
La réponse automatique comme solution à la gestion des relations avec la clientèle

Une application de la question-réponse

Leila Kosseim* — Luc Plamondon** — Guy Lapalme**

* CLAC Laboratory, Computer Science Department
Concordia University, 1400 boul. de Maisonneuve Ouest, LB 903-7
Montréal, Québec, Canada H3G 1M8
kosseim@cs.concordia.ca

** RALI, Département d'informatique et recherche opérationnelle
Université de Montréal, C.P. 6128, succursale Centre-ville
Montréal, Québec, Canada H3C 3J7
{plamondl,lapalme}@iro.umontreal.ca

RÉSUMÉ. De plus en plus d'entreprises font face à une surabondance de questions de la part de leurs clients qui utilisent maintenant le courrier électronique plutôt que le courrier postal ou le téléphone. Dans cet article, nous discutons de l'utilisation de techniques issues du domaine de la question-réponse pour améliorer la gestion des relations avec la clientèle d'une entreprise en répondant automatiquement aux questions de sa clientèle. Nous expliquons comment la question-réponse peut être utilisée pour automatiser l'extraction de réponses, la composition de phrases-réponses naturelles et la gestion de questions et de documents translinguistiques, et présentons nos expériences dans ces domaines. Les enjeux particuliers à considérer dans le cadre de l'utilisation de la question-réponse pour la gestion des relations avec la clientèle sont ensuite détaillés et des solutions sont proposées.

ABSTRACT. Today, more and more companies are facing an over-abundance of questions from their customers who now use an electronic medium to communicate rather than a more traditional medium such as mail or telephone. In this paper, we discuss the use of techniques developed in the area of question-answering to improve customer service management by responding automatically to customer questions. We justify how question-answering can be used to automate the extraction answers, the generation of a natural responses, and be used to deal with translingual questions and documents, and present our experiments in these domains. The challenges of using question-answering specifically for customer service management are finally presented along with our preliminary work toward some solutions.

MOTS-CLÉS : traitement automatique des langues naturelles, réponse automatique, question-réponse, extraction de réponse, génération de réponse, gestion des relations avec la clientèle.

KEYWORDS: natural language processing, automatic response, question-answering, answer extraction, answer generation, customer relations management.

1. Introduction à la réponse automatique

De plus en plus d'entreprises font face à une surabondance de requêtes de la part de leurs clients qui naviguent sur le site web de l'entreprise et utilisent le courrier électronique pour s'informer sur l'entreprise et ses produits. La gestion des relations avec la clientèle doit maintenant s'adapter aux communications textuelles électroniques. Les techniques issues du domaine du traitement automatique des langues naturelles (TALN) sont devenues essentielles au développement d'outils efficaces pour faciliter le traitement rapide de ces requêtes électroniques.

1.1. Le problème

Il est difficile de trouver des sources fiables faisant état de la gestion des relations avec la clientèle à cause d'intérêts commerciaux et parce que les données disponibles sont généralement offertes par des entreprises vendant des solutions informatiques de gestion des relations avec la clientèle. Cependant, l'étude réalisée en 2000 par la compagnie Jupiter Communications [JUP 00] et rapportée par [BAN 01] semble être typique : 55 % des clients des 125 sites web les plus importants s'attendent à recevoir des réponses correctes à leur courriel dans un laps de temps d'au plus 6 heures. Cependant, seulement 20 % des compagnies satisfont à leurs attentes et 42 % des sites n'ont jamais répondu aux courriels, ont pris plus de cinq jours à répondre aux questions ou n'avaient pas d'adresse électronique indiquée sur leur site.

Devant ce constat alarmant, il devient nécessaire pour le domaine du traitement automatique de la langue de développer de nouvelles approches pour assister, voire automatiser, la gestion des demandes électroniques d'une entreprise.

Pour une entreprise, les communications électroniques avec la clientèle présentent de nombreux avantages. Que ce soit par courriel ou par système de requête en ligne, l'entreprise peut facilement garder trace des communications avec ses clients soit à des fins statistiques, soit à des fins de contrôle de qualité. De plus, par voie électronique, il est plus facile d'envoyer des réponses complexes ; des instructions longues et détaillées, des images, des clips audio ou vidéo sont choses possibles par voie électronique, mais impensables à travers des médias plus traditionnels comme le téléphone. Les avantages des communications électroniques pour la gestion des relations avec la clientèle sont nombreux, mais des outils de TALN sont absolument nécessaires pour maintenir un service rapide et efficace.

1.2. La question-réponse, une solution intéressante

Il y a de nombreuses années, pour stocker et accéder facilement et rapidement à leurs données, les entreprises se sont tournées vers les bases de données (BD). Un préposé utilisant un langage spécial pouvait alors formuler une requête et questionner ces bases pour pouvoir répondre aux demandes de la clientèle. Pour faciliter l'accès à

ces bases, de nombreux travaux ont tenté de développer des interfaces en langue naturelle (LN). En fait, un des domaines les plus anciens et les plus étudiés en TALN est le développement d'interfaces en LN aux BD (ex. [WOO 77]). Grâce à ces interfaces, tout utilisateur peut interroger la base de données sans en connaître la structure interne et sans en apprendre le langage de requête.

Les premiers travaux en question-réponse remontent eux aussi aux premiers jours de l'intelligence artificielle [LEH 78]. À cette époque, les systèmes de question-réponse travaillaient sur des *mondes-jouets* et n'étaient pas assez robustes pour s'adapter à des applications réelles.

Depuis quelques années cependant, deux facteurs importants ont ravivé l'intérêt de la communauté du TALN pour la question-réponse. D'une part, les techniques de base en recherche d'information, étiquetage grammatical et extraction d'entités nommées sont maintenant bien maîtrisées et suffisamment disponibles pour permettre le développement de systèmes plus performants. D'autre part, le besoin de l'industrie pour des systèmes intelligents de recherche d'information est devenu pressant.

Le but d'un système de question-réponse (*question-answering*) est de trouver des réponses exactes à des questions exprimées en langue naturelle en recherchant dans une grande collection de documents, et non dans une base de données [VOO 01]. Les interfaces en LN aux BD et les systèmes de question-réponse partagent un point commun intéressant : tous deux prennent en entrée une question formulée en langue naturelle et tentent de l'interpréter afin d'y répondre correctement. Cependant, les deux types de systèmes diffèrent fondamentalement dans leurs buts scientifiques et leurs contraintes techniques. Les interfaces en LN aux BD traitent des connaissances structurées, c'est-à-dire des données qui ont déjà été décomposées et organisées sémantiquement en entités et relations, tandis que les systèmes de question-réponse cherchent la réponse dans une collection de textes libres non structurés. À cet égard, la question-réponse peut être vue comme une extension des interfaces en LN aux BD où le domaine de discours serait ouvert et où la recherche serait effectuée dans une base non structurée. Les difficultés inhérentes à l'analyse de textes libres, telles la résolution de coréférences, la désambiguïsation sémantique et l'analyse syntaxique complexe, rendent le problème très intéressant pour la communauté du TALN et en ont fait un domaine aujourd'hui très populaire.

Dans le cadre de la gestion des relations avec la clientèle, la majorité des réponses aux requêtes se trouvent dans de la documentation non structurée : documentation corporative, pages web, courriels, mémos, etc. De plus, l'utilisation d'une langue naturelle permet aux clients de questionner eux-mêmes ces bases documentaires dans un langage qu'ils maîtrisent déjà. Pour ces raisons, nous avons investigué en profondeur le domaine de la question-réponse dans l'optique de l'appliquer au problème de la gestion des relations avec la clientèle.

Les expériences en question-réponse présentées dans les prochains paragraphes ont été développées dans le cadre d'une collaboration entre le laboratoire RALI (Université de Montréal), le laboratoire ClaC (Université Concordia) et *Bell Canada En-*

terprises (BCE). BCE est une grande entreprise canadienne offrant des services de communication et de divertissement, tels que la téléphonie et l'accès à l'internet, à des clients commerciaux et individuels. Pour que l'entreprise demeure compétitive, son service à la clientèle doit être efficace ; pour atteindre ce but, BCE a mandaté nos laboratoires afin d'étudier le problème de l'automatisation de la réponse aux demandes de ses clients, en particulier pour alléger le travail des préposés du service à la clientèle en ce qui concerne les questions fréquentes. En effet, après une étude des courriels envoyés à BCE, nous avons identifié qu'un grand nombre d'entre eux sont en fait des demandes d'informations de routine, dont la réponse se trouve dans une base documentaire de l'entreprise (site web, rapports internes, etc.). Dans les prochains paragraphes, nous présenterons nos travaux en question-réponse dans le cadre de cette collaboration avec BCE. En particulier, nous discuterons d'extraction de réponse, de composition de réponse et de question-réponse translinguistique.

