

Traitement des données temporelles certaines et incertaines en OWL 2 : Approche basée sur la théorie des probabilités

Nassira Achich^{1,2}, Fatma Ghorbel^{1,2}, Fayçal Hamdi², Elisabeth Métais², Faïez Gargouri¹

¹ Laboratoire MIRACL, Université de Sfax, Tunisie

² Laboratoire CEDRIC, Conservatoire National des Arts et Métiers,
Paris, France

Résumé

Les données temporelles fournies par les patients atteints d'Alzheimer sont sujettes à l'incertitude. De nombreuses approches ont été proposées pour traiter des données temporelles certaines, mais non pas celles qui sont incertaines. Cet article propose une approche pour représenter et raisonner sur des intervalles et des points de temps quantitatifs certains et incertains et les relations qualitatives entre eux. Elle inclut trois volets : (1) une extension de l'approche 4D-fluents avec de nouvelles composantes ontologiques pour représenter des données temporelles certaines et incertaines. (2) une extension de l'algèbre des intervalles d'Allen pour raisonner sur des intervalles de temps certains et incertains. Une adaptation de relations cette algèbre pour relier un intervalle de temps et un point de temps, et deux points de temps. (3) Enfin une ontologie qui intègre toutes ces extensions. Un prototype a été implémenté et intégré dans une prothèse de mémoire pour les patients atteints de la maladie d'Alzheimer afin de gérer des entrées de données incertaines.

Mots-clés

Approche 4D-Fluents, Algèbre d'intervalle d'Allen, Réseaux bayésiens, OWL 2, Raisonnement temporel, Représentation temporelle, Données temporelles incertaines.

Abstract

Temporal data given by Alzheimer's patients are subject to uncertainty. Many approaches have been proposed to deal with certain temporal data and neglect uncertain ones. This paper proposes an approach to representing and reasoning about certain and uncertain quantitative intervals and time points and the qualitative relations between them. It includes three contributions. (1) We extend the 4D-fluent approach with new ontological components to represent certain and uncertain temporal data. (2) We extend Allen's interval algebra to reason about certain and uncertain time intervals. We adapt these relations to relate a time interval and a time point, and two time points. (3) We propose an ontology based on our extensions. A prototype is implemented and integrated into an ontology-

based memory prosthesis for patients with Alzheimer's disease to handle uncertain data inputs.

Keywords

4D-Fluents Approach, Allen's Interval Algebra, Bayesian Network, OWL 2, Temporal Reasoning, Temporal Representation, Uncertain Temporal Data.

1 Introduction

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 [31], qui est une prothèse de mémoire pour les patients d'Alzheimer. Cette prothèse offre un ensemble de services tels que l'aide aux utilisateurs pour se souvenir de leurs proches. Elle est basée sur une ontologie OWL 2, appelée « *PersonLink*¹ » [23]. Cette dernière permet de modéliser et de raisonner sur des relations interpersonnelles (e.g. parent, voisin) et de décrire des personnes. Un patient Alzheimer peut saisir dans CAPTAIN MEMO des données incertaines telles que « *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 entre eux, 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.

Dans cet article, nous présentons notre approche proposée dans [1], 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. Cette approche inclut trois volets. Le premier concerne la représentation des données temporelles certaines et incertaines

¹<http://cedric.cnam.fr/isid/ontologies/PersonLink.owl#>

en OWL 2. Pour cela nous avons étendu l'approche 4D-fluents [40], qui ne couvre que les intervalles de temps certains dans OWL, avec de nouvelles composantes ontologiques certaines pour représenter des données temporelles quantitatives certaines et incertaines, et des relations temporelles qualitatives entre les intervalles de temps et les points. Nous avons utilisé les réseaux bayésiens pour calculer les mesures de certitudes. Le deuxième volet concerne le raisonnement sur les données temporelles certaines et incertaines en étendant l'algèbre d'intervalle d'Allen [3]. 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. Enfin le dernier volet est la proposition d'une ontologie OWL 2 appelée « *UncertTimeOnto* » qui peut être intégrée dans d'autres ontologies pour traiter des données temporelles certaines et incertaines. Elle a été implémentée sur la base des extensions proposées. Les inférences sont effectuées à l'aide de règles SWRL intégrées dans l'ontologie.

Le reste de cet article est organisé comme suit. Les travaux proches sont passés en revue dans la section 2. La section 3 présente notre extension de l'approche 4D-fluents pour représenter les données temporelles incertaines dans OWL 2. La section 4 présente notre extension de l'algèbre d'Allen pour raisonner sur les données temporelles incertaines. La section 5 présente notre ontologie « *UncertTimeOnto* » basée sur nos extensions. La section 6 introduit les expérimentations de notre approche dans le cadre de CAPTAIN MEMO. Dans la dernière section, nous terminons par une conclusion.

