

L'extraction des réponses dans un système de question-réponse

Anne-Laure Ligozat, Brigitte Grau, Isabelle Robba, Anne Vilnat

LIMSI-CNRS

{anne-laure.ligozat ; brigitte.grau ; isabelle.robba ; anne.vinat}@limsi.fr

Résumé

Les systèmes de question-réponse sont la plupart du temps composés de trois grands modules : l'analyse de la question, la sélection des documents et l'extraction de la réponse. Dans cet article, nous nous intéressons au troisième module, plus particulièrement dans le cas plus délicat où la réponse attendue n'est pas du type entité nommée. Nous décrivons comment l'analyseur Cass est employé pour marquer la réponse dans les phrases candidates et nous évaluons les résultats de cette approche. Au préalable, nous décrivons et évaluons le module dédié à l'analyse de la question, car les informations qui en sont issues sont nécessaires à notre étape finale d'extraction.

Mots-clés : systèmes de question-réponse.

Abstract

Question-answering systems are usually composed of three main modules, namely, question analysis, document selection and answer extraction. In this paper, we focus on the third module, more specifically when the expected answer is not a named entity because this kind of answer is more difficult to extract. We describe how the parser Cass is employed to tag the answer in the candidate sentence. Then, we evaluate the results of this approach. Beforehand, we describe and evaluate the question analysis module, as the information it produces is used during our final answer extraction step.

Keywords: question answering systems.

1. Introduction

Un système de question-réponse peut habituellement être décomposé en plusieurs étapes : analyse des questions, sélection des documents ou des phrases, et extraction de la réponse. De nombreuses stratégies ont été implémentées pour ce dernier module, mais celles-ci ont rarement été évaluées isolément. (Moldovan *et al.*, 2003) par exemple présentent cependant une analyse de leur système module par module. Nous avons choisi de présenter ici notre propre stratégie d'extraction de la réponse, qui a été utilisée lors de l'évaluation de systèmes de question-réponse multilingues CLEF 2005. Nous nous consacrerons tout particulièrement à l'extraction des réponses non entités nommées, qui présentent des difficultés particulières. Le module d'extraction de la réponse étant fortement dépendant des performances du module d'analyse de la question qui lui en fournit les caractéristiques, nous examinerons dans un premier temps ce module.

2. Architecture du système

Notre système de question-réponse peut traiter des questions et des documents en anglais comme en français, et permet d'obtenir la réponse dans une langue différente de la question. Il a participé à la campagne d'évaluation CLEF 2005 ¹, et notamment à la tâche français vers anglais, pour laquelle les questions sont en français et les documents à traiter en anglais.

Ce système est composé de plusieurs modules présentés figure 1 et dont une description plus détaillée pourra être trouvée dans (Ferret *et al.*, 2002). Le premier module est celui de l'analyse des questions, dont le but est de détecter un certain nombre de leurs caractéristiques qui permettront de trouver les réponses dans les documents. Puis la collection est parcourue grâce au moteur de recherche MG ². Les documents retournés sont ensuite réindexés en fonction de la présence des termes de la question ou de leurs variantes, et plus précisément de leur nombre et de leur type ; puis un module permet de reconnaître les entités nommées, et les phrases des documents sélectionnés sont pondérées en fonction des informations sur la question. Enfin, différents processus dépendant du type attendu de la réponse sont appliqués dans le but d'extraire les réponses des phrases.

