

УДК 81'27

CALCUL DE LA PRESUPPOSITION ENTRE PROPOSITIONS TEXTUELLES A L'AIDE DES CONNECTEURS: APPROCHE LOGIQUE ET SEMANTIQUE

Omar Larouk

PhD en Sciences de l'Information et de la Communication, Linguistique et Informatique,
Dr, Maître de Conférences des Universités Hors-Classe (MCFU-HC) de l'École
Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB)
Université de Lyon

17, boulevard de 11 Novembre 1918, Villeurbanne-Lyon, 69623, France. larouk@enssib.fr

Marina V. Garanovich

Candidate ès Lettres, Maître de conférences
du Département de la Linguistique théorique et pratique
Université nationale d'Etat de la Recherche scientifique de Perm

15, Boukirev, Perm, 614990, Russie. gamarina@yandex.ru

L'augmentation de la production scientifique et des informations non structurées sur le web nécessitent des analyses plus poussées concernant la composante linguistique des systèmes de recherche d'information. Le problème concerne l'extraction des connaissances issues des bases de données en permanente évolution. Cet accroissement des contenus multilingues scientifiques ou masses de données hétérogènes génèrent du bruit documentaire. Les solutions proposées par les systèmes de recherches d'informations ne sont pas complétement couronnées de succès. En effet, le principal problème de l'indexation est lié aux facteurs linguistiques et notamment à la représentation des connaissances contextuelles. Le problème de la présupposition entre propositions est réel. En effet, la demande d'informations portée par la question (Q) et la réponse (R) fournie par un système automatique (base de données ou moteur de recherche) pose le problème de la relation implicative (inductive) entre deux propositions (Q et R) dans un système de recherche d'informations.

Les mots clés: présupposition; implication; connecteurs pour la recherche d'information.

Introduction

L'augmentation de la production scientifique et des informations non structurées sur le web nécessitent des analyses plus poussées concernant la composante linguistique des systèmes de recherche d'information. Le problème concerne l'extraction des connaissances issues des bases de données en permanente évolution. Cet accroissement des contenus multilingues scientifiques ou masses de données hétérogènes génèrent du bruit documentaire. Les solutions proposées par les systèmes de recherches d'informations ne sont pas complétement couronnées de succès. En effet, le principal problème de l'indexation est lié aux facteurs linguistiques et notamment à la représentation des connaissances contextuelles. Le problème de la présupposition entre propositions est réel. En effet, la demande d'informations portée par la question (Q) et la réponse (R) fournie par une

système automatique (base de données ou moteur de recherche) pose le problème de la relation implicative (inductive) entre deux propositions (Q et R) dans un système de recherche d'informations.

1. Systèmes de Question/Réponse et Analyse linguistique: Présupposition, Implication, Connecteurs pour la recherche d'information

La problématique des systèmes de Question/Réponse se situe ainsi à l'intersection de plusieurs domaines, dont notamment la recherche d'information et le traitement de la langue naturelle. Un système de Question/Réponse se doit aussi de mettre en oeuvre des fonctionnalités nouvelles liées aux accès en langues naturelles sur les réseaux sociaux, et la terminologie de la documentation électronique. Ces systèmes de recherche d'informations documentaires sont basés sur le modèle logique (présence ou absence du mot dans le texte.). Or l'évaluation de l'existence d'une

réponse à une question dans les bases de données documentaires (textuelle ou multimédia) où les services web passe par la reconnaissance des mots, phrases et des textes.

Un système de Question/Réponse est décomposé en plusieurs étapes: analyse de la question avec un choix des mots pleins (descripteurs), des syntagmes de mots, des propositions et enfin des documents contenant des unités textuelles. Le module d'extraction de la réponse dépend des performances du module algorithmique d'analyse morphologique et syntaxique de la question qui lui en fournit les caractéristiques linguistiques.

Les systèmes de Question/Réponse se différencient des moteurs de recherche par des données de saisie différentes. En effet, les moteurs de recherche renvoient des liens vers des documents pertinents en regard d'une requête composée de mots-clés, les systèmes de Question/Réponse renvoient une réponse précise à une question exprimée et saisie en langue naturelle.

2. Présupposition, Implication et Connecteurs

Frege et Russell sont les premiers logiciens à discuter de la notion de présupposition quand ils analysent "l'exemple classique" suivant:

(1) /Le roi de France est chauve/

Premier logicien à donner des bases à la théorie de présupposition, Strawson a introduit la présupposition quand il critique justement la théorie des descriptions définies¹ de Russell et les travaux de Frege. Les inférences de l'exemple (1) en logique des prédicats sont donc:

(1') "il y a un roi de France" – $\{\exists x / x \text{ est roi de France}\}$

(1'') "il y a un roi de France **et** il est chauve" – $\{\exists x / x \text{ est roi de France ET } x \text{ est chauve}\}$

La proposition (1) peut être considérée comme la conjonction logique des propositions dérivées. La fausseté (*en 2017!*) de l'une des propositions dérivées entraîne la fausseté de la proposition (1). Strawson établit une distinction entre la proposition et la phrase [Strawson 1977]. On ne peut parler de valeur de vérité que quand on analyse la phrase en termes de propositions. La présupposition est une relation entre propositions: une proposition **S** présuppose **S'** si et seulement si la vérité de **S'** est une précondition de la vérité ou de la fausseté de **S**.

Si on reprend l'exemple précédent, on peut dire la proposition (1) présuppose (1') et (1'').