2 Travaux proches

Les données temporelles peuvent être certaines ou incertaines. Elles sont également caractérisées en termes quantitatifs ou qualitatifs. Des données temporelles quantitatives certaines signifient des intervalles de temps et des points de temps certains. Des données temporelles quantitatives incertaines signifient des intervalles de temps et des points de temps incertains. Les intervalles de temps incertains sont des intervalles de temps normaux caractérisés par des limites de début et/ou de fin incertaines (par exemple, « *Peut-être de 2015 à 2018* »). Les points de temps incertains sont des points de temps qui sont définis de manière incertaine (par exemple, « *Je ne suis pas sûr si c'était en 2010* »). Les relations temporelles qualitatives lient deux intervalles de temps (Intervalle-Intervalle), un intervalle de temps et un point de temps (Intervalle-Point et Point-Intervalle) ou deux points de temps (Point-Point) ; où les intervalles de temps et les points peuvent être certains ou incertains. Les données temporelles qualitatives peuvent également être certaines ou incertaines et peuvent être déduites de données quantitatives.

Dans cette section, nous introduisons les travaux connexes relatifs à la représentation et au raisonnement sur les données

temporelles dans le web sémantique.

2.1 Représentation des données temporelles dans le web sémantique

La représentation des données temporelles sous forme d'ontologie est un besoin crucial. Cependant, les langages ontologiques 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 dans le domaine du web sémantique.

Nous avons identifié dans la littérature plusieurs approches permettant la modélisation de la dimension temporelle dans le domaine du web sémantique : les logiques de description temporelles [5], le versioning [28], la réification [10], les relations N-aires [33] et l'approche 4D-fluents. Nous avons classé 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 [21].

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 conduit pas à une redondance au niveau de la représentation des données. Cependant, elle nécessite l'extension de OWL ou RDF, qui est une tâche fastidieuse. Les domaines concrets [30] est aussi une approche qui nécessite l'introduction de types de données et d'opérateurs supplémentaires dans OWL. OWL-MeT [11] et TL-OWL [27] sont des implémentations de cette approche. Le RDF temporel, qui appartient également à la première catégorie, n'utilise que des triplets RDF. Il n'a pas toute l'expressivité du langage OWL et ne permet pas d'exprimer des relations qualitatives. Dans [26], les auteurs présentent un cadre complet pour incorporer le raisonnement temporel dans RDF. [29] introduisent un modèle de données de contraintes pour représenter les données temporelles, qui étend RDF, nommé stRDF. Ils étendent SPARQL pour interroger stRDF.

Le versioning, la réification, « *N-ary relations* », « *4D-Fluents* » et les graphes nommés [39] sont des approches qui appartiennent à la deuxième catégorie. Le versioning est décrit comme la capacité de gérer les changements dans l'ontologie en créant différentes versions de l'ontologie. [41] proposent une approche pour la gestion des versions de schéma dans OWL. Son principal problème est que toutes les versions sont indépendantes les unes des autres ce qui nécessite des recherches exhaustives dans chacune d'elles. La réification est une technique de représentation des relations N-aires lorsque seules les relations binaires sont autorisées. Un nouvel objet est créé chaque fois qu'une relation temporelle doit être représentée. « *N-ary relations* » est une approche qui propose de représenter une relation N-aire comme deux propriétés liées chacune à un nouvel objet tout en maintenant la sémantique des propriétés. Cette approche souffre d'un problème de redondance des données. [35] proposent un plug-in pour

PROTÉGÉ, nommé CHRONOS, qui utilise cette approche pour gérer les données temporelles. Une solution aux problèmes de redondance, dont souffre « N-ary relations », a été proposée par l'approche 4D-fluents qui représente les intervalles de temps et leur évolution dans OWL. Dans cette approche les changements ne se produisent que dans les parties temporelles et les concepts variant dans le temps sont représentés comme des objets à 4 dimensions avec la 4ème dimension étant les données temporelles. Plusieurs travaux ont utilisé l'approche 4D-fluents. La plupart d'entre eux, qui sont [34], [7], [22], [4], [8] et [24], l'ont utilisé dans le contexte de données temporelles considérées comme précises et certaines, ce qui ne correspondent pas à notre contexte. D'autres l'ont étendue pour représenter des données temporelles imparfaites, en particulier des données temporelles imprécises comme [16], qui ne correspondent pas non plus à notre contexte. A notre connaissance, il n'existe aucune approche pour représenter des données temporelles incertaines dans OWL 2. L'approche des graphes nommés représente chaque intervalle de temps par exactement un graphe nommé, où tous les triplets appartenant partagent la même période de validité.

