

Du Chatbot au TAL, du mot clé au sens principal

Questions :

- 1) Les chatbot sont des programmes généralement conçus pour répondre au mieux à un ou plusieurs besoins de celui qui va les utiliser. Dans le cas de Léa, la chatbot de la sncf, il s'agit de répondre à toutes les questions possibles d'un internaute ayant un rapport avec la sncf, ceci allant de la date de sa création au prix d'un billet pour un trajet donné par exemple. Dans le cas d'Eliza, il s'agit simplement d'encourager l'utilisateur à raconter ce qu'il a sur le cours, sans forcément avoir à converser avec la personne à laquelle il parle. Généralement, les chatbot sont plutôt gentilles : elles sont en effet destinées à entrer en contact et à tenir une conversation avec des humains ; si elles étaient impolies, ou vexantes, voire méchantes, plus personne ne voudrait les utiliser. Mais gare à vous : si on les agresse, elles savent se défendre. Les chatbot sont donc au service de l'utilisateur : en cela, elles sont surtout utiles, qu'elles apportent des informations ou non, et parfois amusante dans certaines de leurs réponses (à leurs dépens, par une réponse décalée et peu appropriée à la question posée, ou par volonté du programmeur).
- 2) Si l'on veut qu'une chatbot puisse être utile, il faut qu'elle soit à même de fournir à un utilisateur les renseignements qu'il est susceptible de demander. Il faut donc qu'elle possède des données suffisantes à ce sujet, ou, à défaut, qu'elle soit capable de rediriger l'utilisateur vers un site, ou une personne à contacter qui détient et pourra fournir cette information. Pour qu'elle soit amusante, il faut encore une fois qu'elle possède les moyens de l'être, et ce par exemple, grâce à des réponses préparées par le programmeur à certaines questions. On peut ici penser à SIRI, sur l'iphone. A la question « qui est le meilleur smartphone ? » SIRI répond « Ah bon, il y a d'autres smartphones ? », mais cela seulement parce que la question a été prévue et a permis une réponse humoristique en prévision. Enfin, si l'on veut qu'une chatbot soit gentilles avec ses usagers, il faut penser à lui intégrer les formules de politesse, ainsi que les formules de présentation qui peuvent favoriser dès le départ une bonne communication entre les usagers et elle. N'importe qui pourrait alors s'en servir (même s'il est évident que cela ne serait pas pour la même utilisation) : les enfants pour s'amuser, les adultes pour avoir des informations sur un domaine particulier ou simplement pour se distraire.
- 3) Il y a en partie adéquation entre ces qualités et celles qui permettraient à un système de passer le test de Turing. Les formules de politesses ou formules de présentations, telles que « ça va ? comment tu t'appelles ? » sont des réactions semblables à celles que pourraient avoir deux humains lorsqu'ils se rencontrent pour la première fois. De plus, des réponses drôles et amusantes, moins stéréotypées que celles qu'on peut attendre d'une machine, nous font penser qu'il s'agit là de quelqu'un capable de faire de l'esprit, donc de raisonner, et donc qui pencherait davantage vers l'humain que vers le robot.
- 4) Les autres critères envisagés pour une chatbot sont qu'elle doit pouvoir répondre à propos, ce qui n'est pas toujours le cas.
- 5) **Eliza (english)** : elle ne répond à aucune question (ni l'heure, la date, ne dit rien quel que soit le sujet et ne donne aucune information sur elle). Elle ne fait que poser des questions. Du coup, le côté informationnel n'est pas rempli. Elle respecte les formules de politesse, mais n'est pas drôle.
Elizia : elle ne donne pas non plus les informations qu'on lui demande, quel que soit le sujet. Elle incite seulement le locuteur à lui parler. Les formules de politesse et la gentillesse sont

respectées, tant qu'on ne la tutoie pas (ou qu'on ne l'insulte pas). Elle n'a pas de réponses amusantes.

Cleverbot : Il renvoie des réponses aux questions posées le plus souvent mais celles-ci ne sont pas toujours correctes, car dépassées parfois (exemple : « Who is the president of france ?**Jacques Chirac**.No, it not anymore. It's Nicolas sarkozy.**Thank you.** »). Il n'est pas capable de garder en mémoire certaines informations données par l'utilisateur, ce qui peut parfois causer certaines confusions au fil de la conversation (comme savoir si son interlocuteur est une fille ou un garçon...)