Voyons pourquoi il faut intégrer l'analyse des connecteurs dans l'analyse de la présupposition? En effet, nous allons voir que le comportement des connecteurs asymétriques génère la notion de cause comme dans l'exemple suivant:

(2) /Le gardien a vu le feu **et** il a donné l'alerte/
P₁ C P₂

et non pas la réciproque:

(2)⁻¹ /Le gardien a donné l'alerte **et** il a vu le feu/^{*}

En appliquant la définition de Strawson, on peut avoir comme présupposée (2'):

(2') "il y a un gardien qui a vu le feu"

Nous constatons que dans l'exemple (2), la conjonction de ces deux propositions {P₁} et {P₂} présente un phénomène de succession, de temporalité ou de conséquence. D'où le rejet de la séquence {P₂ et P₁}. Cependant, on peut paraphraser l'exemple (2) ainsi:

Si le gardien a vu le feu, alors il donne l'alerte

Si le gardien a vu le feu, alors il a dû donner l'alerte

Les deux propositions P₁ et P₂ se retrouvent dans une relation d'implication. Un critère pour mesurer la symétrie des connecteurs est celui de la commutativité comme le montre la règle suivante:

Si {P₁ **ET** P₂} = {P₂ **ET** P₁}

alors "**ET**" est commutatif (symétrie);

sinon "**ET**" est asymétrique (causalité).

On peut constater que le connecteur "**OU**" véhicule la notion déductive d'implication dans l'exemple suivant:

(3) /Donnez-moi l'argent de la caisse **ou** je tire/²

3. Implication liée aux questions en langage ordinaire

Dans le calcul propositionnel, les valeurs de vérité de l'implication logique "A → B" sont vraies dans tous les cas sauf celui où A est vrai et B est faux. Pour exprimer la notion de cause ou de nécessité, l'implication présente quelques insuffisances. En effet, l'affirmation "A → B" chez les logiciens classiques [Gauthier 1991] possède des valeurs fixes {0,1}. Cependant, cette expression cache beaucoup de notions "sémantiques" et l'affirmation "A → B" peut avoir les équivalences suivantes:

- a) B a été rendu nécessaire par A;
- b) B est impossible sans A;
- c) A est une condition favorable de B.

Dans le cas (a), on peut dire que "A est (**la**) cause de l'**effet** B" c'est-à-dire que l'événement B se produit à cause de l'événement A.

(4) /Comme il avait faim, il a volé/
A, B

Soit sa variante "sémantique" suivante:

(4_a) /il a volé **parce qu'**il avait faim/
B Q A

Q: conjonction de subordination.

**CALCULATION OF THE PRESUPPOSITION
BETWEEN TEXTUAL PROPOSITIONS USING CONNECTORS:
LOGICAL AND SEMANTIC APPROACH**

Omar Larouk

Professor, High School of Information Science and Librarianship
Lyon University

Marina V. Garanovich

Associate Professor, Theoretical and Applied Linguistics Department
Perm State University

The increase in scientific production and unstructured information on the web requires further analysis of the linguistic component of information retrieval systems. The problem concerns the extraction of knowledge from databases in permanent evolution. This increase in multilingual scientific content or masses of heterogeneous data generates documentary noise. The solutions offered by information retrieval systems are not complementary to successful ones. Indeed, the main problem of indexing is related to linguistic factors and in particular to the representation of contextual knowledge. The problem of presupposition between propositions is real. Indeed, the request for information carried by the question (Q) and the answer (R) provided by an automatic system (database or search engine) raises the problem of the implicative (inductive) relation between two propositions (Q and R) in an information retrieval system. A Question/Answer system is decomposed into several stages: analysis of the question with a choice of full words (descriptors), word phrases, propositions and finally documents containing textual units.

Keywords: presupposition; implication; information retrieval.

**ИСПОЛЬЗОВАНИЕ ЛОГИЧЕСКИХ ОПЕРАЦИЙ
ПРИ АНАЛИЗЕ ПРЕСУППОЗИЦИИ МЕЖДУ ВЫСКАЗЫВАНИЯМИ:
ЛОГИЧЕСКИЙ И СЕМАНТИЧЕСКИЙ ПОДХОДЫ**

Омар Ларук

д. информац. наук, компьютерной лингвистики и информатики, профессор Высшей школы информатики и библиотечного дела
Лионский университет

Марина Владимировна Гаранович

к. филол. н., доцент кафедры теоретического и прикладного языкознания
Пермский государственный национальный исследовательский университет

Увеличение количества научных продуктов и не структурируемой информации в Интернете требует анализа, более основательного относительно лингвистической составляющей систем информационного поиска. Эта проблема касается извлечения информации и документов из постоянно изменяющихся баз данных. Появление научных многоязычных источников и баз разнородных данных создают проблемы, связанные с поиском научной информации. Технические средства, предложенные системами информационного поиска, не решили существующие проблемы. Сегодня главная проблема индексации связана с лингвистическими факторами, а именно с представлением контекстуальных знаний. Проблема presupposition между предложениями реальна. Действительно, информационный запрос, сформулированный в форме вопроса (Q), и ответ (R) предоставленный автоматической системой (базой данных или поисковой системой), ставит проблему порядка следования (индуктивно-го) между двумя предложениями (Q и R) в системе информационного поиска.

Ключевые слова: presupposition; импликация; средства поиска информации.