



HAL
open science

La logique temporelle et ses applications dans le traitement du langage naturel

Djallel Bouneffouf

► **To cite this version:**

| Djallel Bouneffouf. La logique temporelle et ses applications dans le traitement du langage naturel.
| [Rapport de recherche] Télécom SudParis. 2013. hal-00799121

HAL Id: hal-00799121

<https://hal.science/hal-00799121>

Submitted on 11 Mar 2013

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.

La Logique Temporelle et ses Applications dans le Traitement du Langage Naturel

Djallel Bouneffouf,

Department of Computer Science, Télécom SudParis, UMR CNRS Samovar, 91011 Evry
Cedex, France

{Djallel.Bouneffouf}@it-sudparis.eu

Abstract. Dans ce papier, nous clarifions quelques concepts linguistiques, en étudiant les différents niveaux de représentation et de traitement des énoncés linguistiques. Nous avons aussi illustré avec quelques travaux les différents types d'inférences temporelles existants, Nous avons remarqué que ces travaux ne se basent que sur l'amélioration des détections des relations temporelles existantes entre événements et expressions temporelles.

Keywords: Logique Temporelle; Traitement du Langage Naturel.

1 Introduction

Le mot « temps » recouvre plusieurs significations en français, et il est nécessaire pour la compréhension de distinguer le temps grammatical du temps notionnel. Le second est représenté en logique par une ligne droite et infinie, avec un point marquant le présent et séparant le passé du futur. Le temps grammatical désigne les marques linguistiques utilisées pour exprimer le temps notionnel dans le langage (l'imparfait, le présent de l'indicatif, etc....).

Dans ce qui suit nous allons explorer ces deux notions du temps, du point de vu logique avec ces différentes représentations (structure de points, structure d'intervalles, événement et Allen) ensuite au point de vu langage.

Nous terminons avec une étude sur l'inférence temporelle élaborée par l'un des groupes les plus abouti dans le domaine.

2 La structure de points

La conception du temps est couramment reliée à la notion de point ou d'instant sur un axe temporel. Les points permettent en effet d'utiliser les structures de nombres (entiers, rationnels ou réels). Cette conception est largement utilisée dans la modélisation de phénomènes évoluant dans le temps.

Cette structure temporelle doit être manipulée avec un langage logique ; la logique du temps, historiquement très liée au développement des logiques modales. Elle est basée sur les connecteurs logiques habituels ($\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow$) et les opérateurs temporels P (passé) et F (futur). Ainsi, si l'action de chanter effectuée par John est notée p, on aura les représentations suivantes :

- John chante : p
- John chanta : Pp
- John chantera : Fp
- John avait chanté : PPp (on se place dans le passé d'un point situé au passé lui-même)
- John aura chanté : FPPp

Ces formules seront enrichies avec de nouveaux opérateurs similaires à ceux utilisés en logique modale (Bras, 1990).

Toutes les logiques dérivées de la logique du temps sont basées sur une ontologie de points. Nous allons maintenant nous intéresser à des ontologies d'intervalles.

3 La structure d'intervalles

Du point de vue philosophique, il semble que le concept de point dépourvu de durée ne correspond pas à la réalité :

Du point de vue linguistique, il est encore plus évident qu'une entité ponctuelle est mal adaptée pour l'expression de la référence temporelle. Même les expressions dites ponctuelles se réfèrent à des périodes étendus, comme dans les exemples suivants :

A six heures précises, Harry quitta son bureau.

Une structure d'intervalle est définie par $\langle I, <, \subseteq \rangle$, avec I un ensemble non vide d'entités temporelles, des relations de précédence ($<$) et d'inclusion (\subseteq). Voici quelques propriétés de cette structure :

\subseteq est un ordre partiel, elle est en effet :

Réflexive : $(\forall x \in I) x \subseteq x$.

Antisymétrique : $(\forall x \in I) (\forall y \in I) (x \subseteq y \wedge y \subseteq x \Rightarrow x = y)$.

Transitive : $(\forall x \in I) (\forall y \in I) (x \subseteq y \wedge y \subseteq z \Rightarrow x \subseteq z)$.

Nous pouvons également remplacer la relation \subseteq par la relation O (overlap) qui exprime que deux événements ont une partie commune, et définie par rapport à l'inclusion:

$xOy \Leftrightarrow (\exists z) (z \subseteq x \wedge z \subseteq y)$

La mise en place des logiques temporelles basées sur les sémantiques d'intervalles amènent à des résultats relativement complexes, qu'il n'est pas nécessaire d'exposer ici.