2. Extraction de réponse

L'extraction de réponse consiste à identifier et à extraire d'une grande base de documents la chaîne de caractères qui constitue la réponse exacte à une question. Pour explorer cet axe de recherche, nous avons développé le système de question-réponse Quantum.

2.1. Quantum

Quantum permet à un client d'écrire sa question dans un formulaire web et d'obtenir immédiatement une réponse. La question, posée en anglais, doit être courte et syntaxiquement bien formée, comme par exemple *Who is BCE Emergis' founding CEO ? How many viewers does CTV's «Who wants to be a millionaire» attract ? What revenue did Telesat report for 2002 ?* Quantum se limite aux questions factuelles, c'est-à-dire dont la réponse est un fait explicitement mentionné dans le texte ; les questions nécessitant des capacités de jugement ou d'inférence sont exclues car les difficultés qu'elles posent vont bien au-delà du simple traitement de la langue. Les réponses données au client sont de courts extraits de textes (les textes sont en l'occurrence des communiqués de presse publiés sur le site web de BCE), en général de la longueur d'un syntagme nominal. La version en ligne de Quantum¹ donne cinq suggestions de réponse par question, ordonnées de la meilleure à la moins plausible. Les suggestions de réponse sont accompagnées d'un lien vers le document d'origine afin de permettre au client de prendre connaissance du contexte de la réponse.

Quantum a d'abord été développé en fonction des campagnes d'évaluation TREC. Ces campagnes fournissent un cadre d'évaluation standard, c'est pourquoi nous présenterons d'abord la teneur de ces campagnes avant de décrire l'architecture de Quantum et d'évaluer ses performances.

1. <http://www-rali.iro.umontreal.ca/ProjetLUB.fr.html>

2.2. Les campagnes d'évaluation TREC

Les campagnes TREC (*Text Retrieval Conference*), sous les auspices du *National Institute for Standards and Technology* des États-Unis (NIST), ont donné le coup d'envoi aux recherches en question-réponse, en 1999, en lançant la première campagne d'évaluation de systèmes de question-réponse. Depuis, chaque année, NIST met à la disposition des équipes participantes un ensemble de questions et une collection de textes afin qu'elles mettent leur système à l'épreuve. Les réponses sont ensuite évaluées par NIST suivant une méthodologie standard. Les questions et les textes fournis par NIST sont exclusivement en anglais, c'est pourquoi des campagnes d'évaluation de systèmes traitant d'autres langues ont récemment vu le jour : EQueR (dans le cadre du projet EVALDA de Technolanguage/ELDA) pour la question-réponse en français et CLEF pour la question-réponse en n'importe quelle langue européenne et la question-réponse multilingue (textes en anglais et questions en d'autres langues, dont le français).

À ce jour, les campagnes TREC continuent d'être le foyer des principaux avancements en question-réponse. On y a vu l'apparition de systèmes dérivés des moteurs de recherche classiques, de systèmes à apprentissage statistique, de systèmes faisant de l'analyse linguistique de surface et de systèmes faisant appel à des bases de connaissances et à des réseaux sémantiques tels que WordNet [MIL 95]. Bien que les réponses suggérées par les systèmes doivent être présentes dans la collection de textes standard à TREC, de plus en plus de systèmes consultent aussi le web : parfois pour départager plusieurs réponses prometteuses trouvées dans la collection de textes [MAG 02], d'autres fois pour y puiser directement des réponses (quitte à les apparier ensuite avec des réponses similaires présentes dans la collection) [BRI 01, BRI 02, CLA 01].

TREC-11 (en 2002) est la plus récente campagne à laquelle Quantum ait participé. La collection de textes était constituée d'environ un million de communiqués de presse en anglais diffusés par six fils de presse, dont le *New York Times* et *Associated Press*, pour un total de 3 gigaoctets de données (600 millions de mots). Les systèmes devaient extraire les réponses de ces documents et les réponses devaient être *exactes*, c'est-à-dire que la chaîne de caractères suggérée en guise de réponse ne devait contenir rien de plus ni de moins que la réponse. Les systèmes ne pouvaient soumettre qu'une seule suggestion de réponse par question. Pour de plus amples détails, le lecteur est invité à consulter la description de la piste *Question Answering* de TREC-11 [VOO 02]. Nous disposons des 500 questions de l'épreuve ainsi que de la liste des réponses acceptées pour chacune des questions par les évaluateurs humains ; cela nous permet d'évaluer les performances de Quantum et de ses composants de façon automatique.

2.3. Architecture de Quantum

Quantum, ainsi que la plupart des systèmes de question-réponse rencontrés dans les campagnes TREC, procède en trois étapes : analyse de la question, recherche des passages les plus pertinents et extraction de la réponse. La figure 1 schématise le

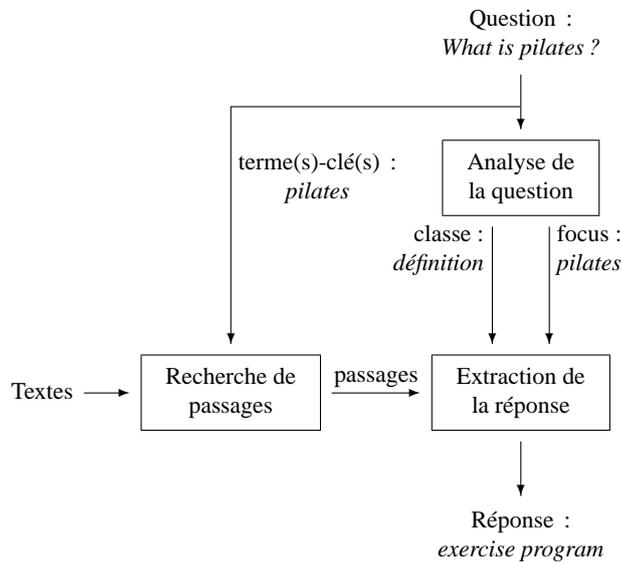


Figure 1. Architecture de Quantum. La question et les textes sont les deux entrées du système ; la réponse en est la sortie. Trois modules distincts sont nécessaires

traitement d'une question à l'aide d'un exemple de question tirée d'une campagne TREC.

2.3.1. Analyse de la question

Le but de l'analyse de la question est de déterminer le type de la réponse à chercher, donc le mécanisme d'extraction — ou fonction d'extraction — à utiliser. Parfois, il est nécessaire de transmettre à cette fonction un paramètre appelé *focus* de la question : il s'agit d'un mot ou groupe de mots présents dans la question et ayant un rapport étroit avec la réponse. Par exemple, la réponse à la question *What card game uses only 48 cards ?* doit être un hyponyme du focus *card game*, c'est pourquoi il est nécessaire de transmettre ce paramètre à la fonction d'extraction choisie. Quant à la réponse à la question *How many black keys are on the piano ?*, elle doit contenir un nombre suivi d'une répétition du focus *black keys*. Cependant, certaines autres questions comme *When was water found on Mars ?* ne nécessitent pas l'identification d'un focus puisqu'il s'agit ici de trouver une entité nommée de temps pour compléter le concept désigné par *when*. Peu importe qu'ils soient focus ou non, tous les mots de la question contribuent à l'identification de la réponse par le biais du score de la recherche d'information (section 2.3.2). Nous présentons au tableau 1 les 12 fonctions d'extraction de Quantum ainsi que des exemples de leur fonctionnement. Il est important de souligner que dans la classification des questions que nous avons retenue, la nécessité ou non de passer un focus à telle ou telle fonction d'extraction ainsi que les critères de

