

Natural Language Processing

Introduction

Qu'est-ce que le traitement du langage naturel (natural language processing¹) ?

Il faut tout d'abord faire attention à ne pas confondre « Natural language understanding » et « Understanding natural language » :

1. *Le Natural Language Understanding (NLU) (@Je ne vois pas comment je pourrais traduire ça en français de sorte à le nuancer avec le point suivant)* est un sous-domaine du NLP, il consiste en la création d'applications mettant en œuvre le langage naturel comme données en entrée dans l'application mais également comme résultat en sortie. On ne calque pas nécessairement le fonctionnement du cerveau mais nous avons plutôt recours à des techniques d'ingénierie.
2. Le Understanding Natural Language qui est plutôt de l'ordre des neurosciences. Pouvons-nous comprendre comment le cerveau humain, en tant que mécanisme biologique, est capable de traiter le langage. Il existe des modèles dans les sciences cognitives.

Dans ce cours, nous nous concentrons principalement sur le NLU et le NLP, pas le second point.

Quelques exemples de NLU

- Il y a un jeu télévisé diffusé depuis 1964 qui s'appelle « Jeopardy! ». Le principe est simple : à partir de réponses communément appelés des indices, trois candidats doivent trouver la question correspondante. Chaque bonne réponse (c'est-à-dire chaque bonne question) rapporte une somme, chaque erreur la fait perdre. Ils peuvent choisir entre six catégories et cinq valeurs d'indices par catégorie.

En février 2010, IBM Watson bat le meilleur joueur humain au monde dans ce jeu !

Voici la vidéo (<https://www.youtube.com/watch?v=P18EdAKuC1U>) :

¹ On va abréger en NLP dans la suite du document