Des critiques ont été adressées aux sémantiques d'intervalles, notamment en ce qui concerne la difficulté de définir la vérité d'une proposition (vraie sur toutes intervalles ? sur au moins l'un deux ?). Ces problèmes ont provoqué la nécessité de concevoir une entité plus globale et plus complète.

4 La structure d'événements

L'événement est une nouvelle entité primitive, de durée non nulle et fini, correspondant intuitivement à des fragments de notre perception du monde. Pour les linguistes comme pour les philosophes, les logiciens et les spécialistes de l'intelligence artificielle, la tendance est de préférer les événements aux intervalles car les événements ont une structure à portée non seulement temporelle, mais aussi spatiale.

Davidson a proposé de traiter les événements comme des objets, ajoutant à l'ensemble des individus d'un modèle, un ensemble d'événements, par exemple, la phrase Marie aime Paul n'est plus représentée par aimer (Paul, Marie), mais par :

Aimer(e, Paul, Marie)

Une structure d'événement est définie par Kamp par le triplet $\langle E, \alpha, O \rangle$, où E est un ensemble d'entités de base non nulle, α est la relation de précédence, et O la relation de recouvrement si $e1 O e2$ alors une partie de $e1$ au moins a eu lieu en même temps que $e2$.

α est asymétrique : $(e1 \alpha e2) \Rightarrow (e2 \alpha e1)$

α est transitive : $(e1 \alpha e2) \wedge (e2 \alpha e3) \Rightarrow (e1 \alpha e3)$

O est symétrique : $(e1 O e2) \Rightarrow (e2 O e1)$

O est réflexive : $(e1 O e2)$

Principe de séparation : $(e1 \alpha e2) \Rightarrow \neg (e1 O e2)$

Transitive mixte : $(e1 \alpha e2) \wedge (e2 O e3) \wedge (e3 \alpha e4) \Rightarrow (e1 \alpha e4)$

Principe de linéarité : $(e1 \alpha e2) (e1 O e2) (e2 \alpha e1)$

Ces conditions minimales sont dictées par l'intuition lorsque nous avons des événements et des relations qui les lient.

Nous avons présenté trois ontologies (structures de points, structures d'intervalles et structures d'événements). Il est fondamental de séparer le niveau temporel (points et intervalles) du niveau relatif à l'expérience du monde (événements). En effet, si les relations définies dans les structures d'événements sont des relations temporelles, les événements sont également des expériences, des « faits » qui ont lieu et qui déterminent la structure du temps.

C'est pourquoi nous pouvons dire que la logique d'Allen, que nous allons présenter, permet de rattacher les deux notions.

5 La théorie d'Allen

Selon Allen, deux intervalles peuvent être liés entre eux par les 13 relations primitives suivantes (Bras, 1990). Où X et Y sont des termes de types intervalles de temps (on appelle « relation inverse » la relation correspondante entre Y et X) :

Relation	Symbole	Symbole relation inverse
X before Y	<	>
X equals Y	=	=
X meets Y	M	Mi
X overlaps Y	O	Oi
X during Y	D	Di
X starts Y	S	Si
X finishes Y	F	Fi

Fig. 1. Tableau 2.1: Les relations d'Allen

Les relations sont mutuellement exclusives : une seule relation est possible entre deux intervalles.

Il est possible de composer les relations. Ainsi, la transitivité des relations entre intervalles est définie par:

$$\forall I \forall j \forall k ((I < j) \wedge (j < k) \Rightarrow I < k)$$

$$\forall I \forall j \forall k ((I mj) \wedge (j dk) \Rightarrow (I ok) \vee (I dk) \vee (I sk))$$

Il existe 169 relations de transitivité de ce type.

