie « *UncertPossibTimeOnto* ». Nous conduisons une étude de cas en l'intégrant dans CAPTAIN MEMO. Enfin, nous comparons nos deux approches qui traitent l'incertitude des données temporelles.

6.5.1 Prototype implémenté

Nous implémentons en JAVA un prototype basé sur notre ontologie « *UncertPossibTimeOnto* ». Ce prototype est dédié aux utilitaires pour exploiter notre approche proposée. Pour manipuler les ontologies, nous utilisons JENA-API.

L'utilisateur peut instancier « *UncertPossibTimeOnto* » via l'interface utilisateur illustrée par la Figure 18. Il peut saisir des intervalles de temps et des points temporels incertains et les enregistrer. Après chaque nouvelle entrée de données temporelles, des inférences sont automatiquement exécutées. Les relations temporelles sont exécutées sur la base des règles SWRL que nous avons ajoutées à l'ontologie.

The screenshot shows a software interface for managing uncertain time intervals. It includes sections for adding uncertain and certain time intervals, adding qualitative temporal relations, and viewing saved temporal data. The interface is titled 'UncertPossibTimeOnto Population' and 'Configuration'.

Figure 18. Prototype basé sur « UncertPossibTimeOnto »

6.5.2 Étude de cas : application à l'ontologie PersonLink

Nous intégrons le prototype implémenté dans la prothèse de mémoire CAPTAIN MEMO pour gérer les données incertaines entrées dans le contexte de l'ontologie « PersonLink ».

Soit l'exemple suivant : « *Je pense qu'Anna vivait en Allemagne entre 1992 et 2000. Ensuite, elle a déménagé en Australie avant de revenir peut-être entre 2009 et 2013* ». Soit $A = [A^-_{Nba}, A^+_{Nea}]$ et $B = [B^-_{Nbb}, B^+_{Neb}]$ être deux intervalles de temps incertains ; où $A = [1992_{Nba}, 2000_{Nea}]$ représentant la durée de vie d'Anna en Allemagne et $B = [2009_{Nbb}, 2013_{Neb}]$ représentant le retour d'Anna en Allemagne.

Supposons que nous ayons les distributions de possibilité des bornes de A et B, comme le montrent les Figures 19 et 20. Nous calculons les degrés de possibilité et de nécessité chaque borne de A et B puis nous déduisons le degré de nécessité de A selon les équations que nous définies dans ce chapitre. Cela signifie que « $N_A = \min(N_{ba}, N_{ea}) = \min(0.65, 0.52) = 0.52$ ». De la même manière, nous déduisons le degré de nécessité de B comme suit : « $N_B = \min(N_{bb}, N_{eb}) = \min(0.74, 0.61) = 0.61$ ». Enfin nous déduisons le degré de nécessité de la relation « Before » entre les deux intervalles A et B comme suit : « $Before_N(A, B)$ » donne « $(A^{+N_{ba}} < B^{-N_{eb}})$ » ce qui signifie que « $Before_N(A, B) = \min(A^{+N_{ba}}, B^{-N_{eb}}) = \min(0.65, 0.61) = 0.61$ ».

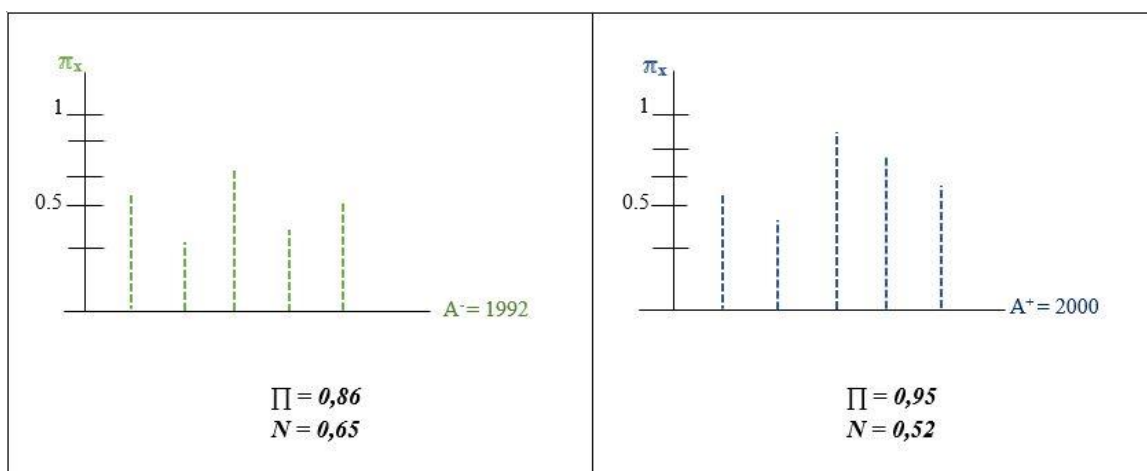


Figure 19. Distributions de possibilité de chaque borne de l'intervalle de temps incertain A

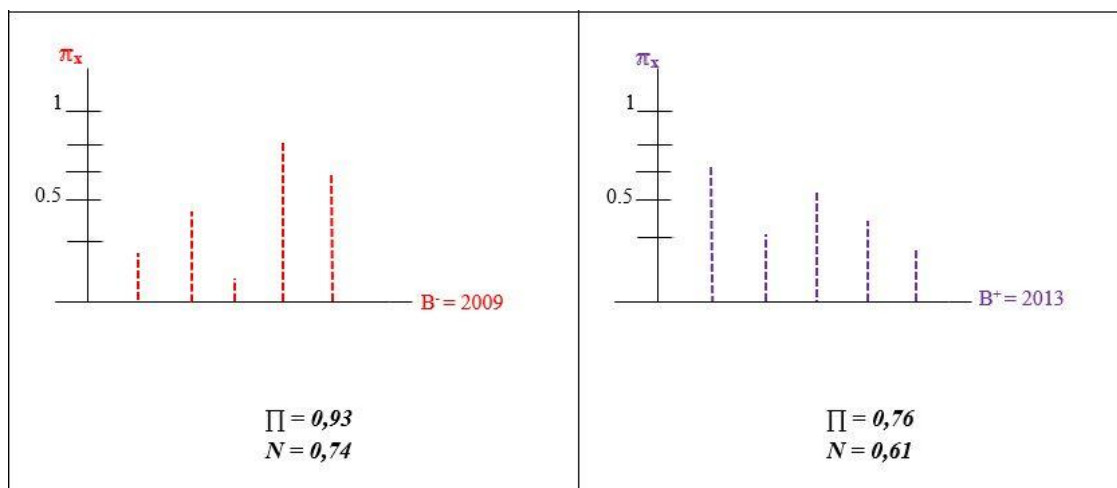


Figure 20. Distributions de possibilité de chaque borne de l'intervalle de temps incertain B

La représentation sémantique de l'exemple en utilisant les degrés obtenus sur la base de notre extension 4D-fluents est illustrée dans la Figure 21. Nous représentons les intervalles de temps incertains A et B associés aux degrés de nécessité. La relation qualitative « BeforeIntervalsNecessity » entre la durée de vie d'Anna en Allemagne, est représentée par l'intervalle A et son retour en Allemagne par B.

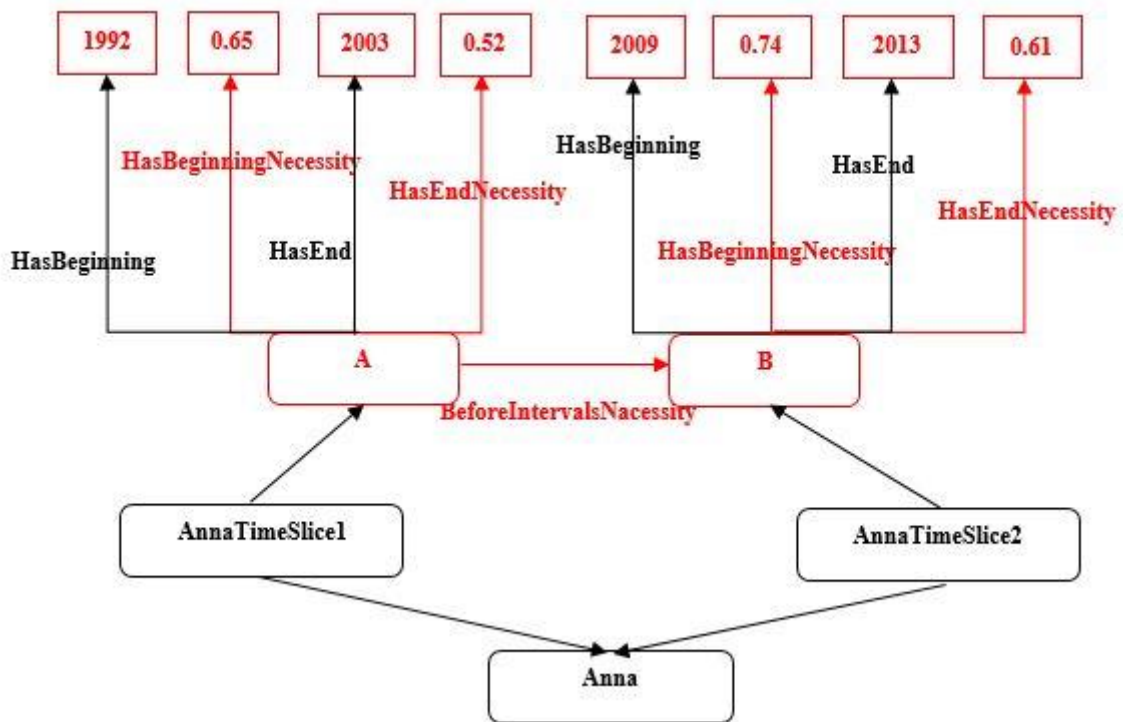


Figure 21. Un exemple d'ontologie PersonLink basée sur notre approche

6.5.3 Comparaison entre l'approche possibiliste et l'approche probabiliste

Afin d'évaluer la performance de notre approche proposée, nous comparons nos résultats avec ceux de l'approche que nous avons présentée dans le chapitre précédent et qui traite des données temporelles incertaines en utilisant la théorie des probabilités. Nous calculons les mesures de certitude à l'aide de réseaux bayésiens.

- En ce qui concerne la représentation, les deux approches ont pu être représentées sémantiquement en étendant l'approche 4D-fluents.
- En ce qui concerne le raisonnement, les deux approches garantissent des inférences correctes et des mesures précises de certitude des relations temporelles qualitatives.

- En ce qui concerne la mesure de certitude obtenue, les résultats montrent que la théorie des probabilités attribue des valeurs de certitude plus élevées que la théorie des possibilités. Cela signifie que la théorie des probabilités est plus optimiste en termes de degrés de certitude que la théorie des possibilités. Cependant, la théorie des possibilités donne des résultats plus explicites et réalistes, comme le montre le Tableau 26.

Tableau 26. Degrés de certitude obtenus à partir des théories de possibilité et de probabilité

		Approche possibiliste (Degré de nécessité N)	Approche probabiliste (Degré de certitude C)
A	A ⁻	0.65	0.91
	A ⁺	0.52	0.73
B	B ⁻	0.74	0.75
	B ⁺	0.61	0.86
Relation « Before »		0.61	0.78

Par conséquent, nous considérons que les résultats obtenus par la théorie des possibilités, dans ce cas, sont plus fiables.

6.6. Conclusion

Nous avons présenté dans ce chapitre une nouvelle approche basée sur **la théorie des possibilités** pour traiter l'incertitude qui touche les données temporelles dans OWL 2. Notre première contribution consistait à **étendre l'approche 4D-fluents** avec de nouveaux composants ontologiques **pour représenter les données temporelles incertaines et les relations qualitatives associées**. Notre deuxième contribution consistait à **étendre l'intervalle d'Allen algèbre pour raisonner sur les données manipulées**. Nous avons proposé 13 relations temporelles entre des intervalles de temps incertains. Nous avons également proposé des relations temporelles entre un intervalle de temps et un point de temps et entre deux points de temps. Notre extension conserve les propriétés de réflexivité/irréflexivité, de symétrie/asymétrie et de transitivité. La troisième contribution consistait à proposer **une ontologie temporelle possibiliste « *UncertPossibTimeOnto* »** qui implémente nos extensions. Enfin nous avons **comparé** notre approche proposée avec notre approche traitant les données temporelles incertaines en **utilisant la théorie des probabilités**.

Notre approche peut être appliquée à d'autres contextes tels que les données historiques. De plus, notre extension de l'algèbre d'Allen peut être appliquée à d'autres domaines de recherche tels que les bases de données.

Dans le prochain chapitre, nous proposons une approche qui traite les données temporelles à la fois incertaines et imprécises en utilisant la théorie de possibilité. Nous conduisons une étude de cas dans le contexte de CAPTAIN MEMO.

CHAPITRE 7 : Traitement des données temporelles à la fois incertaines et imprécises dans OWL 2

(Approche basée sur la théorie de possibilité)

7.1. Introduction

Les données peuvent souffrir **de plusieurs imperfections en même temps**. Par exemple, une information comme « *je pense que j'ai vécu à Paris entre les années 60 et les années 70* » englobe deux sortes d'imperfections qui sont l'incertitude exprimée par « *je pense* » et l'imprécision exprimée par « *entre les années 60 et les années 70* ». Représenter et raisonner sur ce type de données est ce que nous avons spécifiquement abordé dans ce chapitre.

Dans le domaine du web sémantique, plusieurs approches ont été proposées pour traiter des données temporelles parfaites. Cependant, **traiter des données temporelles à la fois incertaines et imprécises est un autre problème qui doit être résolu. À notre connaissance, aucune approche n'a abordé ce problème dans le contexte du web sémantique. La théorie des possibilités** fournit des modèles capables de gérer à la fois différents types d'informations et différentes formes d'imperfections. L'adaptation de cette théorie au contexte du web sémantique tire parti de l'utilisation de technologies sémantiques sur des données qui ne pourraient pas être pleinement exploitées en raison de leur imprécision et de leur incertitude ou de toute autre imperfection.

Dans ce chapitre, nous proposons **une approche basée sur la théorie des possibilités pour représenter et raisonner sur des intervalles de temps et des points de temps simultanément incertains et imprécis et les relations qualitatives associées**. Nous commençons par montrer comment calculer **les degrés de possibilité et de nécessité liés à l'imprécision et à l'incertitude des données temporelles manipulées**. Ensuite, **nous proposons une représentation sous forme d'ontologie des données temporelles, de leurs degrés de possibilité et de nécessité et les relations qualitatives qui les lient**. Pour le **raisonnement, nous étendons l'algèbre des intervalles d'Allen** pour traiter à la fois l'imprécision et l'incertitude. Les relations d'intervalle résultantes préservent les propriétés de l'algèbre d'origine. Nous adaptons ces relations pour relier un intervalle de temps et un point de temps, et deux points de temps. Toutes les relations proposées peuvent être utilisées pour un raisonnement temporel au moyen de tables de transitivité. Nous proposons **une ontologie temporelle possibiliste OWL 2 basée sur la représentation sémantique proposée et l'extension des relations d'Allen nommée « PossibilisticTemporalOnto »**. Les relations temporelles qualitatives proposées sont inférées via un ensemble de **règles SWRL**. Enfin, nous implémentons **un prototype basé sur cette ontologie** et nous menons **une étude de cas appliquée à des données temporelles saisies par des patients Alzheimer dans le cadre de CAPTAIN MEMO**.

Dans le reste du chapitre, les sections 7.2, 7.3 et 7.4 sont consacrées à la présentation de nos contributions de représentation et de raisonnement basées sur les mesures de possibilité et de nécessité. La section 7.5 présente notre ontologie « *PossibilisticTemporalOnto* ». Dans la section 7.6, nous présentons l'expérimentation de notre approche. Enfin, nous terminons par une conclusion du chapitre.

7.2. Mesures de possibilité et de nécessité

L'approche que nous proposons dans cette partie traite les points de temps et les intervalles de temps qui sont à la fois incertains et imprécis, en utilisant la théorie des possibilités.

Nous calculons les degrés de possibilité et de nécessité liés à l'imprécision des différents intervalles de temps possibles ($I_{i=1..n}$) (respectivement des points de temps $P_{p=1..n}$) en nous basant sur les équations définies par (Zadeh, 1978). En conséquence, nous obtenons un ensemble de mesures pour chaque intervalle de temps I_i (respectivement point de temps P_p), puis nous évaluons ces mesures et sélectionnons l'intervalle le plus possible (qui a le plus grand degré de possibilité).

Nous calculons également les degrés de possibilité et de nécessité liés à l'incertitude des informations temporelles données (intervalles de temps et points de temps). En effet, selon (Zadeh, 1978), « Un événement probable (ou incertain) doit être possible ». Selon (Dubois & Prade, 1990), un événement incertain doit être probable et un événement probable doit être possible. La nécessité « N » représente le degré de certitude de l'information. D'après (Ammar, 2014), si « $N = 1$ », alors l'information est certaine (certitude absolue). Dans notre cas, l'information temporelle est incertaine, donc N doit être strictement inférieur à 1 ($N < 1$).

7.2.1 Mesures de possibilité et de nécessité des intervalles de temps

Soit $[L, U]$ un intervalle de temps représentant les bornes extrêmes de l'information temporelle conjointement incertaine et imprécise, où « L » est la borne inférieure extrême et « U » est la borne supérieure extrême de l'intervalle. Par exemple, si on a l'information « Je pense que j'habitais à Paris entre les années 60 et 70 », $[L, U] = [60, 70]$ précisément [1960, 1979].

Soit $I_i = [I_i^-, I_i^+]$ représentant les différents intervalles possibles pouvant représenter cette information, définis par un expert du domaine, où : $I_i = [I_1, I_2, \dots, I_n]$ et « $L \leq U$ ».

Soit $I_i = [I_i^-, I_i^+]$ chaque intervalle de temps possible représenté par ses bornes inférieure et supérieure imprécises ; où $I_i^- = [I_i^{-(1)}, I_i^{-(N)}]$ et $I_i^+ = [I_i^{+(1)}, I_i^{+(N)}]$.

Pour calculer les mesures de possibilité et de nécessité, nous proposons cinq règles. Les quatre premières règles concernent l'imprécision des intervalles de temps. La cinquième règle concerne l'incertitude de l'ensemble de l'information donnée.

RègleIntevalle1 : Soit « RI_1 » une règle spécifiée à l'aide d'une relation binaire « \geq » et d'une borne inférieure extrême fixe de l'information temporelle « L ». Cette règle signifie que la borne inférieure de départ « $I_i^{-(l)}$ » doit être supérieure ou égale à « L » :

$$RI_1 (\geq I_i^{-(l)}, L) : I_i^{-(l)} \geq L$$

Les degrés de possibilité et de nécessité liés à chaque I_i et satisfaisant RI_1 , peuvent être exprimés et calculés, respectivement, comme suit :

$$\Pi_{(RI_1(I_i))}(True) = \sup_{I_i^{-(l)} \leq w} \pi_L(w)$$

$$1 - \Pi_{(RI_1(I_i))}(False) = \inf_{I_i^{-(l)} > w} 1 - \pi_L(w)$$

Pour simplifier la notation, on va utiliser les notations suivantes :

$$P(RI_1(I_i)) = \sup_{I_i^{-(l)} \leq w} \pi_L(w)$$

$$N(RI_1(I_i)) = \inf_{I_i^{-(l)} > w} 1 - \pi_L(w)$$

RègleIntevalle2 : Soit « RI_2 » une règle spécifiée à l'aide d'une relation binaire « \leq » et d'une borne supérieure fixe extrême de l'information « U ». Cette règle signifie que la borne supérieure finale « $I_i^{+(N)}$ » doit être inférieur ou égale à « U » :

$$RI_2 (\leq I_i^{+(N)}, U) : I_i^{+(N)} \leq U$$

Les degrés de possibilité $P(RI_2(I_i))$ et de nécessité $N(RI_2(I_i))$ liés à chaque I_i et satisfaisant RI_2 , peuvent être exprimés, respectivement, comme :

$$P(RI_2(I_i)) = \sup_{I_i^{+(N)} \geq w} \pi_U(w)$$

$$N(RI_2(I_i)) = \inf_{I_i^{+(N)} < w} 1 - \pi_U(w)$$

RègleIntevalle3 : Soit « RI_3 » une règle, spécifiée à l'aide d'une relation binaire « \leq ». Cette règle signifie que la borne supérieure de départ « $I_i^{+(1)}$ » doit être supérieure ou égale à la borne inférieure de fin « $I_i^{-(N)}$ ».

$$RI_3 (\leq I_i^{-(N)}, I_i^{+(1)}) : I_i^{-(N)} \leq I_i^{+(1)}$$

Les degrés de possibilité $P(RI_3(I_i))$ et de nécessité $N(RI_3(I_i))$ liés à chaque I_i et satisfaisant RI_3 , peuvent être exprimés, respectivement, comme :

$$P(RI_3(I_i)) = \sup_{I_i^{-(N)} \leq w \leq I_i^{+(1)}} \pi_Z(w)$$

$$N(RI_3(I_i)) = \inf_{I_i^{-(N)} > w > I_i^{+(1)}} 1 - \pi_Z(w)$$

RègleIntevalle4 : Pour calculer les degrés de possibilité et de nécessité de « I_i » au moyen des règles précédentes, nous proposons une quatrième « RI_4 » qui combine les trois premières. Cela signifie que cette règle présente la conjonction des trois règles :

$$RI_4 = RI_1 \wedge RI_2 \wedge RI_3$$

La conjonction fournit un résultat plus précis que si chaque source est considérée séparément (Dubois & Prade, 1990). La possibilité et les degrés de nécessité dans ce cas peuvent être calculés en utilisant la fonction « min » (Dubois & Prade, 1998) :

$$P(RI_4(I_i)) = \min (P(RI_1(I_i)), P(RI_2(I_i)), P(RI_3(I_i)))$$

$$N(RI_4(I_i)) = \min (N(RI_1(I_i)), N(RI_2(I_i)), N(RI_3(I_i)))$$

RègleIntevalle5 : L'incertitude de l'ensemble de l'information signifie que tous les intervalles possibles « I_i » représentant l'intervalle de temps imprécis sont incertains. Ainsi, nous proposons une cinquième règle « RI_5 » pour estimer les degrés de possibilité et de nécessité liés à l'incertitude de l'ensemble de l'information temporelle. P et N sont déduits des résultats obtenus de « RI_4 » liés à l'imprécision. Nous nous appuyons sur les équations proposées par (Dubois & Prade, 2006). Dans notre cas :

$$P(RI_5(I_i)) = P(RI_4(I_1)) \cup \dots \cup P(RI_4(I_n))$$

$$N(RI_5(I_i)) = N(RI_4(I_1)) \cap \dots \cap N(RI_4(I_n))$$

Cela signifie que :

$$P(RI_5(I_i)) = \max (P(RI_4(I_1)), \dots, P(RI_4(I_n)))$$

$$N(RI_5(I_i)) = \min (N(RI_4(I_1)), \dots, N(RI_4(I_n)))$$

7.2.2 Mesures de possibilité et de nécessité des points de temps

Soit $[P_l, P_u]$ un intervalle temporel représentant les limites extrêmes du point de temps à la fois incertain et imprécis, où « P_l » est le point de temps le plus bas possible et « P_u » est le plus grand point de temps possible qui représente un point de temps imprécis. Par exemple, si nous avons l'information « Peut-être qu'ils viendront vers 10 heures du matin », $[P_l, P_u] = [09:45, 10 :15]$. Soit P_p représentant les différents instants possibles pouvant représenter cette information, définis par un expert du domaine, où : $P_p = \{P_1, P_2, \dots, P_3\}$ et $P_l \leq P_p \leq P_u$.

Soit $P_p = [P_p^l, P_p^n]$ tous les points temporels possibles représentés par ses bornes supérieures et inférieures imprécises.

Pour calculer les mesures de possibilité et de nécessité des points de temps possibles, nous adaptons les règles proposées pour les intervalles de temps. Nous proposons quatre règles. Les trois premières règles concernent l'imprécision des points de temps. La quatrième règle concerne l'incertitude de l'ensemble de l'information donnée.

