

Análisis y simplificación de oraciones

Práctica 2 – Conocimiento y Razonamiento Automatizado

2020-2021

1. Objetivos

El objetivo de la práctica es doble:

- Por un lado, crear una Gramática de Cláusulas Definidas que permita analizar sintácticamente una oración en español y crear el árbol de constituyentes.
- Por otro, simplificar oraciones complejas convirtiéndolas en oraciones simples.

Dada la complejidad de la gramática de la lengua española se trabajará con una versión reducida de la misma que permita analizar un conjunto limitado de frases.

2. Práctica a entregar

Una vez que se han realizado con éxito los ejercicios 2.1 y 3.1, se puede abordar la creación de una **gramática** más compleja que **valide, represente el árbol de constituyentes y simplifique oraciones complejas** de los siguientes tipos (entre paréntesis):

- Oración Simple (o).
- Oración Coordinada (oc).
- Oración Subordinada de Relativo (or).
- Oraciones Compuestas, es decir combinaciones de oraciones simples, coordinadas y de relativo (ocm).

En cuanto a los grupos sintácticos, se deben incluir reglas que soporten los siguientes tipos:

- Grupo Nominal (gn).
- Grupo Adjetival (gadj).
- Grupo Adverbial (gadv).
- Grupo Preposicional (gp).
- Grupo Verbal (gv).

lo que implica la definición de los siguientes tipos de términos:

- Determinantes (det).
- Nombres (n).
- Nombres propios (np).
- Verbos (v).
- Adjetivos (adj).
- Adverbios (adv).
- Conjunciones (conj).
- Preposiciones (prep).

La simplificación de oraciones coordinadas, subordinadas de relativo o compuestas consistirá en la conversión de cualquiera de estas oraciones en tantas oraciones simples como se componga la original. A continuación se muestran varios ejemplos:

- “*Juan estudia Filosofía pero María estudia Derecho*” se debe descomponer en “*Juan estudia Filosofía*” y “*María estudia Derecho*”.
- “*Juan come y bebe mientras María lee*” se debe descomponer en “*Juan come*”, “*Juan bebe*” y “*María lee*”.
- “*Juan, que es muy alto, tiene el pelo rubio*” se debe descomponer en “*Juan es muy alto*” y “*Juan tiene el pelo rubio*”.
- “*Juan, que es muy alto, tiene el pelo rubio pero María es morena*” se debe descomponer en “*Juan es muy alto*” , “*Juan tiene el pelo rubio*” y “*María es morena*”.

El código de la práctica se entregará en dos ficheros: el **programa** propiamente dicho y la utilidad **draw.pl** que deberá ser consultada por el primero para representar gráficamente los árboles de constituyentes de las oraciones analizadas y simplificadas.

3. Conjunto de oraciones

Además de las recogidas en los apartados 2 y 3, la gramática debe **analizar** y **simplificar** correctamente las siguientes oraciones tipo:

1. *Juan es moreno y María es alta.*
2. *Juan estudia Filosofía pero María estudia Derecho.*
3. *María toma un café mientras Juan recoge la mesa.*
4. *Juan toma café y lee el periódico.*
5. *Juan y Héctor comen patatas fritas y beben cerveza.*
6. *Juan come patatas fritas pero María prefiere paella aunque Héctor toma café e Irene lee una novela.*
7. *Irene canta y salta mientras Juan estudia.*
8. *Héctor come patatas fritas y bebe zumo mientras Juan canta y salta aunque María lee una novela.*
9. *Juan, que es ágil, escala en el rocódromo por las tardes.*
10. *Juan, que es muy delicado, come solamente manzanas rojas.*
11. *El procesador de textos, que es una herramienta bastante potente, sirve para escribir documentos.*
12. *El procesador de textos es una herramienta muy potente que sirve para escribir documentos pero es bastante lento.*
13. *El ratón que cazó el gato era gris.*
14. *El hombre que vimos ayer era mi vecino.*

4. Mejoras del programa

Se valorará positivamente cualquier mejora de la gramática sobre lo expuesto anteriormente. Algunas ideas son las siguientes:

- Introducir algún preprocesamiento de la oración que agilice el proceso de análisis de una estructura sintáctica compleja.
- Implementar algún mecanismo para flexionar las palabras del diccionario, añadiendo el plural a las palabras en singular o el femenino para las palabras de género masculino (o viceversa).

- Asignar funciones a cada uno de los grupos gramaticales indicando si funcionan como Complemento Directo, C. Indirecto, Atributo, C. Circunstancial, etc.
- Cualquier otra que cuente con el visto bueno del profesor.

Para refrescar los conceptos de análisis sintáctico (tipos de sintagmas, tipos de oración, funciones de los sintagmas, ejemplos resueltos,...) se recomienda consultar la página <https://sintaxis.org>. El analizador que incluye no siempre es capaz de analizar las oraciones propuestas en este enunciado, pero puede servir para resolver los ejemplos de la propia página web.

5. Detalles de la entrega

La práctica se podrá realizar en equipos de hasta 3 personas (los mismos que en la práctica anterior). Además del **código fuente del programa (.pl)**, convenientemente estructurado y comentado, se deberá incluir una **breve memoria (PDF)** indicando los miembros del equipo, reparto de tareas entre los mismos, conjunto de frases analizadas/simplificadas con éxito, mejora realizada, errores y/o aspectos no implementados, así como las fuentes consultadas para la resolución de la práctica.

La entrega se realizará **a través de la plataforma** en un **único fichero .zip** antes de las **23:59 horas del día 27 de abril de 2021**. El nombre del fichero será NombreApellido1Apellido2.zip de uno de los integrantes del grupo. Una vez recibidos los trabajos y establecido el número de equipos, se publicará una fecha con la hora de defensa para cada equipo.

La entrega de prácticas **copiadas, total o parcialmente**, supondrá el **suspenso** del laboratorio para todos los alumnos implicados.