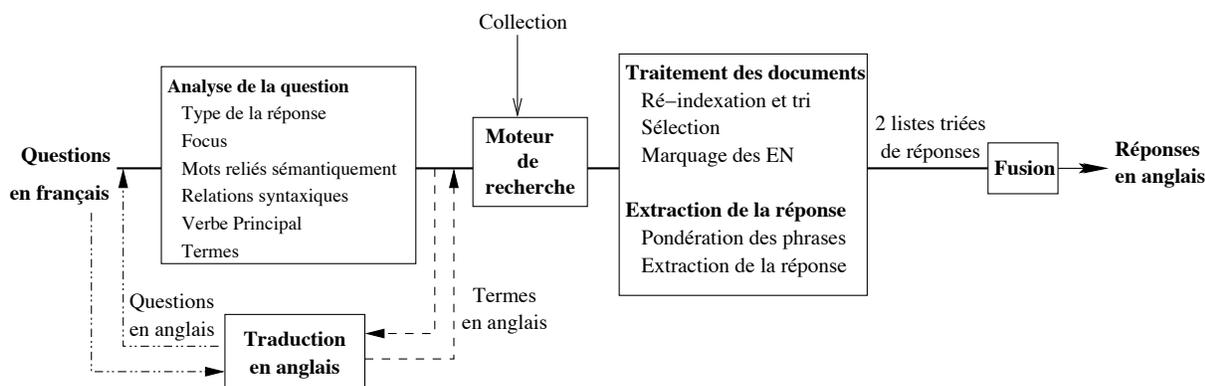


Figure 1. Architecture du système de question-réponse multilingue

3. L'extraction des réponses dans les systèmes de question-réponse

L'étape d'extraction de la réponse à partir des phrases sélectionnées dans la phase de recherche d'information, est généralement fortement liée à l'étape d'analyse des questions : en effet, des informations déduites de la question comme sa catégorie ou son type attendu, dépendra la stratégie d'extraction possible. Ces deux modules peuvent empiéter plus ou moins l'un sur l'autre. Ainsi, (Hartrumpf, 2005) effectue une analyse sémantique des questions, afin d'obtenir leur représentation sous la forme de réseaux sémantiques, puis infère autant que possible des reformulations de chaque question, qu'il comparera ensuite avec chaque phrase du corpus. Le travail d'analyse puis de reformulation de la question est donc très important dans ce système, et permet de réduire considérablement la distance sémantique entre la question et les phrases réponses.

Afin d'apparier les questions et les phrases candidates, diverses techniques peuvent être mises en place, utilisant des outils ou des ressources variés, et des représentations plus ou moins

¹ Multilingual Question Answering task at the Cross Language Evaluation Forum, <http://clef-qa.itc.it/>

² MG for Managing Gigabytes, <http://www.cs.mu.oz.au/mg/>

linguistiques. Ainsi, (Moldovan *et al.*, 2003) utilisent une méthode d'appariement fondée sur la logique, tandis que (Brill *et al.*, 2002) se fondent sur des règles surfaciques de reformulation et sur la redondance des réponses. Certaines caractéristiques sont néanmoins communes à la plupart des systèmes, comme la recherche du type attendu lors de l'analyse de la question, en particulier dans le cas d'une entité nommée. Cependant, si cette caractéristique est commune, la hiérarchie des types attendus varie d'un système à l'autre, et dans la campagne d'évaluation CLEF 2005, les systèmes affichaient de 3 (Pérez-Coutiño *et al.*, 2005) à 86 (Laurent *et al.*, 2005) types attendus. De nombreux systèmes utilisent également des patrons sous forme d'expressions régulières, notamment pour les questions n'attendant pas une réponse entité nommée, car le lien entre les éléments de la question et la réponse n'est pas toujours explicite (par exemple dans le cas d'une question portant sur la date de naissance d'un homme célèbre). De façon générale néanmoins, lorsque la question n'attend pas une entité nommée comme réponse, peu de systèmes ont une stratégie bien définie d'extraction de la réponse. (Ahn *et al.*, 2005) ont défini des règles de correspondance entre leurs classes de questions et les types attendus, même dans le cas où la réponse n'est pas une entité nommée : ainsi, pour une question *CAUSE_REASON*, la réponse attendue doit être une phrase. (Laurent *et al.*, 2005) marquent quant à eux les types de questions auxquelles les entités des documents pourraient répondre au moment de l'indexation, selon des patrons surfaciques.

4. Analyse des questions

Dans notre système, l'analyse des questions est composée de plusieurs modules séquentiels. Tout d'abord, un étiquetage morpho-syntaxique est effectué par le TreeTagger³ ; puis l'analyseur syntaxique Cass⁴ détermine leur segmentation en constituants et les relations de dépendance syntaxique à l'aide d'une grammaire que nous avons écrite pour le traitement spécifique des questions. A partir de ces analyses, certaines caractéristiques de la question sont déduites : type attendu de la réponse, catégorie de la question, contextes temporels... Ces informations vont ensuite être utilisées par les différents modules du système.