Toutes les approches passées en revue ne traitent que les données temporelles certaines et négligent celles qui sont incertaines. Elles ne sont pas destinées à gérer des points de temps et des relations temporelles qualitatives entre un intervalle de temps et un point de temps ou même deux points de temps. Un critère de base dans le choix de notre approche est qu'elle doit s'appuyer sur des constructeurs déjà définis dans OWL. Par conséquent, nous excluons les approches « *Temporal Description Logic* », « *Concrete Domain* » et « *Temporal RDF* ». Nous excluons également l'approche « *Named Graphs* », car elle ne prend pas en charge OWL et n'est pas une solution conforme aux standards W3C.

Notre choix s'est porté sur l'approche 4D-fluents que nous avons étendue pour représenter des données temporelles quantitatives incertaines et les relations temporelles qualitatives associées. Comparée aux approches de réification, de « N-ary relation » et de versioning, l'approche 4D-fluents minimise le problème de redondance des données, car les changements ne se produisent que sur les parties temporelles et maintiennent la partie statique inchangée.

2.2 Raisonnement sur les données temporelles dans le web sémantique : L'Algèbre d'intervalle d'Allen

13 relations temporelles qualitatives entre des intervalles de temps classiques sont proposées par Allen. Elles sont définies en termes d'ordre des bornes de début et de fin des intervalles correspondants (voir tableau 1). Une particularité de l'algèbre d'Allen est que l'on peut déduire de nouvelles relations à travers la composition d'autres relations (par exemple, « Before (A, B) » et « Equals (B,C) » donne « Before(A, C) »). Cependant cette algèbre ne traite pas des intervalles de temps incertains. De plus, elle ne permet pas de lier un point de temps et un intervalle de temps, ni deux points de temps.

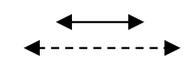
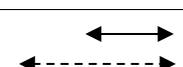
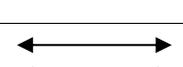
Plusieurs approches ont été proposées pour étendre cette algèbre. Certaines proposent des relations temporelles entre des intervalles de temps précis comme [13], [45], [46] et [47]. Dans [20], les auteurs proposent des relations temporelles floues vues

comme des ensembles flous de relations d'Allen comme la relation « *Fuzz-Meets* » qui concerne la relation d'Allen « *Meets* ». Dans [36] les auteurs proposent une approche basée sur la théorie des possibilités pour modéliser des relations d'intervalles temporels incertains en attribuant un degré de préférence à chaque relation d'Allen de base. Cette approche est proposée dans le cadre d'un réseau probabiliste d'intervalles temporels ; où les nœuds sont des intervalles de temps et les arêtes sont des relations d'intervalle incertaines. [6] propose une extension floue de l'algèbre d'Allen. Dans ce travail, les auteurs associent un degré de préférence à chaque relation temporelle.

Certaines approches proposent des relations temporelles entre des intervalles de temps imprécis comme [32], [37] et [14]. Cependant, ces extensions sont basées sur des théories liées aux données imparfaites et ne peuvent être supportées dans le cadre de certaines ontologies. Par exemple, les approches proposées dans [33] et [14], 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 [33], la relation « *Equals* » n'est pas réflexive. Les compositions des relations résultantes ne sont pas étudiées par les auteurs. [14] 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 de tableau de composition des relations proposées. La représentation et le raisonnement sur des données temporelles existent également dans d'autres domaines tels que les bases de données, les images de télédétection et la récupération d'informations existent [48] et [49].

Dans nos travaux précédents [2], nous avons étendu l'algèbre d'intervalle d'Allen pour raisonner sur des dates et des horloges précises et imprécises dans le cadre d'une ontologie dite « certaine ».

Tableau 1. 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

Les extensions indiquées sont basées sur des théories liées aux données imparfaites. Par conséquent, les relations résultantes ne peuvent pas être traitées dans le contexte d'une ontologie « certaine ». De plus, la plupart de ces approches ne conservent pas les propriétés de l'algèbre d'origine et n'étudient pas la composition des relations résultantes.

3 Représentation des données temporelles certaines et incertaines dans OWL 2

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

3.1 Données temporelles quantitatives

Nous avons étendu l'approche 4D-fluents pour représenter (i) des intervalles de temps incertains et (ii) des points de temps certains et incertains. La Figure 1 illustre notre extension.

3.1.1 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 ».

3.1.2 Points de temps certains et incertains

Nous étendons l'approche 4D-fluents pour représenter des points de temps certains et incertains en introduisant une classe nommée « *TimePoint* ».