RèglePoint1 : Soit « RP_l » une règle spécifiée à l'aide de la relation binaire « \geq » et d'un point de temps fixe le plus bas possible de l'information temporelle « P_l ». Cette règle signifie qu'une borne de début d'un point temporel possible « P_p^l » doit être supérieure ou égale à « P_l » :

$$RP_l (\geq, P_p^l, P_l) : P_p^l \geq P_l$$

Degrés de possibilité et de nécessité liés à chacun P_p et satisfaisant RP_l , peuvent être exprimés et calculés, respectivement, comme :

$$\Pi_{(RP_l(P_p))}(True) = \sup_{P_p^l \leq w} \pi_{P_l}(w)$$

$$1 - \Pi_{(RP_l(P_p))}(False) = \inf_{P_p^l > w} 1 - \pi_{P_l}(w)$$

Pour simplifier la notation, on va utiliser les notations suivantes :

$$P(RP_1(P_p)) = \sup_{P_p^1 \leq w} \pi_{P_1}(w)$$

$$N(RP_1(P_p)) = \inf_{P_p^1 > w} 1 - \pi_{P_1}(w)$$

RèglePoint2 : Soit « RP₂ » une règle spécifiée à l'aide de la relation binaire « ≤ » et un point de temps fixe le plus grand possible de l'information temporelle « P_u ». Cette règle signifie que la limite de fin du point de temps possible « P_pⁿ » doit être supérieure ou égale à « P_u » :

$$RP_2 (\leq P_p^n, P_u) : P_p^n \leq P_u$$

Le degré de possibilité $P(RP_2(P_p))$ et le degré de nécessité $N(RP_2(P_p))$ de P_p satisfaisant RP_2 peuvent être calculés comme suit :

$$P(RP_2(P_p)) = \sup_{P_p^n \geq w} \pi_{P_u}(w)$$

$$N(RP_2(P_p)) = \inf_{P_p^n < w} 1 - \pi_{P_u}(w)$$

RèglePoint3 : Pour calculer les degrés de possibilité et de nécessité de P_p au moyen des règles précédentes, nous en proposons une troisième « RP₃ » qui combine la première et la deuxième règle :

$$RP_3 = RP_1 \wedge RP_2$$

Les degrés de possibilité et de nécessité dans ce cas peuvent être calculés en utilisant la fonction « min » :

$$P(RP_3(P_p)) = \min(P(RP_1(P_p)), P(RP_2(P_p)))$$

$$N(RP_3(P_p)) = \min(N(RP_1(P_p)), N(RP_2(P_p)))$$

RèglePoint4 : L'incertitude de l'ensemble de l'information signifie que tous les points de temps possibles « P_p » représentant le point de temps imprécis, sont incertains. Ainsi, nous proposons une autre règle « RP₄ » pour estimer les degrés de possibilité et de nécessité liés à l'incertitude de l'ensemble de l'information temporelle. P et N sont déduits des résultats obtenus de « RP₃ » liés à l'imprécision.

$$P(RP_4(P_p)) = P(RP_3(P_1)) \cup \dots \cup P(RP_3(P_n))$$

$$N(RP_4(P_p)) = N(RP_3(P_1)) \cap \dots \cap N(RP_3(P_n))$$

Cela signifie que :

$$P(RP_4(P_p)) = \max(P(RP_3(P_1)), \dots, P(RP_3(P_n))) \quad (34)$$

$$N(RP_4(P_p)) = \min(N(RP_3(P_1)), \dots, N(RP_3(P_n))) \quad (35)$$

7.3. Représentation des données temporelles incertaines et imprécises dans OWL 2

Pour représenter des données temporelles quantitatives qui sont à la fois incertaines et imprécises et les relations temporelles qualitatives associées, nous définissons un ensemble de composants ontologiques, comme le montre la Figure 22.

Pour représenter une information, nous définissons deux instances « ClassInst1 » et « ClassInst2 » liées via une propriété d'objet « Relation ». Pour représenter la partie temporelle, nous introduisons une classe nommée « TimeFrame » liée à la fois à « ClassInst1 » et « ClassInst2 » via une propriété d'objet nommée « HasTimeFrame ».

7.3.1 Données temporelles quantitatives

Nous proposons une représentation basée sur une ontologie pour décrire : (1) des intervalles de temps qui sont incertains et imprécis et (2) des points de temps incertains et imprécis.

7.3.1.1 Intervalles de temps

Soit $I = [I^-, I^+]$ un intervalle de temps à la fois incertain et imprécis ; où $I^- = [I^{-(I)}, I^{-(N)}]$ et $I^+ = [I^{+(I)}, I^{+(E)}]$ associés aux degrés de possibilité et de nécessité liés à l'imprécision et aux degrés de possibilité et de nécessité liés à l'incertitude.

Nous introduisons une classe « Uncert_Imp_Time_Interval » pour représenter I. C'est une sous-classe de la classe « TimeFrame ». Nous introduisons deux classes nommées « Lower_Bound » et « Upper_Bound » pour représenter « I^- » et « I^+ ». Nous définissons deux propriétés d'objet « Has_Lower_Bound » et « Has_Upper_Bound » pour relier respectivement

la classe « Uncert_Imp_Time_Interval » aux classes « Lower_Bound » et « Upper_Bound ». Nous définissons quatre propriétés précises de type de données « Starts_At », « Starts_To », « Ends_At » et « Ends_To », où « Starts_At » a comme co-domaine « $I^{-(1)}$ », « Starts_To » a comme co-domaine « $I^{-(N)}$ ». « Ends_At » a comme co-domaine « $I^{+(1)}$ » et « Ends_To » a comme co-domaine « $I^{+(N)}$ ».

Nous définissons deux propriétés précises de type de données « Pos_Imp_Degree » et « Nec_Imp_Degree » liées à « Uncert_Imp_Time_Interval » pour représenter respectivement le degré de possibilité et de nécessité associé à I (c'est-à-dire l'intervalle le plus possible I_i , qui a le plus grand degré de possibilité). « P_{Im} » est le co-domaine de « Pos_Imp_Degree » et « N_{Im} » est le co-domaine de « Nec_Imp_Degree ». Nous introduisons deux propriétés de type de données précises « Poss_Uncert_Degree » et « Nec_Uncert_Degree » pour représenter à la fois les degrés de possibilité et de nécessité liés à l'incertitude. Elles permettent de relier « Uncert_Imp_Time_Interval » avec, respectivement, « P_{Un} » et « N_{Un} » associés à l'intervalle de temps I.

7.3.1.2 Points de temps

Soit $P = [P^{(1)}, P^{(n)}]$ un point de temps à la fois incertain et imprécis associé à des degrés de possibilité et de nécessité liés à l'imprécision et des degrés de possibilité et de nécessité liés à l'incertitude.

Nous introduisons une classe « Uncert_Imp_Time_Point » pour représenter P. C'est une sous-classe de la classe « TimeFrame ». Nous ajoutons deux propriétés de type de données « Point_Begins » et « Point_Ends » ; où « Point_Begins » a comme co-domaine « $P^{(1)}$ » et « Point_Ends » a comme co-domaine « $P^{(n)}$ ».

Nous définissons deux propriétés de type de données nommées « PPos_Imp_Degree » et « PNec_Imp_Degree » liées à « Uncert_Imp_Time_Point » pour représenter respectivement la possibilité et degré de nécessité associé à P (c'est-à-dire le point de temps le plus possible « P_p », qui a le plus grand degré de possibilité). « PPos_Imp_Degree » a comme co-domaine « PP_{Im} » et « PNec_Imp_Degree » a comme co-domaine « PN_{Im} ». Nous introduisons également deux propriétés de type de données nommées « PPos_Uncert_Degree » et « PNec_Uncert_Degree » pour représenter à la fois les degrés de possibilité et de nécessité liés à l'incertitude. Il associe la classe « Uncert_Imp_Time_Point » avec, respectivement, « PP_{Un} » et « PN_{Un} » associés à l'instant P.

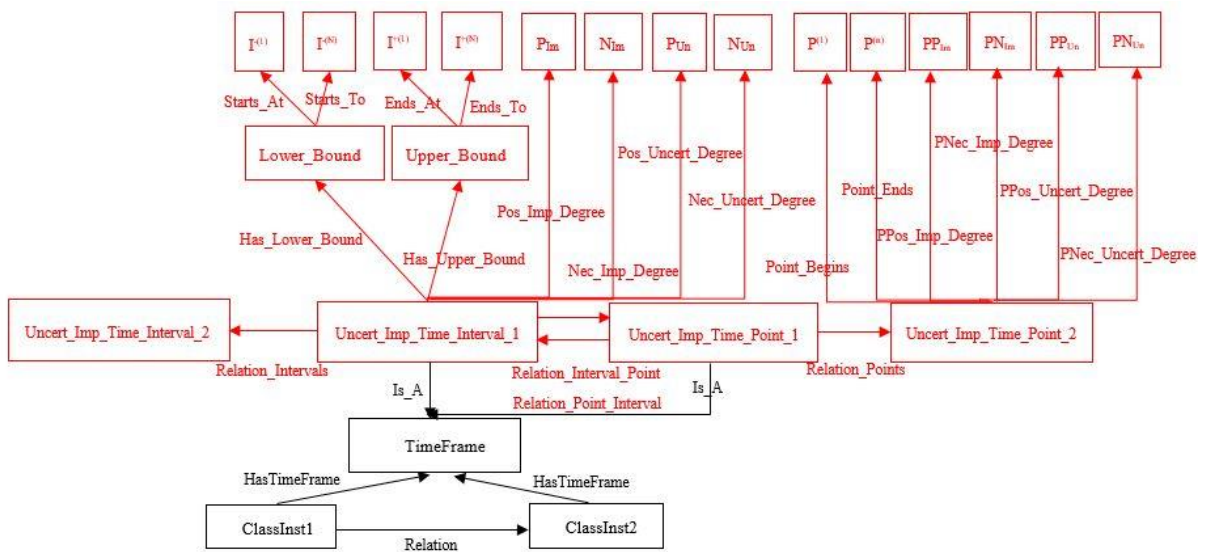


Figure 22. Représentation sémantique des données à la fois certaines et imprécises

7.3.2 Données temporelles qualitatives

Quatre relations temporelles peuvent exister entre des points temporels et des intervalles de temps : les relations Point-Point, Intervalle-Point, Point-Intervalle et Intervalle-Intervalle. Pour représenter les relations qualitatives entre deux points temporels à la fois incertains et imprécis, nous introduisons la propriété d'objet « Relation_Points » qui relie deux instances de « Uncert_Imp_Time_Point ». Pour représenter les relations qualitatives entre un intervalle de temps à la fois incertain et imprécis et un point de temps à la fois incertain et imprécis, nous introduisons la propriété d'objet « Relation_Interval_Point » qui relie une instance de la classe « Uncert_Imp_Time_Interval » et une instance de la classe « Uncert_Imp_Time_Point ». Pour représenter les relations qualitatives entre un point de temps incertain et imprécis et un intervalle de temps incertain et imprécis, nous introduisons une propriété d'objet nommée « Relation_Point_Interval » qui relie une instance de la classe « Uncert_Imp_Time_Point » et une

instance de la classe « Uncert_Imp_Time_Interval ». Pour représenter les relations qualitatives entre deux intervalles de temps incertains et imprécis, nous introduisons la propriété d'objet « Relation_Intervals » qui relie deux instances de la classe « Uncert_Imp_Time_Interval ».

7.4. Raisonnement sur des données temporelles incertaines et imprécises dans OWL 2

Nous étendons l'algèbre d'Allen pour : (1) raisonner sur des données temporelles quantitatives qui sont à la fois incertaines et imprécises pour déduire des relations temporelles qualitatives et (2) raisonner sur les relations temporelles qualitatives pour en déduire de nouvelles.

7.4.1 Relations temporelles qualitatives

Nous proposons des relations temporelles qualitatives entre des intervalles de temps incertains et imprécis simultanément. Nous prouvons que les relations résultantes préservent de nombreuses propriétés de l'algèbre d'Allen. Ensuite, nous adaptons ces relations pour relier un intervalle de temps et un point de temps ou deux points de temps.

7.4.1.1 Relations entre deux intervalles de temps

Lorsque l'on considère des intervalles de temps précis, notre approche se réduit à l'algèbre d'intervalle d'Allen. Nous redéfinissons les relations d'Allen pour raisonner sur des intervalles de temps qui sont conjointement incertains et imprécis sur la base des règles définies et des fonctions booléennes correspondantes ($<$, $>$ et $=$).

Par exemple, soit $I = [I^-, I^+]$ et $J = [J^-, J^+]$ deux intervalles de temps incertains et imprécis, où $I^- = [I^{-(1)}, I^{-(N)}]$, $I^+ = [I^{+(1)}, I^{+(N)}]$, $J^- = [J^{-(1)}, J^{-(E)}]$, $J^+ = [J^{+(1)}, J^{+(E)}]$. On redéfinit « I Before J » comme suit :

$$I \text{ Before } J \rightarrow RI(<, I^{+(N)}, J^{(1)})$$

Cela signifie que « I Before J » est vérifié si « $RI(<, I^{+(N)}, J^{(1)})$ » est satisfait. Cela signifie que la relation booléenne « $<$ » doit être respectée entre la fin de la borne inférieure $I^{+(N)}$ de I et la borne supérieure initiale $J^{(1)}$ de J. De même, nous redéfinissons les autres relations d'Allen comme indiqué dans le Tableau 27.

Tableau 27. Relations temporelles entre les intervalles de temps à la fois incertains et imprécis I et J

Relation(A, B)	Définition	Inverse(B, A)
Before(I, J)	$RI(<, I^{+(N)}, J^{(I)})$	After(J, I)
Meets (I, J)	$RI(=, I^{(I)}, J^{(I)}) \wedge RI(=, I^{+(N)}, J^{+(N)})$	Met-by(J, I)
Overlaps(I, J)	$RI(<, I^{(N)}, J^{(I)}) \wedge RI(<, J^{(N)}, I^{(I)}) \wedge RI(<, I^{+(N)}, J^{+(I)})$	Overlapped-by(J, I)
Starts(I, J)	$RI(=, I^{(I)}, J^{(I)}) \wedge RI(=, I^{(N)}, J^{(N)}) \wedge RI(<, I^{+(N)}, J^{+(I)})$	Started-by(J, I)
During(I, J)	$RI(<, J^{(N)}, I^{(I)}) \wedge RI(<, I^{+(N)}, J^{+(I)})$	Contains(J, I)
Ends(I, J)	$RI(<, J^{(N)}, I^{(I)}) \wedge RI(=, I^{(I)}, J^{+(I)}) \wedge RI(=, I^{+(N)}, J^{+(N)})$	Ended-by(J, I)
Equals(I, J)	$RI(=, I^{(I)}, J^{(I)}) \wedge RI(=, I^{(N)}, J^{(N)}) \wedge RI(=, I^{+(I)}, J^{+(I)}) \wedge RI(=, I^{+(N)}, J^{+(N)})$	Equals(J, I)

Nos relations temporelles préservent de nombreuses propriétés de l'algèbre d'Allen. Soit $I = [I^-, I^+]$, $J = [J^-, J^+]$ et $K = [K^-, K^+]$ être à la fois des intervalles de temps incertains et imprécis.

- **Réflexivité/ irréflexivité.** Les relations temporelles {Before, After, Meets, Met-by, Overlaps, Overlapped-by, Starts, Started-by, During, Contains, Ends et Endedby} sont irréflexives, c'est-à-dire soit « TR » une de ces relations. Il tient que : $TR(I, I) = 0$. Par exemple, $Before(I, I) = RI(<, I^{+(N)}, I^{(I)}) = 0$ comme $I^{+(N)} > I^{(I)}$. De plus, la relation « Equals » est réflexive. Il tient que :

$$Equals(I, I) = RI(=, I^{(I)}, I^{(I)}) \wedge RI(=, I^{(N)}, I^{(N)}) \wedge RI(=, I^{+(N)}, I^{+(N)}) = 1.$$

- **Symétrie/asymétrie.** Les relations temporelles {Before, After, Meets, Met-by, Overlaps, Overlapped-by, Starts, Started-by, During, Contains, Ends et Endedby} sont asymétriques, c'est-à-dire soit « TR » une de ces relations. Il tient que : $TR(I, J)$ et $TR(J, I) \rightarrow I = J$. Par exemple, nous pouvons déduire de « $Before(I, J)$ » et « $Before(J, I)$ » que « $I = J$ » tient. En effet « $RI(<, I^{+(N)}, J^{(I)})$ » est déduit à partir de « $Before(I, J)$ » et « $RI(<, J^{+(N)}, I^{(I)})$ » est déduit à partir de « $Before(J, I)$ ». Avec « $(I^{+(N)} < J^{(I)})$ » et « $(J^{+(N)} < I^{(I)})$ » nous concluons que $I = J$. De plus, la relation « Equals » est symétrique. Il tient que : $Equals(I, J) = Equals(J, I)$.

- **Transitivité.** Les relations temporelles {Before, After, Meets, Met-by, Overlaps, Overlapped-by, Starts, Started-by, During, Contains, Ends, Endedby et Equals} sont transitives, c'est-à-dire soit « TR » une des relations. Il tient que :

$$TR(I, J) \text{ et } TR(J, K) \rightarrow TR(I, K)$$

Par exemple, on peut déduire de « $Before(I, J)$ » et « $Before(J, K)$ » que « $Before(I, K)$ » est vrai. En effet « $RI(<, I^{+(N)}, J^{(I)})$ » est déduit de « $Before(I, J)$ » et « $RI(<, J^{+(N)}, K^{(I)})$ » est déduit de « $Before(J, K)$ ». Par conséquent « $RI(<, I^{+(N)}, K^{(I)})$ » est vérifié. Cela signifie que « $Before(I, K)$ » tient.

7.4.1.2 Relations entre un intervalle de temps et un point de temps

Nous adaptons les relations temporelles qualitatives entre les intervalles de temps pour proposer des relations entre un intervalle de temps et un point de temps, c'est-à-dire des relations Intervalle-Point et Point-Intervalle, comme le montre le Tableau 28.

Tableau 28. Relations temporelles entre un intervalle de temps I conjointement incertain et imprécis et un point de temps P conjointement incertain et imprécis

Relation(I, P)	Définition	Inverse(P, I)
Before(I, P)	$R (<, I^{+(N)}, P^{(n)})$	After (P, I)
After(I, P)	$R (>, I^{(I)}, P^{(n)})$	Before(P, I)
Starts(I, P)	$R (=, I^{(I)}, P^{(I)})$	Started-by(P, I)
During (I, P)	$R (<, I^{(I)}, P^{(I)}) \wedge R (<, P^{(n)}, I^{+(N)})$	Contains(P, I)
Ends(I, P)	$R (=, I^{(I)}, P^{(I)}) \wedge R (=, I^{+(N)}, P^{(n)})$	Ended-by(P, I)

7.4.1.3 Relations entre deux points de temps

Les relations temporelles qualitatives entre intervalles de temps sont adaptées pour les relations entre points de temps, comme le montre le Tableau 29.

Par exemple, soit P et Q deux points de temps qui sont incertains et imprécis ; où $P = [09 : 45, 10 : 15]$ et $Q = [12 : 00, 12 : 30]$. Nous concluons que $Q \text{ After } P = 1$. En effet, $Q \text{ After } P = RP (<, P^{(n)}, Q^{(I)})$.

Tableau 29. Relations temporelles entre les points temporels conjointement incertains et imprécis P et Q

Relation(I, P)	Définition	Inverse(P, I)
Before(I, P)	$R(<, I^{+(N)}, P^{(n)})$	After(P, I)
After(I, P)	$R(>, I^{(l)}, P^{(n)})$	Before(P, I)
Starts(I, P)	$R(=, I^{(l)}, P^{(l)})$	Started-by(P, I)
During(I, P)	$R(<, I^{(l)}, P^{(l)}) \wedge R(<, P^{(n)}, I^{+(N)})$	Contains(P, I)
Ends(I, P)	$R(=, I^{+(l)}, P^{(l)}) \wedge R(=, I^{+(N)}, P^{(n)})$	Ended-by(P, I)

7.4.2 Table de transitivité

Nous introduisons des tables de transitivité pour dériver de nouvelles données temporelles à partir des relations temporelles qualitatives entre des intervalles de temps et des points qui sont conjointement incertains et imprécis.

Relations temporelles qualitatives entre les intervalles de temps. L'algèbre des intervalles d'Allen nous permet de déduire la relation temporelle $TR_3(I, K)$ à partir de $TR_1(I, J)$ et $TR_2(J, K)$, où $I = [I^-, I^+]$, $J = [J^-, J^+]$ et $K = [K^-, K^+]$ sont des intervalles de temps précis et TR_1 , TR_2 et TR_3 sont des relations d'intervalle d'Allen. Par exemple, en utilisant les définitions originales d'Allen, nous pouvons déduire de « During (I, J) » et « Meet (J, K) » que « Before(I, K) » est vrai. En effet, de « During (I, J) » nous avons « $(I^- > J^-)$ et $(I^+ < J^+)$ », et de « Meet (J, K) » nous avons « $(J^+ = K^-)$ ». De $(I^+ < J^+)$ et $(J^+ = K^-)$, nous déduisons Before(I, K).

Nous généralisons ces déductions en utilisant les trois intervalles de temps qui sont à la fois incertains et imprécis $I = [I^-, I^+]$, $J = [J^-, J^+]$ et $K = [K^-, K^+]$. En nous basant sur le Tableau 25, nous pouvons déduire de During (I, J) et de Meet (J, K) que Before(I, K) est vrai. En effet de During (I, J) nous avons $RI(<, J^{-(N)}, I^{-(l)}) \wedge RI(<, I^{+(N)}, J^{+(l)})$, et de Meet (J, K) nous avons $RI(=, J^{+(l)}, K^{-(l)})$ $RI(=, J^{+(N)}, K^{-(N)})$. De $RI(<, I^{+(N)}, J^{+(l)})$ et $RI(=, J^{+(l)}, K^{-(l)})$ nous déduisons Before(I, K). Notre table de transitivité coïncide avec celle d'Allen.

Relations temporelles qualitatives entre un intervalle de temps et un point dans le temps.

Nous introduisons deux tables de transitivité pour raisonner sur les relations temporelles qualitatives entre un intervalle de temps et un point de temps.

Soit I et J des intervalles de temps et P et L des points de temps. Ils peuvent être tous précis ou tous incertains et imprécis. Le Tableau 30 montre la table de transitivité qui permet d'inférer

$TR_3(P, L)$ à partir de $TR_1(P, I)$ et $TR_2(I, L)$. TR_1 est une relation Point-Intervalle ; TR_2 est une relation Intervalle-Point et TR_3 est une relation Point-Point.

Certaines relations de composition sont « indécidables ». Par exemple, les relations « After(P, I) » et « Before(I, L) » donnent trois relations possibles : « Before(P, L) », « Equals(P, L) » et « After(P, L) ».

Tableau 30. Table de transitivité permettant de raisonner $TR_3(P, L)$ à partir de $TR_1(P, I)$ et $TR_2(I, L)$

TR₁(P, I) TR₂(I, L)	Before(P, I)	After(P, I)	Starts(P, I)	During(P, I)	Ends(P, I)
Before(I, L)	Before(P, L)	Une des relations temporelles	Before(P, L)	Before(P, L)	Before(P, L)
After(I, L)	Une des relations temporelles	After(P, L)	After(P, L)	After(P, L)	After(P, L)
Started-by(I, L)	Before(P, L)	After(P, L)	Equals(P, L)	After(P, L)	After(P, L)
Contains(I, L)	Before(P, L)	After(P, L)	Before(P, L)	Une des relations temporelles	After(P, L)
Ended-by(I, L)	Before(P, L)	After(P, L)	Before(P, L)	Before(P, L)	Equals(P, L)

Le Tableau 31 permet d'inférer $TR_3(I, J)$ à partir de $TR_1(I, P)$ et $TR_2(P, J)$. TR_1 est une relation Intervalle-Point, TR_2 est une relation point-intervalle et TR_3 est une relation intervalle-intervalle.

