

**Subiecte Master anul I** (Se poate folosi orice tool de pe Net, din cele discutate la laborator etc.)

**Fisierul reprezentand teza se transmite la adresa de email:** [curs.ia.nlp@fmi.unibuc.ro](mailto:curs.ia.nlp@fmi.unibuc.ro)

Fisierul cu frazele (pentru exercitiile 1,2 si 5):

<https://docs.google.com/spreadsheets/d/18VBp7bYCY7CMI8o3lX-1vMDZjD-XkwkfPwMU5FmmwZw> --- **fiecare student preia frazele din dreptul numelui sau.**

### **1. Left-corner parser (2p)**

Sa se aplice algoritmul left-corner parser sirului de intrare preluat din fisierul cu frazele.

Pentru aplicarea algoritmului concepeti o gramatica CFG minimala (fara reguli in plus) si explicati semnificatia notatiilor folosite pentru neterminale (Ex.: VP = verb phrase).

Mod de rezolvare: Pentru gramatica data realizati tabelul "left corner table" si desenati arborele de structura a frazei (parse tree). Numerotati muchiile in ordinea aparitiei lor in arbore pe parcursul analizei sintactice. In caz de backtracking, se vor preciza muchiile care dispar si se va redesena arborele din momentul aplicarii unei alte reguli (atunci cand se reia mersul inainte).

### **2. Dependency parsing (1p)**

Aplicati analiza sintactica de dependenta (dependency parsing), utilizand Stanford parser, sirului de intrare din fisierul cu fraze.

Desenati (sub forma de graf) structura de dependenta rezultata.

### **3. Algoritmul Lesk extins (2p)**

Explicati modul de utilizare a masurii de inrudire calculate de algoritmul Lesk extins pentru realizarea dezambiguizarii propriu-zise a sensului unui cuvant tinta.

### **4. Dezambiguizare cu modelul Bayes (2p)**

a) Explicati in ce consta diferenta esentiala in estimarea parametrilor - in cazul folosirii modelului Bayesian naiv – intre dezambiguizarea supervizata si cea nesupervizata.

b) Prezantati, pe scurt, o modalitate de alegere a caracteristicilor – sintactice SAU semantice – in cazul dezambiguizarii nesupervizate cu modelul Bayesian naiv.

## 5. Ant Colony Algorithm (2p)

- a) Pentru textul din fisierul cu frazele desenati graful folosit de algoritmul Ant Colony (doar nodurile si muchiile grafului). Pe nodurile corespunzatoare sensurilor se va pune id-ul synsetului (de exemplu cel obtinut prin nltk). Pentru id-urile synseturilor, puteti scrie prescurtat: de exemplu, in loc de cate un nod pentru fiecare sens: word.n.01, word.n.02, word.n.03, word n.04 puteti scrie un singur nod word.n.01-04
- b) Indicati cazurile in care se modifica energia intr-un cuib si respectiv intr-un nod simplu.
- c) Explicati rolul feromonului in deplasarea furnicilor. Explicati cand si cum se modifica acesta.