Le système devant pouvoir travailler à la fois sur de l'anglais et du français, son implémentation respecte une quasi-complète symétrie entre les langues, et toutes les informations dépendant de la langue ont été regroupées dans des fichiers de données pour une adaptation plus aisée.

Le module d'extraction de la réponse utilise les informations suivantes pour détecter une réponse dans une phrase candidate : catégorie de la question, et type attendu de la réponse dans tous les cas, et verbe principal de la question, ainsi que focus dans le cas d'une réponse non entité nommée. Pour déterminer ces informations, le module d'analyse de la question utilise l'étiquetage morpho-syntaxique et l'analyse syntaxique pour construire une forme simplifiée de la question. Ainsi, la question *Qu'est-ce que la LSPN ?* sera convertie en : *qu_est-ce_queNPI*.

Puis cette forme simplifiée est comparée à une liste de tels patrons sous forme d'expressions régulières, ce qui permettra dans l'exemple précédent de reconnaître une question de définition. Un lexique nous permettra ensuite de reconnaître "LSPN" comme un acronyme, et par conséquent la question comme une question de définition d'un acronyme, qui, dans le cadre des évaluations question-réponse, devra avoir pour réponse le nom complet associé à cet acronyme.

Les catégories de questions déterminées ont été établies notamment pour répondre aux besoins du module d'extraction de la réponse, lorsque la réponse attendue n'est pas une entité nommée.

³ <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

⁴ <http://www.sfs.nphil.uni-tuebingen.de/~abney/>

	Questions EN		Questions non EN	
	Rappel	Précision	Rappel	Précision
Français	87%	94%	90%	90%
Anglais	83%	99%	88%	92%

Tableau 1. Résultats de l'analyse des questions

Nous avons ainsi mené une étude sur un corpus de questions et de réponses longues associées, afin de déterminer des catégories de questions attendant des patrons spécifiques, et l'analyse des questions a été adaptée à ces catégories empiriques.

Notre module d'analyse des questions a pu être testé dans les deux langues à l'occasion de plusieurs campagnes d'évaluation de systèmes de question-réponse. Nous présentons tableau 1 une évaluation de notre module sur les questions de la campagne d'évaluation de systèmes multilingues CLEF 2005. Cette évaluation a été effectuée sur le fichier des questions de la tâche français vers anglais pour la partie française (première ligne de résultats dans le tableau), et sur leur traduction manuelle en anglais (seconde ligne) pour la partie anglaise. Cette traduction qui nous est donnée par les organisateurs de CLEF à l'issue de la campagne, est intéressante pour nous puisqu'elle nous permet de tester nos modules sur une traduction correcte, puisque manuelle, des questions.

Étiquetage morpho-syntaxique

What is the G7? - G7 est étiqueté adjectif

Analyse syntaxique

How many people are diagnosed as suffering from colon cancer each year?
- *each year* est attaché au groupe prépositionnel précédent *from colon cancer*

Incomplétion des listes d'entités nommées

Name a traffic free resort in Switzerland. - *resort* n'est pas reconnu comme un lieu

Règles erronées

Who was the Norwegian Prime Minister when the referendum on Norway's possible accession to the EU was held? - l'analyse des questions donne *hold* comme verbe principal

Figure 2. Exemples d'erreurs du module d'analyse des questions

Les erreurs de ce module peuvent avoir des raisons diverses, nous en présentons quelques exemples figure 2. Une partie des erreurs d'étiquetage ou d'analyse syntaxique ont été corrigées grâce à l'écriture de règles de post-traitement pour l'étiquetage, et de grammaires spécifiques aux questions pour l'analyse. En ce qui concerne la complétion des listes de déclencheurs d'entités nommées, une solution adoptée par de nombreux systèmes est de se fier à la hiérarchie d'une base de connaissances comme WordNet. Cependant, sans désambiguïsation des termes de la question, qui est de toute façon rendue difficile par le faible contexte d'une question, cette méthode n'est pas toujours fiable ; en effet, lorsqu'un mot possède plusieurs sens, on peut alors lui faire correspondre un type totalement inapproprié, et qui risque fortement d'induire en erreur l'extraction de la réponse. Par exemple, pour une question comme *What party did Andrei*

Brezhnev found ?, le mot *party* pourra être considéré comme un hyponyme de *person* (dans le sens d'une personne impliquée dans une affaire légale). Il faut alors tenir compte de critères d'utilisation ou d'heuristiques sur le sens pour choisir le bon type d'entité nommée attendu par la question.