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* » 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* ».

3.2 Données temporelles qualitatives

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

Pour les relations temporelles qualitatives incertaines, quatre propriétés d'objet ont également été 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 Intervalle-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. Et enfin « *RelationPointsCertainty* » qui relie deux instances de la classe « *TimePoint* » pour représenter un degré de certitude associé à une relation point-point.

4 Raisonnement sur des données temporelles certaines et incertaines en OWL 2

Notre approche étend 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.

par des non-experts et permettent une mise en correspondance de l'ontologie avec le modèle probabiliste grâce à son modèle graphique convivial [44]. Il fournit un modèle à base d'arcs nous permettant de représenter des ontologies [25].

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 2.

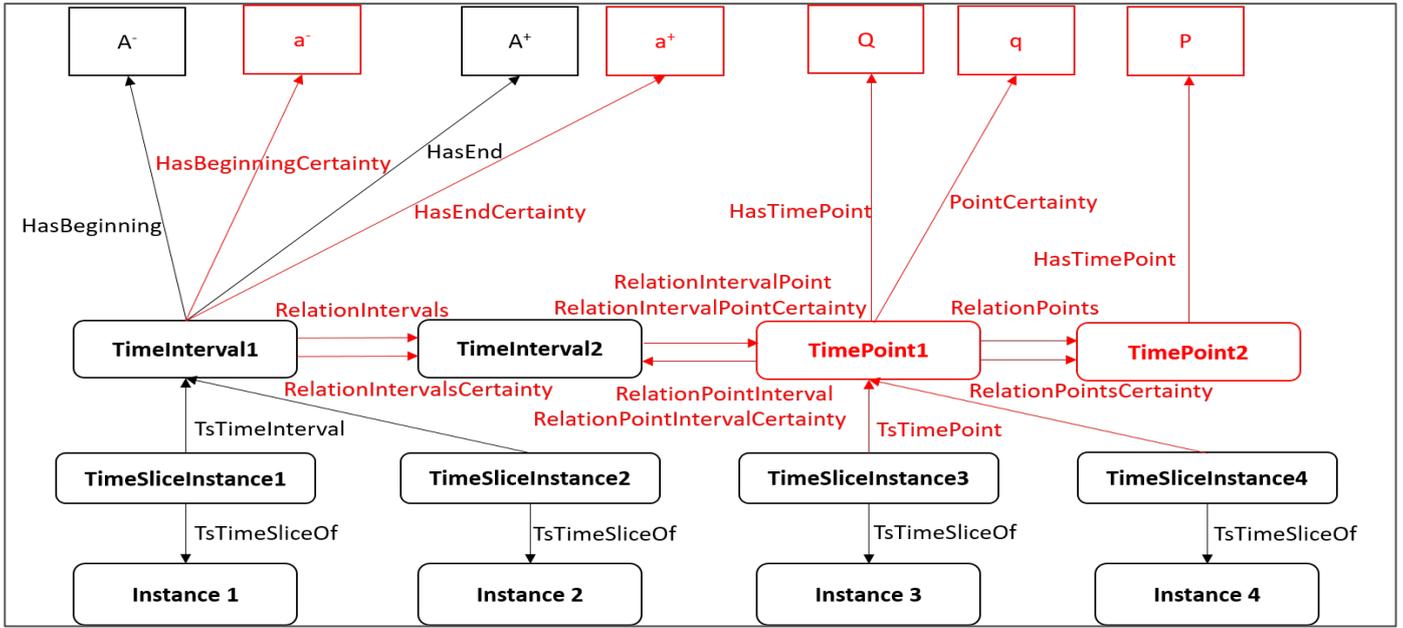


Figure 1. Notre extension de l'approche 4D-Fluents pour représenter des données temporelles certaines et incertaines

4.1 Relations entre les intervalles de temps

En considérant 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 $A = [A^-, A^+]$ et $B = [B^-, B^+]$ 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 [38], [25] et [12]. Selon [42], 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 [43]. Les réseaux bayésiens peuvent être analysés

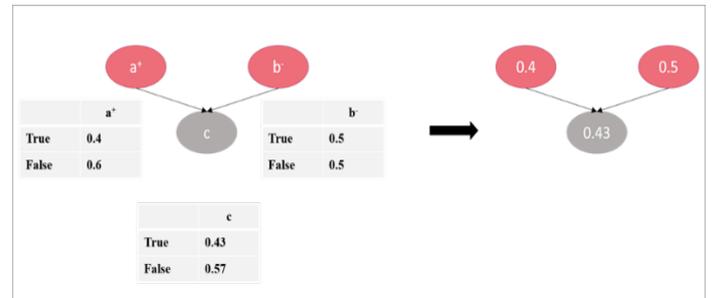


Figure 2. Réseau bayésien associé à la relation « Before_c » entre deux intervalles.