Fonction	Exemple de question et de forme de réponse
<i>définition</i>	Q : <i>What is an atom ?</i> (focus = <i>atom</i>) R : <hypéronyme de <i>atom</i> > , < <i>atom</i> ou hyponyme de <i>atom</i> > R : < <i>atom</i> ou hyponyme de <i>atom</i> > (<hypéronyme de <i>atom</i> >) R : < <i>atom</i> ou hyponyme de <i>atom</i> > is <hypéronyme de <i>atom</i> >
<i>spécialisation</i>	Q : <i>What card game uses only 48 cards ?</i> (focus = <i>card game</i>) R : <hyponyme de <i>card game</i> >
<i>cardinalité</i>	Q : <i>How many black keys are on the piano ?</i> (focus = <i>black keys</i>) R : <nombre> < <i>black keys</i> ou hyponyme de <i>black key</i> >
<i>measure</i>	Q : <i>How much vitamin C should you take in a day ?</i> (focus = <i>vitamin C</i>) R : <nombre> <hyponyme de <i>unit</i> > of < <i>vitamin C</i> ou hyponyme de <i>vitamin C</i> >
<i>attribut</i>	Q : <i>How high is Mount Kinabalu ?</i> (focus = <i>high</i>) R : Patrons divers
<i>synonyme</i>	Q : <i>What is another name for the North Star ?</i> (focus = <i>North Star</i>) R : <i>the North Star</i> , also known as <SN> R : <SN> , also known as <i>the North Star</i>
<i>personne</i>	Q : <i>Who won the Oscar for best actor in 1970 ?</i> R : <entité nommée PERSON>
<i>temps</i>	Q : <i>When was water found on Mars ?</i> R : <entité nommée TIME> R : <hyponyme de <i>time period</i> >
<i>lieu</i>	Q : <i>Where is the highest point on earth ?</i> R : <entité nommée LOCATION>
<i>manière</i>	Q : <i>How did Mahatma Gandhi die ?</i> R : Non implémenté pour TREC
<i>raison</i>	Q : <i>Why does the moon turn orange ?</i> R : Non implémenté pour TREC
<i>objet</i>	Fonction par défaut R : <SN>

Tableau 1. Fonctions d'extraction, exemples de questions TREC et échantillon de patrons de réponse. Les hypéronymes et hyponymes sont obtenus avec WordNet, les entités nommées sont obtenues avec GATE et les étiquettes SN sont obtenues avec un segmenteur de syntagmes nominaux. Lorsque nous indiquons qu'un patron de réponse contient le focus de la question, nous sous-entendons aussi les variantes du focus ou un syntagme nominal plus grand ayant pour tête le focus

sélection du focus sont motivés par des considérations purement techniques et propres au fonctionnement interne de Quantum. Une classification plus générale basée sur des critères psycho-linguistiques rigoureux a été faite par Graesser [GRA 92].

L'analyse de la question s'effectue à l'aide d'un segmenteur de mots, d'un étiqueteur grammatical statistique ainsi que d'un segmenteur de syntagmes nominaux basé sur les étiquettes grammaticales. Nous avons élaboré une soixantaine de règles d'analyse utilisant à la fois des mots, des étiquettes grammaticales et des étiquettes

de syntagmes nominaux [PLA 02]. Par exemple, la règle d'analyse qui s'applique à la question de la figure 1 est la suivante :

what BE <syntagme nominal SN₁> → type = *définition*, focus = SN₁

2.3.2. Filtrage des passages les plus pertinents

Les mécanismes d'extraction de réponse sont trop complexes pour être appliqués à la totalité de la collection de textes à chaque question. Afin de réduire l'espace de recherche de 3 milliards de caractères à quelques milliers, nous utilisons un moteur de recherche conventionnel qui extrait et ordonne les passages les plus pertinents. Nous utilisons à cet effet le moteur Okapi [ROB 98] car il peut raffiner sa recherche jusqu'à suggérer les paragraphes les plus pertinents, à la différence des autres moteurs de recherche qui s'en tiennent à une liste de documents complets. Il nous suffit ensuite de conserver les 20 meilleurs paragraphes et le score de pertinence qu'Okapi leur attribue. La requête est formée de tous les mots de la question ; Okapi se charge de les tronquer et d'éliminer les mots vides (*stopwords*) avant d'effectuer la recherche. La requête de l'exemple de la figure 1 est composée du seul mot de la question qui ne soit pas vide de sens : *pilates*.

2.3.3. Extraction de la réponse

La fonction d'extraction choisie lors de l'analyse de la question, paramétrée ou non avec le focus, est appliquée aux paragraphes les plus pertinents [PLA 02]. Trois techniques ou outils peuvent être utilisés par l'une ou l'autre des 12 fonctions d'extraction : des expressions régulières, le réseau sémantique WordNet et l'extracteur d'entités nommées Annie de la suite de développement GATE [CUN 02]. Par exemple, il est souhaitable que la réponse à une demande de définition contienne un hypéronyme du focus ; ainsi, pour répondre à la question *What is ouzo ?*, WordNet nous permet de vérifier que la réponse *liquor* est bien un hypéronyme du focus *ouzo*. Cependant, dans l'exemple à la figure 1, la réponse doit être trouvée à l'aide d'expressions régulières car WordNet n'a pas d'entrée *pilates*, le focus de la question.

Chaque syntagme nominal dans les paragraphes retenus se voit attribuer un score d'extraction selon qu'il satisfait aux critères de la fonction d'extraction. Ce score d'extraction est combiné au score de pertinence du paragraphe afin de tenir compte de la densité des mots de la question se trouvant autour du syntagme nominal examiné. Le meilleur syntagme nominal est sélectionné pour constituer la réponse. Nous avons choisi de considérer les syntagmes nominaux comme unités de base de réponse car il s'avère que seulement 2 % des questions des campagnes TREC ne peuvent être répondues à l'aide d'un syntagme nominal.

2.4. Évaluation de Quantum

La méthodologie d'évaluation utilisée lors des campagnes TREC est devenue un standard d'évaluation *de facto* pour les systèmes de question-réponse. Nous allons

utiliser 454 des 500 questions de TREC-11 (les autres n'ont pas de réponse connue dans la base de textes) pour évaluer la plus récente version de Quantum. Pour chaque question, nous disposons de la liste de toutes les réponses correctes rencontrées par les évaluateurs humains lors de l'évaluation des systèmes participants. Nous les comparons ensuite de façon automatique avec les réponses données par Quantum afin d'obtenir le pourcentage de questions correctement répondues.

Le module d'analyse de la question de Quantum a été en mesure de choisir la bonne fonction d'extraction et d'identifier le focus correctement pour 86 % des questions de l'ensemble de test. Environ la moitié des erreurs sont dues à des formes de questions inattendues pour lesquelles nous n'avions pas prévu de règle d'analyse. Les autres erreurs découlent d'un mauvais étiquetage grammatical : l'étiqueteur que nous utilisons est probabiliste et son corpus d'entraînement contenait très peu de questions (la syntaxe des phrases interrogatives est différente des phrases déclaratives composant la majorité du corpus).

Une question incorrectement analysée par le module d'analyse ne compromet pas nécessairement les chances de trouver une bonne réponse ; ce n'est cependant pas le cas du module de recherche de passages pertinents. En effet, à cause du temps considérable requis pour l'extraction des réponses, nous ne conservons que les 20 passages les plus pertinents et si aucun d'entre eux ne contient la réponse, le module d'extraction échouera forcément. Pour 79 % des questions de notre ensemble de test, un des 20 passages suggérés par Okapi contenait la bonne réponse.

Pour obtenir une évaluation du module d'extraction de la réponse qui soit indépendante des erreurs causées par les autres modules, nous réduisons l'ensemble de test aux questions qui ont été correctement analysées par le module d'analyse et pour lesquelles nous avons vérifié au préalable qu'une réponse était présente dans les passages proposés par le module de recherche de passages. Nous avons mesuré que le module d'extraction était en mesure de répondre correctement à 28 % des 313 questions de ce nouvel ensemble de test. Le système complet est parvenu à répondre correctement à 20 % des 454 questions de l'ensemble de test original.

3. Composition de réponses

À ce jour, la majorité des travaux en question-réponse ont porté sur *l'extraction de réponse* ; par contre, le problème de la *composition* ou *génération de réponse* n'a pas reçu beaucoup d'attention. La composition de réponse est une étape essentielle dans le cadre de la réponse automatique si l'on désire maintenir un niveau de communication naturel et acceptable entre l'homme et la machine. Au fur et à mesure que les systèmes de question-réponse s'attaqueront à des problèmes de plus en plus difficiles, notamment dans un contexte de dialogue homme-machine, une réponse courte deviendra vite insuffisante. De plus, si la communication se fait par lot (*batch mode*), comme c'est le cas de la communication par courriel, envoyer à l'utilisateur uniquement la réponse exacte, sans aucun contexte, est inadéquat. Une phrase naturelle et complète, voire un

message personnalisé [KOS 01], ainsi que des renseignements supplémentaires sont nécessaires. Cependant, à notre connaissance, peu de travaux ont été effectués dans le cadre de la question-réponse pour formuler des réponses bien formées.