Tableau 31. Table de transitivité permettant de raisonner $TR_3(P, L)$ à partir de $TR_1(P, I)$ et $TR_2(I, L)$

TR₁(I,P) TR₂(P, J)	Before(I, P)	After(I, P)	Started-by(I, P)	Contains(I, P)	Ended-by(I, P)
Before(P, J)	Before(I, J)	Une des relations d'intervalle de temps	Contains(I, J) ∨ Ended-by(I, J) ∨ Met-by(I, J) ∨ Before(I, J) ∨ Overlaps(I, J)	Contains(I, J) ∨ Ended-by(I, J) ∨ Overlapped-by(I, J) ∨ Meets(I, J) ∨ Before(I, J)	Before(I, J)

After(P, J)	Une des relations d'intervalle de temps	After(I, J)	After(I, J)	After(I, J) ∨ Met-by(I, J) ∨ During(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J)	During(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J) ∨ Met-by(I, J) ∨ After(I, J)
Starts(P, J)	Before(I, J)	After(I, J) ∨ Met-by(I, J) ∨ Overlapped-by(I, J) ∨ Ended-by(I, J) ∨ Contains(I, J)	Started-by(I, J) ∨ Equals(I, J) ∨ Starts(I, J)	Contains(I, J) ∨ Ended-by(I, J) ∨ Overlaps(I, J)	Meets(I, J)
During(P, J)	During(I, J) ∨ Starts(I, J) ∨ Overlaps(I, J) ∨ Meets(I, J) ∨ Before(I, J)	After(I, J) ∨ Met-by(I, J) ∨ Overlapped-by(I, J) ∨ Ended-by(I, J) ∨ Contains(I, J)	Contains(I, J) ∨ Ends(I, J) ∨ Overlapped-by(I, J)	Overlaps(I, J) ∨ Overlapped-by(I, J) ∨ Starts(I, J) ∨ Started-by(I, J) ∨ During(I, J) ∨ Contains(I, J) ∨ Ends(I, J) ∨ Ended-by(I, J) ∨ Equals(I, J)	Overlapped-by(I, J) ∨ Starts(I, J) ∨ During(I, J)
Ends(P, J)	During(I, J) ∨ Starts(I, J) ∨ Overlaps(I, J) ∨ Meets(I, J) ∨ Before(I, J)	After(I, J)	Met-by(I, J)	Contains(I, J) ∨ Started-by(I, J) ∨ Overlapped-by(I, J)	Equals(I, J) ∨ Ended-by(I, J) ∨ During(I, J)

Relations temporelles qualitatives entre les points de temps. Nous introduisons une table de transitivité (Tableau 32) pour raisonner sur les relations temporelles qualitatives entre les points temporels. Cette table permet d'inférer $TR_3(P, M)$ à partir de $TR_1(P, L)$ et $TR_2(L, M)$, où P, L et M sont des points de temps précis ou à la fois incertains et imprécis. TR_1 , TR_2 et TR_3 sont des relations Point-Point.

Tableau 32. Table de transitivité qui permet de raisonner $TR_3(P, M)$ à partir de $TR_1(P, L)$ et $TR_2(L, M)$.

TR₁(P, L)	Before(P, L)	Equals(P, L)	After(P, L)
TR₂(L, M)			
Before(L, M)	Before(P, M)	Before(P, M)	Une des relations temporelles

Equals(L, M)	Before(P, M)	Equals(P, M)	After(P, M)
After(L, M)	Une des relations temporelles	After(P, M)	After(P, M)

Exemple : Soit P, Q et L les trois points temporels incertains. De « Before(P, Q) = 1 », on déduit « After(Q, P) = 1 ». En nous basant sur la relation de transitivité qui permet de déduire « After(Q, L) » de « After(Q, P) » et « After(P, L) », nous déduisons que : After(Q, L) = 1.

7.5. « PossibilisticTemporalOnto » : notre ontologie temporelle possibiliste pour les données temporelles à la fois incertaines et imprécises

Nous proposons une ontologie temporelle possibiliste appelée « *PossibilisticTemporalOnto* »⁷ (Figure 23) qui implémente notre approche de représentation sémantique qui est instanciée sur la base de notre extension des relations d'Allen.

Notre ontologie temporelle possibiliste est une ontologie classique contenant de nouveaux composants représentant des mesures de possibilité et de nécessité de données temporelles qui sont à la fois incertaines et imprécises. Nous l'avons créée avec l'éditeur d'ontologies Protégé. Elle contient 6 classes, 68 propriétés d'objet et 15 propriétés de données.

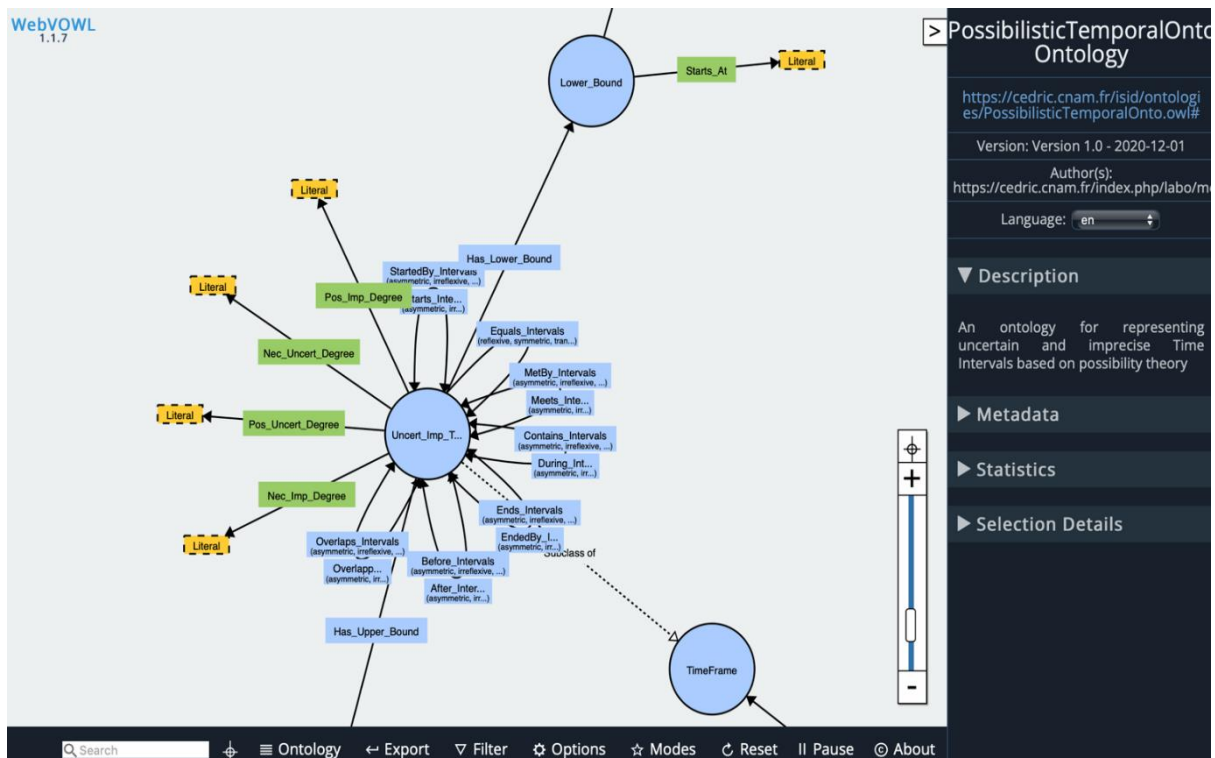


Figure 23. Notre ontologie des données temporelles à la fois incertaines et imprécises "PossibilisticTemporalOnto"

⁷ <https://cedric.cnam.fr/isid/ontologies/PossibilisticTemporalOnto.owl#>

Nous avons inféré via un ensemble de règles SWRL et à l'aide du raisonneur Pellet notre extension de l'algèbre d'Allen.

7.6. Expérimentations

Nous avons expérimenté notre approche en implémentant un prototype basé sur l'ontologie « *PossibilisticTemporalOnto* ». Nous avons mené une étude de cas dans le cadre de la prothèse CAPTAIN MEMO.

7.6.1 Prototype implémenté

Notre prototype, implémenté en JAVA, propose des interfaces utilisateur pour explorer notre approche (cf. Figure 24). Il permet aux utilisateurs de saisir et d'enregistrer tous les intervalles de temps possibles et de calculer les degrés de possibilité et de nécessité liés à l'imprécision et à l'incertitude pour prendre des décisions suivant les mesures obtenues. Des relations d'intervalle temporelles qualitatives sont exécutées sur la base des règles SWRL proposées.

The screenshot shows a software interface titled "A Possibility Theory-Based Approach to Deal with Uncertain and Imprecise Time Intervals in OWL 2". The interface is divided into several sections:

- Add all the possible time intervals:** Contains input fields for "Name of the Time Interval" (I1), "Starts at" (1961), "Starts to" (1967), "Ends at" (1969), and "Ends to" (1974). There are buttons for "Add another time interval" and "Save".
- Saved Time Intervals:** A table listing saved intervals with their start and end ranges.

Interval	Start Range	End Range
I1	[1961, 1967]	[1969, 1974]
I2	[1965, 1972]	[1975, 1978]
I3	[1963, 1969]	[1973, 1979]
I4	[1966, 1968]	[1969, 1973]
I5	[1969, 1973]	[1975, 1977]
- Saved Possibility and Necessity Degrees:** A table showing the calculated possibility (P) and necessity (N) degrees for each interval.

Interval	P	N
I1	0,71	0,41
I2	0,55	0,58
I3	0,16	0,37
I4	0,21	0,35
I5	0,13	0,39
- Add Qualitative temporal interval relation:** Contains input fields for "Interval 1" (I), "Qualitative Temporal Relation" (Before), and "Interval 2" (J). There is a "Save" button.

At the bottom of the interface, there are buttons for "Generate OWL File" and "Visualize".

Figure 24. Prototype implémenté basé sur notre ontologie

7.6.2 Étude de cas : Application aux patients atteints d'Alzheimer

Nous appliquons dans le cadre de CAPTAIN MEMO pour manipuler des données temporelles à la fois incertaines et imprécises.

Par exemple, si nous avons l'information " *Je pense que Françoise vivait en France entre les années 60 et 70. Ensuite, elle a déménagé en Allemagne peut-être depuis environ 5 ans avant de partir au Canada* ". Soit $[L_1, U_1] = [60, 70]$, cela signifie que : $[L_1, U_1] = [1960, 1979]$. Soit $I_i = [I_i^-, I_i^+]$ les différents intervalles de temps possibles dans $[L_1, U_1]$, représentant la durée de vie de Françoise en France, où $I_i^- = [I_i^{-(1)}, I_i^{-(N)}]$ et $I_i^+ = [I_i^{+(1)}, I_i^{+(N)}]$ représenté par le Tableau 33.

Tableau 33. Les différents Intervalles I_i possibles

I_i	$I_i^- = [I_i^{-(1)}, I_i^{-(N)}]$	$I_i^+ = [I_i^{+(1)}, I_i^{+(N)}]$
I_1	$I_1^- = [1961, 1967]$	$I_1^+ = [1969, 1974]$
I_2	$I_2^- = [1965, 1972]$	$I_2^+ = [1975, 1978]$
I_3	$I_3^- = [1963, 1969]$	$I_3^+ = [1973, 1979]$
I_4	$I_4^- = [1966, 1968]$	$I_4^+ = [1969, 1973]$
I_5	$I_5^- = [1969, 1973]$	$I_5^+ = [1975, 1977]$

Soit $[L_2, U_2] = [2013, 2020]$. Soit $J_j = [J_j^-, J_j^+]$ les différents intervalles possibles J dans $[L_2, U_2]$, représentant le déménagement de Françoise en Allemagne, où $J_j^- = [J_j^{-(1)}, J_j^{-(E)}]$ et $J_j^+ = [J_j^{+(1)}, J_j^{+(E)}]$ représenté par le Tableau 34.

Tableau 34. Les différents Intervalles J_j possibles

J_j	$J_j^- = [J_j^{-(1)}, J_j^{-(E)}]$	$J_j^+ = [J_j^{+(1)}, J_j^{+(E)}]$
J_1	$J_1^- = [2013, 2016]$	$J_1^+ = [2016, 2018]$
J_2	$J_2^- = [2015, 2017]$	$J_2^+ = [2018, 2020]$
J_3	$J_3^- = [2014, 2015]$	$J_3^+ = [2017, 2019]$

Mesures de possibilité et de nécessité. Nous prenons « I_1 » comme exemple, et nous calculons les différentes mesures de possibilité de I_1 liées à l'imprécision, sur la base des règles définies (c'est-à-dire R_1, R_2, R_3 et R_4), comme le montre la Figure 25.

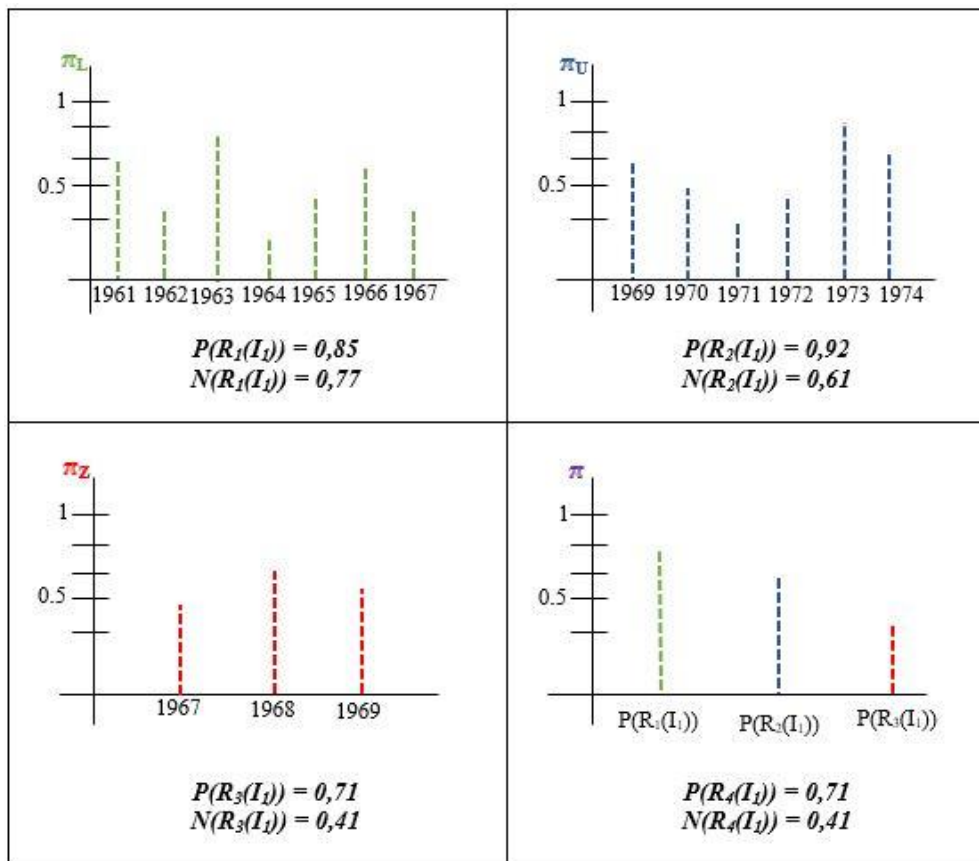


Figure 25. Mesures de possibilité et de nécessité de I_1 basées sur les règles définies

De même, nous calculons les degrés de possibilité et de nécessité liés à l'imprécision de tous les intervalles possibles I_i (c.-à-d., I_2, I_3, I_4 et I_5) ainsi que tous les intervalles J_j (c.-à-d., J_1, J_2 et J_3). Ensuite, nous calculons, à partir des mesures obtenues, les degrés de possibilité et de nécessité liés à l'incertitude sur la base des équations introduites dans la règle 5 (R_5). Les Tableaux 35 et 36 présentent les résultats.

Tableau 35. Mesures de possibilité et de nécessité de l'intervalle de temps conjointement incertain et imprécis I

I_i	$P(R_i)$				$N(R_i)$			
	R_1	R_2	R_3	R_4	R_1	R_2	R_3	R_4
I_1	0.85	0.92	0.71	0.71	0.77	0.61	0.41	0.41
I_2	0.92	0.62	0.55	0.55	0.58	0.66	0.76	0.58
I_3	0.16	0.27	0.26	0.16	0.42	0.37	0.44	0.37
I_4	0.5	0.41	0.21	0.21	0.51	0.35	0.38	0.35
I_5	0.13	0.3	0.4	0.13	0.38	0.4	0.6	0.38
$P(R_5) = 0.71$					$N(R_5) = 0.35$			

Tableau 36. Mesures de possibilité et de nécessité de l'intervalle de temps conjointement incertain et imprécis J

J _j	P(R _i)				N(R _i)			
	R ₁	R ₂	R ₃	R ₄	R ₁	R ₂	R ₃	R ₄
J ₁	0.66	0.85	0.79	0.66	0.58	0.29	0.25	0.25
J ₂	0.62	0.95	0.73	0.62	0.59	0.26	0.59	0.26
J ₃	0.91	0.78	0.96	0.78	0.32	0.29	0.15	0.15
P(R ₅) = 0.78					N(R ₅) = 0.15			

I₁ a le degré de possibilité ($P(R_4(I_2)) = 0.71$) le plus élevé. Cela signifie que I₁ représente l'intervalle, dans l'ensemble des intervalles I_i, qui correspond le plus à l'intervalle imprécis I.

L'information relative à l'intervalle de temps imprécis I est incertaine au moins à un degré égal à « 0.35 » ($N(R_5) = 0.35$) et il est possible au plus à un degré égal à « 0.71 » ($P(R_5) = 0.71$).

J₃ a le degré de possibilité le plus élevé. Cela signifie que J₃ représente l'intervalle, dans l'ensemble des intervalles J_j, qui correspond le plus à l'intervalle imprécis J.

L'information relative à l'intervalle de temps imprécis J est incertaine au moins à un degré égal à « 0.15 » ($N(R_5) = 0.15$) et il est possible au plus à un degré égal à « 0.78 » ($P(R_5) = 0.78$).

Représentation sémantique de l'exemple dans le contexte de l'ontologie PersonLink.

Nous décrivons à l'aide des l'ontologie « *PersonLink* » et « *PossibilisticTemporalOnto* » l'exemple de la précédente section en utilisant les mesures obtenues comme le montre la Figure 26. Nous représentons les intervalles de temps I et J qui sont à la fois incertains et imprécis. Nous représentons la relation qualitative "*BeforeIntervals*" entre la durée de vie de Françoise en France, représentée par l'intervalle I, et son déménagement en Allemagne, représenté par J.

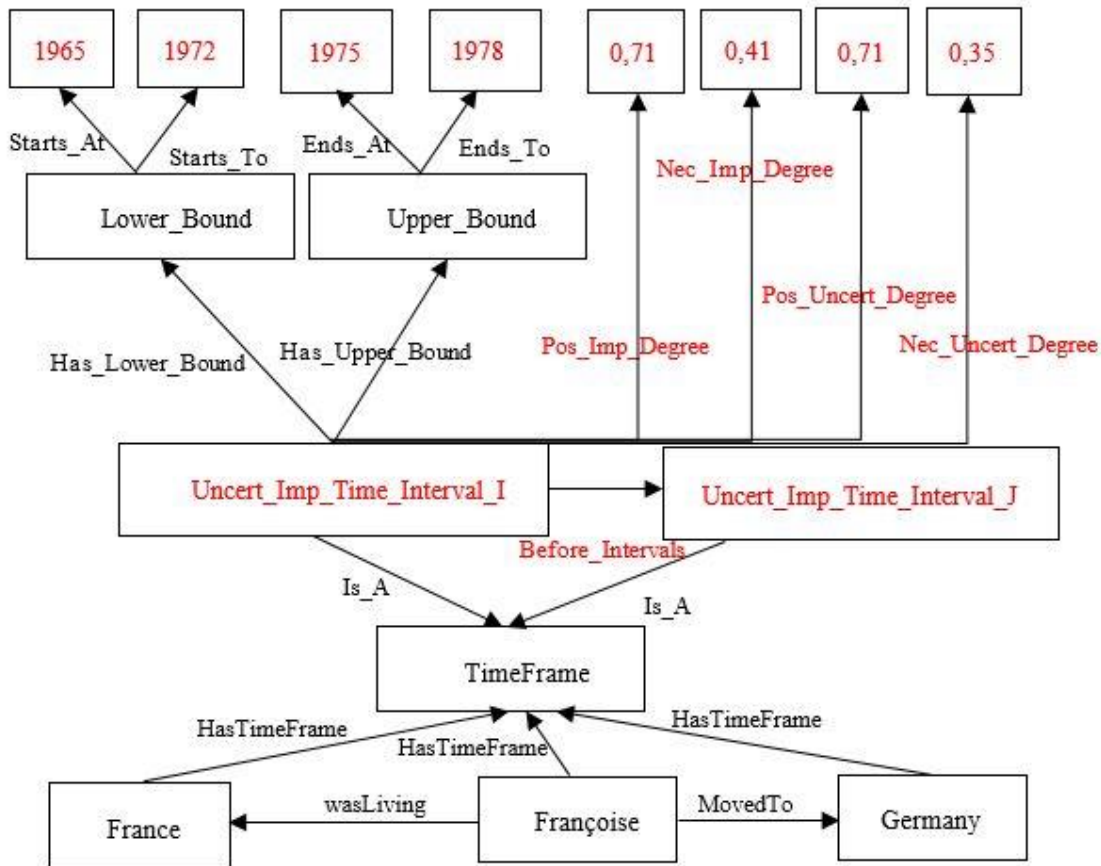


Figure 26. Représentation sémantique de l'exemple dans le contexte de l'ontologie PersonLink

7.7. Conclusion

Nous avons présenté dans ce chapitre une nouvelle approche pour traiter des intervalles de temps à la fois incertains et imprécis dans OWL 2. Cette approche est basée sur la théorie des possibilités. Notre première contribution consiste à calculer les degrés de possibilité et de nécessité à partir d'un ensemble de règles que nous avons définies. Notre deuxième contribution consiste à proposer une représentation sémantique des données temporelles manipulées à partir des mesures obtenues. Notre troisième contribution consiste à étendre l'algèbre des intervalles d'Allen pour raisonner sur les données manipulées. Nous avons proposé 13 relations temporelles entre des intervalles de temps à la fois incertains et imprécis. Notre extension préserve les propriétés de réflexivité/irréflexivité, la symétrie/asymétrie et la transitivité. Notre quatrième contribution concerne la proposition d'une ontologie temporelle possibiliste qui implémente notre représentation sémantique proposée et l'extension d'Allen. Enfin, nous avons implémenté un prototype basé sur notre ontologie. Par rapport aux approches basées sur la théorie floue et de l'incertitude qui donne à chacun intervalle possible le même degré de certitude, notre

approche à associer à chaque intervalle possible un degré de certitude différent des autres. Cela aide les décideurs à avoir plus d'exactitude des informations nécessaires.

Dans le chapitre suivant, nous proposons **notre approche basée sur la théorie de l'évidence** pour traiter **les données temporelles conflictuelles** en mesurant les masses de croyances.

CHAPITRE 8 : Traitement des données temporelles conflictuelles dans OWL 2

(Approche basée sur la théorie de l'évidence)

8.1. Introduction

Selon la typologie que nous avons proposée pour les imperfections des données temporelles, **le conflit** caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires et donc incompatibles. Il peut apparaître dans de nombreux contextes et situations, comme dans notre cas concernant la saisie des données par des patients Alzheimer. En effet, un patient Alzheimer peut répondre différemment plusieurs fois à une même question. Par exemple, lorsque nous lui demandons « *Quand avez-vous obtenu votre diplôme universitaire ?* », il peut donner trois réponses temporelles différentes comme : 1998, 1995 et 1996. Dans ce chapitre, nous traitons cette problématique qui, **à notre connaissance, est la première du genre dans le domaine du Web sémantique**. Nous nous basons sur **la théorie de l'évidence** qui représente un outil puissant pour représenter des informations imparfaites et permet de modéliser toutes les formes d'imperfection, y compris les conflits.