Nous classons les relations 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 », « Meet_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 1.

Nous redéfinissons la relation « Overlaps(A, B) » comme suit :

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

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.

Tableau 2. 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 3.

Nous redéfinissons les autres relations d'Allen de la même manière que la relation « Overlaps_c(A, B) », comme indiqué dans le Tableau 2.

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 :

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

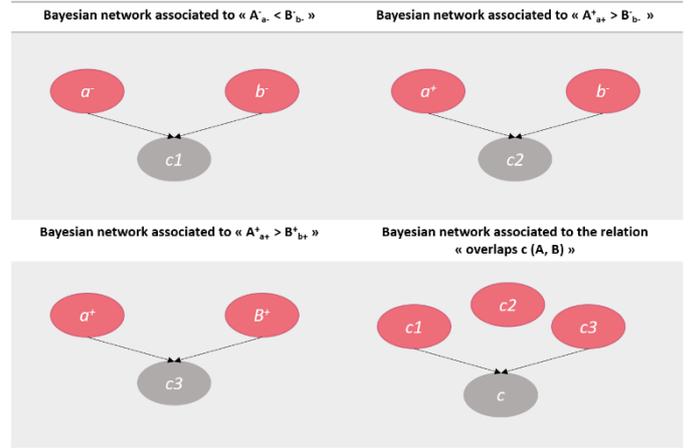


Figure 3. Réseau bayésien associé à la relation « Overlaps_c » entre deux intervalles de temps incertains.

« Equals_c » est réflexif. Il tient que :

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

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 : $R(A, B)$ et $R(B, A) \rightarrow A = B$

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 :

$$\text{Equals}_c(A, B) = \text{Equals}_c(B, A)$$

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: $R(I, J)$ 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.

4.2 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 3.

Tableau 3. 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)	$P < A^-$		After(A, P)
After(P, A)	$A^+ < P$		Before(A, P)
Starts(P, A)	$P = A^-$		Started-by(A, P)
During(P, A)	$(A^- < P) \wedge (P < A^+)$		Contains(A, P)
Ends(P, A)	$P = A^+$		Ended-by(A, P)
Relation temporelle entre un intervalle de temps incertain A = [A⁻, A⁺], et un point de temps incertain P_p			
Before _c (P _p , A)	$P_p < A^-$		After _c (A, P _p)
After _c (P _p , A)	$A^+ < P_p$		Before _c (A, P _p)
Starts _c (P _p , A)	$P_p = A^-$		Started-by _c (A, P _p)
During _c (P _p , A)	$(A^- < P_p) \wedge (P_p < A^+)$		Contains _c (A, P _p)
Ends _c (P _p , A)	$P_p = A^+$		Ended-by _c (A, P _p)

4.3 Relations entre les points de temps

Les relations temporelles qualitatives entre les intervalles de temps ont été adaptées pour proposer des relations entre les points de temps, comme le montre le Tableau 4.

Tableau 4. 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)

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. Les quatre tables sont présentées dans notre article journal [1].

5 « UncertTimeOnto » : une ontologie des données temporelles incertaines en OWL 2

Cette section présente notre ontologie OWL 2, appelée « UncertTimeOnto »². 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 :

« 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 », « MetbyIntervals », « OverlappedbyIntervals », « StartedbyIntervals », « ContainsIntervals », « EndedbyIntervals » et « EqualsIntervals ».

Notre ontologie comporte également un ensemble de règles SWRL pour déduire de nouvelles relations temporelles qualitatives. Pour chaque relation temporelle, nous associons 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 4 montre un exemple d'une règle SWRL.

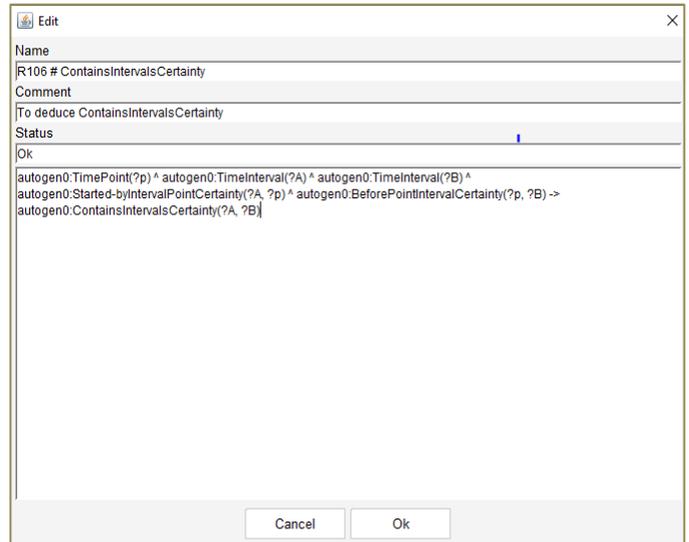


Figure 4. Exemple d'une règle SWRL

« 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 avons utilisé, pour l'implémenter, l'éditeur d'ontologie Protégé, comme la montre la Figure 5.