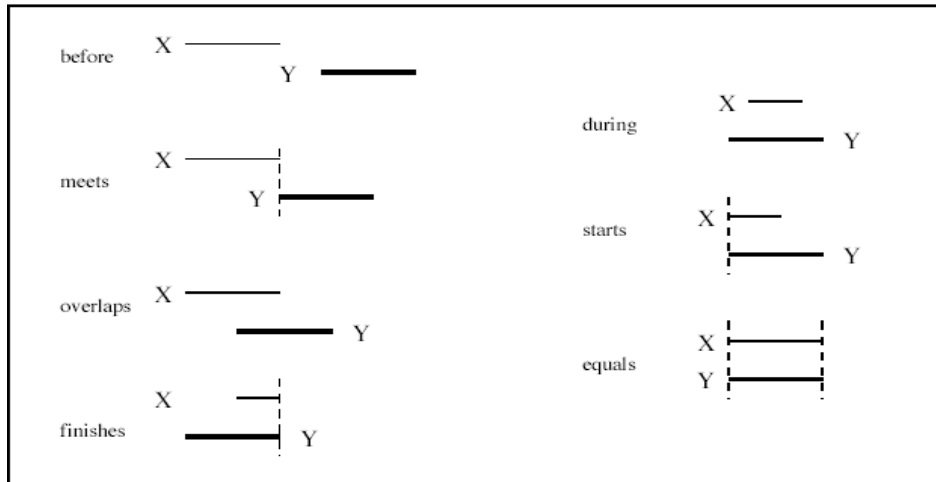


Fig. 2. Représentation des relations d'Allen

Deux intervalles peuvent être reliés par une relation primitive, mais aussi par une relation complexe ; il est ainsi possible de représenter une connaissance incomplète des relations.

La connaissance temporelle sur un ensemble d'intervalles peut être représentée par un réseau de contraintes. Il s'agit d'un graphe orienté dont les nœuds représentent les intervalles et dont les arcs sont étiquetés par la relation entre les intervalles.

L'exemple suivant, très simple, permet d'illustrer rapidement le raisonnement sur les intervalles :

Paul entra dans la pièce (1). Marie regardait la télévision (2) .Elle l'éteint (3).

(1) Et (2) introduisent l'assertion temporelle suivante :

I entre during I regarde_télévision

Puis, en examinant (2) et (3), nous obtenons :

I regarde_télévision meet I enteindre- télévision

6 Le temps dans la langue

Qu'est-ce qui distingue le temps linguistique des autres notions de temps? "Ce que le temps linguistique a de particulier c'est qu'il est organiquement lié à l'exercice de la parole, qu'il se définit et s'ordonne comme fonction du discours. Ce temps a son centre – un centre, à la fois, générateur et axial - dans le présent de l'instance de la parole" (Benveniste, 1974). Le discours instaure un maintenant, moment de l'énonciation. En opposition au maintenant, nous créons un alors. Ce maintenant est donc le fondement des oppositions de la langue.

6.1 Le modèle de Reichenbach

Reichenbach a proposé, pour modéliser la sémantique des temps grammaticaux, les trois repères suivants :

E le moment de l'événement

S le moment de l'énonciation ou de la parole (Speech Time)

R le moment de référence

Les relations possibles entre repères sont la relation de simultanéité notée « , » et la relation de précédence notée « _ ». La nouveauté réside surtout dans l'ajout d'un moment de référence, qui permet de prendre en considération certains temps composés. Ainsi, la représentation de quelques temps grammaticaux à l'aide du modèle de Reichenbach est la suivante :

Passé simple	je vis Paul	E, R_S
Plus-que-parfait	j'avais vu Paul	E_R_S
Futur	je verrai Paul	S_E,R
Futur antérieur	j'aurai vu Paul	S_E_R

6.2 Les adverbiaux temporels

Un adverbial est un élément (mot ou groupe de mots) ayant une fonction similaire à celle d'un adverbe ou d'un complément circonstanciel, c'est-à-dire qu'il modifie le verbe auquel il est rattaché (Charolles, 1997). Nous pouvons le supprimer sans rendre la syntaxe ni la sémantique de la phrase incorrecte. Ainsi, les passages soulignés des exemples suivants ont une fonction adverbiale temporelle :

Paul arrive demain.

Marie est revenue à cinq heures.

Nous pouvons distinguer :

Les adverbiaux de référence temporelle dont le rôle est d'exprimer la localisation d'un événement dans le temps : demain.

Les adverbiaux de durée : pendant une heure, en trois jours

Les adverbiaux de durée : pendent une heure.

Les adverbiaux de fréquence : souvent, tous les mois.

Les adverbiaux itératifs : trois fois, plusieurs fois.

Les adverbiaux de quantification : toujours, quelquefois.

Les adverbiaux présuppositionnels : encore, déjà.

7 L'inférence temporelle

Si l'annotation des marqueurs du temps dans le discours sont l'objet de plusieurs sujets de recherches, l'étude de l'inférence temporelle et ses applications ne sont qu'à ses débuts. Ce problème commence à générer des travaux en informatique linguistique liés aux enjeux que représentent les informations temporelles entre autre pour la recherche d'information et les systèmes questions-réponses.