Nous proposons **une approche pour représenter et raisonner sur les données temporelles quantitatives conflictuelles** (c'est-à-dire les intervalles de temps et les points) **et les relations qualitatives associées** (par exemple, « before »). Pour la représentation, **nous calculons**, à l'aide des mesures de crédibilité proposées dans notre approche DBE_ALZ (Ghorbel, et al., 2020), **les masses de croyance des données temporelles conflictuelles**. Nous proposons ensuite **une représentation sémantique qui associe les mesures calculées aux données traitées et leurs relations qualitatives**. Pour le raisonnement, **notre approche repose sur l'algèbre des intervalles d'Allen**. Nous l'étendons pour raisonner sur des relations temporelles conflictuelles. Les relations d'intervalle résultantes préservent les propriétés de l'algèbre d'origine. Nous adaptons ces relations pour relier un intervalle de temps et un point de temps, et deux points de temps. Toutes les relations proposées peuvent être utilisées pour un raisonnement temporel à travers des tables de transitivité. Nous proposons « *BeliefTimeOnto* », **une ontologie évidentielle basée sur la représentation sémantique et l'extension des relations d'Allen**. Pour valider notre travail, nous implémentons **un prototype basé sur l'ontologie proposée** pour explorer notre approche et nous menons **une étude de cas appliquée à des données temporelles saisies** par des patients Alzheimer dans le cadre de la prothèse CAPTAIN MEMO.

Le reste du chapitre est organisé comme suit. La section 8.2 présente notre contribution pour calculer les masses de croyance en utilisant la théorie de l'évidence. Les sections 8.3 et 8.4 présentent nos contributions pour représenter et raisonner sur des données temporelles conflictuelles. La section 8.5 présente notre ontologie « *BeliefTimeOnto* » et la section 8.6 l'expérimentation de notre approche. Nous finissons ce chapitre par une conclusion.

8.2. Masses de croyance pour les données temporelles quantitatives

Nous adaptons la théorie de l'évidence pour estimer les degrés de masse des intervalles de temps et des points de temps conflictuels.

8.2.1 Intervalles de temps

Nous représentons un événement conflictuel « I » représenté par une masse de croyance sur un cadre de discernement Ω . I est représenté par un nombre d'intervalles de temps conflictuels où $I = \{I_1 \dots I_n\}$ et $I_{i=1..n} = [I_i^-, I_i^+]$. La fonction de masse de I est définie comme suit :

$$m : 2^\Omega \rightarrow [0, 1]$$

$$\sum_{I \subseteq \Omega} m(I) = 1$$

Par exemple, soit la question suivante : « *Quand avez-vous quitté votre emploi ?* ». Supposons que nous ayons 3 intervalles de temps conflictuels répondant à cet événement : « I₁ », « I₂ » et I₃ représentés respectivement par : I₁ = [1989, 1993], I₂ = [1992, 1995] et I₃ = [1998, 1999].

L'état de connaissance d'une personne sur I est décrit comme suit :

$$2^\Omega = \{\emptyset, m(I_1), m(I_2), m(I_3), m(I_1, I_2), m(I_2, I_3), m(I_1, I_3), m(I_1, I_2, I_3)\} = 1.$$

Cela signifie que :

$$m(\emptyset) = 0 \text{ and } m([1989, 1993], [1992, 1995], [1998, 1999], \{[1989, 1993], [1992, 1995]\}, \{[1992, 1995], [1998, 1999]\}, \{[1989, 1993], [1998, 1999]\}, \{[1989, 1993], [1992, 1995], [1998, 1999]\}) = 1.$$

Chaque borne de l'intervalle « I_i » est associée à une masse de croyance « m ». Cela signifie que pour $I_i = [I_{mi}^-, I_{mi}^+]$, m_{i-} et m_{i+} sont respectivement les masses de croyance associées à I_i^- et I_i^+ .

Pour estimer la masse de croyance de chaque borne d'un intervalle de temps conflictuel, nous calculons d'abord la fonction de crédibilité en appliquant notre approche (Ghorbel, et al., 2020) qui permet de mesurer la crédibilité l'aide de la théorie des probabilités et des logiques floues. Ensuite, nous déduisons les masses de croyance.

Notre approche, appelée DBE_ALZ, gère les éventuelles erreurs de saisies de données en estimant la crédibilité de chaque entrée de données. Elle est basée sur un modèle de crédibilité qui définit un ensemble de dimensions et de sous-dimensions permettant une estimation qualitative de la crédibilité d'une saisie de données. Ce modèle inclut :

- deux dimensions principales nommées « *raisonnement des données* » et « *fiabilité de l'auteur* » ;
- deux dimensions « *fiabilité du patient Alzheimer* » et « *fiabilité du soignant* » qui héritent de la dimension « *fiabilité de l'auteur* » ;
- les sous-dimensions « *raisonnabilité de la saisie des données vis-à-vis du bon sens* » et « *le caractère raisonnable de la saisie des données concernant les affirmations personnalisées* » qui permettent d'estimer la dimension « *le caractère raisonnable de la saisie des données* » ;
- les sous-dimensions « *stade de la maladie d'Alzheimer* » et « *état du moment* » qui permettent d'estimer la dimension « *Fiabilité du patient Alzheimer* » ;
- et les sous-dimensions « *âge* » et « *domaine de connaissances* » qui permettent d'estimer la dimension « *fiabilité des aidants* ».

La crédibilité d'une donnée est estimée sur la base de son caractère raisonnable et de la fiabilité de son auteur. En nous basant sur le modèle de crédibilité, il est possible d'estimer quantitativement la crédibilité des données saisies en définissant un ensemble de mesures associées aux dimensions et sous-dimensions proposées. Ces mesures s'appuient sur les théories des probabilités et des ensembles flous pour raisonner sur des connaissances incertaines et imprécises (réseaux bayésiens et systèmes d'inférence floue de Mamdani).

Dans la théorie de l'évidence, la masse de croyance de « I » est définie comme :

$$\text{Bel}(I) = \sum_{I_i \subseteq I, I_i \neq \emptyset} m(I_i)$$

Nous adaptons cette définition à notre cas :

$$\text{Bel}(I_i) = (\text{Bel}(I_i^-), \text{Bel}(I_i^+))$$

Par exemple, si « $\text{Bel}(I_i^-) = 0.5$ » et « $\text{Bel}(I_i^+) = 0.3$ », alors nous pouvons déduire que : « $m_{i^-} = 0.5$ » et « $m_{i^+} = 0.3$ ».

Pour calculer la masse de « I_i », nous utilisons la règle de combinaison conjonctive définie par (Shafer, 1976) comme suit :

$$m(I_i) = m_{i^-} \cap m_{i^+} (I_i) = \sum_{I^- \cap I^+ = I_i} m_{i^-}(I^-) m_{i^+}(I^+)$$

Par exemple, pour $I_1 = [1989, 1993]$, si « m_{i^-} » est égal à « 0.54 » et « m_{i^+} » est égal à « 0.78 » alors « $m(I_1)$ » est égal à « 0.42 ».

8.2.2 Points de temps

Nous représentons un événement conflictuel « P » représenté par une fonction de masse sur un cadre de discernement Ω . « P » est représenté par un certain nombre de points de temps « Pp » en conflit ; où $P = \{P_1, \dots, P_n\}$. En adaptant la théorie de l'évidence, nous aurons :

$$m : 2^\Omega \rightarrow [0, 1]$$

$$\sum_{P \subseteq \Omega} m(P) = 1$$

Par exemple, soit P l'événement exprimant la question suivante « *En quelle année avez-vous visité Paris ?* ». Supposons que nous ayons 4 points de temps conflictuels P_1, P_2, P_3 et P_4 répondant à cet événement représentés respectivement par : $P_1 = 2001, P_2 = 2003, P_3 = 2005$ et $P_4 = 2009$.

L'état de connaissance d'une personne sur P est décrit comme suit :

$$2^\Omega = \{\emptyset, m(P_1), m(P_2), m(P_3), m(P_4), m(P_1, P_2), m(P_1, P_3), m(P_1, P_4), m(P_2, P_3), m(P_2, P_4), m(P_3, P_4), m(P_1, P_2, P_3), m(P_1, P_2, P_4), m(P_1, P_3, P_4), m(P_2, P_3, P_4), m(P_1, P_2, P_3, P_4)\} = 1.$$

Cela signifie que :

$$m(\emptyset) = 0 \text{ and } m(\{2001, 2003, 2005, 2009, \{2001, 2003\}, \{2001, 2005\}, \{2001, 2009\}, \{2003, 2005\}, \{2003, 2009\}, \{2005, 2009\}, \{2001, 2003, 2005\}, \{2001, 2005, 2009\}, \{2003, 2005, 2009\}, \{2001, 2003, 2005, 2009\}\}) = 1.$$

La fonction de croyance de P est définie comme :

$$\text{Bel}(P) = \sum_{P_p \subseteq I, P_p \neq \emptyset} m(P)$$

Cela signifie que :

$$\text{Bel}(P) = (\text{Bel}(P_1), \text{Bel}(P_2), \text{Bel}(P_3), \text{Bel}(P_4))$$

Pour estimer les masses de croyance $m(P_1), m(P_2), m(P_3)$ et $m(P_4)$, nous utilisons la même approche que les intervalles de temps. Nous calculons la fonction de crédibilité de chaque point de temps en conflit. Ensuite, nous déduisons les masses de croyance.

Par exemple, si $\text{Bel}(P_1) = 0.7$, $\text{Bel}(P_2) = 0.1$, $\text{Bel}(P_3) = 0$ et $\text{Bel}(P_4) = 0$, alors nous pouvons déduire que :

$$m(P_1) = 0.7, m(P_2) = 0.1, m(P_3) = 0, m(P_4) = 0 \text{ et } m(\{m(P_1, P_2), m(P_1, P_3), m(P_1, P_4), m(P_2, P_3), m(P_2, P_4), m(P_3, P_4), m(P_1, P_2, P_3), m(P_1, P_2, P_4), m(P_1, P_3, P_4), m(P_2, P_3, P_4), m(P_1, P_2, P_3, P_4)\}) = 0.2.$$

8.3. Représentation des données temporelles conflictuelles dans OWL 2

Pour représenter les données temporelles quantitatives conflictuelles et les relations temporelles qualitatives associées, nous utilisons des composants ontologiques prédéfinis de l'ontologie « *OWL-Time* » et nous en définissons d'autres.

8.3.1 Données temporelles quantitatives

Nous proposons une représentation sémantique, illustrée par la Figure 27, pour représenter : (1) des intervalles de temps conflictuels (2) des points de temps conflictuels.

Soit $I_{mi} = [I_{mi}, I_{mi}^+]$ un intervalle de temps conflictuel associé à la masse de croyance la plus élevée « $m(I)$ ». Nous utilisons la classe prédéfinie dans l'ontologie « *OWL-Time* » nommée « *time:Interval* » qui est une sous classe de « *time:TemporalEntity* ». Nous définissons deux propriétés de type de données nommées « *BeginsAt* » et « *EndsAt* » pour représenter respectivement I et I^+ . Nous définissons deux propriétés de type de données nommées « *BeginMass* » et « *EndsMass* » pour représenter les masses de croyance associées à I et I^+ . « *BeginnMass* » et « *EndingMass* » représentent respectivement la classe « *time:Interval* » avec « m_{i-} » et « m_{i+} ».

Soit P_{mp} un point de temps conflictuel associé à la masse de croyance la plus élevée « m_p ». Nous utilisons la classe définie dans l'ontologie *OWL-Time* nommée « *time:Instant* » pour représenter « P ». Nous définissons une propriété de type de données nommée « *TimePoint* » pour représenter « P ». Nous définissons une propriété de type de données « *PointMassDegree* » pour représenter le degré de fonction de masse associé à « P » et relie la classe « *time:Instant* » à « m_p ».

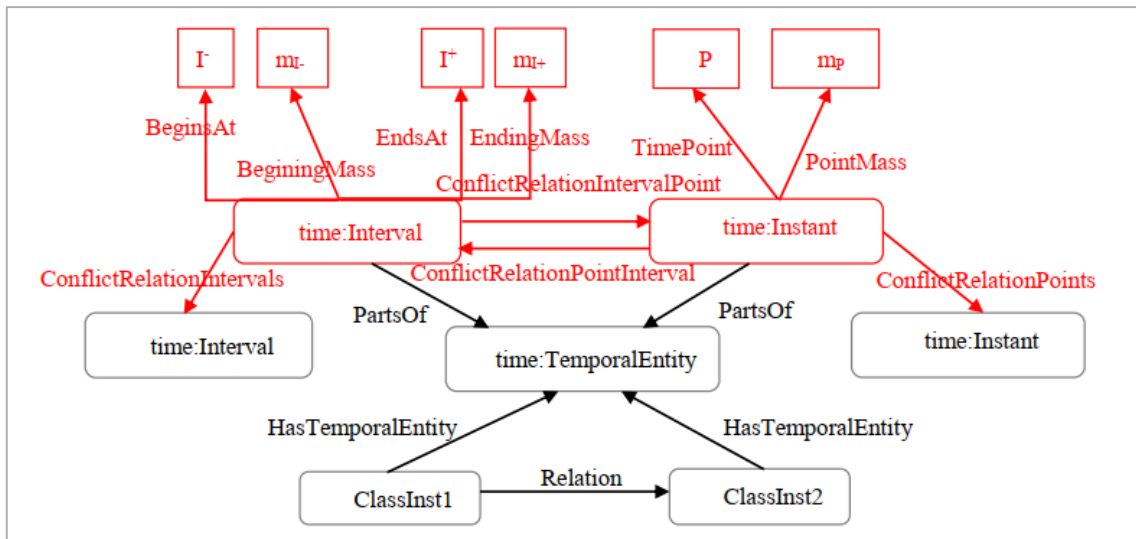


Figure 27. Représentation sémantique des données temporelles conflictuelles

8.3.2 Données temporelles qualitatives

Quatre relations temporelles peuvent exister entre des points temporels et des intervalles de temps : les relations Point-Point, Intervalle-Point, Point-Intervalle et Intervalle-Intervalle. Pour représenter les relations qualitatives entre deux points de temps en conflit, nous introduisons la propriété d'objet « *ConflictRelationPoints* » qui relie deux instances de « *time:Instant* ». Pour représenter les relations qualitatives entre un intervalle de temps en conflit et un point de temps en conflit, nous introduisons la propriété d'objet « *ConflictRelationIntervalPoint* » qui relie une instance de la classe « *time:Interval* » à une instance de la classe « *time:Instant* ». Pour représenter les relations qualitatives entre un point de temps en conflit et un intervalle de temps en conflit, nous introduisons une propriété d'objet nommée « *ConflictRelationPointInterval* » relie une instance de la classe « *time:Instant* » et une instance de la classe « *time:Interval* ». Enfin, pour représenter les relations qualitatives entre deux intervalles de temps conflictuels, nous introduisons la propriété d'objet « *ConflictRelationIntervals* » qui relie deux instances de la classe « *time:Interval* ».

8.4. Raisonnement sur des données temporelles conflictuelles dans OWL 2

Le raisonnement temporel repose sur la composition des relations de base d'Allen. Nous étendons ces relations pour : (1) raisonner sur les données temporelles quantitatives conflictuelles pour en déduire des relations temporelles qualitatives et (2) raisonner sur les relations temporelles qualitatives pour en inférer de nouvelles.

8.4.1 Relations temporelles qualitatives

Nous proposons des relations temporelles qualitatives entre des intervalles de temps conflictuels. Nous prouvons que les relations résultantes préservent de nombreuses propriétés de l'algèbre d'Allen. Ensuite, nous adaptons ces relations pour relier un intervalle de temps et un point de temps ou deux points de temps.

8.4.1.1 Relations entre intervalles de temps.

Dans le cas des intervalles de temps précis, notre approche se réduit à l'algèbre d'intervalles d'Allen. Nous redéfinissons la relation d'Allen pour raisonner sur les intervalles de temps conflictuels.

Par exemple, soit $I = [I_{mi-}, I_{mi+}^+]$ et $J = [J_{mj-}, J_{mj+}^+]$ deux intervalles de temps en conflit. Nous redéfinissons la relation « *Before*(I, J) » comme : « *Before_m*(I, J) » ; où « m » est le degré de masse associé à la relation « *Before* » entre I et J . Cela signifie que la borne « I_{mi-} » de l'intervalle I est inférieure à la borne inférieure conflictuelle « I_{mi+}^+ » de J .

$$Before_m(I, J) \rightarrow I_{mi+}^+ < J_{mj-}$$

La masse de croyance « m » est déduit des degrés de masse « m_i » et « m_{i+} » en utilisant la règle de combinaison conjonctive :

$$m = m_I \cap m_J(IJ) = \sum_{I \cap J = IJ} m(I_{mi+}^+) m(J_{mj-})$$

Par exemple, si nous avons la masse de croyance « m_{i+} » égale à « 0.45 » et la masse de croyance « m_{j+} » égal à « 0.82 ». Nous déduisons alors que « m » est égal à « 0.37 ».

Nous classons les relations proposées en deux catégories. La première couvre les relations qui consistent en une seule formule sans aucune conjonction (c'est-à-dire « *Before_m* », « *after_m* », « *Meet_m* » et « *Met-by_m* »). Ainsi, comme pour la relation « *Before_m* », nous les redéfinissons comme le montre le Tableau 36. La deuxième catégorie couvre le reste des relations. Celles-ci sont définies sur la base d'un ensemble de conjonctions (par exemple, « *Overlaps_m* »).

Nous redéfinissons la relation « *Overlaps*(I, J) » comme « *Overlaps_m*(I, J) » ; où « m » est la masse de croyance associée à cette relation. Cela signifie que la borne inférieure conflictuelle de I est inférieure à la borne de début conflictuelle de J ; la borne inférieure conflictuelle de I est supérieure à la borne inférieure conflictuelle de J ; et la borne supérieure conflictuelle de l'intervalle A est inférieure à la borne supérieure conflictuelle de B .

$$Overlaps_m(I, J) \rightarrow (I_{mi-} < J_{mj-}) \wedge (I_{mi+}^+ < J_{mj-}) \wedge (I_{mi+}^+ < J_{mj+}^+)$$

La masse de croyance « m » se déduit des degrés de masse « m₁ », « m₂ » et « m₃ » associés respectivement aux conjonctions « (I_{mi-} < J_{mj-}) », « (I⁺_{mi+} < J_{mj-}) » et « (I⁺_{mi+} < J⁺_{mj+}) », en utilisant la règle de combinaison conjonctive définie par (Shafer, 1976). La première consiste à définir la masse de croyance « m₁ », la seconde « m₂ » et la troisième « m₃ ». La dernière consiste à définir la masse de croyance « m ». Nous redéfinissons les autres relations d'Allen de la même manière que la relation « *Overlaps_m(I, J)* », comme le montre le Tableau 37.

Les relations temporelles proposées entre des intervalles de temps conflictuels présentent de nombreuses propriétés de l'algèbre d'Allen. Nous obtenons des généralisations des propriétés de réflexivité/irréflexivité, de symétrie/asymétrie et de transitivité. Par exemple, « *Equals* » est réflexif alors : « *Equals_m(I, I) = (I_{mi-} = I_{mi-}) ∩ (I⁺_{mi+} = I⁺_{mi+}) = I}}}}* ».

Tableau 37. Extension de l'algèbre des intervalles d'Allen pour les relations temporelles qualitatives

	ConflictRelationInter- vals(I, J)	ConflictRelation- PointInterval(P, I)/ ConflictRelationIn- tervalPoint(I, P)	ConflictRelation- Points(P, L)
Before _{m}} After _{m}}	$I_{mi-}^+ < J_{mj-}$	$P_{mp} < I_{mi-}$	$P_{mp} < L_{ml}$
Meets _{m}} Met-by _{m}}	$I_{mi-}^+ = J_{mj-}$	$P_{mp} = I_{mi-}^+$	
Overlaps _{m}} Overlapped-by _{m}}	$(I_{mi-} < J_{mj-}) \wedge (I_{mi-}^+ < J_{mj-}^-)$ $\wedge (I_{mi-}^+ < J_{mj-}^+)$		
Starts _{m}} Started-by _{m}}	$(I_{mi-} = J_{mj-}^-) \wedge I_{mi-}^+ < J_{mj-}^+$	$P_{mp} = I_{mi-}$	
During _{m}} Contains _{m}}	$(I_{mi-} < J_{mj-}^-) \wedge (I_{mi-}^+ < J_{mj-}^+)$	$(I_{mi-} < P_{mp}) \wedge (P_{mp} < I_{mi-}^+)$	
Ends _{m}} Ended-by _{m}}	$(I_{mi-} < J_{mj-}^-) \wedge (I_{mi-}^+ = J_{mj-}^+)$	$P_{mp} = I_{mi-}^+$	
Equals _{m}}	$(I_{mi-} = J_{mj-}^-) \wedge (I_{mi-}^+ = J_{mj-}^+)$		$P_{mp} = L_{ml}$

8.4.1.2 Relations entre un intervalle de temps et un point de temps

Nous adaptons les relations temporelles qualitatives entre les intervalles de temps pour proposer des relations entre un intervalle de temps et un point de temps, c'est-à-dire des relations Intervalle-Point et Point-Intervalle, comme le montre le Tableau 37.

8.4.1.3 Relations entre points de temps

Les relations temporelles qualitatives entre les intervalles de temps sont adaptées pour proposer des relations entre les points de temps, comme le montre le Tableau 37.

8.4.2 Tables de transitivité

La table de transitivité d'Allen nous permet d'obtenir à partir de $Rel_1(I, J)$ et $Rel_2(J, H)$ que $Rel_3(I, H)$ est vérifié, où I, J et H sont des intervalles de temps en conflit et Rel_1, Rel_2 et Rel_3 sont des relations d'Allen. Nous généralisons ces déductions en utilisant les trois intervalles de temps conflictuels $I = [I_{mi-}, I_{mi+}]$, $J = [J_{mj-}, J_{mj+}]$ et $H = [H_{mh-}, H_{mh+}]$. D'après le Tableau 35, « $Before_m(I, H)$ » peut être déduit de « $During_m(I, J)$ » et « $Meet_m(J, H)$ ». En effet, « $(J_{mj-} < I_{mi-}) \wedge (I_{mi+} < J_{mj+})$ » est déduit de « $During_m(I, J)$ » et « $(J_{mj+} = H_{mh-})$ » de « $Meet_m(J, H)$ ». De « $(I_{mi+} < J_{mj+})$ » et « $(J_{mj+} = H_{mh-})$ » nous concluons que « $(I_{mi+} < H_{mh-})$ » ce qui signifie que « $Before_m(I, H)$ » est valide. La table de transitivité proposée coïncide avec celle d'Allen.

Nous introduisons deux tables de transitivité pour raisonner sur les relations temporelles qualitatives entre un intervalle de temps et un point temporel et une table de transitivité pour raisonner sur les relations temporelles.

Le Tableau 38 montre le tableau de transitivité qui permet de déduire $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$. R_1 est une relation Point-Intervalle ; R_2 est une relation Intervalle-Point et R_3 est une relation Point-Point. Certaines relations de composition sont « indécidables ». Par exemple, les relations « $After_m(P, I)$ » et « $Before_m(I, L)$ » donnent trois relations possibles : « $Before_m(P, L)$ », « $Equals_m(P, L)$ » et « $After_m(P, L)$ ».

Tableau 38. Table de transitivité qui permet de raisonner $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$.

$R_1(P, I)$ $R_2(I, L)$	$Before_m(P, I)$	$After_m(P, I)$	$Starts_m(P, I)$	$During_m(P, I)$	$Ends_m(P, I)$
$Before_m(I, L)$	$Before_m(P, L)$	L'une des relations de points de temps	$Before_m(P, L)$	$Before_m(P, L)$	$Before_m(P, L)$
$After_m(I, L)$	L'une des relations de points de temps	$After_m(P, L)$	$After_m(P, L)$	$After_m(P, L)$	$After_m(P, L)$
$Started-by_m(I, L)$	$Before_m(P, L)$	$After_m(P, L)$	$Equals_m(P, L)$	$After_m(P, L)$	$After_m(P, L)$

Contains_m(I, L)	Before _m (P, L)	After _m (P, L)	Before _m (P, L)	L'une des relations de points de temps	After _m (P, L)
Ended-by_m(I, L)	Before _m (P, L)	After _m (P, L)	Before _m (P, L)	Before _m (P, L)	Equals _m (P, L)

Le Tableau 39 permet de déduire $R_3(I, J)$ à partir de $R_1(I, P)$ et $R_2(P, J)$. R_1 est une relation Intervalle-Point ; R_2 est une relation point-intervalle et R_3 est une relation intervalle-intervalle.