Léa : Elle ne comprend pas toutes les questions (ou en tout cas elle n'est pas capable d'y répondre) et renvoie alors l'interlocuteur vers un conseiller, ce qui pourrait être ben mais est aussi une solution de facilité parfois peu efficace.

Bearbot : répond à propos sauf quand cela touche aux langues étrangères (au-delà de quelques mots les réponses n'ont plus de sens). Il a du mal à donner des informations générales, et ne les retient pas en mémoire (exemple : Human: « who is the president of france ? » Bearbot: « I'm not that into politics. »).

6) Il faudrait qu' Elizia puisse répondre à au moins quelques questions basiques (même si elle élude toute question la concernant). Il faudrait aussi que le tutoiement ne la dérange pas, si on lui demande la permission par exemple. En ce qui concerne les autres

7) Il n'y a pas de déclencheur pour la répétition même si Eliza arrive à savoir qu'on répète la même phrase (exemple d'échange :

- « - hi, how are you ?
- Why are you interested in wether or not I am?
- Hi, how are you?
- Why did you repeat yourself? »)

Les autres chatbots testées, par contre, ne semblent pas la remarquer. Ils se contentent d'essayer d'apporter une réponse appropriée (sans succès, puisque genre de répétition ne fait pas vraiment sens pour eux). Il en va de même si l'on se met à répéter scrupuleusement tout ce que les chatbots « disent ». Eux ne s'aperçoivent de rien et traitent la réponse individuellement, comme s'il s'agissait à chaque fois d'une question posée par l'utilisateur. Par exemple, avec cleverbot :

- « - **What's your favorite color?**
- What's your favorite color ?
- **My favorite color is blue.**
- My favorite color is blue.
- **I thought your favorite color was purple.**
- I thought your favorite color was purple.
- **Nope. I like green.** »

8) Le rôle du TAL est d'utiliser l'informatique pour traiter, utiliser, des données se rapportant au langage. Les chatpot sont une application du TAL, car il s'agit de traiter les différents aspects du langage humain pour permettre à un programme informatique de tenir une conversation en étant en interaction avec quelqu'un.

Un mot clé, c'est le mot qui contient à lui seul tout ce qui est dit dans la phrase (ou dans le texte). C'est ce qui est le plus important, contrairement aux informations qui sont autour ; si le mot clé est supprimé, il devient difficile, voir impossible de savoir de quoi on parle.

Deux mots clés peuvent être valables pour le même cas, notamment en ce qui concerne les homonymes : par exemple, si je dis « j'ai vu mon avocat » et j'ai acheté des avocats au marché », le mot clé sera dans les deux cas « avocat ». Pour autant, il ne renvoie pas du tout à la même réalité.