Avec l'exploration de ce nouveau champ d'action dans le traitement du langage naturel, l'inférence temporelle nous permet d'établir des relations temporelles existantes

entre évènements dans un texte, de détecter les relations existantes entre expressions temporelles et aussi les relations entre expressions temporelles et événements. Dans ce qui suit nous présentons l'une des recherches les plus abouti dans le domaine.

7.1 Le travail du groupe Human Language Technology Research Institut (HLTRI) sur l'inférence temporelle :

HLTRI est un groupe de recherche travaillant sur l'inférence temporelle. Il est aussi membre de l'organisation fondatrice du langage (TimeML¹) qui est TimeML is a robust specification language for events and temporal expressions in natural language. un langage de spécification d'événements et d'expressions temporelles dans le langage naturel. It is designed to address four problems in event and temporal expression markup:

Afin d'étudier l'inférence temporelle dans le langage naturel, le groupe (HLTRI) a établi un grand corpus de questions-réponses qui sont fondées sur la recherche d'information temporelle.

Les questions sont annotées comme ceci:

- Expressions temporelles, annotées par la balise TIMEX3.
- La balise EVENT correspond à un événement.
- LIEN est une balise qui code les relations entre éléments temporels.

Pour découvrir les relations temporelles entre les événements dans un texte, le groupe a utilisé la représentation graphique.

Les nœuds du graphe sont représentés par les événements et les arcs entre les nœuds sont soit des relations TLink, SLink ou ALink.

Pour classer les événements dans un même texte, il utilise les trois relations ALINK, TLINK, SLINK et entre les événements de deux textes différents il n'utilise que le module TLINK.

- TLink, représentant les relations temporelles entre les événements ou entre un événement et une expression temporelle.

- SLink ou relation de subordination, est utilisée pour introduire des contextes et des relations entre deux événements.

- ALink ou relation aspectuelle, représentant la relation aspectuelle entre un événement et son argument (en général c'est un autre événement).

Afin d'avoir toutes les relations temporelles possibles entre les événements des deux textes, le groupe a conçu un module d'inférence temporelle.

À partir des différents liens TLINK, ALINK et SLINK existant entre les expressions temporelles et les événements, le module infère de nouveaux liens non détectés auparavant. Pour cela le groupe a définit plusieurs règles d'inférences qui sont citées ci-dessous :

¹ TimeML a été développé dans le cadre de trois ateliers AQUAINT et des projets.

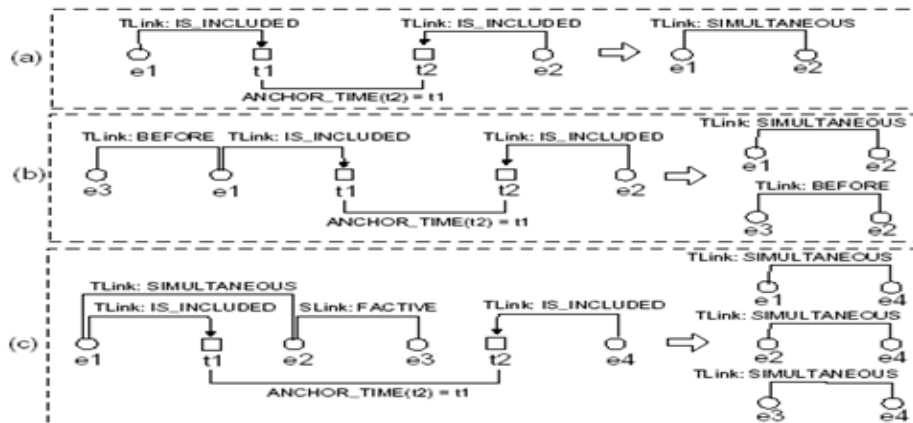


Fig. 3. Règles de l'inférence temporelle

Concrètement le module suit les étapes suivantes :

Etape1:trouver T1 et T2, deux expressions temporelles dans la phrase ou des phrases adjacentes.

Etape2:Rechercher des événements E1 et E2 lié à T1 et T2 respectivement.

Etape3:Trouver une relation CE1 lien entre E1 et d'autres événements.

Etape4:Trouver relation CE2 lien entre E2 et d'autres événements.

Etape5:Utiliser une inférence temporelle reliant CE2 et CE1.