Nous avons quant à nous choisi d'inclure dans nos listes de termes déclencheurs, uniquement les termes dont l'utilisation est peu ambiguë. Nous avons ainsi complété certaines de nos listes en anglais à partir de WordNet, en ne conservant dans les listes d'hyponymes que les termes pour lesquels les fréquences des sens laissent peu d'ambiguïtés. Par exemple, le terme *hero* a été conservé comme déclencheur du type entité nommée *PERSON* car son sens principal est un hyponyme de personne, selon les fréquences d'utilisation fournies par WordNet⁵. Ceci nous a permis d'augmenter considérablement le nombre de termes reconnus pour ces classes ; ces termes doivent cependant encore être transférés au français.

Lors de la campagne CLEF 2005, un nouveau type de questions temporellement restreintes a été introduit. Cette dénomination se réfère à des questions comprenant un contexte temporel, qui peut être soit une date : *Qui est devenu le Premier Ministre de Slovénie en 1992 ?*, soit une période : *Combien de millions de personnes se sont échappés d'Europe de l'Est en RFA entre 1950 et 1992 ?*, soit un événement : *Quel poste occupait Silvio Berlusconi avant qu'il ne démissionne ?*.

Notre module d'analyse des questions a été adapté, afin de pouvoir traiter ces questions. Nous avons donc ajouté une reconnaissance du contexte temporel de la question, qui a permis de déterminer correctement le contexte temporel ainsi que son type pour toutes les questions temporellement restreintes.

5. Extraction des réponses

Le module d'extraction de la réponse prend en entrée des phrases appelées candidates, issues de la sélection des documents, puis de leur ré-indexation par Fastr (Jacquemin, 1996), et enfin d'une seconde sélection s'appuyant sur les mots-clefs et sur certaines caractéristiques de la question. Deux stratégies différentes sont utilisées pour extraire la réponse des phrases sélectionnées, en fonction du type attendu de la réponse. Dans le cas où ce type est une entité nommée, on sélectionne l'entité nommée du type attendu la plus proche du barycentre des mots de la question ou de leurs variantes (pondérés en fonction de divers critères). Les questions de la forme *combien de* sont cependant traitées à part, selon une stratégie plus proche de celle présentée ci-après.

Dans le cas où ce type n'est pas une entité nommée, différents patrons sont appliqués selon la catégorie de la question. Ces patrons ont été écrits sous la forme de règles de l'analyseur Cass ; les phrases candidates sont alors traitées par l'analyseur. Après avoir décrit comment nous utilisons Cass pour écrire des patrons d'extraction, nous donnons les résultats du module d'extraction ; enfin nous examinons les différents types d'erreur qui font échouer ce module.

5.1. L'analyseur Cass

L'analyseur Cass est fondé sur une cascade d'automates à états-finis. Il comprend différents niveaux de reconnaissance des phrases en entrée, qui sont chacun constitués de règles de gram-

⁵ Il existe un sens de *hero* qui fait référence à un type de sandwich.

Niveau 1 : Constitution des groupes

GNFoc -> DT ? RB ? (ADJ (CC ADJ) ?) ? FC RB ? ;

GN -> DT ? RB* ADJ* (NNINNS)+ RB* ADJ* ;

GNNP -> DT ? RB* ADJ* (NPINPS)+ ;

Niveau 2 : Marquage des réponses

RgFsep -> b= (GNIGNNP) SEP GNFoc ;

Figure 3. Exemples de règles pour l'analyseur Cass

maires sous formes d'expressions régulières. Son fonctionnement est le suivant : il prend en entrée des phrases étiquetées par un étiqueteur morpho-syntaxique, et applique chaque niveau de grammaire à ces phrases. Un avantage important de Cass est que sa grammaire est modifiable ; ainsi, la grammaire utilisée pour l'analyse des questions a été adaptée pour prendre en compte un certain nombre de spécificités des formes interrogatives.