- 9) Le rôle de la syntaxe ou de l'analyse syntaxique est d'aller au-delà du seul sens des mots, pour se pencher sur le sens de la phrase, beaucoup plus complexe. Il faut pour cela observer la catégorie des mots, les fonctions syntaxiques qu'ils occupent dans la phrase, ainsi que les relations qu'ils entretiennent entre eux (accords, conjugaison lorsqu'il s'agit de verbe...). Le rôle de la ponctuation doit aussi rentrer en compte.
- 10) Pour détecter le sens clé d'une question, il faut arriver à détacher ce que l'on dit de l'information qu'on a sur ce que l'on dit (ce que l'on dit étant le sens clé de la question). Souvent, le sens de la question est contenu dans la fonction complément (qui est tu ? Comment vas-tu ? Quel âge as-tu ?), notamment pour les questions usant d'un mot interrogatifs. Cela dit, une telle approche peut faire prendre en compte seulement certains aspects de la question et en oublier d'autres. Par exemple dans « le 14 juillet, iras-tu au feu d'artifice ? » le fait que ce soit à un feu d'artifice en particulier peut ne pas avoir été pris en compte dans la question.
- 11) Une bonne discussion peut être une discussion qui a un sens, un fil, c'est-à-dire une logique qui se construit tout au long de la conversation, et pas plusieurs bouts de dialogues mis bout à bout (pleins de sujets sans aucun rapport) comme c'est souvent le cas avec les chatbox. Il faut également que les deux participants soient également impliqués, c'est-à-dire à part égale, et non pas un qui ne fait que poser des questions et l'autre qui ne fait qu'y répondre comme c'est le cas avec Eliza.
- 12) Pour afficher les mots reconnus dans l'historique, il faut commencer la phrase par ::: .Pour obtenir une liste de tous les mots qu'Elizia teste contre les mots clés, il suffit de taper « all : » en début de phrase. Pour afficher les lemmes, il suffit d'utiliser « lemmas : ». Enfin, pour obtenir les mots, il faut taper « words : ».
- 13) On peut se servir de ces fonctionnalités pour mieux communiquer avec Elizia, en utilisant en priorité les mots qu'elle va reconnaître, ou pour revenir à un sujet dont on a déjà parlé. Cela est aussi utile pour savoir comment fonctionne Elizia, dans son traitement des énoncés (lorsqu'on lui demande les lemmes par exemples, ou les mots clés).
- 14) Dans l'interface, en ce qui concerne la négation, on ne retrouve que des réponses concernant le fait d'être négatif, ce qui est certes pertinent puisque simulant un entretien avec un psychologue, mais cela ne prend pas en compte tous les énoncés comme « je n'ai rien mangé », « je n'ai jamais vu l'ami de ma sœur » qui sont des faits. Elizia devrait alors essayer de détecter la négation (quand ces mots sont associés à n' ou ne ou ni...) et proposer une réponse différente, comme « vraiment ? pourquoi ? » ou « vraiment ? quel dommage ! ». En ce qui concerne les mots d'anglais proposés, ils sont trop peu pour qu'Elizia puisse véritablement reconnaître cette langue. Par ailleurs, « hi » est accepté dans les salutations mais n'est pas reconnu comme mot anglais. Il faudrait donc rajouter du vocabulaire à Elizia. Enfin, en ce qui concerne la question du temps, il y a quantité de réponses concernant le temps qui passe mais aucune pour le temps qu'il fait. Par conséquent, lorsqu'on lui pose une question sur la météo, Elizia n'est pas capable de fournir une réponse adaptée. Il faudrait ajouter un cas spécial, pour que quand Elizia rencontre le mot « météo » ou « quel » associé à « temps » elle renvoie une réponse telle que « pour moi, il fait toujours beau ».
- 15) On peut ajouter des termes correspondant à différentes langues, un peu sur le même principe que Léa ou Cleverbot, afin qu'à défaut de « parler » ces langues, la chatbot puisse être à même de les reconnaître et d'informer l'utilisateur de sa méconnaissance. Elizia n'est pas non plus capable de faire des calculs, même simples, même si elle reconnaît les chiffres.

Initiation au TAL LZ021 Devoir 2 ROBIN Amélie

Pour le reste, elizia possède déjà un vocabulaire important. Le problème est plutôt qu'elle fonctionne beaucoup en mot clé et pas assez sur le sens général de la phrase.

16)

	Elizia	Eliza (english)	Léa	Cleverbot	Bearbot
Formules de salutation	TB	TB	TB	TB	TB
Maîtrise langues étrangères (anglais en priorité, allemand ensuite)	Nul	Nul	Moyen (reconnait des langues)	Bien (reconnait et répond en anglais, allemand pour ce qui est basique)	Moyen (reconnait, réponses minimales avant qu'elles ne soit correctes syntaxiquement mais non intelligibles)
Informations sur un sujet quelconque (Qui est le président français ? Quel jour sommes nous ? Quelle heure est il ?)	Nul	Nul	Bien	Moyen	Moyen
Capacité à faire des calculs simples	Nul	Nul	Nul	Moyen	Bien
Capacité à garder des informations en mémoire (autre que le nom)	Nul	nul	Nul	Moyen	Nul
Capacité à apprendre de nouvelles informations	Nul	Nul	Nul	Bien	Nul
Réaction face à la répétition	Bien (détecte)	Bien	Moyen	Moyen	Moyen
Réaction face aux smileys	Nul	Nul	Moyen	Bien	Bien

Une comparaison a été établie face aux questions posées aux différents chatbots au cours de ce travail.

Un graphique a également été réalisé pour rendre plus claire ces différences de capacités et d'efficacité entre les chatbots. On voit qu'Elizia et son homologue anglais sont peu adaptées à une

Initiation au TAL LZ021 Devoir 2 ROBIN Amélie

interaction, tandis que chez Cleverbot et Bearbot, cela est mieux développé. Quant à Léa, elle est ciblée sur un domaine plus spécifique ce qui explique parfois son manque d'efficacité dans une interaction basique avec un individu.