Pour générer des réponses complètes, nous avons doté Quantum d'un module de génération de phrases-réponses. Le but est de générer une phrase-réponse bien formée à partir de la question et de sa réponse exacte. Par exemple, à partir de la question :

Who was the president and chief operating officer of BCE in 1999 ?

et de la réponse exacte :

Jean Monty

nous générons des phrases-réponses telles que :

- *Jean Monty was the president and chief operating officer of BCE in 1999.*
- *The president and chief operating officer of BCE in 1999 was Jean Monty.*
- *In 1999, Jean Monty was the president and chief operating officer of BCE.*

3.1. Compositions de phrases-réponses par pattern matching

Notre première tentative de génération de phrases-réponses se fonde sur une approche pauvre en connaissances linguistiques [KOS 03, PLA 03a]. L'avantage est que l'approche est rapide et facilement adaptable à d'autres langues pour lesquelles des outils d'analyse linguistique ne sont pas disponibles. Cependant, comme nous le verrons plus tard, les résultats de cette approche laissent à désirer, surtout dans le cadre des relations avec la clientèle : une approche plus riche en connaissances linguistiques est nécessaire.

3.1.1. Approche

Pour composer une phrase-réponse à partir de la question et de sa réponse exacte, la question est transformée en une forme déclarative à l'aide de règles de *pattern matching*. Par exemple, la question *Where is the Statue of Liberty ?* est reformulée en «*the Statue of Liberty is <RÉPONSE>*». Pour développer les règles, nous nous sommes inspirés de 198 questions de TREC-8 et de 682 questions de TREC-9 pour bâtir les règles de transformation et nous avons utilisé 447 questions de TREC-10 et 454 de TREC-11 pour les évaluer.

Avant de procéder à la transformation, la forme grammaticale de la question est normalisée afin de limiter le nombre de cas à traiter. Par exemple, une question commençant par *What's...* est normalisée en *What is...*, *What is the name of...* est changée en *Name...*, etc.

Le tableau 2 illustre un exemple de règle de transformation. La règle est formée de 2 catégories de patrons : des patrons qui déterminent à quelle forme de question

<p>Question : When did SYNTAGME VERBE-SIMPLE ? <i>When did the Jurassic Period end ?</i></p> <p>Phrase-réponse : SYNTAGME VERBE-PASSÉ <RÉPONSE> <i>the Jurassic Period ended</i> <RÉPONSE></p>
--

Tableau 2. Exemple de patron de transformation pour la composition par pattern matching

s'applique la règle, et des patrons qui définissent la forme de la phrase-réponse. Les patrons sont des combinaisons des éléments suivants :

- des mots spécifiques (ex. when) ;
- des chaînes de caractères quelconques (représentées dans le tableau par SYNTAGME) ;
- des verbes identifiés à l'aide d'un dictionnaire créé à partir de WordNet (ex. VERBE-SIMPLE).

La transposition d'un verbe de sa forme simple lorsqu'il est utilisé dans une question avec l'auxiliaire du passé *did* (ex. *did . . . end*) à sa forme passée dans la réponse, sans auxiliaire de temps (ex. *ended*), est un phénomène particulier de l'anglais qui ne peut être traité par la seule application de patrons. Afin d'effectuer cette correspondance rapidement, nous avons au préalable extrait tous les verbes répertoriés par le réseau sémantique WordNet et nous avons construit une table de hachage permettant de faire correspondre la forme simple d'un verbe à sa forme passée.

3.1.2. Évaluation

Au total, 76 patrons de question ont été élaborés. Ils sont essayés à tour de rôle sur une même question et tous ceux qui s'appliquent déclenchent la création des phrases-réponses correspondantes. Les 76 patrons de question couvrent 93 % des 198 questions de TREC-8 et 89 % des 682 questions de TREC-9. Par couverture, nous entendons qu'au moins un patron de question est applicable.

Pour évaluer la qualité linguistique des phrases-réponses, nous avons généré les phrases-réponses à partir des questions de TREC-10 et TREC-11. Au total, 1510 phrases-réponses ont été générées pour 1000 questions. Nous avons ensuite demandé à 3 juges humains d'évaluer ces phrases en fonction de leur niveau grammatical.

Le tableau 3 illustre les résultats de l'évaluation. Pour 7 % des questions, le système n'a pas pu produire de phrase-réponse. Parmi les phrases-réponses générées, 60 % ont été considérées grammaticalement correctes (ex. «*the first American to walk in space was* <RÉPONSE>»), alors que 40 % des phrases-réponses ont été considérées incorrectes (ex. «*the first man to fly across the Pacific Ocean,* <RÉPONSE>» ou «*it from Denver to Aspen is* <RÉPONSE> *away*»).

Couverture (rappel)	Qualité grammaticale (précision)
93 %	60 %

Tableau 3. *Évaluation des phrases-réponses produites par pattern matching*

Bien que le nombre de phrases-réponses grammaticalement correctes était plus élevé que ce à quoi nous nous attendions, un taux de formulations grammaticales correctes de seulement 60 % est inacceptable pour assurer une communication homme-machine naturelle, surtout dans le cadre de la communication avec les clients d'une entreprise. Des phrases-réponses de meilleure qualité sont nécessaires et des traits linguistiques plus avancés doivent être pris en considération.

3.2. Composition de phrases-réponses par schémas syntaxiques

Pour produire des phrases-réponses de meilleure qualité linguistique, nous avons alors tenté de développer des schémas de transformation qui prendraient en considération des traits syntaxiques de la question. Dans le cas où plus d'une phrase-réponse peut être générée, seule la plus naturelle est produite [ANA 03].

3.2.1. Approche

Pour produire des phrases-réponses grammaticalement correctes et naturelles, nous avons tout d'abord constitué un corpus de phrases *naturelles*. Nous avons sélectionné 150 questions des campagnes TREC-8, TREC-9 et TREC-10. Ensuite, nous avons donné les réponses exactes à 40 personnes et leur avons demandé de formuler des phrases-réponses naturelles pour chaque question. Une fois le corpus construit, nous avons étiqueté les questions et les phrases-réponses pour identifier :

- les étiquettes grammaticales (nom, adverbe, etc.) avec l'étiqueteur de Brill [BRI 95] ;
- deux syntagmes syntaxiques de base (syntagme nominal et syntagme prépositionnel) avec l'analyseur syntaxique de Bergler [BER 03].

Le résultat de l'analyse est un ensemble de patrons de réponse correspondant à chaque patron de question. Pour ce projet, les patrons ne sont basés que sur la structure grammaticale et ne tiennent pas compte des choix lexicaux des répondants. Le tableau 4 illustre des exemples de patrons de reformulation. Tout comme pour l'approche précédente, 2 types de patrons sont spécifiés : la forme de la question et la forme de la réponse. Par contre, ici, les patrons prennent en considération plus de connaissances linguistiques : des étiquettes grammaticales (ex. VBD/verbe au passé) et des groupes syntaxiques (ex. SN/syntagme nominal, SP/syntagme prépositionnel) ; de plus, à chaque patron de réponse est associé un poids indiquant à quelle fréquence ce patron est apparu dans le corpus d'entraînement. Par exemple, pour 61 % des répon-

Question	
Who VBD SN SP <i>Who released the Internet worm in the late 1980s ?</i>	
Phrases-réponses	
61 %	RÉPONSE VBD SN SP <i>RÉPONSE released the Internet worm in the late 1980s.</i>
14 %	SP RÉPONSE VBD SN <i>In the late 1980s RÉPONSE released the Internet worm.</i>
10 %	SN +to beAux VBD SP +by RÉPONSE <i>The Internet worm was released in the late 1980s by RÉPONSE.</i>
Question	
Who [VBD Z] SN SP VBD SP SP SP <i>Who was the leader of the Branch Davidian Cult confronted by the FBI in Waco, Texas in 1993 ?</i>	
Phrases-réponses	
66 %	RÉPONSE VBD SN1 SP1 VBD SP2 SP3 SP4 <i>RÉPONSE was the leader of the Branch Davidian Cult confronted by the FBI in Waco, Texas in 1993.</i>
14 %	SN SP1 VBD SP2 SP3 SP4 VBD RÉPONSE <i>The leader of the Branch Davidian Cult confronted by the FBI in Waco, Texas in 1993 was RÉPONSE.</i>
10 %	RÉPONSE VBD SN1 SP1 SP2 <i>RÉPONSE was the leader of the Branch Davidian Cult.</i>

Tableau 4. Exemples de patrons de transformation par schémas syntaxiques.

dants, le patron RÉPONSE VBD SN SP est le plus *naturel* comme réponse à une question de la forme Who VBD SN SP, alors que le patron SP RÉPONSE VBD SN semble le plus naturel pour 14 % des répondants et le patron SN +to beAux VBD SP +by RÉPONSE pour 10 % des répondants.

