

**TP : Conception d'une application de TALN**

**(activité d'équipe de 2 à 4 membres)**

## Travail pratique : questionnements logiques en langage naturel

*Ce travail compte pour 5% de la note finale.*

*Échéance : le dimanche 16 Décembre 2018 (23h59)*

### Objectifs d'apprentissage :

À l'issue de ce travail, un étudiant devrait être capable de :

- résoudre un problème concret de traitement automatique du langage naturel;
- identifier les enjeux du traitement du langage naturel sur un problème concret;
- analyser les limites d'une solution à un problème concret de traitement du langage naturel.

### Description du contexte du travail pratique :

Écrire un programme Prolog qui devra comprendre (par une analyse sémantique) des phrases en français de la forme suivante et répondre correctement à la question logique :

- a) ..... est un ....
- b) tout ..... est un ....
- c) est ce que ..... est un ....

Les phrases de la forme (a) seront analysées pour créer des faits qui seront insérés dans la base par des assertions (en utilisant le prédicat `assert/1`). Par exemple, la phrase «rex est un chien » donnera le fait : `chien(rex)`.

Les phrases de la forme (b) seront analysées pour créer des règles, elle aussi insérées par des assertions. Par exemple, la phrase « tout chien est un canidé » donnera la règle : `canide(X) :- chien(X)`.

Les phrases de la forme (c) seront analysées comme des requêtes auxquelles le programme répondra oui ou non selon l'état de la base de faits. Ainsi la phrase « Est-ce-que Rex est un canidé? » reviendra à poser la requête :

`?- canide(rex).`

Le programme répondra oui ou non selon si la réponse à cette requête.

Voici deux exemples de fonctionnement :

?- go. rex est un chien ok tout chien est un canidé ok est ce que rex est un canidé oui yes	?- go. rex est un chien ok tout chien est un canidé ok est ce que bill est un canidé non yes
--	---

Le programme à construire en langage Prolog consiste en un prédicat `go/0` qui fait appel à 3 autres prédicats, chacun analysant une des 3 formes données à la suite l'une de l'autre : `analyseA`, `analyseB`, `analyseC`.

Ces 3 prédicats analysent la phrase saisie et donne la réponse :

- ok pour analyseA et analyseB,
- la réponse à la requête pour analyseC, soit oui ou non.

La saisie d'une phrase se fait à l'aide du prédicat prédéfini readln/1. Par exemple :

? - readln(P).

| : rex est un chien

P=[rex, est, un, chien]

### **Travail à réaliser :**

- 1- Ensemble de phrases utilisées (10 %) :** Constituer 5 (au minimum) ensembles de phrases contenant les 3 formes (a), (b) et (c). Il est possible de faire des variantes du type : « est une ».
- 2- Analyse de la forme (a) (15 %) :** Programmer en Prolog analyseA en utilisant une grammaire attribuée et l'opérateur DCG.
- 3- Analyse de la forme (b) (15 %) :** Programmer en Prolog analyseB en utilisant une grammaire attribuée et l'opérateur DCG.
- 4- Analyse de la forme (c) (20 %) :** Programmer en Prolog analyseC en utilisant une grammaire attribuée et l'opérateur DCG.
- 5- Interface (10 %) :** Programmer le prédicat go/0 en langage Prolog selon les directives données précédemment.
- 6- Résultats (20%) :**
  - Faire des tests pour montrer les résultats d'analyse de phrases (à insérer dans le rapport).
  - Discuter les résultats obtenus : Êtes-vous satisfait de votre travail ? Comment pourriez-vous l'améliorer ? Est-ce que de nouvelles phrases pourraient facilement être interprétées ?

Les 10% restants sont dédiés à l'appréciation globale du travail remis (présentation et expression écrite).

Ne pas oublier de valider votre travail à l'aide de la grille d'autoévaluation fournie à l'annexe 1.

### **Modalités de remise de ce travail :**

En plus du programme Prolog (fichier PL), vous devez remettre un rapport. Ce rapport doit avoir comme d'habitude une page couverture, une introduction et une conclusion. Le rapport contiendra 2 parties : les 5 ensembles de phrases utilisés et la partie Résultats.

L'ensemble du travail est à remettre en format électronique via le Portail des cours. Les travaux remis en retard ne seront pas corrigés.

## Annexe 1. Grille d'autoévaluation

Conception (90%)				
<i>Phrases (10%)</i>	Un ensemble d'au moins 5 ensembles de phrases sont indiquées et utilisables dans le programme.	Un ensemble d'au moins 5 ensembles de phrases sont indiquées mais ne sont pas toutes utilisables dans le programme.	Moins de 5 ensembles de phrases sont indiquées et utilisables dans le programme.	On ne sait pas vraiment quelles sont les phrases utilisables par le programme.
<i>AnalyseA (15%)</i>	Le prédicat permet d'analyser les phrases de la forme (a) et faire l'assertion du fait correspond.	Le prédicat permet d'analyser les phrases de la forme (a) mais l'assertion du fait correspondant ne se fait pas.	Le prédicat ne permet d'analyser les phrases de la forme (a) au complet et ne fait pas l'assertion du fait correspondant.	Aucune analyse des phrases de la forme (a).
<i>AnalyseB (15%)</i>	Le prédicat permet d'analyser les phrases de la forme (b) et faire l'assertion du fait correspond.	Le prédicat permet d'analyser les phrases de la forme (b) mais l'assertion du fait correspondant ne se fait pas.	Le prédicat ne permet d'analyser les phrases de la forme (b) au complet et ne fait pas l'assertion du fait correspondant.	Aucune analyse des phrases de la forme (b).
<i>AnalyseC (20%)</i>	Le prédicat permet d'analyser les phrases de la forme (c) et répond à la question correctement.	Le prédicat permet d'analyser les phrases de la forme (c) mais ne répond pas à la question correctement.	Le prédicat permet d'analyser les phrases de la forme (c) mais ne donne pas de réponse	Le prédicat ne permet d'analyser les phrases de la forme (c) ou est absent.
<i>Interface (10%)</i>	Le prédicat go/0 est implanté selon les directives et fonctionne sans erreur.	Le prédicat est implanté selon les directives mais fonctionne avec erreur.	Le prédicat est implanté mais ne fonctionne pas du tout.	Le prédicat n'est pas implanté

<i>Résultats (20%)</i>	Les tests sont présentés ainsi qu'une discussion de l'expérience et des améliorations possibles.	Les tests sont présentés mais il manque une discussion.	Les tests ne sont pas présentés mais il y a une discussion	Il n'y a pas de tests, ni de discussion.
----------------------------	--	---	--	--

<b>Appréciation globale (10%)</b>				
<i>Expression écrite (5%)</i>	Le rapport ne contient aucune faute (vocabulaire, grammaire, syntaxe, etc.).	Le rapport ne contient pas plus d'une dizaine de fautes (vocabulaire, grammaire, syntaxe, etc.).	Le rapport ne contient pas plus de 5 fautes par page (vocabulaire, grammaire, syntaxe, etc.).	Le rapport contient plus de 5 fautes par page (vocabulaire, grammaire, syntaxe, etc.).
<i>Présentation du rapport (5%)</i>	Tous les éléments demandés sont présents et le rapport est paginé et contient une page de garde.	Tous les éléments demandés sont présents, mais le rapport n'est pas paginé ou ne contient pas de page de garde.	Il manque 1 élément demandé.	Il y a au moins 2 éléments demandés manquants.