Tableau 39. Table de transitivité qui permet de raisonner $R_3(P, L)$ à partir de $R_1(P, I)$ et $R_2(I, L)$.

R₁(I, P) R₂(P, J)	Before_m(I, P)	After_m(I, P)	Started-by_m(I, P)	Contains_m(I, P)	Ended-by_m(I, P)
Before_m(P, J)	Before _m (I, J)	L'une des relations d'intervalles de temps	Contains _m (I, J) ∨ Ended-by _m (I, J) ∨ Met-by _m (I, J) ∨ Before _m (I, J) ∨ Overlaps _m (I, J)	Contains _m (I, J) ∨ Ended-by _m (I, J) ∨ Overlapped-by _m (I, J) ∨ Meets _m (I, J) ∨ Before _m (I, J)	Before _m (I, J)
After_m(P, J)	L'une des relations d'intervalles de temps	After _m (I, J)	After _m (I, J)	After _m (I, J) ∨ Met-by _m (I, J) ∨ During _m (I, J) ∨ Started-by _m (I, J) ∨ Overlapped-by _m (I, J)	During _m (I, J) ∨ Started-by _m (I, J) ∨ Overlapped-by _m (I, J) ∨ Met-by _m (I, J) ∨ After _m (I, J)
Starts_m(P, J)	Before _m (I, J)	After _m (I, J) ∨ Met-by _m (I, J) ∨ Overlapped-by _m (I, J) ∨ Ended-by _m (I, J) ∨ Contains _m (I, J)	Started-by _m (I, J) ∨ Equals _m (I, J) ∨ Starts _m (I, J)	Contains _m (I, J) ∨ Ended-by _m (I, J) ∨ Overlaps _m (I, J)	Meets _m (I, J)
During_m(P, J)	During _m (I, J) ∨ Starts _m (I, J) ∨ Overlaps _m (I, J) ∨ Meets _m (I, J) ∨ Before _m (I, J)	After _m (I, J) ∨ Met-by _m (I, J) ∨ Overlapped-by _m (I, J) ∨ Ended-by _m (I, J) ∨ Contains _m (I, J)	Contains _m (I, J) ∨ Ends _m (I, J) ∨ Overlapped-by _m (I, J)	Overlaps _m (I, J) ∨ Overlapped-by _m (I, J) ∨ Starts _m (I, J) ∨ Started-by _m (I, J) ∨ During _m (I, J)	Overlapped-by _m (I, J) ∨ Starts _m (I, J) ∨ During _m (I, J)

				Contains _m (I, J) ∨ Ends _m (I, J) ∨ Ended-by _m (I, J) ∨ Equals _m (I, J)	
Ends_m(P, J)	During _m (I, J) ∨ Starts _m (I, J) ∨ Overlaps _m (I, J) ∨ Meets _m (I, J) ∨ Before _m (I, J)	After _m (I, J)	Met-by _m (I, J)	Contains _m (I, J) ∨ Started-by _m (I, J) ∨ Overlapped-by _m (I, J)	Equals _m (I, J) ∨ Ended-by _m (I, J) ∨ During _m (I, J)

Nous introduisons une table (Tableau 40) de transitivité pour raisonner sur les relations temporelles qualitatives entre les points temporels. Elle permet de déduire $R_3(P, M)$ à partir de $R_1(P, L)$ et $R_2(L, M)$, où P, L et M sont des points de temps en conflit. R_1, R_2 et R_3 sont des relations Point-Point.

Tableau 40. Table de transitivité qui permet de raisonner $R_3(P, M)$ à partir de $R_1(P, L)$ et $R_2(L, M)$.

R₁(P, L)	Before_m(P, L)	Equals_m(P, L)	After_m(P, L)
R₂(L, M)			
Before_m(L, M)	Before _m (P, M)	Before _m (P, M)	L'une des relations de point de temps
Equals_m(L, M)	Before _m (P, M)	Equals _m (P, M)	After _m (P, M)
After_m(L, M)	L'une des relations de point de temps	After _m (P, M)	After _m (P, M)

8.4. « BeliefTimeOnto » : notre ontologie temporelle évidentielle pour les données temporelles conflictuelles

A partir de notre représentation sémantique et de l'extension des relations d'Allen, nous proposons une ontologie nommée « *BeliefTimeOnto* » avec une extension qui permet de représenter **les masses de croyance des données temporelles conflictuelles**. Nous créons notre ontologie à l'aide de l'éditeur d'ontologies Protégé comme le montre la Figure 28. « *BeliefTimeOnto* » contient 4 classes, 38 propriétés d'objets, 6 propriétés de données. Nous représentons l'extension de l'algèbre Allen avec des règles SWRL.

Name	Rule	Comment
R1	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingBeforeIntervals
R10	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingOverlappedByIntervals
R11	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingStartedByIntervals
R12	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingContainsIntervals
R13	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingEndedByIntervals
R14	autogen0:TimeInterval(?p) ^ autogen0:TimePoint(?p, ?pt) ^ autogen0:PointMass(?p, ?pm) ^ autogen0:TimeInterval(?q) ^ autogen0:TimePoint(?q, ?qt) ^ autogen0:Pol...	To deduce ConflictingBeforePoints
R15	autogen0:TimeInterval(?p) ^ autogen0:TimePoint(?p, ?pt) ^ autogen0:PointMass(?p, ?pm) ^ autogen0:TimeInterval(?q) ^ autogen0:TimePoint(?q, ?qt) ^ autogen0:Pol...	To deduce ConflictingAfterPoints
R16	autogen0:TimeInterval(?p) ^ autogen0:TimePoint(?p, ?pt) ^ autogen0:PointMass(?p, ?pm) ^ autogen0:TimeInterval(?q) ^ autogen0:TimePoint(?q, ?qt) ^ autogen0:Pol...	To deduce ConflictingEqualsPoints
R17	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingBeforeIntervalPoint
R18	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingAfterIntervalPoint
R19	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingBeforePointInterval
R2	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingIleIsIntervals
R20	autogen0:TimeInterval(?i) ^ autogen0:BeginsAt(?i, ?ib) ^ autogen0:EndsAt(?i, ?ie) ^ autogen0:BeginningMass(?ib, ?mib) ^ autogen0:EndingMass(?ie, ?mie) ^ auto...	To deduce ConflictingAfterPointInterval

Control | Rules | Asserted Axioms | Inferred Axioms | OWL 2 RL

OWL axioms successfully transferred to rule engine.
 Number of SWRL rules exported to rule engine: 32
 Number of OWL class declarations exported to rule engine: 4
 Number of OWL individual declarations exported to rule engine: 0
 Number of OWL object property declarations exported to rule engine: 39
 Number of OWL data property declarations exported to rule engine: 6
 Total number of OWL axioms exported to rule engine: 283
 The transfer took 2259 millisecond(s).
 Press the 'Run Drools' button to run the rule engine.
 Successful execution of rule engine.
 Number of inferred axioms: 191
 The process took 158 millisecond(s).
 Look at the 'Inferred Axioms' tab to see the inferred axioms.
 Press the 'Drools->OWL' button to translate the inferred axioms to OWL knowledge.

Figure 28. Règles SWRL de notre ontologie

8.5. Expérimentation

Pour valider notre approche, nous menons une étude de cas dont le but est de montrer l'intérêt de notre approche dans le cadre de la prothèse CAPTAIN MEMO.

Nous implémentons un prototype basé sur l'approche proposée et l'ontologie « *BeliefTimeOnto* ». Ensuite, nous l'intégrons dans CAPTAIN MEMO pour gérer la dimension temporelle.

8.5.1 Prototype implémenté

Notre approche a été implémentée en JAVA. Il est présenté sous forme d'une interface conviviale comme illustré par la Figure 29. Tout d'abord, l'utilisateur instancie l'ontologie « *BeliefTimeOnto* » via l'interface utilisateur qui lui offre la possibilité de saisir des données temporelles conflictuelles. Après chaque nouvelle saisie de données temporelles, le composant « Inférence de données temporelles qualitatives » est automatiquement exécuté pour déduire les données manquantes et les degrés de masse associés. Ce composant est basé les règles

SWRL. Le troisième composant permet aux utilisateurs d'interroger l'ontologie via des requêtes SPARQL.

The screenshot shows a software interface for configuring temporal data. It has a title bar with 'BeliefTimeOnto Population' and 'Configuration'. The main area is divided into four panels:

- Add Time Interval:** Contains a 'Conflicting Time Interval' section with a text input for 'Time Interval's Name' (value: A), and two sub-sections: 'Beginning Bound' (From: 1970, Mass Degree: 0.72) and 'Ending Bound' (To: 1975, Mass Degree: 0.78). Buttons for 'Add Time Interval' and 'Save' are at the bottom.
- Conflicting Time Point:** Contains a 'Time Point Name' input field, 'From' and 'To' input fields, and buttons for 'Add Time Point' and 'Save'.
- Add Qualitative Temporal Relation:** Contains two 'Time Interval/Point' dropdown menus (values: A and B), a 'Qualitative Temporal Relation' dropdown (value: Before_m), a 'Mass Degree' input field (value: 0.64), and a 'Save' button.
- Saved Temporal Data:** Contains a 'Search Criterion' dropdown (value: All Time Points) and a table of results.

Result	Item	Mass Degree
A	Conflicting Time Interval	0,72
B	Conflicting Time Interval	0,81

Figure 29. Prototype implémenté basé sur "BeliefTimeOnto"

8.5.2 Etude de cas : application dans CAPTAIN MEMO

Nous avons mené une étude de cas dans le cadre CAPTAIN comme suit :

- Nous avons mené une étude de cas dans le cadre CAPTAIN comme suit : Nous avons posé à un patient Alzheimer, 5 fois pendant 5 jours, la question suivante : « *En quelle année avez-vous donné naissance à votre fille ?* ».
- Nous avons posé au même patient Alzheimer, 3 fois pendant 3 jours, la question suivante : « *Quand avez-vous quitté votre emploi ?* ». Nous avons obtenu 3 réponses (1 réponse/jour) (cf. Tableau 41).

Tableau 41. Réponses du patient Alzheimer

	<i>Jour 1</i>	<i>Jour 2</i>	<i>Jour 3</i>	<i>Jour 4</i>	<i>Jour 5</i>
<i>En quelle année avez-vous donné naissance à votre fille ?</i>	1970	1973	1975	1973	1971
<i>Quand avez-vous quitté votre emploi ?</i>	[1989,1993]	[1992,1995]	[1998,1999]		

Soit « P » l'événement exprimant la première question. Supposons que nous ayons 4 points de temps conflictuels P_1 , P_2 , P_3 et P_4 répondant à cet événement représentés respectivement par : $P_1 = 1970$, $P_2 = 1973$, $P_3 = 1975$ et $P_4 = 1971$. Nous excluons la réponse du jour 4, car elle est égale à jour 1.

Notre état de connaissance sur « P » est défini comme suit :

$$m(\{1970, 1973, 1975, 1971, \{1970, 1973\}, \{1970, 1975\}, \{1970, 1971\}, \{1973, 1975\}, \{1973, 1971\}, \{1975, 1971\}, \{1970, 1973, 1975\}, \{1970, 1973, 1971\}, \{1973, 1975, 1971\}, \{1970, 1973, 1975, 1971\}) = 1.$$

La fonction de croyance de P est définie comme :

$$\text{Bel}(P) = (\text{Bel}(P_1), \text{Bel}(P_2), \text{Bel}(P_3), \text{Bel}(P_4))$$

Pour estimer la masse de $m(P_1)$, $m(P_2)$, $m(P_3)$ et $m(P_4)$, nous calculons les mesures de crédibilité des points temporels à l'aide de notre prototype basé sur DBE_ALZ. Les résultats sont décrits dans le Tableau 42.

Tableau 42. Mesures de crédibilité des points de temps basées sur DBE_ALZ

Points de temps P_p	Mesures de crédibilité $\text{Bel}(P_p)$
P_1	0.42
P_2	0.54
P_3	0.01
P_4	0

Nous déduisons les masses de croyance comme indiqué dans le Tableau 43.

Tableau 43. Les masses de croyance des points de temps

	Mass m
$m(P_1)$	0.42
$m(P_2)$	0.54
$m(P_3)$	0.01
$m(P_4)$	0
$m(\{P_1, P_2\}, \{P_1, P_3\}, \{P_1, P_4\}, \{P_2, P_3\}, \{P_2, P_4\}, \{P_3, P_4\}, \{P_1, P_2, P_3\}, \{P_1, P_2, P_4\}, \{P_1, P_3, P_4\}, \{P_2, P_3, P_4\}, \{P_1, P_2, P_3, P_4\})$	0.03

P_2 est le point de temps ayant la masse de croyance la plus élevée.

Soit « I » l'événement exprimant la seconde question. Supposons que nous ayons 5 intervalles de temps conflictuels I_1, I_2, I_3 et I_4 répondant à cet événement représentés respectivement par : $I_1 = [1989, 1993]$, $I_2 = [1992, 1995]$ et $I_3 = [1998, 1999]$.

Notre état de connaissance sur I est défini comme :

$$m(\{[1989, 1993], [1992, 1995], [1998, 1999], \{[1989, 1993], [1992, 1995]\}, \{[1992, 1995], [1998, 1999]\}, \{[1989, 1993], [1998, 1999]\}, \{[1989, 1993], [1992, 1995], [1998, 1999]\}) = 1.$$

Chaque borne de l'intervalle I_i est associée à un degré de masse « m ». Cela signifie que pour $I_{i=1..3} = [I_i^-, I_i^+]$, m_i^- et m_i^+ sont respectivement les degrés de masse associés à I_i^- et I_i^+ . Pour calculer la masse de chaque borne, nous utilisons la règle de combinaison conjonctive et la fonction de croyance de I_i . La fonction de croyance de chaque borne de I_i est calculée à l'aide d'un prototype basé sur DBE_ALZ. Les résultats sont décrits dans le Tableau 44.

Par exemple, pour I_1 , la fonction de crédibilité est définie comme :

$$\text{Bel}(I_1) = (\text{Bel}(I_1^-), \text{Bel}(I_1^+))$$

Tableau 44. Fonction de crédibilité des intervalles de temps I_i

Intervalles de temps I_i	$\text{Bel}(I_i^-)$	$\text{Bel}(I_i^+)$	$\text{Bel}(I_i)$
I_1	0.54	0.78	0.4
I_2	0.21	0.37	0.07
I_3	0.63	0.81	0.51

Nous déduisons ensuite les masses de croyance indiquées dans le Tableau 45.

Tableau 45. Masses de croyance des intervalles de temps I_i

	Mass m
$m(I_1)$	0.4
$m(I_2)$	0.07
$m(I_3)$	0.51
$m(\{(I_1, I_2), m(I_2, I_3), m(I_1, I_3), m(I_1, I_2, I_3)\})$	0.02

I_3 est l'intervalle de temps ayant la masse de croyance la plus élevée.

La relation qualitative « *ConflictingBeforePointInterval* » existe entre P et I. Nous calculons la masse de croyance « m » associé à la relation « *Before* » :

$$\text{Before}_m(P_2, I_3) \rightarrow P_{mp} < I_{mi}$$

Cela signifie que : $\text{Before}_m(P_2, I_3) \rightarrow 1973_{0.54} < 1998_{0.63}$. Nous déduisons que « $m = 0.33$ ».

Nous représentons à l'aide de notre ontologie « *BeliefTimeOnto* » comme le montre la Figure 30.

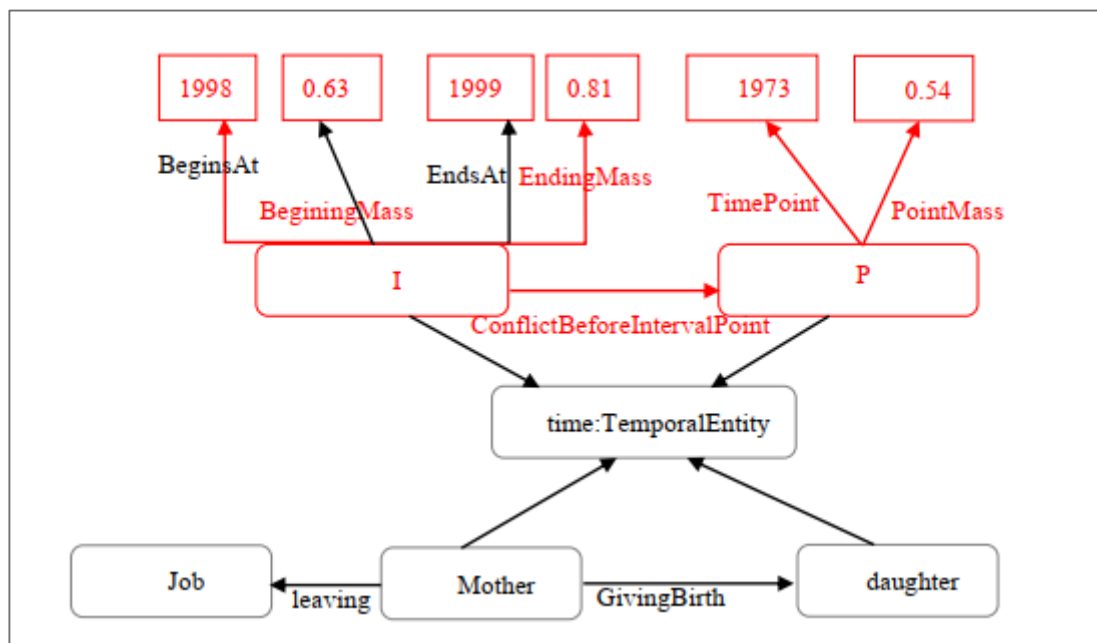


Figure 30. Représentation sémantique de l'exemple proposé basée sur notre approche

8.6. Conclusion

Dans ce chapitre, nous avons proposé une approche basée sur la théorie des évidences pour représenter et raisonner sur des données temporelles conflictuelles dans OWL 2. Cette approche inclut trois contributions. La première concerne le calcul des masses de croyance des données temporelles conflictuelles en calculant les mesures de crédibilité basées sur notre précédente approche DBE_ALZ. Nous avons proposé une représentation sémantique qui associe les mesures calculées aux données traitées et leurs relations qualitatives. La seconde consiste à raisonner sur des données temporelles conflictuelles en étendant l’algèbre d’intervalle d’Allen. Nous avons proposé 13 relations temporelles entre des intervalles de temps conflictuels. Notre extension conserve les propriétés de réflexivité/irréflexivité, de symétrie/asymétrie et de transitivité. Ces relations sont adaptées pour relier un intervalle de temps et un instant ou deux instants. Nous avons introduit quatre tables de transitivité pour raisonner sur les relations temporelles. La troisième contribution consistait en la proposition d’une l’ontologie « *BeliefTimeOnto* » qui implémente la représentation sémantique proposée et l’extension d’algèbre d’intervalle d’Allen. Les inférences sont effectuées à l’aide de règles SWRL. Pour valider notre travail, nous avons mis en place un prototype permettant aux utilisateurs d’explorer notre approche, basée sur notre ontologie « *BeliefTimeOnto* ».

Dans le chapitre suivant, nous présentons notre ontologie de l’imperfection des données temporelles et l’évaluation effectuée dans le cadre de CAPTAIN MEMO.

CHAPITRE 9 : Ontologie de l'imperfection des données temporelles et évaluation dans CAPTAIN MEMO

9.1. Introduction

Sur la base de **nos ontologies proposées** dans les chapitres précédents, nous fondons **une ontologie** qui **englobe toutes les imperfections des données temporelles traitées** comme illustré dans la Figure 31. Cette ontologie, nommée « *TimeOntoImperfection* », est une ontologie supérieure (Top-level ontology), qui peut être **fusionnée avec d'autres ontologies** d'application ou de domaine qui doivent être étendues pour **représenter et raisonner sur des données temporelles imparfaites** (comme nous l'avons montré pour l'ontologie PersonLink). **Nous rappelons que dans la littérature, il n'existe pas, à notre connaissance, une ontologie de l'imperfection des données temporelles.**

Dans ce chapitre, nous présentons dans la première section **les éléments de notre ontologie** « *TimeOntoImperfection* » ainsi que **le prototype implémenté** basé sur cette dernière. Dans la deuxième section, nous présentons **l'évaluation réalisée** dans le cadre de CAPTAIN MEMO.

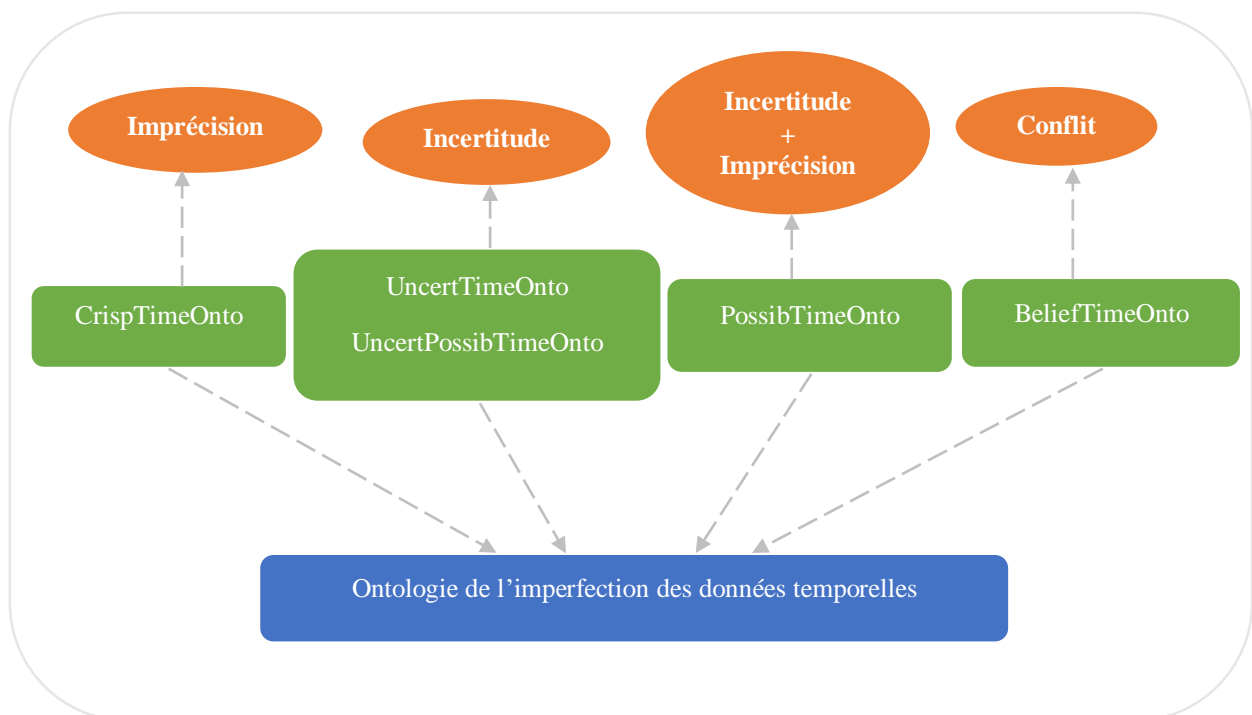


Figure 31. Composants de l'ontologie de l'imperfection des données temporelles

9.2. Ontologie de l'imperfection des données temporelles

L'ontologie de l'imperfection des données temporelles (Figure 39) se base sur nos ontologies « *CrispTimeOnto* », « *UncertTimeOnto* », « *UncertPossibTimeOnto* », « *PossibTi-*

meOnto » et « BeliefTimeOnto », représentant respectivement des données temporelles **imprécises, incertaines, à la fois incertaines et imprécises et conflictuelles**. Nous rappelons que nos ontologies sont basées sur **l'extension de l'approche 4D-fluents** avec des éléments prédéfinis de **l'ontologie OWL-Time**, des éléments que nous avons définis pour **représenter les imperfections ciblées**, et **l'extension de l'algèbre d'intervalles d'Allen** pour **raisonner sur les données temporelles imparfaites**. Nous créons notre ontologie « *TimeOntoImperfection* » à l'aide de l'éditeur d'ontologies Protégé.