²<https://cedric.cnam.fr/isid/ontologies/UncertTimeOnto.owl#>

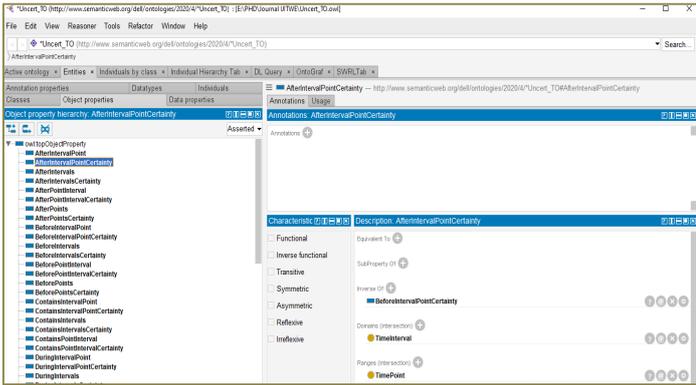


Figure 5. Notre ontologie créée par l'éditeur Protégé

6 Expérimentation

Pour valider notre approche, nous avons mené 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 avons implémenté un prototype basé sur l'ontologie « *UncertTimeOnto* ». Ensuite, nous l'avons intégré dans CAPTAIN MEMO pour gérer la dimension temporelle des données décrites par l'ontologie « *PersonLink* ».

6.1 Prototype implémenté

Pour valider notre approche, nous avons implémenté avec le langage Java, un prototype basé sur notre ontologie « *UncertTimeOnto* ». Nous avons utilisé Netica-J³ pour implémenter les réseaux bayésiens et OWL API pour manipuler les ontologies.

Ce prototype inclut trois composants : « *UncertTimeOnto Population* », « *Qualitative Temporal Data Inference* » et « *Querying* ». Tout d'abord, l'utilisateur instancie « *UncertTimeOnto* » via une interface utilisateur. Cette interface permet à l'utilisateur de saisir des données temporelles certaines et/ou incertaines. Après chaque nouvelle saisie de données temporelles, le composant « *Qualitative Temporal Data Inference* » 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.

6.2 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 séjour d'Helena en Chine. Soit $Q = 1996_{0.8}$ un point de temps incertain représentant le retour d'Helena en Chine. Soit « $P = 2000$ » le point de temps certain représentant l'année de séjour d'Helena à Hong Kong. La Figure 7 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.

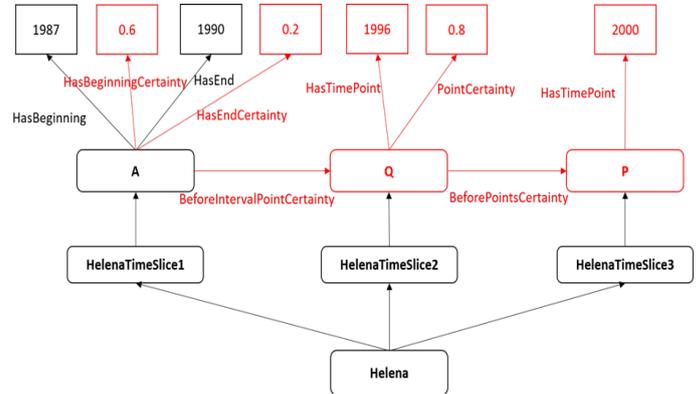


Figure 6. Un exemple d'ontologie « *PersonLink* » basée sur notre approche

7 Conclusion

Dans cet article, 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 point de temps ; 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. Enfin, la troisième contribution est la création d'une 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 basé sur « *UncertTimeOnto* » permettant aux utilisateurs d'explorer notre approche.

8 Références

- [1] Achich, N., Ghorbel, F., Hamdi, F., Métais, E., & Gargouri, F. (2021). Certain and Uncertain Temporal Data Representation and Reasoning in OWL 2. *International Journal on Semantic*

³ <https://www.norsys.com/netica-j.html#download>

Web and Information Systems (IJSWIS), 17(3), 51-72.

[2] 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.

[3] Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. *Commun*, 26(11), 832-843.