Dans le module d'extraction de la réponse, Cass n'est pas utilisé dans un but d'analyse syntaxique, mais pour marquer les réponses dans les phrases candidates, un dernier module s'occupant de l'extraction finale. Cass prend en entrée d'une part les phrases candidates étiquetées par le TreeTagger, d'autre part un second fichier, lui aussi dans le format 3 colonnes du TreeTagger et où certains mots ont reçu une étiquette particulière différente de celle attribuée par le TreeTagger. Ce second fichier nous permet de marquer d'une étiquette spécifique les informations issues de l'analyse de la question telles que : le focus (FC), le verbe principal (VP), le type général de la réponse attendu (TG) ; les synonymes de toutes ces données sont également marqués (FS, VS, TS).

Ensuite les règles de grammaire écrites ici pour Cass se comportent comme des patrons permettant l'extraction finale : les places possibles des réponses sont marquées à proximité des termes caractéristiques de la question. Le focus est recherché en priorité, et à défaut de sa présence le verbe principal et le type général le sont ensuite.

Pour donner un exemple d'utilisation de Cass, prenons la question *Which genes cause cancer ?*, qui porte la catégorie *Quel* et à laquelle le module d'extraction répond avec succès au rang 1. Dans cette question, le focus est *gene* (FC), le verbe principal est *cause* (VP), il n'y a pas de type général. Seul le verbe possède des synonymes : *do, make, induce, stimulate, have, get*, qui sont tous marqués VS.

Comme il a été dit plus haut, les règles sont organisées par niveaux ; un premier niveau permet de construire les groupes syntaxiques. Ci-dessus, figure 3, les 3 règles du niveau 1 construisent le groupe nominal contenant le focus, celui contenant le nom commun et celui contenant le nom propre. Ensuite la règle de niveau 2 permet de marquer comme réponse finale candidate tout GN qui serait à gauche du GN contenant le focus et qui en serait séparé à l'aide d'un séparateur quelconque (la virgule, le tiret...).

Grâce à ces 4 règles, la réponse correcte *oncogenes*, est marquée dans la phrase candidate : *G proteins have assumed critical importance in recent years as researchers have discovered that several of them are produced by oncogenes, genes that cause cancer*. Et le module final d'extraction, retourne le groupe marqué de l'étiquette *b=* :

[RgFsep **b**=[GN [NNS **oncogenes**]] [VIRG ,] [GNFoc [FC genes]]]

5.2. Résultats du module d'extraction

Comme dans l'évaluation des questions, nous avons utilisé, pour évaluer ce module, la traduction manuelle des questions du français vers l'anglais, proposée par les organisateurs de CLEF 2005. C'est en effet pour cette série de questions, que la perte de réponses au moment de l'extraction est la plus grande, comme le montre la table 2. Dans cette table, nous avons également indiqué, afin de permettre la comparaison, les chiffres sur les questions attendant une entité nommée. Ainsi, au rang 1, pour les réponses non EN on passe de 41,17 % de bonnes réponses à 30,6 %. En ce qui concerne les questions EN, on perd beaucoup moins de bonnes réponses puisque l'on passe de 29,56 % à 26,08 %.

	Questions EN	Questions non EN	Total
Nombre de questions	115	85	200
Réponses courtes correctes			
- au rang 1	30 - 26,08 %	26 - 30,6 %	28 %
- dans les 5 premiers rangs	39 - 33,9 %	35 - 41,17 %	37 %
Réponses longues correctes			
- au rang 1	34 - 29,56 %	35 - 41,17 %	34,5 %
- dans les 5 premiers rangs	57 - 49,56 %	48 - 56,47 %	52,5 %

Tableau 2. Pertes lors du passage de la réponse longue à la réponse courte : difficultés de l'extraction

Lors de l'analyse de cette série de questions, six catégories ont été détectées. La catégorie *DéfinitionAutre* attend pour réponse ce que l'on pourrait appeler un extrait de définition, comme cela est le cas dans *What is Jari Litmanen's profession?*. La catégorie *Instance*, attend comme réponse une instance particulière choisie parmi d'autres qui auraient convenu également, comme dans *Name a building wrapped by Christo*. La catégorie *Quel* attend elle aussi une instance mais cette fois une seule réponse est possible *Which dynasty rules Jordan?* Les patrons permettant de répondre aux questions de ces deux dernières catégories sont très vraisemblablement les mêmes. Enfin pour 4 questions le module d'analyse des questions n'est pas parvenu à attribuer de catégorie, des patrons par défaut seront alors appliqués.

