**¿Qué es la Lingüística Computacional o PLN?**

La Lingüística Computacional (Computational Linguistics) puede considerarse una disciplina de la linguística aplicada y la Intelingencia Artificial, y tiene como objetivo la realización de aplicaciones informáticas que imiten la capacidad humana de hablar y entender. A la Lingüística Computacional se le llama a veces Procesamiento del Lenguaje Natural (PLN), o Natural Language Processing (NLP). Ejemplos de aplicaciones de PLN son, por ejemplo, los programas que reconocen el habla, los traductores automáticos, ...

Todavía se sabe relativamente poco sobre el lenguaje humano. Los linguistas llevan décadas intentando descifrar cómo funciona esta capacidad única de la especie humana. Muchos animales tienen formas complejas de comunicación pero, que se sepa, ninguno de estos "lenguajes" cumple la característica más significativa del lenguaje humano natural: **la infinitud discreta**.

*El lenguaje humano natural es discreto en cuanto a sus unidades, pero infinito en cuanto a las combinaciones que pueden hacerse con estas unidades.*

Por ejemplo, las palabras son unidades discretas y finitas de la lengua. Sin embargo, combinando un número limitado de palabras podemos construir infinitas frases.

Y esa es la razón por la que un niño o un adulto construye contínuamente frases que no ha escuchado jamás a partir de palabras que sí ha tenido que escuchar y memorizar con anterioridad. Así, hablar es inventar continuamente nuevas combinaciones.

**Lenguaje natural vs lenguaje artificial**

Tanto el lenguaje natural como el lenguaje artificial son humanos. El primero es natural porque se aprende (o adquiere) inconsciente e involuntariamente. Ningún bebé decide aprender o no la lengua que hablan sus padres, y ningún padre sienta a su hijo y le enseña las reglas sintácticas de su lengua, entre otras razones porque ni él mismo las sabría describir. Las personas hablan y se entienden, pero no se cuestionan las reglas que utilizan al hablar. Los linguistas sí lo hacen, y a veces llegan a conclusiones interesantes.

Los lenguajes artificiales sí que se aprenden voluntária y conscientemente. Aunque algunos tienen la característica de la infinitud discreta, son muy diferentes a los lenguajes naturales. Un ejemplo de lenguaje artificial son los lenguajes de programación utilizados para desarrollar programas informáticos. Un ejemplo de lenguaje artificial finito sería el menú de los cajeros automáticos, el sistema operativo MS Windows,...

**¿Podremos hacer que una máquina hable algún día?**

Bien, es hora de ponerse futurista.La máxima aspiración de la PLN sería conseguir que la computadora nos hablara y nos entendiera (en el sentido estrictamente lingüístico, claro).

Es decir que fuera capaz de generar infinitas frases como las nuestras. Pero no solo infinitas, sino también **aceptables**. Una oración es aceptable cuando los hablantes nativos de esa lengua la reconocen como oración bien formada. Así, aunque cualquier hablante de español entienda 'perro comer hueso', ninguno la consideraría aceptable.

Se supone que en el lenguaje intervienen múltiples factores cognitivos y psicológicos, pero no hace falta representar toda la estructura mental y cognitiva humana para empezar a trabajar. Cada programa informático, según sea su función, se ocupará de unos aspectos u otros del lenguaje y sus estrechas relaciones con los demás  componentes cognitivos. Se trabaja, pues, modularmente.

Por ejemplo, imaginemos que queremos crear un programa que genere infinitas oraciones bien formadas sintácticamente sin tener en cuenta el significado de éstas. Para hacerlo, solo haría falta un lexicón y un conjunto de reglas combinatorias. El lexicón es algo así como una lista de palabras que están almacenadas en nuestro cerebro. Estas palabras se relacionan de manera compleja con sus respectivos significados,  pero eso no nos interesaría para crear el programa . Al conjunto de reglas que nos permite combinar las palabras que hay en el lexicón se le llama sintaxis. Gracias a estas reglas el cerebro es capaz de crear infinitas oraciones.

Pues bien, la reglas sintácticas y el lexicón es lo que habría que hacer explícito en nuestro programa.  Formalizando ambas cosas, y precindiendo del resto, muchos lingüistas piensan que es suficiente para hacer que la computadora genere infinitas oraciones sintácticamente correctas.

Hoy día se están realizando muchas investigaciones para conseguir "explicar" a una computadora la manera que tenemos de comunicarnos los humanos, pero todavía queda mucho camino. Una de las mayores dificultades es el hecho de que todavía no hemos conseguido descifrar y explicitar totalmente las reglas inconscientes que rigen nuestra capacidad lingüística. De esto se encarga la linguistica teórica, que junto a la informática, son las dos herramientas con las que trabaja el PLN.

*Xavier López Morrás (2004)*
prolepsi@yahoo.es