[4] Anagnostopoulos, E. Batsakis, S., & Petrakis, E.(2013). CHRONOS: A Reasoning Engine

for Qualitative Temporal Information in OWL. *Procedia Computer Science*,(pp. 70-77).

[5] Artale, A., & Franconi, E. (2000). A Survey of Temporal Extensions of Description Logics. *Annals of Mathematics and Artificial Intelligence*, 30(1-4), 171-21.

[6] Badaloni, S., & Giacomini, M. (2006). The Algebra IAfuz: a Framework for Qualitative Fuzzy Temporal Reasoning. *Artificial intelligence*, 170(10), 872-908.

[7] Batsakis, S., & Petrakis, E. G. M. (2011). SOWL: A Framework for Handling Spatio-Temporal Information in OWL 2.0. In International Workshop on Rules and Rule Markup Languages for the Semantic Web (pp. 242–249).

[8] Batsakis, S., Tachmazidis, I., & Antoniou, G. (2017). Representing Time and Space for the Semantic web. *International Journal on Artificial Intelligence Tools*, 26(03).

[9] Baumann, R., Loebe, F., & Herre, H. (2012). Ontology of Time in GFO. In International Conference on Formal Ontologies and Information Systems (pp. 293–306).

[10] Buneman, P., & Kostylev, E. (2010). Annotation Algebras for RDFS. In Workshop on the Role of Semantic Web in Provenance Management.

[11] Ermolayev, V., Jentzsch, E., Karsayev, O., Keberle, N., Matzke, W.-E., Samoylov, V., & Sohnius, R. (2006). An Agent-Oriented Model of a Dynamic Engineering Design Process. *Agent-Oriented Information Systems III* (pp. 168–183). LNCS 3529.

[12] Fareh, M. (2019). Modeling incomplete knowledge of semantic web using Bayesian networks. *Applied Artificial Intelligence*, 33(11), 1022-1034.

[13] Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. *A.I* (pp. 199–227). 54.

[14] Gammoudi A., Hadjali A., & Yaghlane B. B. (2017). Fuzz-TIME: an intelligent system for managing fuzzy temporal information. *Intelligent Computing and Cybernetics*, 10(2), 200-222.

[15] Ghorbel, F., Métais, E., Ellouze, N., Hamdi, F., & Gargouri, F. (2017, March). Towards accessibility guidelines of interaction and user interface design for Alzheimer's disease patients. In Tenth International Conference on Advances in Computer-Human Interactions.

[16] Ghorbel, F., Hamdi, E. & Métais, E. (2019). Ontology-

Based Representation and Reasoning about Precise and Imprecise Time Intervals. In IEEE International Conference on Fuzzy Systems.

[17] Ghorbel, F., Hamdi, F., Metais, E. (2019, June). Estimating the Believability of Uncertain Data Inputs in Applications for Alzheimer's Disease' Patients. In International Conference on Applications of Natural Language to Information Systems (pp. 208-219). Springer, Cham.

[18] Ghorbel, F., Hamdi, F., & Métais, E. (2020). Dealing with Precise and Imprecise Temporal Data in Crisp Ontology. *International Journal of Information Technology and Web Engineering (IJITWE)*, 15(2), 30-49.

[19] 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.

[20] Guesgen, H. W., Hertzberg, J., & Philpott, A. (1994). Towards Implementing Fuzzy Allen Relations. In ECAI-94 Workshop on Spatial and Temporal Reasoning (pp. 49-55).

[21] Gutierrez, C., Hurtado, C., & Vaisman, A. (2005). Temporal RDF. In European Semantic Web Conference (pp. 93-107).

[22] Harbelot, B. A. (2013). Continuum: A Spatiotemporal Data Model to Represent and Qualify Filiation Relationships. *ACM SIGSPATIAL International Workshop*, (pp. 76-85).

[23] Herradi, N., Hamdi, F., Métais, E., Ghorbel, F., & Soukane, A. (2015). PersonLink: an ontology representing family relationships for the CAPTAIN MEMO memory prosthesis. In International Conference on Conceptual Modeling (pp. 3-13).

[24] Herradi, N., Hamdi, F., & Métais, E. (2017). A Semantic Representation of Time Intervals in OWL 2. In KEOD (pp. 269-275).

[25] Hlel, E., Jamoussi, S., & Hamadou, A. B. (2018). A new method for building probabilistic ontology (prob-ont). In *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications* (pp. 1409-1434). IGI Global.

[26] Hurtado, C. &. (2006). Reasoning with Temporal Constraints in RDF. *Principles and Practice of Semantic Web Reasoning*, (pp. 164–178).