9.2.1 Eléments de « TimeOntoImperfection »

- **Classes.** L'ontologie de l'imperfection des données temporelles contient cinq classes (Figure 32). « **TimeSlice** » qui est la classe de domaine pour les entités représentant des parties temporelles. « **time:TemporalEntity** » qui est une sous-classe de la classe « **TimeSlice** ». « **time:TimeInterval** » et « **time:TimeInstant** » qui sont respectivement les classes représentant les intervalles et les points de temps. « **time:DateTimeDescription** » qui est la classes représentant les dates et les horloges de temps. « **time:TimeInterval** », « **time:TimeInstant** » et « **time:DateTimeDescription** » qui sont des sous-classes de la classe « **time:TemporalEntity** ».

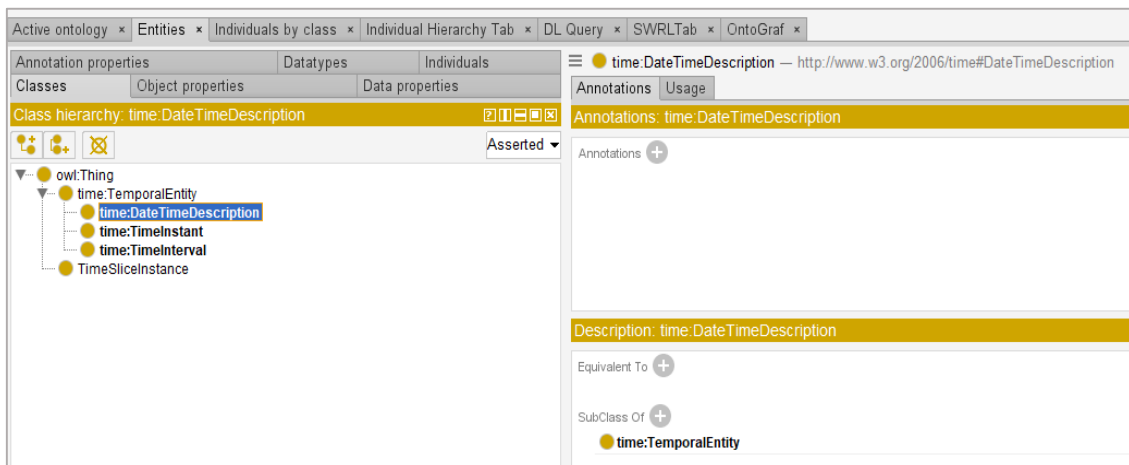


Figure 32. Les classes de l'ontologie "TimeOntoImperfection"

- **Propriétés d'objets.** L'ontologie de l'imperfection des données temporelles contient 132 propriétés d'objets (Figure 33). La propriété d'objet « **TsTimeSliceInstance** » relie les classes « **TimeSlice** » et « **time:TemporalEntity** ». « **Is_a** » connecte « **time:TemporalEntity** » avec « **time:TimeInterval** », « **time:TimeInstant** »

et « time:DateTimeDescription ». Par ailleurs, nous définissons **les propriétés d'objets qui représentent les relations qualitatives** entre deux intervalles (Intervalle-Intervalle), un intervalle de temps et un point de temps (Intervalle-Point), un point de temps et un intervalle de temps (Point-Intervalle) et deux points de

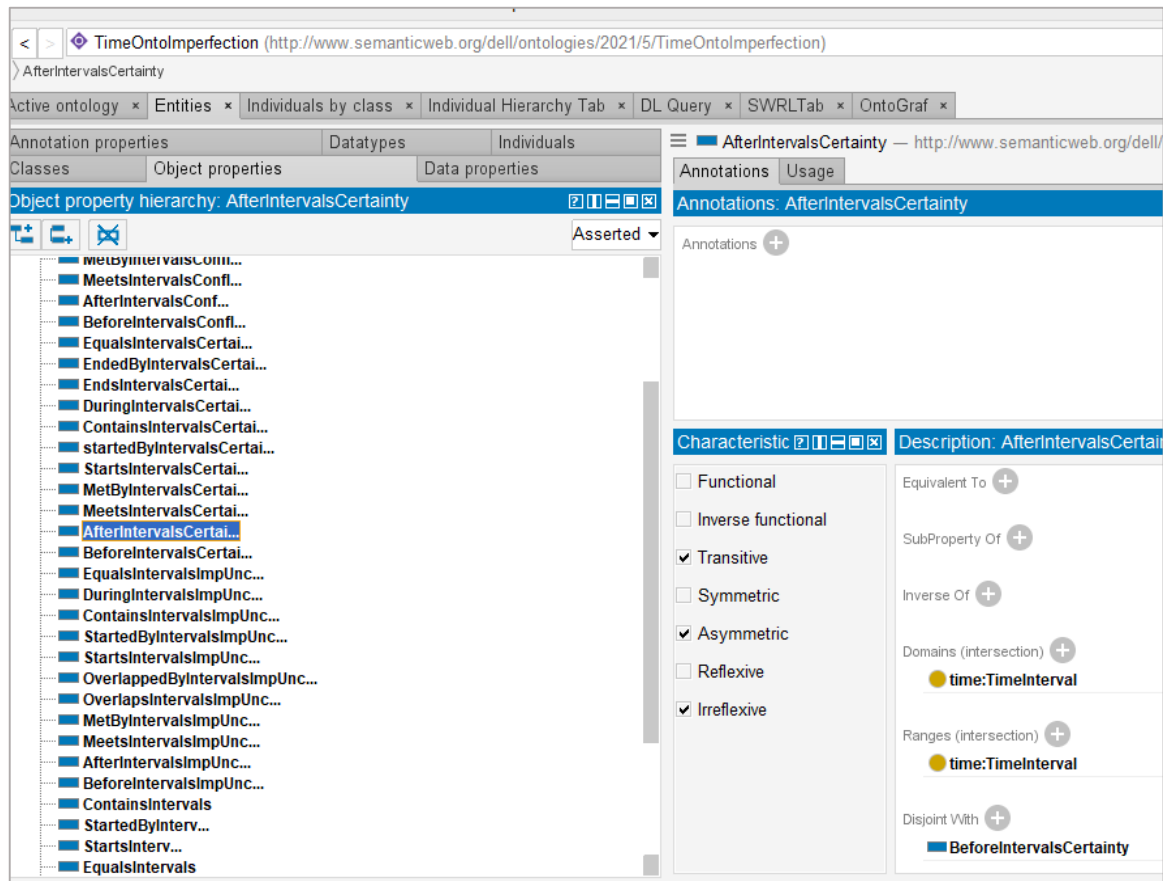


Figure 33. Propriétés d'objet de l'ontologie "TimeOntoImperfection"

temps (Point-Point). Nous instancions les propriétés d'objet **sur la base de nos extensions de l'algèbre d'Allen**. Par exemple, « RelationIntervals » peut être l'une des relations d'Allen.

En d'autres termes, 13 propriétés d'objet sont associées : « BeforeIntervals », « Meet-intervals », « OverlapsIntervals », « StartsIntervals », « DuringIntervals », « EndsIntervals », « AfterIntervals », « Met-byIntervals », « Overlapped-byIntervals », « Started-byIntervals », « ContainsIntervals », « EndedbyIntervals » et « EqualsIntervals ». Par conséquent, nous trouvons 26 propriétés d'objet qui représentent **les relations qualitatives précises** (c.-à-d., ne contenant aucun type d'imperfection), 26 propriétés d'objet qui représentent **les relations qualitatives imprécises**, 26 propriétés d'objet qui représentent **les relations qualitatives incertaines**, 26 propriétés d'objet qui représentent les

relations qualitatives **à la fois incertaines et imprécises** et 26 propriétés d'objet qui représentent **les relations qualitatives conflictuelles**.

- **Propriétés de type de données.** L'ontologie de l'imperfection des données temporelles contient 40 propriétés de type de données qui représentent **les bornes des intervalles de temps, les points de temps, les dates et les horloges de temps** ainsi que toutes les mesures des différents types d'imperfections qui sont **les mesures de certitudes, les mesures de possibilités, les mesures de nécessités et les masses de croyance**. La Figure 34 illustre ces propriétés.

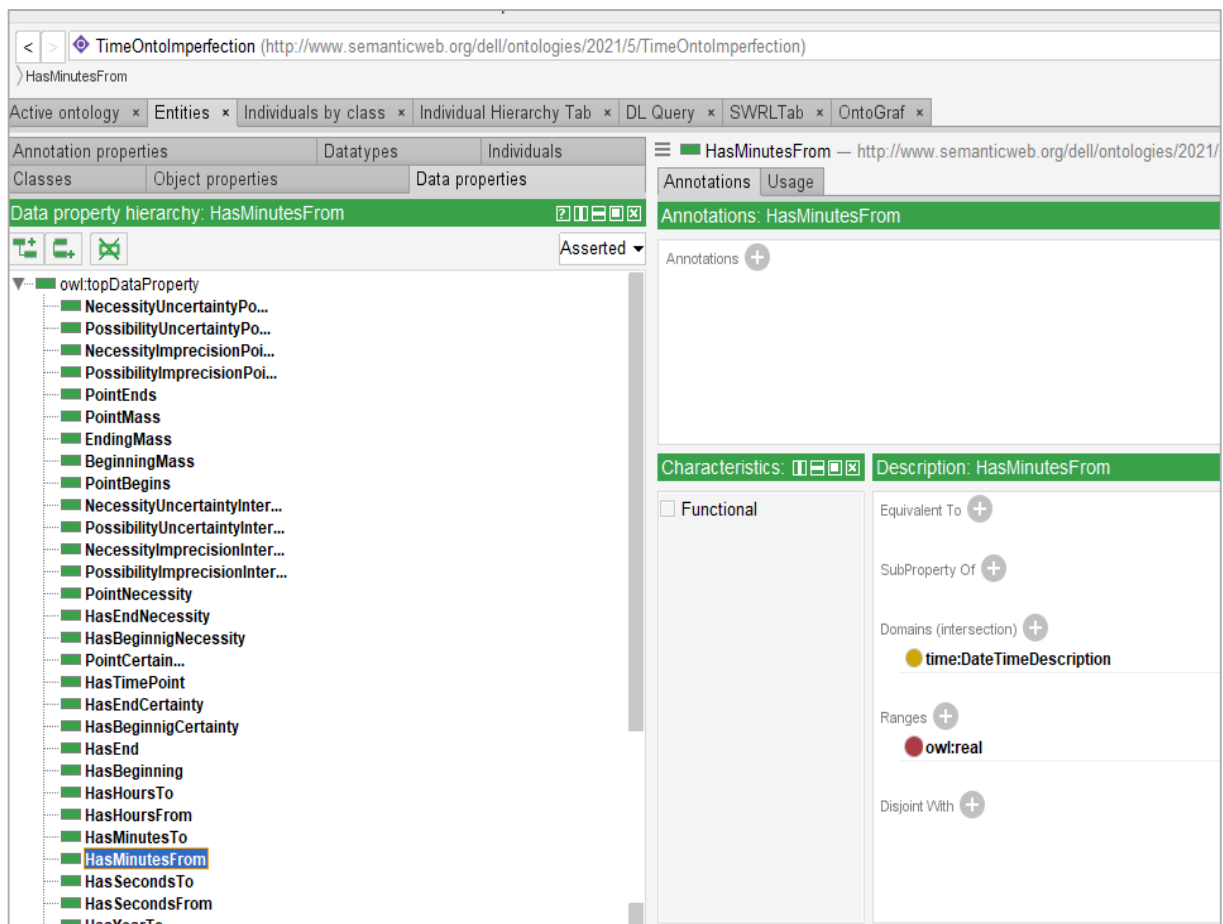


Figure 34. Propriétés de type de données de "TimeOntoImperfection"

- **Règles SWRL.** Nous inférons via un ensemble **de règles SWRL** notre **extension de l'algèbre d'Allen**. Figure 35 illustre un exemple d'une règle SWRL.

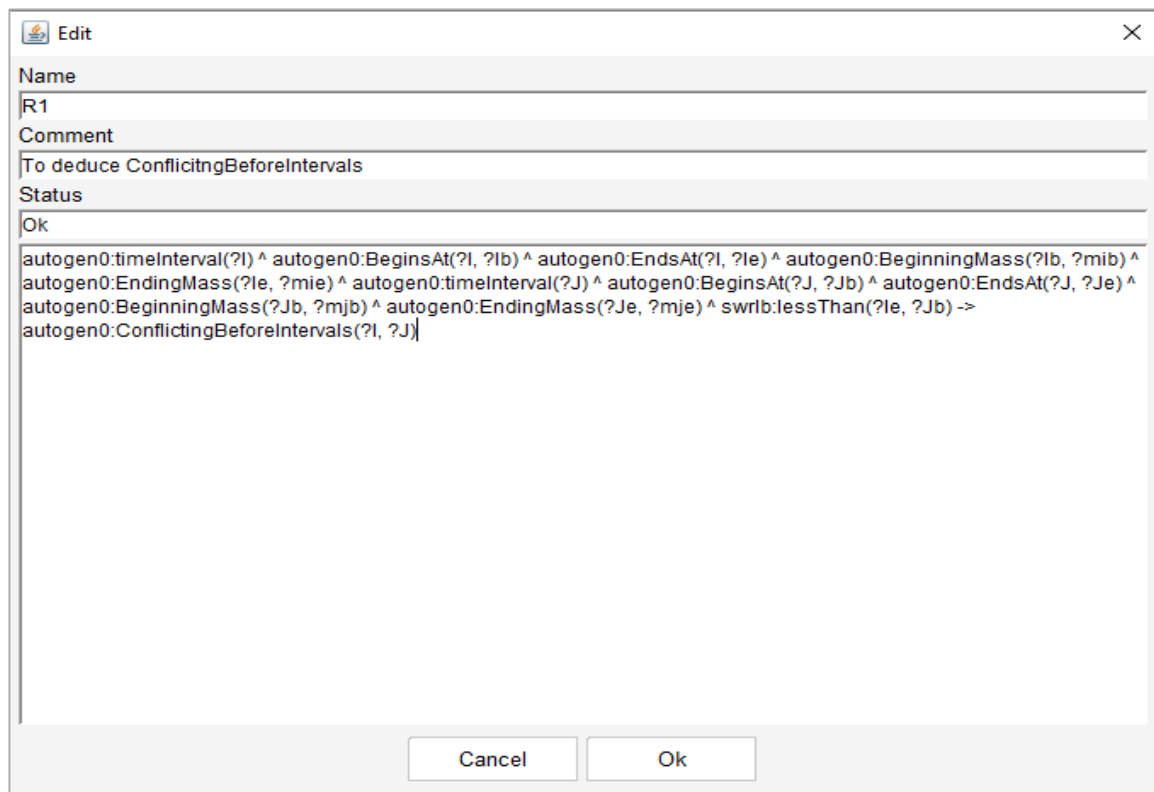


Figure 35. Exemple d'une règle SWRL

9.2.2 Prototype basé sur notre ontologie « TimeOntoImperfection »

Nous proposons un **prototype basé sur notre ontologie « TimeOntoImperfection »** que nous intégrons dans CAPTAIN MEMO. Ce prototype a été implémenté en Java.

L'interface principale du prototype permet à l'utilisateur de **saisir des intervalles de temps et des points de temps parfaits et/ou imparfaits** et de calculer **les degrés de certitudes, de possibilité, de nécessité et les degrés de crédibilités et les masses de croyance** en se basant sur nos approches proposées comme la montre la Figure 36.

"TimeOntoImperfection" Population Configuration

Add Time Interval

Time Interval: Imprecise

Time Interval's Name:

Beginning Bound: From To

Ending Bound: From To

Add Time Point: From To

Buttons: Add Time Interval, Save, Add Time Point, Save

Bottom Buttons: Certainty Degree, Necessity Degree, Mass/Belief Degree, Possibility Degree

Figure 36. Prototype basé sur "TimeOntoImperfection"

Après chaque nouvelle saisie de données temporelles, le composant « *Add Qualitative Temporal Relation* » est automatiquement exécuté pour déduire les données manquantes, notamment **les relations qualitatives associées et les mesures associées en se basant sur les règles SWRL**. La Figure 37 montre ce composant de l'interface utilisateur du prototype intégré.

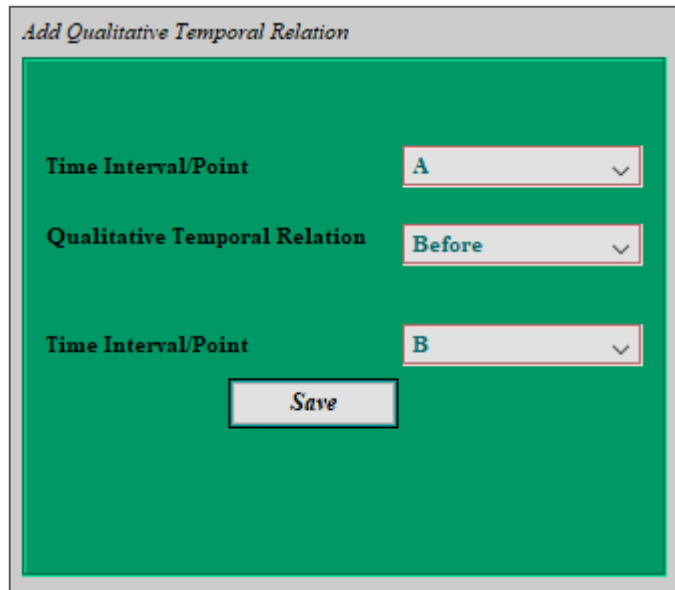


Figure 37. Composant des relations qualitatives du prototype

Notre prototype permet aussi d'effectuer une recherche sur les données saisies et enregistrées en utilisant un filtre intégré dans une barre de choix. Ceci est implémenté avec des requêtes SPAQL. La Figure 38 montre le composant décrit.

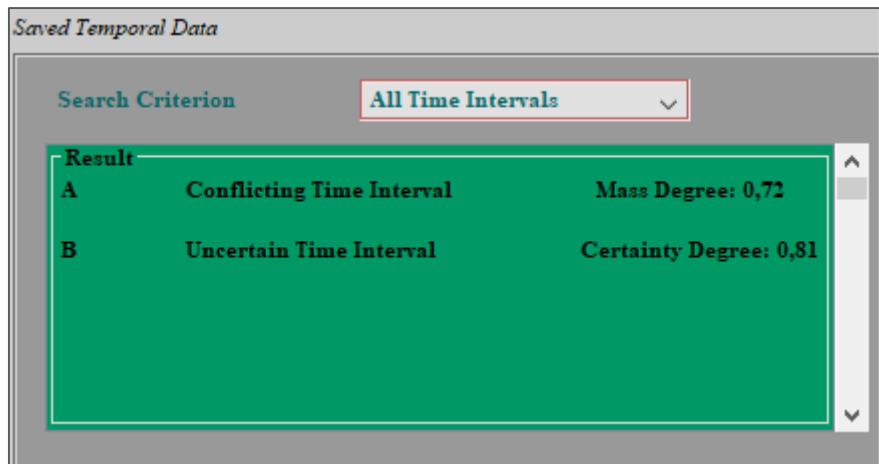


Figure 38. Composant permettant de chercher dans les données temporelles enregistrées

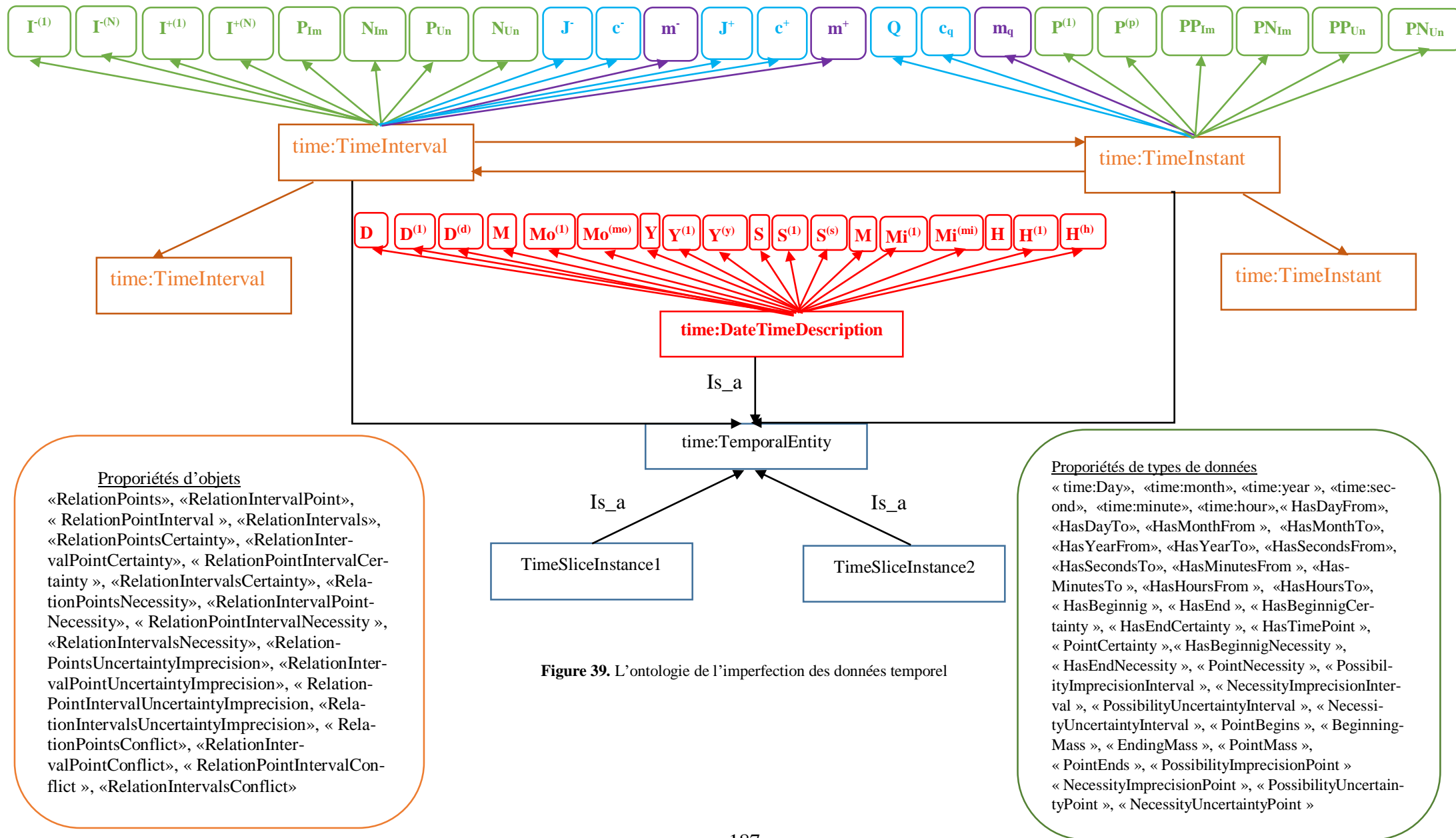


Figure 39. L'ontologie de l'imperfection des données temporel

9.3. Evaluation dans le cadre de CAPTAIN MEMO

Dans cette section, nous présentons les évaluations effectuées dans le cadre de CAPTAIN MEMO. Nous avons intégré **notre ontologie de l'imperfection des données temporelles** dans CAPTAIN MEMO et nous évaluons **nos approches traitant les différents types d'imperfection**. Cette évaluation fait intervenir **des personnes atteintes de la maladie d'Alzheimer** qui séjournent dans **la maison de retraite de Sfax en Tunisie** (rue Manzel-Chaker km. 8).

La présente section vise à évaluer l'efficacité de nos approches des imperfections des données temporelles traitées. Cette évaluation est effectuée dans le cadre de CAPTAIN MEMO. Nous avons recours aux mesures de la **Précision**.

