

Examen semestriel

Module Optionnel : Traitement
Automatique du Langage Naturel

Corrigé

Exercice 1 :

Question 1 : Quelles difficultés rencontre-t-on au niveau d'analyse phonologique. Expliquez.

Réponse :

Il s'agit de reconnaître dans le flot sonore les unités acoustiques élémentaires (phonèmes). La difficulté est que la forme acoustique d'un phonème varie selon plusieurs facteurs : le sexe, l'âge, la région, la fatigue, la peur, l'intensité de la parole, etc...

(2 points)

Question 2 : Décrire comment un analyseur morpho-lexical traite une phrase qui lui a été soumise.

Réponse :

- 1. Découper la phrase en éléments (mots).**
- 2. Pour chaque élément, rechercher ses différentes interprétations possibles ; le rechercher dans le dictionnaire des noms propres, les noms communs, les abréviations, les formes fléchies, les formes dérivées, les noms composés, ...etc.**
- 3. Afficher chaque élément avec ses interprétations possibles en choisissant un format approprié (matrice attributs-valeurs , etc...).**
- 4. Envoyer le résultat à l'étape d'analyse suivante (analyse syntaxique).**

(2 points)

Question 3 : Soit la phrase « Pierre a acheté des livres ». Quel résultat produirait un analyseur morpho-lexical après le traitement de cette phrase ?.

Réponse :

Pierre : Nom propre / Nom Commun Féminin singulier.

a : Auxiliaire "avoir " , temps présent, indicatif , 3^{ème} personne singulier / préposition / adverbe

acheté : Participe passé verbe "acheter"

des : Article contracté / article partitif

livres : nom commun masculin pluriel / nom commun féminin pluriel

(2 points)

Question 4 : Expliquez brièvement pourquoi les grammaires formelles sont inadaptées pour modéliser le langage naturel.

Réponse :

Les grammaires formelles ne peuvent pas prendre en compte quelques phénomènes complexes comme l'accord (genre/nombre) et la coordination. Exemple : Elle/Il/ Ils est(sont) heureux(euses).

Elles peuvent générer des phrases syntaxiquement correctes mais dénuées de tous sens. Exemple : "la table mange le chat"

(1 point)

Question 5 : Expliquez brièvement le principe d'une analyse en HPSG.

Réponse :

Détecter la "tête" de la phrase, et s'appuyer sur ce résultat pour trouver les éléments connexes.

(1 point)

Question 6 : Décrivez brièvement les principaux avantages de HPSG.

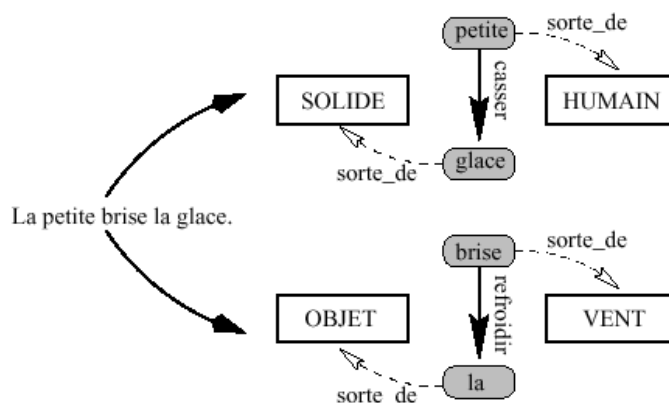
Réponse :

- **Formalisme universel, qui peut s'adapter en théorie à toutes les langues.**
- **Se prête à l'implémentation informatique, dont il reprend quelques concepts (unification, enregistrement, ...etc).**
- **Formalisme accordant beaucoup d'importance à la richesse du lexique**

(1 point)

Question 7 : Montrez comment la phrase "La petite brise la glace" est ambiguë en utilisant les réseaux sémantiques.

Réponse :



(1 point)

Question 8 : Qu'appelle-t-on "le bruit" dans un SRI ?

Réponse :

C'est un ensemble d'informations non pertinentes.

(1 point)

Question 9 : Expliquez pourquoi la génération de paraphrases est intéressante dans un système de recherches d'informations (SRI).

Réponse :

La génération de paraphrases permet de présenter à l'utilisateur un éventail de requêtes correspondant à son besoin d'informations. L'utilisateur pourra ainsi affiner sa recherche (choisit la requête la plus indiquée à son besoin).

(1 point)

Question 10 : Qu'est ce que "la trace" dans un SRI. En quoi est-elle intéressante ?

Réponse :

La trace : description du fonctionnement du moteur de recherche. L'affichage de cette information permet à l'utilisateur de mieux exprimer ses requêtes,

(1 point)

Exercice 2 :

On veut concevoir un programme qui reçoit un nom commun du français et donne son pluriel.
Question 1 : Ecrire l'algorithme correspondant. Décrivez en détail les structures de données utilisées.

Réponse :

Indications (l'étudiant devait écrire l'algorithme en détail):

Déclaration des structures de données :

Dico : Dictionnaire des noms communs,

Tab-al : Table des noms du français se terminant par "al" mais qui admettent quand même un "s" au pluriel (exemple : festival).

- Lire un mot
- Tester si ce mot a une longueur inférieure ou égal à 3.
- Rechercher si le mot se termine par un 'z'
- Rechercher si le mot appartient à la table Tab-al .
- Traiter les différents cas de sortie : 's', 'aux', 'z' .

(5 points)

Question 2 : Si on veut faire le même travail pour les verbes arabes, décrivez les outils nécessaires et les étapes à suivre (on ne vous demande pas d'écrire d'algorithme).

Réponse :

Pour répondre au même problème en langue arabe, il y'a obligation d'utiliser la notion de "schème" d'un nom (wazn). On considère ainsi, une ou plusieurs tables qui fait correspondre à chaque schème, le schème du pluriel correspondant.

Etapes :

- Lire le mot
- Extraire son schème.
- Rechercher dans la table des schèmes le schème du pluriel.

(2 points)

La correction et les résultats seront publiés sur le site : <http://www.loukam.net>