3.2.2. Évaluation

La même approche que pour la composition par *pattern matching* a été utilisée pour évaluer cette nouvelle technique. Cependant, le corpus de test a été réduit à 120 questions choisies aléatoirement. Les patrons de reformulation ont permis de générer des phrases-réponses pour seulement 47 % des questions du corpus de test. Nous croyons que ce bas niveau de couverture est dû au fait que notre corpus d'entraînement était plutôt petit (150 questions) et que les patrons syntaxiques sont bien plus restrictifs. Par contre, comme le montre le tableau 5, 88 % des phrases-réponses générées étaient considérées grammaticalement correctes, contre 60 % pour la technique par *pattern matching*.

Couverture (rappel)	Qualité grammaticale (précision)
47 %	88 %

Tableau 5. *Évaluation des phrases-réponses produites par schémas syntaxiques*

Les résultats de cette nouvelle approche sont intéressants car, même si le système se tait plus souvent (couverture plus basse), quand il génère une phrase-réponse, elle est plus souvent correcte. Dans le cadre de la réponse automatique pour le service à la clientèle d'une entreprise, une précision élevée est nécessaire, quitte à ce que le rappel soit plus bas. En fait, une réponse incorrectement formulée envoyée à un client pourrait avoir des répercussions négatives sur la relation client-entreprise.

4. Question-réponse translinguistique

Les entreprises actives dans tout le Canada, comme BCE, se doivent généralement d'assurer un service à la clientèle en anglais et en français. De plus, même si la majorité des communications électroniques faites *via* le web sont en anglais, une entreprise qui tient compte de la réalité multiethnique de son marché présente un avantage sur ses concurrents. Cependant, les systèmes de question-réponse ne sont pas indépendants de la langue pour laquelle ils ont été conçus, notamment en ce qui a trait à l'analyse des questions et aux ressources linguistiques utilisées pour extraire la réponse. Un système différent doit donc être conçu pour chaque langue. De plus, la base documentaire doit être traduite dans chacune de ces langues. Le coût d'un service à la clientèle multilingue devient vite prohibitif.

Un système de question-réponse translinguistique constitue une solution intermédiaire à ce problème. Au lieu de maintenir des bases documentaires équivalentes mais dans des langues différentes, le système de question-réponse original et la base documentaire initiale sont conservés, et ce sont les questions et les réponses qui sont traduites de et vers la langue du client. Par exemple, chez BCE, la plupart des documents sont rédigés en anglais et nous disposons déjà d'un système de question-réponse pour l'anglais ; au lieu de traduire les documents en français et de concevoir un nouveau système de question-réponse pour le français, nous avons modifié Quantum afin que le client puisse poser sa question en français même si l'extraction de la réponse se fait dans la base de textes en anglais. La réponse peut ensuite être traduite en français.

4.1. Architecture de Quantum translinguistique

Pour Quantum comme pour bien d'autres systèmes de question-réponse, l'étape d'extraction de la réponse est l'étape la plus complexe. Elle est donc déterminante dans le choix de l'approche à favoriser lorsqu'il s'agit d'adapter un système mono-

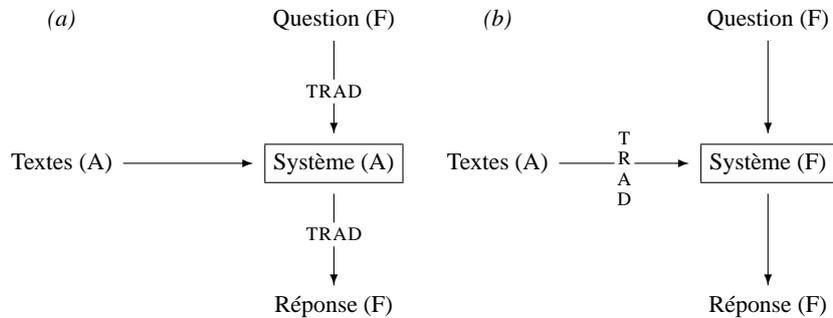


Figure 2. Deux approches pour l'adaptation d'un système monolingue anglais (A) pour répondre à des questions en français (F). En (a), le coeur du système demeure intact et des ressources linguistiques en anglais peuvent être utilisées. En (b), le coeur du système est transposé dans l'autre langue et les textes sont traduits

lingue à des questions ou à une collection de textes d'une autre langue. Deux facteurs sont à considérer plus particulièrement : la disponibilité de ressources linguistiques nécessaires à l'extraction de la réponse et la facilité de réalisation du nouveau système.

Dans le cas de Quantum [PLA 03b], ces deux facteurs ont joué en faveur de la conservation du module d'extraction en anglais (figure 2a) et, par conséquent, en faveur de la traduction de la question et de la réponse plutôt que des textes. D'une part, les ressources linguistiques sont généralement plus nombreuses et plus poussées en anglais qu'en français, si ce n'est qu'elles sont disponibles gratuitement (c'est le cas de WordNet et de l'extracteur d'entités nommées Annie auxquels Quantum fait appel). D'autre part, en laissant le module d'extraction dans sa langue d'origine, la réalisation du système bilingue s'en trouve facilitée. En effet, si nous avions plutôt choisi de transposer le module d'extraction dans la même langue que celle de la question (figure 2b), il aurait fallu trouver de nouvelles ressources linguistiques, adapter le système à leur interface et traduire des textes entiers de l'anglais au français, ce qui pour le moment demeure une entreprise longue et ardue. Par contre, ne traduire que la question et la réponse est plus facile ; nous verrons plus loin qu'il n'est pas nécessaire de parvenir à une traduction syntaxiquement correcte de la question et que la traduction de la réponse peut tirer parti du contexte particulier de la question-réponse.

Dans les scénarios présentés, le coeur du système est considéré comme une boîte noire hermétique au processus de traduction ; ceci suppose une traduction parfaite des questions ou des textes. Les méthodes de traduction actuelles ne le permettent évidemment pas. Il peut s'avérer avantageux d'*ouvrir* le système afin de tirer parti d'un maximum d'informations pouvant faciliter le processus de traduction. Cela permet

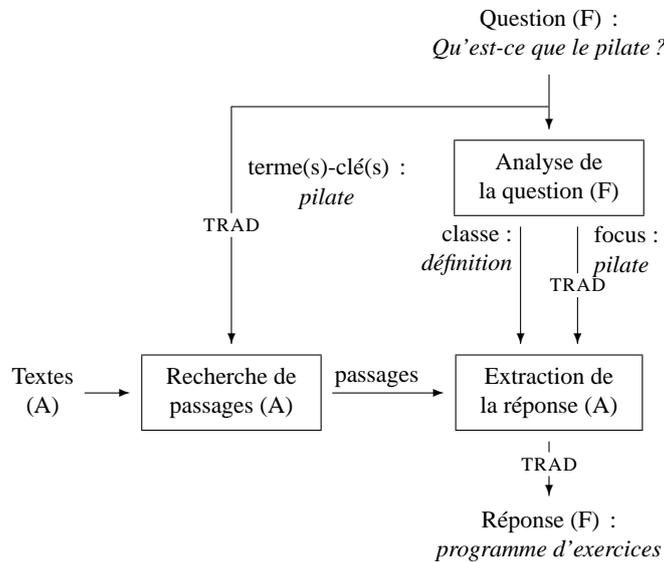


Figure 3. Architecture de la version translinguistique français/anglais de Quantum. Les textes sont en anglais (A), mais la question et la réponse sont en français (F). Le module d'analyse des questions a été francisé et les autres modules opèrent en anglais. La traduction se fait en 3 points : termes clés, focus et réponse

entre autres de contourner le problème complexe de traduction de la question en une question anglaise syntaxiquement bien formée. Qui plus est, des techniques de traduction différentes peuvent être employées selon la tâche.

