

# Processament del Llenguatge Humà (GIA-PLH)

## Exercicis sobre anàlisi de dependències

### Ejercicio 1.

Dada la siguiente oración: *John quit his job*.  
Dibuja los siguientes análisis de dependencias

- a) (2,1), (0,2), (1,3), (3,4)
- b) (2,1), (0,2), (2,3), (3,4)
- c) (2,1), (0,2), (2,4), (3,4)
- d) (2,1), (0,2), (2,4), (4,3)
- e) (0,1), (1,2), (2,3), (3,4)

- ¿Cuales son inválidos y porqué?
- ¿Cuales son análisis proyectivos?

### Ejercicio 2.

En un modelo lineal de análisis de dependencias. el vector de atributos  $f(x, y)$  para una oración  $x$  y su árbol de dependencias  $y$  se define como:

$$f(x, y) = \sum_{(h,m) \in y} \mathbf{f}(x, h, m)$$

donde  $\mathbf{f}(x, h, m)$  es una función que proyecta una dependencia  $(h, m)$  y una oración  $x$  a un vector de atributos locales.

Queremos que el vector  $f(x, y)$  tenga dos dimensiones:

- $f_1(x, y)$  = frecuencia en que una dependencia con *head car* y modificador *the* aparece en  $(x, y)$
- $f_2(x, y)$  = frecuencia en que una dependencia con PoS del *head* es NN, PoS del modificador es DT, y no hay un adjetivo (JJ) entre el DT y el NN en  $(x, y)$

Asumiendo que cada elemento  $x_i$  de la oración es un par (*word, PoS*) y que las funciones  $word(x_i)$  y  $pos(x_i)$  devuelven *word* y "PoS" del par, respectivamente:

1. Define la función  $\mathbf{f}(x, h, m) = \langle \mathbf{f}_1(x, h, m), \mathbf{f}_2(x, h, m) \rangle$  que produce la definición anterior de  $f(x, y)$ .
2. Computa el valor de  $f(x, y)$  para el siguiente par  $(x, y)$ :

$$x = \textit{The/DT car/NN with/IN the/DT red/JJ hood/NN won/VBD the/DT car/NN race/NN}$$
$$y = \{(2, 1), (7, 2), (2, 3), (3, 6), (6, 4), (6, 5), (0, 7), (7, 10), (10, 8), (10, 9)\}$$

### Ejercicio 3.

Dada la oración "*natural language technology courses are fun*",

1. Dibuja los árboles de dependencias no etiquetado para las siguientes interpretaciones:
  - (a) *technology courses about natural language are fun* (los cursos de tecnología sobre lenguaje natural son divertidos)
  - (b) *courses about technology on natural language are fun* (los cursos sobre tecnología del lenguaje natural son divertidos)
  - (c) *natural courses about language technology are fun* (los cursos naturales sobre la tecnología del lenguaje son divertidos)
  - (d) *courses about natural technology for language are fun* (los cursos sobre tecnología natural para el lenguaje son divertidos)
2. Simula el comportamiento de un analizador de dependencias basado en transiciones que use el modelo arco-estandar, es decir, con operaciones *shift*, *left-arc* y *right-arc*). Lista el contenido intermedio en la pila y el buffer, así como la acción seleccionada a cada paso requerido para obtener el árbol de cada una de las interpretaciones anteriores.