Un total de **20 patients atteints de la maladie d'Alzheimer** $P = \{P_1 \dots P_{20}\}$ et leurs **soignants associés** $C = \{C_1 \dots C_{20}\}$ ont été recrutés pour participer à cette étude. Tous les aidants sont des parents au premier degré, par exemple un fils ou une femme. Tous les patients atteints de la maladie d'Alzheimer sont à **un stade précoce de la maladie**. Ils étaient âgés de 68 à 75 ans. Nous avons demandé au parrain légal de chaque patient **la lettre de consentement**. Nous avons exclu les participants présentant des troubles manifestes du comportement, une aphasie sévère et une perte auditive et/ou visuelle sévère.

Chaque patient a été invité à répondre aux mêmes **10 questions** relatives à sa vie privée et à ses antécédents à chaque session de test. Les bases de connaissances associées aux participants sont structurées à l'aide de l'ontologie PersonLink.

Deux scénarios sont proposés :

- **Premier scénario** « **Sans intégrer notre prototype dans CAPTAIN MEMO** » : Nous n'intégrons pas notre prototype dans la prothèse mémoire CAPTAIN MEMO. Toutes les données saisies par le patient Alzheimer P_i sont enregistrées dans $KBi/sc1$ (base de connaissances correspondant aux données saisies par le patient P_i sur la base du premier scénario). Les $KBi/s1$ sont générés **après 8 semaines d'utilisation** de CAPTAIN MEMO. Il est demandé à chaque soignant C_i de n'identifier que les vraies données saisies par le patient. Les réponses forment la base de connaissances de référence relative au premier scénario $KBi/GS1$.

- **Deuxième Scénario** « Nous intégrons notre prototype basé sur l'ontologie dans CAPTAIN MEMO » : Nous intégrons notre prototype basé sur l'ontologie dans CAPTAIN MEMO. Toutes les données saisies par le patient Alzheimer P_i et ayant une mesure de possibilité, une mesure de nécessité et une masse de croyance sont enregistrées en $KBi/sc2$. $KBi/sc2$ sont générés après avoir utilisé CAPTAIN MEMO.

Nous comparons la base de connaissances générée du second scénario au golden standard du premier scénario. Nous utilisons la métrique d'évaluation **Precision** :

- $P_i@1$ ($(|KBi/sc1 \cap KBi/ GS1| / |KBi/sc1|)$) représente la Précision associée au patient P_i selon le premier scénario.
- $P_i@2$ ($(|KBi/sc2 \cap KBi/ GS1| / |KBi/sc2|)$) représente la Précision associée au patient P_i selon le deuxième scénario.

Tableau 46. Précision des résultats de l'évaluation de notre approche

Early stage Alzheimer's patients	Precision according to the first scenario: $P_{i@1}$	Precision according to the second scenario: $P_{i@2}$
P_1	0.864	0.986
P_2	0.834	0.882
P_3	0.895	0.935
P_4	0.724	0.731
P_5	0.822	0.954
P_6	0.671	0.701
P_7	0.735	0.902
P_8	0.642	0.714
P_9	0.821	0.931
P_{10}	0.771	0.777
P_{11}	0.878	0.954
P_{12}	0.723	0.765
P_{13}	0.699	0.798
P_{14}	0.815	0.932
P_{15}	0.798	0.874
P_{16}	0.801	0.899
P_{17}	0.865	0.945
P_{18}	0.645	0.712
P_{19}	0.705	0.765
P_{20}	0.801	0.862
	Moyenne: 0.775	Moyenne: 0.850

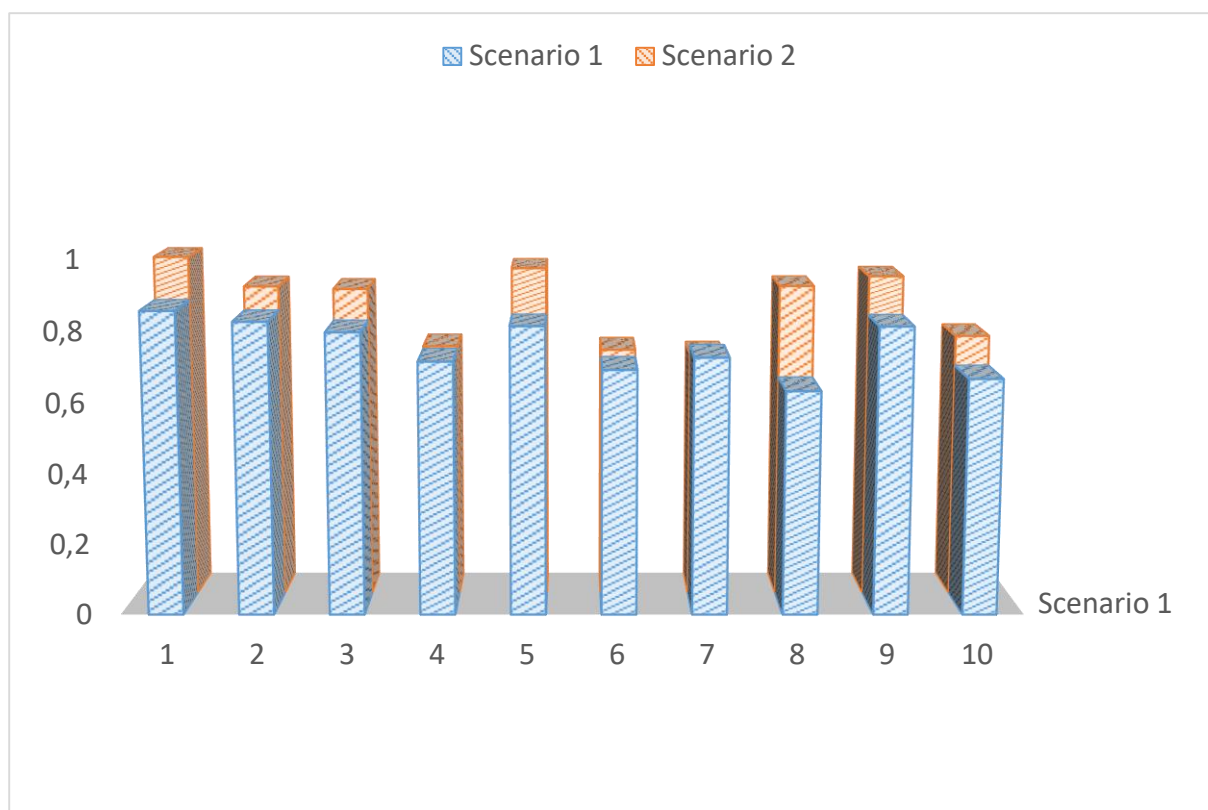


Figure 40. Résultats de l'évaluation

Le Tableau 46 et la Figure 40 montrent que la moyenne globale de la précision associée au deuxième scénario (0.850) est meilleure que la moyenne globale de la précision associée au premier scénario (0.775). Ces résultats prouvent **la pertinence de l'approche proposée**.

9.4. Conclusion

Dans ce chapitre, nous avons présenté **notre ontologie de l'imperfection des données temporelles nommée « *TimeOntoImperfection* »** basée sur nos ontologies créées pour chaque type d'imperfection traitée. Le point crucial est qu'**il n'y a pas d'ontologie de l'imperfection temporelle**. Comme déjà présenté dans le chapitre de l'état de l'art, **les ontologies qui représentent l'aspect temporel ne traitent pas les imperfections**. Nous avons implémenté un prototype basé sur « *TimeOntoImperfection* » que nous avons intégré par la suite dans la prothèse de mémoire CAPTAIN MEMO.

Nous avons évalué notre travail sur des personnes atteintes de la maladie d'Alzheimer qui séjournent dans la maison de retraite de Sfax en Tunisie (rue Manzel-Chaker km. 8). Dans cette évaluation, nous avons proposé deux scénarios : le premier sans intégration du prototype et le deuxième en l'intégrant. Les mesures de précision obtenues montrent **l'apport de nos approches**.

CHAPITRE 10 : Conclusion générale et Perspectives

Pour conclure, nous présentons un résumé de nos travaux réalisés au cours de cette thèse et qui portent sur **le traitement de plusieurs types d'imperfection qui peuvent affecter les données temporelles**, en appliquant ces approches dans le cadre de CAPTAN MEMO, la prothèse de mémoire dédiée aux malades d'Alzheimer. Figure 41 présente les contributions de notre thèse.

10.1. Nos contributions

Dans ce travail de thèse, nous avons commencé par introduire notre typologie de l'imperfection des données temporelles, dans laquelle nous avons distingué des imperfections directes qui peuvent être déduites des données, et d'autres qui sont indirectes et qui peuvent être déduites de celles qui sont directes et des facteurs qui peuvent intervenir à la spécification du type d'imperfection.

Ensuite, nous avons traité certains types d'imperfections directes en utilisant différentes théories d'imperfection :

- En commençant par **l'imprécision**, notre approche consiste à représenter et raisonner sur **les dates et les horloges de temps précis et imprécis** et les relations qualitatives associées avec **une vue précise**. Pour cela, nous avons étendu **l'approche du 4D-Fluents** et **l'Algèbre d'intervalles d'Allen** en utilisant **Vilain et Kautz**. Nous avons considéré que les normes précises sont plus matures et prennent mieux en charge la mise à l'échelle que les outils flous.
- Après, en passant à **l'incertitude**, nous avons distingué **deux approches**. La première consiste à **représenter et raisonner sur les données temporelles** quantitatives (intervalles de temps et points de temps) **certaines et incertaines** et les relations qualitatives associées avec **une vue certaine**. Nous avons utilisé **des réseaux bayésiens** pour calculer **les degrés de certitudes**. La deuxième consiste à traiter **les données temporelles incertaines** en utilisant **la théorie de possibilité**. Nous avons **l'approche du 4D-Fluents** et **l'Algèbre d'intervalles d'Allen** dans les deux approches. Enfin nous avons comparé les deux approches proposées.
- Par la suite, nous avons pris en compte **les types d'imperfections composites**, précisément **l'incertitude et l'imprécision simultanément**. Nous avons proposé

une approche qui **représente et raisonne sur les données temporelles à la fois incertaine et imprécise** et les relations qualitatives associées en utilisant **la théorie des possibilités**. Nous avons proposé **une représentation sémantique** et nous avons étendu **l'Algèbre d'intervalles d'Allen** sur la base des mesures obtenues et des données traitées.

- Nous sommes passés finalement aux données temporelles **conflictuelles**. Nous avons proposé une approche qui représente et raisonne sur ces données en utilisant la **théorie de l'évidence**. Nous avons proposé **une représentation sémantique** et nous étendons **l'Algèbre d'intervalles d'Allen** sur la base des mesures obtenues et des données traitées.

Nous proposons, **pour chaque type d'imperfection, une ontologie** basée sur nos extensions de l'approche 4D-Fluents, nos représentations sémantiques basées sur l'ontologie OWL-Time et nos extensions de l'algèbre d'intervalles d'Allen. Nous avons implémenté **des prototypes basés sur ces ontologies** pour valider nos travaux.

Par ailleurs, nous avons proposé **une ontologie qui englobe toutes les imperfections traitées**. **L'ontologie de l'imperfection des données temporelles**, nommée « *TimeOntoImperfection* », a été ensuite intégrée dans CAPTAIN MEMO, la prothèse de mémoire, et évaluée sur un certain nombre de patients atteints de la maladie d'Alzheimer prouvant la pertinence de nos solutions proposées.

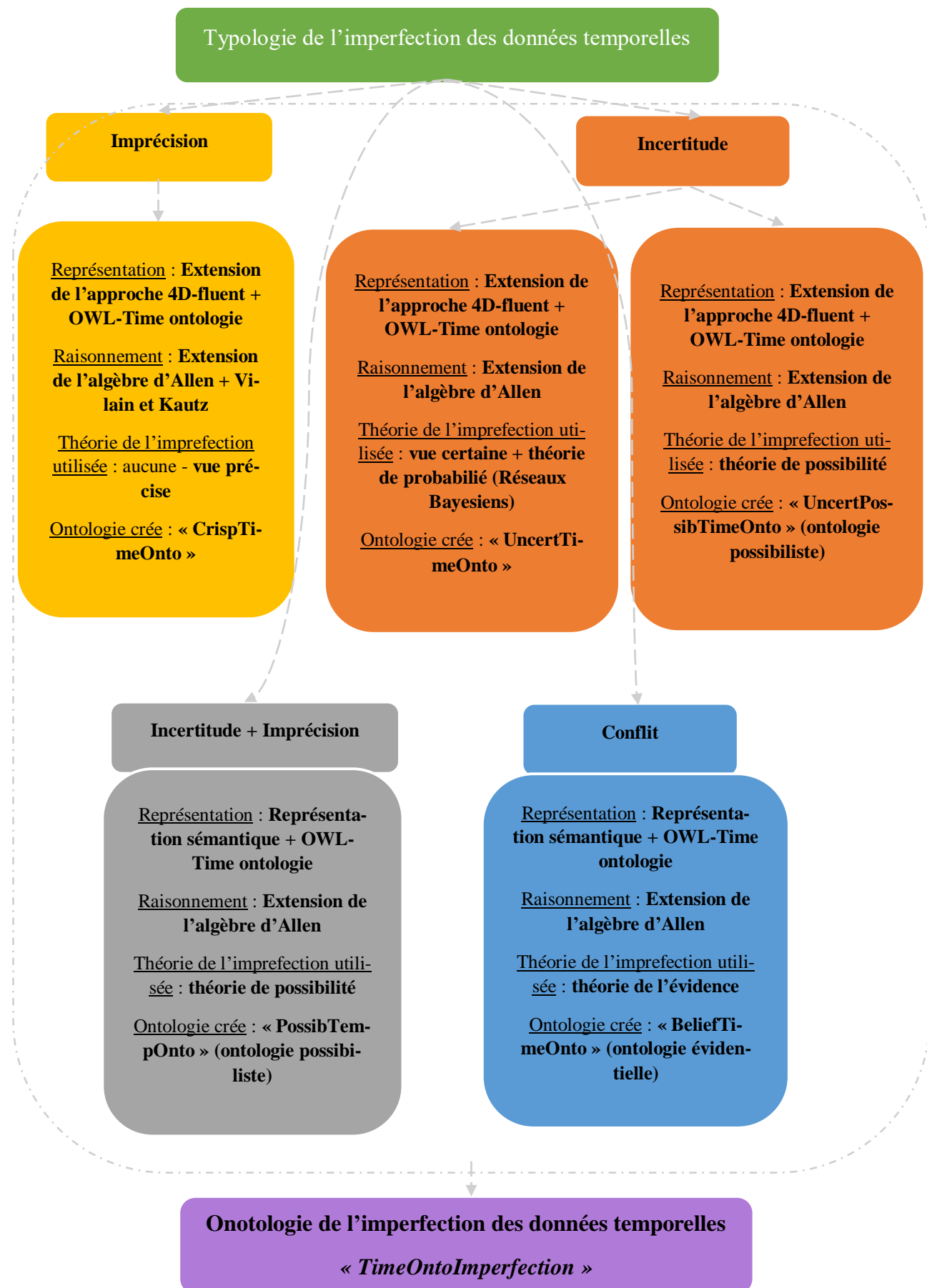


Figure 41. Les différentes contributions proposées dans notre thèse

10.2. Perspectives

Les perspectives suscitées par ces travaux sont multiples.

- Traiter les autres types d'imperfection des données temporelles définies dans la typologie comme la **redondance** et **l'inutilité**. **Exploiter d'autres théories** et notions pour traiter ces imperfections et les appliquer sur d'autres domaines d'application comme les données historiques. Et étendre l'ontologie de l'imperfection des données temporelles « *TimeOntoImperfection* » en se basant sur les résultats.
- Transformer les données temporelles imparfaites basées sur **le langage naturel** en des données temporelles quantitatives en exploitant les « **Big Data** ». Les données temporelles imparfaites sont exprimées en utilisant le langage naturel (par exemple : « A l'ère du Corona », « Pendant la première guerre mondiale » et « Quand j'ai eu mon bac »). Nous désirons représenter quantitativement ces données, c'est-à-dire sous la forme des intervalles de temps et des points de temps. Par exemple, « Pendant la Première Guerre mondiale » peut s'exprimer en l'intervalle de temps [1914, 1918]. Cependant, plusieurs aspects peuvent intervenir comme l'aspect multiculturel et l'aspect multilingue (Sagiroglu & Sinanc, 2013). Nous envisagerons de proposer une approche permettant de **transformer les données temporelles imparfaites basées sur le langage naturel en des données temporelles quantitatives** en exploitant les « **Big Data** » et **l'apprentissage automatique**.
- Représenter quantitativement les adverbes de temps comme par exemple « actuellement », « De nos jours » et « maintenant » et les transformer en des données quantitatives.
- Créer une **version Arabe de notre ontologie de l'imperfection des données temporelles** « *TimeOntoImperfection* » en nous basant sur « *LingOnto* » proposée par (Neji, et al., 2020) et « *Arabic Ontology* » proposée par (Jarrar, 2021), qui est un Wordnet arabe construit en utilisant des principes ontologiques, appelé l'ontologie arabe.

- Traiter les imperfections qui peuvent affecter d'autres types de données (non temporelles) comme **les données géographiques** et **les noms des personnes**. Nous visons à voir la possibilité d'appliquer nos résultats des approches qui traitent les données temporelles sur les autres types de données et dans le « Big Data ».

REFERENCES BIBLIOGRAPHIQUES

<p>(Achich et al., 2019)</p>	<p>Achich, N., Ghorbel, F., Hamdi, F., Métais, E. & Gargouri, F., 2019, September. <i>A Typology of Temporal Data Imperfection</i>. Vienne, s.n., pp. 305-311.</p>
<p>(Achich et al., 2019)</p>	<p>Achich, N., Ghorbel, F., Hamdi, F., Métais, E. & Gargouri, F., 2019. <i>Representing and Reasoning About Precise and Imprecise Time Points and Intervals in Semantic Web: Dealing with Dates and Time Clocks</i>. Linz, Austria, Springer, Cham.</p>
<p>(Achich et al., 2020)</p>	<p>Achich, N., Ghorbel, F., Hamdi, F., Métais, E. & Gargouri, F., 2020. <i>Approach to Reasoning about Uncertain Temporal Data in OWL 2.</i> Verona, Elsevier-Procedia Computer Science.</p>
<p>(Achich et al., 2021)</p>	<p>Achich, N., Ghorbel, F., Hamdi, F., Métais, E. & Gargouri, F., 2021. <i>Dealing with Uncertain and Imprecise Time Intervals in OWL2: A Possibility Theory-Based Approach</i>. s.l., Research Challenges in Information Science. RCIS 2021. Lecture Notes in Business Information Processing, vol 415. Springer, Cham.</p>
<p>(Achich et al., 2021)</p>	<p>Achich, N., Ghorbel, F., Hamdi, F., Métais, E. & Gargouri, F., 2021. <i>Handling Uncertain Time Intervals in Owl 2: Possibility Vs Probability Theories based Approaches</i>. s.l., IEEE.</p>
<p>(Akremi et al., 2016)</p>	<p>Akremi, H., Zghal, S., Jouhet, V. & Diallo, G., 2016. <i>FONTO: Une nouvelle méthode de la fuzzification d'ontologies</i>. s.l., s.n.</p>
<p>(Allen, 1983)</p>	<p>Allen, J., 1983. <i>Maintaining Knowledge about Temporal Intervals..</i> s.l., s.n., pp. 832-843.</p>

(Ammar, 2014)	Ammar, S., 2014. <i>Analyse et traitement possibiliste de signaux ultrasonores en vue d'assistance des non voyants</i> , Sfax: Doctoral dissertation, Ecole Nationale d'Ingénieurs de Sfax (ENIS).
(Anagnostopoulos et al., 2013)	Anagnostopoulos, E., Batsakis, S. & Petrakis, E., 2013. <i>CHRONOS: A Reasoning Engine for Qualitative Temporal Information in OWL</i> . s.l., s.n., pp. 70-77.
(Artale & Franconi, 2000)	Artale, A. & Franconi, E., 2000. <i>A Survey of Temporal Extensions of Description Logics</i> . s.l., s.n., pp. 70-77.
(Badaloni et al., 2006)	Badaloni, S. & Giacomini, M., 2006. <i>The Algebra IAfuz: a Framework for Qualitative Fuzzy Temporal Reasoning</i> . s.l., s.n., pp. 872-908.
(Baget, 2004)	Baget, J.-F., 2004. <i>Improving the Forward Chaining Algorithm for Conceptual Graphs Rules</i> . s.l., International Conference on Principles of Knowledge, pp. 407-414.
(Bal-Bourai & Mokhtari, 2016)	Bal-Bourai, S. & Mokhtari, A., 2016. π -SROIQ(D): <i>Possibilistic Description Logic for Uncertain Geographic Information</i> . s.l., Springer, Cham, pp. 818-829.
(Batsakis & Petrakis, 2011)	Batsakis, S. & Petrakis, E., 2011. <i>SOWL: A Framework for Handling Spatio-Temporal Information in OWL 2.0</i> . s.l., s.n., p. 242-249.
(Batsakis et al., 2017)	Batsakis, S., Petrakis, E. G., Tachmazidis, E. G. & Antoniou, G., 2017. Temporal representation and reasoning in OWL 2. <i>Semantic Web</i> , 8(6), pp. 981-1000.

(Bechhofer et al., 2004)	Bechhofer, S., Van Harmelen, F. & Hendler, J. I. H. D. M. P., 2004. OWL web ontology language reference. <i>W3C recommendation</i> , 10(2).
(Bedi & Marwaha, 2007)	Bedi, P. & Marwaha, S., 2007. Versioning OWL ontologies using temporal tags. <i>International Journal of Computer, Control, Quantum and Information Engineering</i> .
(Berners-Lee et al., 2001)	Berners-Lee, T., JHENDLER, A. & LASSILA, O., 2001. The semantic web. <i>Scientific American</i> , 284(5).
(Borst, 1997)	Borst, W., 1997. <i>Construction of Engineering Ontologies for Knowledge Sharing</i> , s.l.: Thèse de doctorat.
(Bouchon-Meunier, 1993)	Bouchon-Meunier, B., 1993. <i>La logique floue</i> . s.l.:s.n.
(Bouchon-Meunier, 1995)	Bouchon-Meunier, B., 1995. <i>La logique floue et ses applications</i> . s.l., s.n.
(Bourai et al., 2014)	Bourai, S. B., Mokhtari, A. & Khellaf, F., 2014. Poss-SROIQ(D): Possibilistic Description Logic Extension toward an Uncertain Geographic Ontology. <i>New Trends in Databases and Information Systems</i> , pp. 277-286.
(Brickley, 2004)	Brickley, D., 2004. RDF vocabulary description language 1.0: RDF schema.
(Carvalho et al., 2011)	Carvalho, R. N. et al., 2011. <i>Modeling a probabilistic ontology for maritime domain awareness</i> . s.l., IEEE, pp. 1-8.
(Casta, 2009)	Casta, J., 2009. <i>Incertitude et comptabilité</i> . s.l., s.n.

(Charlet et al., 2005)	Charlet, J., Bachimont, B. & Troncy, R., 2005. Le Web sémantique. <i>Hors série de la revue Information-Interaction-Intelligence</i> , Volume 4, pp. 7-20.
(Chomicki, 1994)	Chomicki, J., 1994. <i>Temporal query languages: A survey</i> . Berlin, Springer, p. 506–534.
(Costa, 2005)	Costa, P., 2005. <i>Bayesian Semantics for the Semantic Web</i> . s.l., George Mason University, p. 312.
(Costa et al., 2008)	Costa, P. C. et al., 2008. <i>A First-Order Bayesian Tool for Probabilistic Ontologies</i> . Coconut Grove, Florida, USA, s.n.
(Cui et al., 2010)	Cui, . T. et al., 2010. CNTRO: a Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives.. <i>AMIA Annual Symp.</i> , p. 787.
(Dempster, 1967)	Dempster, A. P., 1967. Upper and lower probabilities induced by a multivalued mapping. <i>The annals of mathematical statistics</i> , p. 325–339.
(Desjardin et al., 2012)	Desjardin, E., Nocent, O. & de Runz, C., 2012. <i>Prise en compte de l'imperfection</i> . Paris, Actes des 2èmes Journées d'Informatique et Archéologie de Paris – JIAP, pp. 385-396.
(Devaux & Lamanna, 2009)	Devaux M. & Lamanna M. The rise and early history of the term ontology (1606-1730). <i>Quaestio</i> , 9:173–208, 2009.
(Ding & Peng, 2004)	Ding, Z. & Peng, Y., 2004. <i>A probabilistic extension to ontology language OWL</i> . s.l., Proceedings of the IEEE.