La figure 3 présente l'architecture de Quantum translinguistique avec le même exemple utilisé pour présenter la version monolingue anglaise de Quantum (figure 1). Le cœur de Quantum est demeuré intact, hormis le module d'analyse de la question qui a été francisé car les méthodes de traduction probabilistes dont nous disposons sont peu performantes en regard de l'ordonnement syntaxique des constituants de la phrase et les règles du module d'analyse ne pourraient s'appliquer adéquatement. La fonction d'extraction choisie par le module d'analyse peut être transmise directement au module d'extraction ; seul le focus, lorsque la fonction d'extraction en requiert un, nécessite d'être traduit vers l'anglais pour les besoins du module d'extraction (entre autres parce que le focus doit être situé dans l'ontologie de WordNet).

Pour traduire le focus, nous utilisons un moteur de traduction probabiliste IBM2 [BRO 93] entraîné sur un ensemble de textes parallèles composés de débats de la Chambre des communes du Canada², de bulletins d'information de l'Union eu-

2. http://www.parl.gc.ca/common/Chamber_House_Debates.asp?Language=F&Parl=37&Ses=2

ropéenne en ligne³ et d'un échantillon de questions de TREC. Le modèle IBM2 tient compte de la position des mots dans la phrase source, ce qui permet de déterminer la traduction la plus probable compte tenu du focus et, dans une moindre mesure, des autres mots de la phrase source. Quant au module de recherche de passages, il requiert que les termes clés de la question soient traduits vers l'anglais. Nous utilisons alors un moteur de traduction IBM1 pour obtenir les termes ayant la plus forte probabilité de faire partie d'une traduction de la question. À la différence du modèle IBM2 utilisé pour traduire le focus, le modèle IBM1 ne tient pas compte de la position des mots de la question : chacun des mots de la question contribue de façon égale à produire les termes en français les plus probables. Les termes générés et qui ne sont pas vides de sens constituent les termes clés de la requête. Nos expériences ont montré que la recherche translinguistique est plus fructueuse lorsque la requête en français contient autant de termes clés que la question en anglais en contient, c'est-à-dire en moyenne 5.

Le module d'extraction de la réponse, quant à lui, ne nécessite aucune modification, ce qui constitue le grand avantage de l'approche translinguistique.

La figure 3 met en évidence le fait qu'il n'est pas nécessaire de traduire la question dans son ensemble et de produire une question syntaxiquement bien formée, ceci parce que nous avons réécrit les règles d'analyse de la question en français et parce que nous avons intégré la traduction en différents points du système. Il suffit en effet de traduire le focus de la question et les termes clés. La traduction de la réponse de l'anglais vers le français est le troisième point de traduction, mais nous ne nous sommes pas encore attaqués à ce problème. Le contexte particulier de la question-réponse jouerait ici en notre faveur : en effet, beaucoup de réponses sont des entités nommées qui ne requièrent pas de traduction. Par exemple, sur un échantillon de 200 questions prises au hasard, 25 % ont pour réponse un nom de personne ou de lieu identique dans les deux langues, un nombre, une date, une raison sociale ou un titre d'ouvrage. Pour les autres types de réponse, il serait intéressant d'utiliser la question pour désambiguïser les termes avant de les traduire.

4.2. *Évaluation*

Nous avons fait traduire en français les questions des campagnes d'évaluation TREC passées et nous en avons sélectionné 789 au hasard pour évaluer Quantum translinguistique. Étant donné que les réponses extraites par le système sont pour l'instant en anglais, il est possible d'utiliser la même procédure d'évaluation automatique que pour les systèmes monolingues ayant participé aux campagnes TREC. Le système translinguistique, testé sur la version française des questions de notre ensemble de test, obtient 44 % moins de bonnes réponses que le système monolingue testé sur les questions originales en anglais [PLA 03b]. Nous avons comparé avec une alternative plus simple consistant à traduire les questions du français vers l'anglais à l'aide

3. http://europa.eu.int/index_fr.htm

d'un moteur de traduction tel que Babelfish⁴ puis à utiliser le système monolingue, mais cette approche a entraîné une perte de performance de 53 % ; c'est dire que cela vaut donc la peine de fonder traduction automatique et question-réponse ensemble plutôt que d'en faire des étapes distinctes. Quantum s'est classé en deuxième position à la piste *Cross-lingual QA* de la campagne d'évaluation CLEF 2003, lors de laquelle 200 questions en français ont été distribuées aux participants ainsi qu'un ensemble de communiqués de presse en anglais pour y puiser les réponses [MAG 03, PLA 03c].

5. Travaux futurs en question-réponse pour la réponse automatique

La question-réponse utilisée dans le cadre de la gestion des relations avec la clientèle doit prendre en considération des défis supplémentaires qui n'ont pas été traités ci-dessus, en particulier l'extraction de la question et les questions contextuelles. Dans les prochains paragraphes, nous décrivons ces défis particuliers ; cependant, comme ils sont actuellement en cours de recherche, seules des pistes de solution sont présentées.

5.1. Identification de la question

Les travaux en question-réponse présentés précédemment se basent sur l'hypothèse que la question est donnée en entrée au système. Cependant, dans le cadre de la réponse automatique pour le service à la clientèle, la question peut ne pas être directement disponible. Dans la version en ligne de Quantum, les clients écrivent leur question dans un formulaire web. Par contre, si l'interface en ligne laisse plus de liberté d'expression aux clients ou si les clients utilisent le courrier électronique pour envoyer leur question, celle-ci doit d'abord être identifiée. Une étape de prétraitement est alors nécessaire pour analyser ce texte plus long et identifier la (ou les) question(s). Cette étape ne consiste pas tout simplement à identifier, par exemple, les phrases se terminant par un point d'interrogation, mais plutôt à saisir le sens du message et l'intention du client afin de comprendre le but de sa requête.

Aussi, les travaux antérieurs en question-réponse se sont essentiellement concentrés sur des questions de types bien précis tels que *qui*, *quand* et *où*, pour lesquels une réponse formée d'un syntagme nominal est suffisante. Dans le cadre du service à la clientèle d'une entreprise, les questions peuvent être de type ouvert (*pourquoi*, *comment*, *etc.*) et dans ce cas, la réponse doit être formée d'une explication ou d'une justification de plusieurs phrases.

5.2. Questions contextuelles

Les questions contextuelles sont une problématique intéressante posée par la gestion des relations avec la clientèle. En effet, un client peut se satisfaire d'un échange

4. <http://world.altavista.com>

unique de type question-réponse, mais le plus souvent, une suite d'échanges sera nécessaire pour satisfaire les besoins informationnels du client. Au cours de ces échanges multiples, une question peut faire référence (explicitement ou non) au contexte de l'échange précédent. Alors que les systèmes de question-réponse actuels tentent de répondre à des questions indépendantes les unes des autres, la question-réponse contextuelle prend en considération les questions et les réponses précédentes pour répondre à la question courante. Cette approche construit ainsi un profil du client pour mieux comprendre la question courante. Par exemple, sans contexte, il est impossible de répondre aux questions *Quand a-t-il quitté son poste ?* ou *Pourquoi ?* Dans le premier cas, le pronom *il* ne peut pas être déréférencé car son antécédent se trouve dans la conversation antérieure ; dans le second cas, tout le corps de la question est éliminé et son antécédent se retrouve lui aussi dans la conversation antérieure. Si l'on sait que ce même client avait précédemment demandé *Est-ce que M. Monty est encore président du conseil de BCE ?*, il est alors possible de répondre à ces questions.

La campagne TREC-10 proposait en 2002 une piste de questions contextuelles [VOO 02]. Les systèmes devaient répondre à une suite de questions comme :

- 1) *Which museum in Florence was damaged by a major bomb explosion in 1993 ?*
- 2) *On what day did this happen ?*
- 3) *Which galleries were involved ?*
- 4) *How many people were killed ?*
- 5) *Where were those people located ?*
- 6) *How much explosive was used ?*

Étonnamment, peu d'équipes ont participé à cette piste et les résultats obtenus semblent indiquer que la capacité de répondre à une question est indépendante du fait que l'on ait répondu correctement ou non aux questions précédentes. Chercher la réponse à une question dans les passages extraits pour les questions précédentes suffit, même si le système n'a pas trouvé la bonne réponse aux questions précédentes.