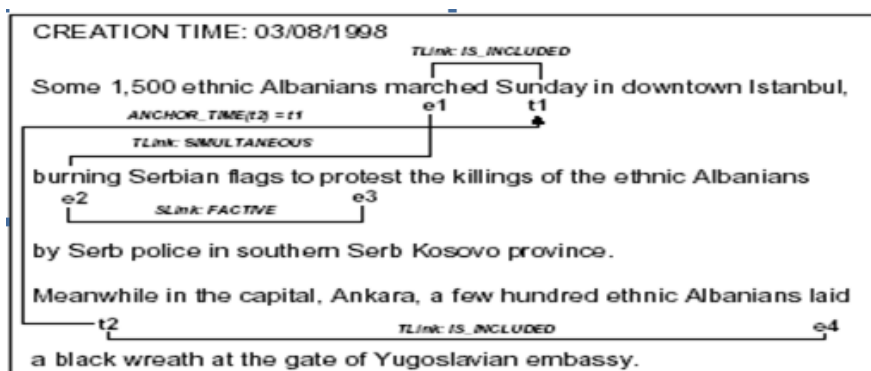


Fig. 4. Application des règles de l'inférences sur un exemple du corpus RTE

Lors de l'application de la procédure à l'exemple illustré dans la figure 2.3 le module suit les étapes suivantes :

Etape 1: trouver les expressions temporelles t1 et t2.

Etape 2: événements e1 et e4 reliées à t1 et t2 avec TLink : is_includes.

Etape 3: détecter la chaîne des événements (e1, e2, e3).

Etape 5: T1 et T2 sont liés (par un ANCHORTIME (t2)=t1), ce qui signifie que e1, e2, e4 et e3 sont simultanés.

Ce groupe s'est focalisé sur la déduction de nouvelles relations entre événements. Cette étude sur l'inférence temporelle a permis d'établir plusieurs règles d'inférences reliant les événements et les expressions temporelles.

8 Conclusion

Dans ce papier nous avons exploré la logique temporelle et ses applications dans le traitement du langage Naturel. Nous avons aussi illustré avec les travaux du groupe (HLTRI) les différents types d'inférences temporelles existants, Nous avons remarqué que ce groupe ne se base que sur l'amélioration des détections des relations temporelles existantes entre évènements et expressions temporelles.

Nous nous sommes inspirés de ces travaux dans notre façon de procéder pour élaborer le corpus et concevoir nos inférences.

Nous allons montrer dans les prochains chapitres comment nous avons concrétisé cet objectif.

References

1. Bouneffouf, D.: Role of temporal inference in the recognition of textual inference. arXiv preprint arXiv:1302.5645 (2013).
2. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The berkeley framenet project. In Proceedings of the COLING-ACL, Montreal.
3. Myriam Bras. Calcul des Structures Temporelles du Discours. PhD thesis, IRIT, 1990.
4. Benveniste Emile Problèmes de linguistique générale. Paris, Gallimard, vol. II.
5. Charolles M. « L'encadrement du discours – univers, champs, domaines et espaces », Cahier de recherche linguistique, 6, p. 1-73. 1997.
6. Moldovan and Rus Logic Forms can be utilized by a wide variety.2001.
7. P. Miller. 1995. "Notes on phonology and orthography in several Katic Mon-Khmer groups in Northeast Thailand." Mon-Khmer Studies 24: 27-51.
8. Pierre Nugues. An Introduction to Language Processing with Perl and Prolog. Springer Verlag, 2006.
9. E. Nyberg, T.Mitamura, J. Carbonnell, J. Callan, K. Clins-Thompson, K Czuba, M. Duggan, L. Hiyakumoto, N. Hu, Y. huang, J. Ko, L.V. Lita, S.Muratagh et V. Pedro. The JAVELIN Question-Answering System at TREC 2002. In Proceeding of the 11th
10. Text Retrieval conference (TREC-11), 2002.
11. Paul Kiparsky and Carol Kiparsky. In Manfred Bierwisch and Karl Erich Heidolph, editors, Progress in Linguistics. A collection of Papers, pages 143–173. Mouton, Paris.
12. A. Rodrigo, A. Peñas, J. Herrera and F. Verdejo. The Effect of Entity Recognition on Answer Validation. In Lecture Notes in Computer Science. In press 2007.
13. workshop on recognizing textual entailment, 2005 .

14. M. Venhagen, I. Mani , R. Sauri, R. Knippen, J .Littman and J. Pustejovsky. 2005. Automating Temporal Annotation With TARSQI. In Proceedings of ACL 2005. demo session.