(Dubois & Prade, 1988)	Dubois, D. & Prade, H., 1988. Possibility Theory : An Approach to Computerized Processing of Uncertainty (traduction revue et augmentée de "Théorie des Possibilités"). <i>Plenum Press, New York.</i>
(Dubois & Prade, 1989)	Dubois, D. & Prade, H., 1989. <i>Processing Fuzzy Temporal Knowledge</i> . s.l., s.n., pp. 729-744.
(Dubois & Prade, 1990)	Dubois, D. & Prade, H., 1990. <i>Teoria Vozmojnosti (Traduction Russe de" Théorie des Possibilités", Masson, Paris, 1985/1987)</i> . Radio i Sviaz, Moscou. éd. s.l.:s.n.
(Dubois & Prade, 1998)	Dubois, D. & Prade, H., 1998. Possibility theory: qualitative and quantitative aspects. <i>Quantified representation of uncertainty and imprecision - Springer,Dordrecht.</i> , pp. 169-226.
(Dubois & Prade, 2006)	Dubois, D. & Prade, H., 2006. <i>Théorie des possibilités</i> . REE éd. s.l.:s.n.
(Ermolayev et al., 2014)	Ermolayev, V. et al., 2014. Ontologies of Time: Review and Trends. <i>International Journal of Computer Science & Applications</i> , 11(3).
(Ermplayev et al., 2008)	Ermolayev, V., Keberle, N., Matzke, W.-E. & Sohnius, . R., 2008. <i>Fuzzy Time Intervals for Simulating Actions</i> . s.l., LNBIP, p. 429–444.
(Essaid & Yaghlane, 2009)	Essaid, A. & Yaghlane, B. B., 2009. <i>BeliefOWL: An Evidential Representation in OWL Ontology</i> . s.l., s.n., pp. 77-80.

(Etzion, 1998)	Etzion, O., 1998. <i>Temporal databases: research and practice</i> . s.l.:Springer Science & Business Media.
(Fareh, 2019)	Fareh, M., 2019. Modeling incomplete knowledge of semantic web using Bayesian networks. <i>Applied Artificial Intelligence</i> , 33(11), pp. 1022-1034.
(Fisher, 2003)	Fisher, P., 2003. Models of uncertainty in spatial data. <i>Geographical information systems</i> , p. 191–203.
(Fisher et al., 2005)	Fisher, P., Comber, A. & Wadsworth, R., 2005. <i>Nature de l'incertitude pour les données spatiales</i> . Paris, Qualité de l'information géographique, pp. 49-64.
(Frasincar et al., 2010)	Frasincar, F., Milea, V. & Kaymak, U., 2010. tOWL: Integrating Time in OWL. <i>Semantic Web Information Management: A Model-Based Perspective</i> , p. 225–246.
(Freska, 1992)	Freksa, C., 1992. <i>Temporal Reasoning Based on Semi-Intervals</i> . s.l., 54, p. 199–227.
(Gammoudi et al., 2017)	Gammoudi, A., Hadjali, A. & Yaghlane, B. B., 2017. <i>Fuzz-TIME: an intelligent system for managing fuzzy temporal information</i> . s.l., s.n., p. 200–222.
(Gavignet et al., 2016)	Gavignet, E., Leclercq, E., Cullot, N. & Savonnet, M., 2016. Raisonner en logique modale sur l'incertitude liée aux données - application en archéologie. <i>Revue Internationale de Geomatique</i> , 26(04), p. 467–490.
(Ghorbel, 2018)	Ghorbel, F., 2018. <i>Dialogue graphique intelligent, fondé sur une ontologie, pour une prothèse de mémoire</i> , Paris: Doctoral dissertation, Paris, CNAM.

<p>(Ghorbel et al., 2020)</p>	<p>Ghorbel, F. Hamdi F. Achich N. Métais E., 2020. Handling data imperfection—False data inputs in applications for Alzheimer’s patients. <i>Data & Knowledge Engineering</i>, Volume 130.</p>
<p>(Ghorbel et al., 2018)</p>	<p>Ghorbel, F., Métais, E. & Hamdi, F., 2018. <i>A Crisp-Based Approach for Representing and Reasoning on Imprecise Time Intervals in OWL 2.</i> s.l., Springer, Cham, pp. 640-649.</p>
<p>(Ghorbel H. et al., 2013)</p>	<p>Ghorbel, H., Maalej, S. & Bouaziz, R., 2013. <i>Un framework pour la génération semi-automatique d’ontologies floues : Text2FuzzyOnto.</i> s.l., s.n., p. 671–699.</p>
<p>(Ghorbel et al., 2018)</p>	<p>Ghorbel, F. et al., 2018. <i>A Fuzzy-Based Approach for Representing and Reasoning on Imprecise Time Intervals in Fuzzy-OWL 2 Ontology.</i> Paris, Springer, Cham., pp. 167-178.</p>
<p>(Grandi, 2013)</p>	<p>Grandi, F., 2013. <i>Dynamic multi-version ontology-based personalization.</i> s.l., s.n., pp. 224-232.</p>
<p>(Grandi, 2016)</p>	<p>Grandi, F., 2016. Dynamic class hierarchy management for multi-version ontology-based personalization. <i>Journal of Computer and System Sciences</i>, 82(1), pp. 69-90.</p>
<p>(Gruber, 1993)</p>	<p>Gruber, T., 1993. A translation approach to portable ontology specifications. <i>Knowledge Acquisition Journal Academic Press</i>, pp. 199-220.</p>

(Guesgen et al., 1994)	Guesgen, H. W., Hertzberg, . J. & Philpott, . A., 1994. <i>Towards Implementing Fuzzy Allen Relations</i> . s.l., s.n., pp. 49-55.
(Gutierrez et al., 2006)	Gutierrez, C., Hurtado, C. A. & Vaisman, A., 2006. Introducing time into RDF.. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 19(2), pp. 207-218.
(Gutierrez et al., 2005)	Gutierrez, C., Hurtado, C. & Vaisman, A., 2005. <i>Temporal RDF</i> . s.l., Springer, p. 93–107.
(Gu et al., 2004)	Gu, T., Pung, H. K., Zhang, D. Q. & Pung, . H. K., 2004. <i>A bayesian approach for dealing with uncertain contexts</i> . s.l., s.n., pp. 205-210.
(Harbelot et al., 2013)	Harbelot, B., Arenas, H. & Cruz, C., 2013. <i>Continuum: A Spatiotemporal Data Model to Represent and Qualify Filiation Relationships</i> . s.l., s.n., pp. 76-85.
(Herradi, 2018)	Herradi, N., 2018. <i>Représentation sémantique multilingue, multiculturelle et temporelle des relations interpersonnelles, appliquée à une prothèse de mémoire</i> , Paris: Doctoral dissertation, Conservatoire national des arts et metiers-CNAM.
(Herradi et al., 2017)	Herradi, N., Hamdi, F. & Métails, E., 2017. <i>A Semantic Representation of Time Intervals in OWL2</i> . Madeira, Portugal, s.n., pp. 1-8.
(Hlel et al., 2015)	Hlel, E., Jamoussi, S. & Hamadou, A., 2015. <i>A probabilistic ontology for the prediction of author's Interests</i> . s.l., Springer, Cham, pp. 492-501.

(Hlel et al., 2018)	Hlel, E., Jamoussi, . S. & Hamadou, A. B., 2018. A new method for building probabilistic ontology (prob-ont). <i>Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications</i> , pp. 1409-1434.
(Hobbs & Pan, 2004)	Hobbs, J. R. & Pan, F., 2004. An Ontology of Time for the Semantic Web. <i>ACM Transactions on Asian Language Processing (TALIP): Special issue on Temporal Information Processing</i> , 3(1), p. 66–85.
(Hurtado & Vaisman, 2006)	Hurtado, C. & Vaisman, A., 2006. <i>Reasoning with Temporal Constraints in RDF</i> . s.l., s.n., p. 164–178.
(Jarrar, 2021)	Jarrar, M., 2021. The Arabic ontology—an Arabic wordnet with ontologically clean content. <i>Applied ontology</i> , pp. 1-26.
(Kaur & Kaur, 2012)	Kaur, T. & Kaur, A., 2012. Extension of a crisp ontology to fuzzy ontology. <i>International Journal Of Computational Engineering Research</i> , pp. 201-207.
(Keberle et al., 2007)	Keberle, N., Litvinenko, Y., Gordeyev, Y. & Ermolay, V., 2007. <i>Ontology evolution analysis with OWL-MeT</i> . s.l., s.n., p. 1–12.
(Kim et al., 2008)	Kim, S.-K. et al., 2008. <i>Temporal Ontology Language for Representing and Reasoning Interval-Based Temporal Knowledge</i> . s.l., 5367, pp. 31-45.
(Klein & Fensel, 2001)	Klein, M. C. A. & Fensel, D., 2001. <i>Ontology Versioning on the Semantic Web</i> . California, s.n., p. 75–91.

(Klir & Yuan, 1995)	Klir, G. J. & Yuan, B., 1995. Fuzzy sets and fuzzy logic: theory and applications. <i>Prentice Hall New Jersey</i> , Volume 574.
(Koubarakis & Kyzirakos, 2010)	Koubarakis, M. & Kyzirakos, K., 2010. <i>Modeling and Querying Metadata in the Semantic Sensor Web: The model stRDF and the query language stSPARQL</i> . s.l., s.n., p. 425–439.
(Krieger, 2010)	Krieger, H.-U., 2010. <i>A General Methodology for Equipping Ontologies With Time</i> . s.l., s.n., p. 3165–3172.
(Lamine & Omri, 2012)	Lamine, R. B. & Omri, M. N., 2012. Information Retrieval Model Based on Possibilistic Ontology. <i>International Journal of Computer Science</i> , 3(9).
(Lamma et al., 2006)	Lamma, E., Riguzzi, F. & Storari, S., 2006. Improving the K2 algorithm using association rule parameters. <i>Modern Information Processing</i> , pp. 207-217.
(Lee et al., 2009)	Lee, J., Chae, H., Kim, C.-H. & Kim, K., 2009. Design of product ontology architecture. <i>Expert Systems with Applications</i> , p. 2300–2309.
(Li et al., 2020)	Li, F. et al., 2020. Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events.. <i>Journal of the American Medical Informatics Association</i> , 27(7), pp. 1046-1056.
(Loiseau, 2004)	Loiseau, Y., 2004. <i>Recherche flexible d'information par filtrage flou qualitatif</i> , Université Paul Sabatier, Toulouse: Thèse de doctorat.

(Lutz et al., 2008)	Lutz, C., Wolter, F. & Zakharyashev, M. K., 2008. <i>Temporal Description Logics: A Survey</i> .. s.l., s.n., p. 3–14.
(Madkour et al., 2016)	Madkour, M., Benhaddou, D. & Tao, C., 2016. Temporal data representation, normalization, extraction, and reasoning: A review from clinical domain. <i>Computer methods and programs in biomedicine</i> , Volume 128, pp. 52-68.
(Messaoudi et al., 2019)	Messaoudi, W., Farah, M. & Farah, I. R., 2019. Fuzzy Spatio-Spectro-Temporal Ontology for Remote Sensing Image Annotation and Interpretation: Application to Natural Risks Assessment.. <i>International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems</i> , 27(05), pp. 815-840.
(Métais et al., 2012)	Métais, E. et al., 2012. Memory prosthesis. <i>Non-pharmacological Therapies in Dementia</i> , 3(2), p. 177.
(Milea et al., 2012)	Milea, V., Frasinca, F., Kaymak, U. & Houben, G. J., 2012. Temporal optimisations and temporal cardinality in the tOWL language.. <i>International Journal of Web Engineering and Technology</i> , 7(1), pp. 45-64.
(Motik, 2010)	Motik, B., 2010. <i>Representing and Querying Validity Time in RDF and OWL: a Logic-Based Approach</i> . s.l., s.n., p. 550–565.
(Motik, 2012)	Motik, B., 2012. Representing and querying validity time in RDF and OWL: A logic-based approach.. <i>Web Semantics: Science, Services and Agents on the World Wide Web</i> , Volume 12, pp. 3-21.

(Nagypál & Motik, 2003)	Nagypál, G. & Motik, B., 2003. <i>A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies</i> . s.l., Springer, p. 906–923.
(Neji et al., 2020)	Neji, M. et al., 2020. <i>A Multilingual Linguistic Domain Ontology</i> . s.l., s.n., pp. 199-206.
(Nidkanen, 1989)	Niskanen, V. A., 1989. <i>Introduction to imprecise reasoning. Uncertainty, decision-making and knowledge engineering</i> . Finland, Publications of the Society for Artificial Intelligence.
(Njah et al., 2016)	Njah, H., Jamoussi, S., Mahdi, W. & Elati, M., 2016. <i>A Bayesian approach to construct Context-Specific Gene Ontology: Application to protein function prediction..</i> s.l., IEEE, pp. 1-6.
(Noy et al., 2006)	Noy, N., Rector, A., Hayes, P. & Welty, C., 2006. <i>Defining N-Ary Relations on the Semantic-Web</i> . s.l., s.n.
(Noy & Rector, 2006)	Noy, N. & Rector, A., 2006. <i>Defining N-Ary Relations on the Semantic-We. W3C Working Group</i> .
(Nys et al., 2018)	Nys, G. A., Van Ruymbeke, M. & Billen, R., 2018. <i>Spatio-temporal reasoning in CIDOC CRM: an hybrid ontology with GeoSPARQL and OWL-Time</i> . RWTH Aachen University, CEUR Workshop Proceedings.
(O’Connor & Das, 2011)	O’Connor, M. J. & Das, A., 2011. <i>A Method for Representing and Querying Temporal Information in OWL</i> . s.l., s.n., pp. 97-110.

(Oltenau et al., 2006)	Olteanu, A. M., Mustière, S., Ruas, A. & Desveaux, E., 2006. <i>L'apport de données spatiales pour une base de données ethnographique</i> . s.l., Actes de Conférence SAGEO.
(Papadakis et al., 2011)	Papadakis, N. et al., 2011. <i>PROTON: A prolog reasoner for temporal ontologies in OWL</i> . s.l., s.n., p. 14660–14667.
(Pearl, 1988)	Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. <i>Morgan Kaufmann-Elsevier</i> .
(Pons et al., 2014)	Pons, . J., Billiet, C., Pons, O. & De Tré, G., 2014. Aspects of dealing with imperfect data in temporal databases. Dans: <i>Flexible Approaches in Data, Information and Knowledge Management</i> . s.l.:Springer, Cham, pp. 189-220.
(Preventis et al., 2012)	Preventis, A., Polyxeni, M., Petrakis, E. G. M. & Batsakis, S., 2012. <i>CHRONOS: A Tool for Handling Temporal Ontologies in Protégé</i> . s.l., s.n., p. 460–467.
(Rodríguez et al., 2014)	Rodríguez, N. D., Cuéllar, M. P., Lilius, J. & Calvo-Flores, M. D., 2014. A fuzzy ontology for semantic modelling and recognition of human behavior. <i>Knowledge-Based Systems</i> , pp. 46-60.
(Ryabov & Trudel, 2004)	Ryabov, V. & Trudel, A., 2004. <i>Probabilistic Temporal Interval Networks</i> . s.l., s.n., pp. 64-67.
(Sadeghi & Goertzel, 2014)	Sadeghi, K. M. M. & Goertzel, B., 2014. <i>Uncertain Interval Algebra via Fuzzy/Probabilistic Modeling</i> . s.l., s.n., pp. 591-598.

(Sagiroglu & Sinanc, 2013)	Sagiroglu, S. & Sinanc, D., 2013. <i>Big data: A review</i> . s.l., s.n., pp. 42-47.
(Salem et al., 2017)	Salem, Y. B. et al., 2017. <i>Ontology based possibilistic reasoning for breast cancer aided diagnosis</i> . s.l., Springer, Cham, pp. 353-366.
(Santos et al., 2011)	Santos, J., Braga, L. & Cohn, A. G., 2011. FONTE: A Protege Plug-in for Engineering Complex. <i>Enterprise Information Systems</i> , p. 222–236.
(Shafer, 1976)	Shafer, G., 1976. A mathematical theory of evidence. <i>Princeton university press</i> .
(Smets, 1997)	Smets, P., 1997. <i>Imperfect Information: Imprecision and Uncertainty</i> . uncertainty management in information systems: from needs to solutions éd. s.l.:Springer.
(Snoussi & Davoine, 2014)	Snoussi, M. & Davoine, P.-A., 2014. Methodological proposals to handle imperfect spatial and temporal information in the context of natural hazard studies. <i>International Journal of Geomatics and Spatial Analysis</i> , 23(3-4), pp. 495-517.
(Sta, 2017)	Sta, H. B., 2017. Quality and the efficiency of data in “Smart-Cities. <i>Future Generation Computer Systems</i> , Volume 74, pp. 409-416.
(Studer et al., 1998)	Studer R., Benjamins VR., & Fensel D.. Knowledge engineering: principles and methods. <i>Data & knowledge engineering</i> , 25(1-2):161–197, 1998.

(Swartout et al., 1999)	Swartout W., Russ T., & Macgregor R.. Ontologies. <i>Intelligent Systems and their Applications, IEEE</i> , 1094:18–19, 1999.
(Tappolet & Bernstein, 2009)	Tappolet, J. & Bernstein, A., 2009. <i>Applied Temporal RDF: Efficient Temporal Querying of RDF Data with SPARQL</i> . s.l., Springer, pp. 308-322.
(Vetro et al., 2016)	Vetrò, A. et al., 2016. Open data quality measurement framework: Definition and application to Open Government Data. <i>Quarterly, Government Information</i> , Volume 33, pp. 325-337.
(Vilain & Kautz, 1986)	Vilain, M. & Kautz, . H. A., 1986. <i>Constraint Propagation Algorithms for Temporal Reasoning</i> . Philadelphia, s.n., p. 377–382.
(Vila, 1994)	Vila, . L., 1994. A survey on temporal reasoning in artificial intelligence. <i>AICOM (Artificial Intelligence Communications)</i> , Issue 7, pp. 4-28.
(Völkel & Groza, 2006)	Völkel, M. & Groza, T., 2006. <i>SemVersion: An RDF-based ontology versioning system</i> . s.l., s.n., p. 44.
(Welty & Fikes, 2006)	Welty, C. & Fikes, R., 2006. <i>A Reusable Ontology for Fluents in OWL</i> . s.l., s.n., p. 226–336.
(Zadeh, 1978)	Zadeh, L., 1978. Fuzzy sets as a basis for a theory of possibility. <i>Fuzzy Sets and Systems</i> , Issue 1, pp. 3-28.
(Zekri et al., 2017)	Zekri, A., Brahmia, Z., Grandi, F. & Bouaziz, R., 2017. Temporal schema versioning in τ OWL: a systematic approach for the management of time-varying knowledge. <i>Journal of Decision systems</i> , 26(2), pp. 113-137.

(Zekri et al., 2015)	Zekri, F., Bouaziz, R. & Turki, E., 2015. <i>A fuzzy-based ontology for Alzheimer's disease decision support</i> . s.l., IEEE, pp. 1-6.
-----------------------------	---

Nassira ACHICH

Traitement de l'imperfection des données temporelles saisies par l'utilisateur - Application aux logiciels destinés à des patients malades d'Alzheimer

Résumé

Notre thèse comprend plusieurs contributions. La première porte sur la définition d'une typologie de l'imperfection des données temporelles dans laquelle nous listons les différentes imperfections qui peuvent affecter les données temporelles. Nous les classifions en imperfections directes et imperfections indirectes. Par la suite, nous traitons les différentes imperfections définies dans la typologie. Nous proposons une approche qui traite l'imprécision des données temporelles en représentant et raisonnant sur les dates et les horloges du temps dans l'ontologie avec une vue précise. Pour cela, nous étendons l'approche du 4D-Fluents et l'algèbre d'intervalles d'Allen. Nous proposons une ontologie basée sur nos extensions nommée « CrispImpTimeOnto ». Nous traitons l'incertitude des données temporelles dans l'ontologie, tout d'abord, en utilisant les réseaux bayésiens pour calculer les mesures d'incertitude et la deuxième en utilisant la théorie des possibilités pour calculer les mesures de nécessité. Nous étendons l'approche du 4D-Fluents pour la représentation et l'Algèbre d'intervalle d'Allen pour le raisonnement pour les deux approches. Nous proposons deux ontologies nommées « UncertTimeOnto » et « UncertPos-sibTimeOnto ». Nous conduisons une comparaison entre les deux approches proposées. Nous passons par la suite aux types d'imperfection composite. Nous présentons une contribution qui consiste à traiter, à la fois, l'incertitude et l'imprécision des données temporelles dans l'ontologie en utilisant la théorie des possibilités. Nous proposons une ontologie, nommée « UncertPossibTimeOnto ». Nous traitons le conflit des données temporelles dans l'ontologie en utilisant la théorie de l'évidence pour calculer les masses de croyances des données. Nous proposons une représentation sémantique en nous basant sur les mesures. Nous étendons l'algèbre d'Allen pour raisonner sur les données temporelles conflictuelles. Nous proposons une ontologie, nommée « BeliefTimeOnto » basée sur nos extensions. Sur la base de nos ontologies proposées, nous proposons une ontologie de l'imperfection des données temporelles nommée « TemporalOntoImperfection ». Nous l'intégrons dans la prothèse de mémoire CAPTAIN MEMO et nous évaluons nos approches traitant les différents types d'imperfection. Cette évaluation fait intervenir des personnes atteintes de la maladie d'Alzheimer.

Mots-clés : Ontologie, Imperfection des données temporelles, théorie de possibilité, théorie de l'évidence, réseaux bayésiens, l'approche 4D-Fluents, l'algèbre d'intervalles d'Allen, Alzheimer.

Résumé en anglais

In our thesis, we propose the typology of temporal data imperfection in which we elaborate the various imperfections that can affect temporal data. Then, we move to dealing with these imperfections in ontology. We propose an approach to deal with imprecision by representing and reasoning about dates and time clocks and associated qualitative relations in ontology with a crisp view. Therefore, we extend the 4D-Fluents approach for the representation and the Allen's interval algebra for the reasoning. Then, we create an ontology based on our extensions. When dealing with uncertain temporal data in ontology and associated qualitative relations, we propose an approach in which we first calculate certainty degrees using Bayesian networks, then using possibility theory. We extend the 4D-Fluents approach to represent the handled data and we extend Allen's interval algebra for the reasoning part. Then, we conduct a small comparison between the two proposed approaches. We propose an ontology based on our extensions. Temporal data may, also, be affected by multiple types of imperfections at the same time. Thus, we present a contribution which consists in dealing with simultaneously uncertain and imprecise temporal data in ontology using possibility theory. We propose an ontology based on our extensions. We deal with conflicting temporal data in ontology by using the evidential theory to calculate the belief masses of the handled data. We propose a semantic representation based on the obtained measures. We also extend Allen's algebra to reason about conflicting temporal data. We propose an ontology based on our extensions. Based on our ontologies, we propose an ontology of temporal data imperfection called « TemporalOntoImperfection » that encompasses all the treated imperfections. We integrate it within CAPTAIN MEMO, the memory prosthesis based on the PersonLink ontology dedicated to Alzheimer's patients, to evaluate our approaches. This assessment involves people with Alzheimer's disease.

Keywords: Ontology, Temporal Data Imperfection, Possibility Theory, Evidential Theory, Bayesian Networks, 4D-Fluents Approach, Allen's Interval Algebra, Alzheimer's.