Dans le cadre du projet avec BCE, nous avons composé une collection de scénarios à partir de l'information retrouvée sur le site web de BCE. Un scénario est une série de paires question/réponse où une question peut faire référence à une question ou une réponse précédente. Actuellement, le corpus contient plus de 60 scénarios, pour un total de 360 paires question/réponse. Nous sommes maintenant en train d'évaluer Quantum sur ce corpus de scénarios et d'étudier différentes méthodes (ex. [HAR 01]) pour tirer parti du contexte précédent et mieux répondre aux questions.

6. Travaux connexes

Les travaux connexes en question-réponse sont très nombreux. Depuis 1999, avec la campagne TREC-8, les travaux en question-réponse se sont multipliés à un rythme effréné. Certains travaux ont donné lieu à des applications commerciales (ex. [HAR 01]), certains utilisent des méthodes sémantiques (ex. [HOV 00]) alors que

d'autres se penchent essentiellement sur des méthodes statistiques (ex. [CLA 01]). Une revue récente des techniques de question-réponse se trouve dans [HAR 03].

Cependant, à notre connaissance, peu de chercheurs se sont penchés sur l'utilisation de la question-réponse pour la gestion des relations avec la clientèle. Ce genre d'application a essentiellement été étudié par des industriels (ex. [KAN 99, RIG 02, XTR 02]). Les quelques chercheurs ayant exploré ce domaine (ex. [BUS 00, COH 96]) ont privilégié des méthodes de classification de textes pour catégoriser les messages selon des classes prédéfinies (requêtes d'achat en ligne, remerciements, plaintes, etc.). Selon la catégorie du message, une réponse générique est envoyée ou le message est tout simplement dirigé vers le département le plus approprié. Certains travaux sont allés plus loin en tentant de répondre par des messages moins génériques. Ces systèmes font généralement une première étape de classification, suivie soit d'extraction d'information pour extraire du message des données spécifiques, soit un calcul de similitude de textes pour appairer la question à une FAQ. Les travaux de [LAL 99, LAL 00, BAR 03], par exemple, font appel à un classifieur de textes probabiliste pour catégoriser le message du client. Selon sa catégorie, des entités sémantiques bien spécifiques sont extraites du message (nom de personne, nombre d'articles, adresse, etc.) de façon à remplir des patrons à trous pour générer une réponse appropriée au courriel.

Dans le cadre de notre collaboration avec BCE, des membres de notre équipe de recherche se sont aussi attaqués à la classification des messages pour faciliter la gestion des relations avec la clientèle. Plus précisément, [DUB 02] a démontré qu'avec des techniques de classification ne faisant pas appel au TALN, il est possible de classifier correctement les trois quarts des messages destinés au service des relations avec les investisseurs de BCE. Pour répondre à des questions fréquentes mais demandant une réponse plus complexe que ce qu'un système de question-réponse pourrait produire, une approche par raisonnement à base de cas (*case-based reasoning*) est en cours d'étude par [LAM 03]. Cette approche semi-automatique permettrait de répondre aux questions similaires à des questions déjà posées en utilisant la réponse donnée antérieurement par le préposé et en la personnalisant.

7. Conclusion

Nos travaux en question-réponse ont abouti à la mise sur pied d'un système fonctionnel capable de traiter des questions directes, en anglais comme en français. Cependant, beaucoup de travail reste à faire avant d'être en mesure d'implanter cet outil dans le processus de la gestion des relations avec la clientèle de BCE.

Tout d'abord, nous devons constituer une base documentaire plus large à l'aide de données dont le format est hétérogène. La base actuellement utilisée par le prototype en ligne de Quantum est constituée de communiqués de presse de BCE et de ses filiales, téléchargés du site web des relations avec les investisseurs de BCE. Cette source a été utilisée car il s'agit d'une collection suffisamment grande de textes de

même nature et facilement convertibles en format texte brut. Avec une telle base documentaire, un client peut obtenir des réponses sur des sujets variés, comme les projets entrepris par l'entreprise, les nominations de cadres supérieurs, certains résultats financiers importants et divers événements reliés à l'entreprise. Cependant, d'autres renseignements fréquemment demandés se trouvent dans des documents qui ne sont pas en format texte (ex. fichiers PDF) ; de plus, les informations sont parfois présentées sous forme de graphiques ou de tableaux et les techniques de TALN ne peuvent alors être utilisées sans une adaptation préalable considérable. Pour ces raisons, nous ne nous sommes attardés, pour l'instant, que sur les documents en format texte brut. Par contre, l'utilisation de textes balisés peut en fait apporter des indices supplémentaires lors de la recherche d'une réponse car la tâche de segmentation du document en unités sémantiques est facilitée par les balises. Finalement, nous croyons qu'une base de données est plus appropriée qu'une base de textes libres pour certaines catégories de questions, comme par exemple des demandes concernant le cours des actions de la compagnie à une date donnée.

Ensuite, il faut pousser plus avant nos travaux en génération de réponse afin d'obtenir une qualité de communication qui soit à la hauteur de l'image que l'entreprise veut donner au client. Pour l'instant, nous ne nous sommes concentrés que sur la composition de phrases à partir de la question exacte. Une réponse plus complexe et personnalisée est nécessaire si la question est extraite d'un message électronique dont le contexte est important. Il peut s'agir d'excuses personnelles, d'une phrase-réponse, d'une référence pour de plus amples informations. . .

L'inclusion d'un formulaire web sur la page du service à la clientèle permettrait aux utilisateurs de poser autant de questions qu'ils le désirent et d'obtenir des réponses sur-le-champ ; en plus de ces avantages, nous croyons que cette pratique diminuera le nombre de requêtes faites par courriel et allégera le fardeau des préposés du service à la clientèle. Une étude préliminaire nous a montré qu'il serait possible d'utiliser notre système de question-réponse pour répondre automatiquement à un certain nombre de ces courriels lorsque ces derniers renferment une question courte et que nous parvenons à la discerner du reste du message. Les défis considérables que nous avons évoqués à la section 5.1 nous empêchent pour l'instant de considérer la réponse automatique aux courriels plus complexes mais nous espérons que les travaux sur la question-réponse contextuelle que nous avons entrepris apporteront une solution à un de ces défis, les échanges multiples.

En terminant, nous désirons rappeler que la réponse automatique n'est qu'une étape parmi tout le processus de gestion des relations avec la clientèle d'une grande entreprise et que des consultations auprès des préposés et gestionnaires du service sont indispensables.

Remerciements

Certains travaux présentés dans cet article sont le fruit des recherches de nombreux chercheurs et étudiants : Glenda Anaya, Stéphane Beauregard, Luc Bélanger,

Julien Dubois, Louis-Julien Guillemette, Luc Lamontagne, Hai Doan Nguyen et Steven Winnikoff. Nous tenons à remercier nos collègues des laboratoires RALI (U. de Montréal) et ClaC (U. Concordia) pour leurs commentaires tout au long du projet. Ce travail a été partiellement financé par les *Laboratoires Universitaires Bell* et le *Conseil de Recherche en Sciences Naturelles et Génie du Canada* (CRSNG).