Dans la table 3, on a différencié les résultats par catégorie de question, et l'on observe que les résultats les plus satisfaisants sont ceux de la catégorie *Définition*. En effet, pour cette catégorie, des patrons relativement simples sont efficaces : la réponse est très souvent accolée au focus, parfois une virgule ou une parenthèse les sépare. Pour cette catégorie, nous avons également testé des patrons minimaux, qui marquent comme réponse le groupe nominal à gauche ou à droite du focus. Avec ces patrons simplistes, nous obtenons déjà 17 réponses correctes au rang 1, contre 22 avec des patrons plus élaborés. L'écriture des patrons simplistes n'a pas encore été faite pour toutes les catégories de question, mais nous pensons que c'est vraisemblablement pour la catégorie *Définition* qu'ils remporteront les meilleurs résultats.

5.3. Exemples d'erreur lors de l'extraction

Au vu de ces résultats, nous avons souhaité examiner les dysfonctionnements pouvant expliquer de telles pertes et avons regardé un peu plus en détails quelques-unes des questions pour lesquelles la réponse longue correcte est retournée au rang 1 et perdue dans les réponses courtes.

Catégorie des questions	Nombre de questions	Rang des réponses correctes	Nombre de réponses correctes	
			courtes	longues
Combien	2	- rang 1 - 5 1ers rangs	0 0	0 0
Définition	49	- rang 1 - 5 1ers rangs	22 26	23 31
DéfinitionAutre	2	- rang 1 - 5 1ers rangs	0 0	0 0
Instance	3	- rang 1 - 5 1ers rangs	0 1	0 1
Quel	25	- rang 1 - 5 1ers rangs	4 7	10 13
Sans catégorie	4	- rang 1 - 5 1ers rangs	0 1	2 3
Total	85	- rang 1 - 5 1ers rangs	26 35	35 48

Tableau 3. Résultats des patrons par catégorie de question

Nous avons ainsi mis en évidence 3 grandes catégories d'erreur, que nous explicitons ici.

Les erreurs *en amont*, sont les erreurs que l'on ne peut pas imputer au module d'extraction ou qui nécessiteraient des modifications du principe actuel des patrons. Ainsi dans la question, *Which dynasty rules Jordan?*, le module d'analyse de la question n'a pas détecté *rule* en tant que verbe principal ⁶. Ou encore dans la question *Which computer virus was confirmed as a hoax by the US National Computer Security Association?* le mot *virus* est retenu comme focus, mais la phrase réponse ne contient que *hoax*. Ici, le focus est correct du point de vue de l'analyse de la question, mais ne suffit pas à permettre d'extraire la réponse.

Certaines erreurs peuvent néanmoins être attribuées aux patrons, même si elles ne sont pas pour autant facilement corrigibles. Ainsi pour la question *Which symbol has been used to hallmark sterling silver in Scotland since 1473?*, les patrons marquent bien la bonne réponse, *An Edinburgh Castle*, mais en lui attribuant une note ⁷ si basse que c'est une autre réponse erronée qui est finalement retournée. Dans le cas de la question *Which alphabet has only four letters, A, C, G and T?*, aucun patron n'est prévu pour marquer l'adjectif du groupe nominal contenant le focus comme étant la bonne réponse, *The genetic alphabet* n'est donc pas marqué et a fortiori pas retourné.

La troisième et dernière catégorie permet de montrer les limites des patrons : ce sont des exemples où la syntaxe de la phrase réponse est trop complexe pour que les patrons puissent fonctionner. Par exemple, si une incise se glisse entre le focus et la réponse ou encore dans le cas de certaines énumérations alors les patrons sont impuissants.

⁶ Quand bien même il l'aurait fait, la phrase réponse ne contenant pas le terme *dynasty*, aucun patron n'aurait su extraire la bonne réponse.

⁷ Deux notes sont utilisées pour classer les réponses. Le mécanisme des patrons attribue à chaque réponse étiquetée une note en fonction de la fiabilité du patron appliqué. Au préalable, chaque phrase candidate, a reçu une première note tenant compte de la présence des mots de la question.