[27] Kim, S. K., Song, M. Y., Kim, C., Yea, S. J., Jang, H. C., & Lee, K. C. (2008, December). Temporal ontology language for representing and reasoning interval-based temporal knowledge. In *Asian Semantic Web Conference* (pp. 31-45). Springer, Berlin, Heidelberg.

[28] Klein, M. C. A., & Fensel, D. (2001). Ontology Versioning on the Semantic Web. In *Semantic Web Working Symposium* (pp. 75-91), Stanford University, California, USA.

[29] Koubarakis, M., & Kyzirakos, K. (2010). Modeling and Querying Metadata in the Semantic Sensor Web: The model stRDF and the query language stSPARQL. In *The semantic web: research and applications* (pp. 425-439).

- [30] Lutz, C. (2003). Description Logics with Concrete Domains - A Survey. In *Advances in Modal Logics* (pp. 265-296).
- [31] Métais, E., Ghorbel, F., Herradi, N., Hamdi, F., Lammari, N., Nakache, D., Ellouze, N., Gargouri, F. and Soukane, A., (2012). Memory prosthesis. *Non-pharmacological Therapies in Dementia*, 3(2), 177.
- [32] Nagypál, G., & Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems* (pp. 906–923).
- [33] Noy, N., Rector, A., Hayes, P., & Welty, C. (2006). Defining N-Ary Relations on the Semantic-Web. *W3C Working Group Note*, 12(4).
- [34] O'Connor, M. J. (2011). A Method for Representing and Querying Temporal Information in OWL. *Biomedical Engineering Systems and Technologies*, (pp. 97-110).
- [35] Preventis, A., Petrakis, E. G., & Batsakis, S. (2014). Chronos Ed: A Tool for Handling Temporal Ontologies in PROTÉGÉ. *International Journal on Artificial Intelligence Tools*, 23(04).
- [36] Ryabov, V., & Trudel, A. (2004). Probabilistic Temporal Interval Networks. In *Temporal Representation and Reasoning* (pp. 64-67).
- [37] Sadeghi, K. M. M., & Goertzel, B. (2014). Uncertain Interval Algebra via Fuzzy/Probabilistic Modeling. In *IEEE International Conference on Fuzzy Systems* (pp. 591-598).
- [38] Lamma, E., Riguzzi, F., & Storari, S. (2006). Improving the K2 algorithm using association rule parameters. In *Modern Information Processing* (pp. 207-217). Elsevier Science.
- [39] Tappolet, J., & Bernstein, A. (2009). Applied Temporal RDF: Efficient Temporal Querying of RDF Data with SPARQL. In *European Semantic Web Conference* (pp. 308–322).
- [40] Welty, C. A., & Fikes, R. (2006). A Reusable Ontology for Fluents in OWL. In *Formal Ontology in Information Systems* (pp. 226-236).
- [41] Zekri, A., Brahmia, Z., Grandi, F., & Bouaziz, R. (2016). τ OWL: A Systematic Approach to Temporal Versioning of Semantic Web Ontologies. *Journal on Data Semantics*, 5(3), 141-163.
- [42] Ding, Z., & Peng, Y. (2004, January). A probabilistic extension to ontology language OWL. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the* (pp. 10-pp). IEEE.
- [43] Gu, T., Pung, H. K., Zhang, D. Q., Pung, H. K., & Zhang, D. Q. (2004). A bayesian approach for dealing with uncertain contexts (pp. 205-210).
- [44] Njah, H., Jamoussi, S., Mahdi, W., & Elati, M. (2016, October). A Bayesian approach to construct Context-Specific Gene Ontology: Application to protein function prediction. In *2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1-6). IEEE.
- [45] Madkour, M., Benhaddou, D., & Tao, C. (2016). Temporal data representation, normalization, extraction, and reasoning: A review from clinical domain. *Computer methods and programs in biomedicine*, 128, 52-68.
- [46] Li, F., Du, J., He, Y., Song, H. Y., Madkour, M., Rao, G., ... & Tao, C. (2020). Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events. *Journal of the American Medical Informatics Association*, 27(7), 1046-1056.
- [47] Nys, G. A., Van Ruymbeke, M., & Billen, R. (2018, October). Spatio-temporal reasoning in CIDOC CRM: an hybrid ontology with GeoSPARQL and OWL-Time. In *CEUR Workshop Proceedings (Vol. 2230)*. RWTH Aachen University.
- [48] Pons, J., Billiet, C., Pons, O., & De Tré, G. (2014). Aspects of dealing with imperfect data in temporal databases. In *Flexible Approaches in Data, Information and Knowledge Management* (pp. 189-220). Springer, Cham.
- [49] 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. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 27(05), 815-840.