8. Bibliographie

- [ANA 03] ANAYA G. B., KOSSEIM L., « Generation of natural responses through syntactic patterns », *Actes de la Dixième Conférence Nationale sur le Traitement Automatique des Langues Naturelles (TALN 2003)*, Batz-sur-Mer, France, juin 2003, p. 297-302.
- [BAN 01] BANTER INC., « Natural Language Engines for Advanced Customer Interaction », 2001, <http://www.realmart.com/required062801.html>.
- [BAR 03] BARBUCEANU M., FOX M. S., HONG L., LALLEMENT Y., ZHANG Z., « Building Agents for the Customer Service Front », *Innovative Applications of Artificial Intelligence*, Acapulco, Mexico, 2003, AAAI Press.
- [BER 03] BERGLER S., WITTE R., KHALIFE M., LI Z., RUDZICZ F., « Using Knowledge-poor Coreference Resolution for Text Summarization », *Workshop on Text Summarization, Document Understanding Conference (DUC)*, Edmonton, Canada, May 31–June 1 2003, NIST.
- [BRI 95] BRILL E., « Transformation-Based Error-Driven Learning and Natural Language Processing : A Case Study in Part of Speech Tagging », *Computational Linguistics*, vol. 21, n° 4, 1995, p. 543-565.
- [BRI 01] BRILL E., LIN J., BANKO M., DUMAIS S., NG A., « Data-Intensive Question Answering », *Proceedings of The Tenth Text Retrieval Conference (TREC-10)*, Gaithersburg, Maryland, 2001, p. 393-400.
- [BRI 02] BRILL E., DUMAIS S., BANKO M., « An Analysis of the AskMSR Question-Answering System », *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP-2002)*, Philadelphie, Pennsylvanie, 2002.
- [BRO 93] BROWN P. F., PIETRA S. A. D., PIETRA V. J. D., MERCER R. L., « The Mathematics of Statistical Machine Translation : Parameter Estimation », *Computational Linguistics*, vol. 19, n° 2, 1993, p. 263–311.
- [BUS 00] BUSEMANN S., SCHMEIER S., ARENS R., « Message Classification in the Call Center », *Proceedings of ANLP-2000*, Seattle, 2000, p. 159-165.
- [CLA 01] CLARKE C. L. A., CORMACK G. V., LYNAM T. R., LI C. M., MCLEARN G. L., « Web Reinforced Question Answering (MultiText Experiments for TREC 2001) », *Proceedings of The Tenth Text Retrieval Conference (TREC-10)*, Gaithersburg, Maryland, 2001, p. 673-679.
- [COH 96] COHEN W., « Learning Rules that Classify E-Mail », *Proceeding of the 1996 AAAI Spring Symposium on Machine Learning in Information Access*, 1996, <http://www.research.att.com/wcohen/index.html>.
- [CUN 02] CUNNINGHAM H., MAYNARD D., BONTCHEVA K., TABLAN V., « GATE : A Framework and Graphical Development Environment for Robust NLP Tools and Applications », *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphie, Pennsylvanie, juillet 2002.

- [DUB 02] DUBOIS J., « Classification automatique de courrier électronique », Mémoire de maîtrise, Université de Montréal, juin 2002, <http://www-rali.iro.umontreal.ca/LUB/pub.fr.html>.
- [GRA 92] GRAESSER A., PERSON N., HUBER J., « Mechanisms that Generate Questions », LAUER T., PEACOCK E., GRAESSER A., Eds., *Questions and Information Systems*, p. 167–187, Lawrence Erlbaum Associates, Hillsdale, NJ, 1992.
- [HAR 01] HARABAGIU S. M., MOLDOVAN D. I., PASCA M., SURDEANU M., MIHALCEA R., GIRJU R., RUS V., LACTUSU F., MORARESCU P., BUNESCU R. C., « Answering Complex, List and Context Questions with LCC's Question-Answering Server », *Proceedings of The Tenth Text Retrieval Conference (TREC-10)*, Gaithersburg, Maryland, 2001, http://trec.nist.gov/pubs/trec10/t10_proceedings.html.
- [HAR 03] HARABAGIU S., MAIORANO S., M. P., « Open-domain textual question answering techniques », *Natural Language Engineering*, vol. 9, n° 3, 2003.
- [HOV 00] HOVY E., GERBER L., HERMJAKOB U., JUNK M., LIN C., « Question Answering in Webclopedia », *Proceedings of The Ninth Text Retrieval Conference (TREC-9)*, Gaithersburg, Maryland, 2000.
- [JUP 00] JUPITER COMMUNICATIONS, « E-mail Customer Service : Taking Control of Rising Customer Demand », 2000.
- [KAN 99] « www.kana.com », site visité en 1999.
- [KOS 01] KOSSEIM L., BEAUREGARD S., LAPALME G., « Using Information Extraction and Natural Language Generation to Answer E-mail », *Data and Knowledge Engineering Journal*, vol. 38, 2001, p. 85-100, Elsevier.
- [KOS 03] KOSSEIM L., PLAMONDON L., GUILLEMETTE L., « Answer Formulation for Question-Answering », XIANG Y., CHAIB-DRAA B., Eds., *Proceedings of The Sixteenth Conference of the Canadian Society for Computational Studies of Intelligence (AI'2003)*, Lecture Notes in Artificial Intelligence no. 2671, Halifax, Canada, juin 2003, Springer-Verlag, p. 24-34.
- [LAL 99] LALLEMENT Y., FOX M., « IntelliServe : Automating Customer Service », *Artificial Intelligence for Electronic Commerce AAAI workshop*, AAAI Press, 1999.
- [LAL 00] LALLEMENT Y., FOX M., « Interact : A Staged Approach to Customer Service Automation », HAMILTON H., YANG Q., Eds., *Canadian AI 2000*, LNAI 1822, Berlin, 2000, Springer-Verlag, p. 164-175.
- [LAM 03] LAMONTAGNE L., LAPALME G., « Applying Case-Based Reasoning to Email Response », *International Conference on Enterprise Information Systems (ICEIS-03)*, Angers, France, 2003, p. 115–123.
- [LEH 78] LEHNERT W., *The Process of Question Answering : A computer simulation of cognition*, Lawrence Erlbaum, 1978.
- [MAG 02] MAGNINI B., NEGRI M., PREVETE R., TANEV H., « Is It the Right Answer? Exploiting Web Redundancy for Answer Validation », *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphie, Pennsylvanie, juillet 2002, p. 425-432.
- [MAG 03] MAGNINI B., ROMAGNOLI S., VALIN A., HERRERA J., NAS A. P., PEINADO V., VERDEJO F., DE RIJKE M., « The Multiple Language Question Answering Track at CLEF 2003 », *Working notes CLEF 2003*, 2003, http://clef.iei.pi.cnr.it:2002/2003/WN_web/36.pdf.

- [MIL 95] MILLER G., « WordNet : a Lexical Database for English », *Communications of the ACM*, vol. 38, n° 1, 1995, p. 39-41.
- [PLA 02] PLAMONDON L., « Le système de question-réponse QUANTUM », Mémoire de maîtrise, Université de Montréal, mars 2002, <http://www.iro.umontreal.ca/~plamondl>.
- [PLA 03a] PLAMONDON L., KOSSEIM L., « Le Web et la question-réponse : transformer une question en réponse », *Actes des Journées Francophones de la Toile (JFT-2003)*, Tours, France, juin 2003, p. 225-234.
- [PLA 03b] PLAMONDON L., FOSTER G., « Multilinguisme et question-réponse : adaptation d'un système monolingue », *Actes de la Dixième Conférence Nationale sur le Traitement Automatique des Langues Naturelles (TALN 2003)*, Batz-sur-Mer, France, juin 2003, p. 107-116.
- [PLA 03c] PLAMONDON L., FOSTER G., « Quantum, a French/English Cross-language Question Answering System », *Working notes CLEF 2003*, 2003, http://clef.iei.pi.cnr.it:2002/2003/WN_web/43.pdf.
- [RIG 02] RIGHTNOW TECHNOLOGIES, « Revelation Knowledge Engine », 2002, <http://www.rightnow.com>.
- [ROB 98] ROBERTSON S., WALKER S., « Okapi/Keenbow at TREC-8 », *Proceedings of The Eighth Text Retrieval Conference (TREC-8)*, Gaithersburg, Maryland, 1998, p. 151-162, http://trec.nist.gov/pubs/trec8/t8_proceedings.html.
- [VOO 01] VOORHEES E., « Overview of the TREC 2001 Question Answering Track », *Proceedings of The Tenth Text Retrieval Conference (TREC-X)*, Gaithersburg, Maryland, 2001, p. 157-165.
- [VOO 02] VOORHEES E. M., « Overview of the TREC 2002 Question Answering Track », *Notebook Proceedings of The Eleventh Text Retrieval Conference (TREC-11)*, Gaithersburg, Maryland, 2002, p. 115-123.
- [WOO 77] WOODS W., KAPLAN R., « Lunar rocks in natural English : Explorations in natural language question answering. Linguistic Structures Processing. », *Fundamental Studies in Computer Science*, vol. 5, 1977, p. 521-569.
- [XTR 02] XTRAMIND TECHNOLOGIES GMBH, « XM-MailMinder », 2002, <http://www.xtramind.com>.