5.4. Possibilités d'amélioration

Malgré ces limites, l'approche par patrons que nous avons implémentée présente l'avantage principal d'être très robuste, et permet de traiter des questions pour lesquelles la reconnaissance du type attendu est soit difficile soit impossible. Les catégories de questions que nous avons définies nous permettent bien de regrouper des patrons particulièrement pertinents, et l'analyse de nos résultats nous incite à développer les distinctions entre questions.

Plusieurs pistes permettraient d'améliorer les résultats de notre module. Tout d'abord les réponses proposées par les patrons pourraient être vérifiées d'un point de vue sémantique. En effet les patrons d'extraction écrits ici sont des patrons de proximité qui extraient une réponse proche du focus ou du type général mais n'effectuent aucun contrôle quant à la nature de la réponse retournée. Il pourrait donc être intéressant de vérifier que la réponse extraite correspond bien au type attendu.

D'autre part, l'utilisation d'un analyseur syntaxique, appliqué aux phrases candidates permettrait de compléter efficacement les patrons. Un tel apport a été montré dans (Ligozat, 2004). Les deux méthodes, patrons et analyse plus profonde, pourraient être appliquées en parallèle et les résultats seraient finalement fusionnés.

6. Conclusion et perspectives

Cet article a été l'occasion de présenter une méthode possible pour effectuer l'extraction de la réponse finale dans un système de question-réponse, dans le cas plus délicat où la réponse attendue n'est pas du type EN. Nous avons expliqué comment les informations issues de l'analyse des questions nous permettent, à l'aide aussi de l'analyseur Cass, de marquer les réponses possibles dans les phrases candidates, et montré les performances possibles d'une approche surfacique comme celle-ci. Nous avons aussi souligné les limites de cette approche à l'aide de patrons.

Certaines améliorations seraient possibles pour ce module d'extraction des réponses, comme une meilleure prise en compte du type attendu de la réponse lorsqu'il existe ; une autre voie d'amélioration de ce module d'extraction consisterait à utiliser en parallèle une ou plusieurs stratégies différentes.

Références

- AHN D., JIJKOUN V., MÜLLER K., DE RIJKE M., TJONG E. et SANG K. (2005). « The University of Amsterdam at QA@CLEF 2005 ». In *Working Notes, CLEF Cross-Language Evaluation Forum*. Vienna, Austria. www.clef-campaign.org/2005/working_notes/.
- BRILL E., DUMAIS S. et BANKO M. (2002). « An analysis of the AskMSR question-answering system ». In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*. Philadelphia, Pennsylvania, USA, p. 257-264.
- FERRET O., GRAU B., HURAUULT-PLANTET M., ILLOUZ G., JACQUEMIN C., MONCEAUX L., ROBBA I. et VILNAT A. (2002). « How NLP Can Improve Question Answering ». In *Knowledge Organization*, 29 (3-4), 135-155.
- HARTRUMPF S. (2005). « University of Hagen at QA@CLEF 2005 : Extending Knowledge and Deepening Linguistic Processing for Question Answering ». In *Working Notes, CLEF Cross-Language Evaluation Forum*. Vienna, Austria.
- JACQUEMIN C. (1996). « A symbolic and surgical acquisition of terms through variation. ».

In *Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing*, p. 425-438.

- LAURENT D., SÉGUÉLA P. et NÈGRE S. (2005). « Cross Lingual Question Answering using QRISTAL for CLEF 2005 ». In *Working Notes, CLEF Cross-Language Evaluation Forum*. Vienna, Austria.
- LIGOZAT A.-L. (2004). « Système de Question-Réponse : Apport de l'Analyse Syntaxique à l'Extraction de la Réponse ». In *Actes de Récital*. Fes, Maroc.
- MOLDOVAN D., PASCA M., HARABAGIU S. et SURDEANU M. (2003). « Performance Issues and Error Analysis in an Open-Domain Question Answering System ». In *ACM Transactions on Information Systems*, volume 21. p. 133–154.
- PÉREZ-COUTIÑO M., Y GÓMEZ M. M., LÓPEZ-LÓPEZ A. et VILLASEÑOR-PINEDA L. (2005). « Experiments for tuning the values of lexical features in Question Answering for Spanish ». In *Working Notes, CLEF Cross-Language Evaluation Forum*. Vienna, Austria.