

Lenguaje formal

En matemáticas, lógica, y ciencias de la computación, un **lenguaje formal** es un lenguaje cuyos símbolos primitivos y reglas para unir esos símbolos están formalmente especificados.^{[1][2]} Al conjunto de los símbolos primitivos se le llama el alfabeto (o vocabulario) del lenguaje, y al conjunto de las reglas se lo llama la gramática formal (o sintaxis). A una cadena de símbolos formada de acuerdo a la gramática se la llama una fórmula bien formada (o palabra) del lenguaje. Estrictamente hablando, un lenguaje formal es idéntico al conjunto de todas sus fórmulas bien formadas. A diferencia de lo que ocurre con el alfabeto (que debe ser un conjunto finito) y con cada fórmula bien formada (que debe tener una longitud también finita), un lenguaje formal puede estar compuesto por un número infinito de fórmulas bien formadas.

Por ejemplo, un alfabeto podría ser el conjunto $\{a,b\}$, y una gramática podría definir a las fórmulas bien formadas como aquellas que tienen el mismo número de símbolos a que b . Entonces, algunas fórmulas bien formadas del lenguaje serían: ab , ba , $abab$, $ababba$, etc.; y el lenguaje formal sería el conjunto de todas esas fórmulas bien formadas.

Para algunos lenguajes formales existe una semántica formal que puede interpretar y dar significado a las fórmulas bien formadas del lenguaje. Sin embargo, una semántica formal no es condición necesaria para definir un lenguaje formal, y eso es una diferencia esencial con los lenguajes naturales.

En algunos lenguajes formales, la *palabra vacía* (esto es, la cadena de símbolos de longitud cero) está permitida, notándose frecuentemente mediante ϵ , e o λ .

EJEMPLOS DE LENGUAJES FORMALES

- Un conjunto de todas las palabras sobre $\{a,b\}$.
- El conjunto $\{a^n : n \text{ es un número primo}\}$.
- El conjunto de todos los programas sintácticamente válidos en un determinado lenguaje de programación.
- El conjunto de todas las fórmulas bien formadas en la lógica de primer orden.

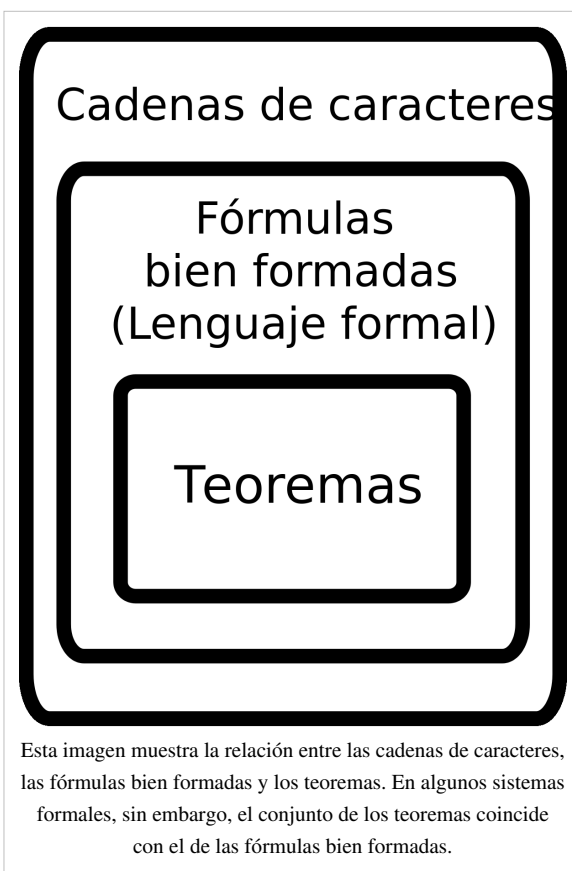
Especificación de lenguajes formales

Los lenguajes formales se pueden especificar de una amplia variedad de formas, como por ejemplo:

- Cadenas producidas por una gramática formal (véase la jerarquía de Chomsky).
- Cadenas producidas por una expresión regular.
- Cadenas aceptadas por un autómata, tal como una máquina de Turing.

Las cadenas están formadas por un conjunto de símbolos que pertenecen a un mismo lenguaje, existen dos formas de componer una sentencia o función con los símbolos:

- Sintaxis



- Semántica

Operaciones

Se pueden utilizar varias operaciones para producir nuevos lenguajes a partir de otros dados. Supóngase que L_1 y L_2 son lenguajes sobre un alfabeto común. Entonces:

- La *concatenación* $L_1 L_2$ consiste de todas aquellas palabras de la forma vw donde v es una palabra de L_1 y w es una palabra de L_2
- La *intersección* $L_1 \cap L_2$ consiste en todas aquellas palabras que están contenidas tanto en L_1 como en L_2
- La *unión* $L_1 \cup L_2$ consiste en todas aquellas palabras que están contenidas ya sea en L_1 o en L_2
- El *complemento* $\sim L_1$ consiste en todas aquellas palabras producibles sobre el alfabeto de L_1 que no están ya contenidas en L_1
- El *cociente* L_1 / L_2 consiste de todas aquellas palabras v para las cuales existe una palabra w en L_2 tales que vw se encuentra en L_1
- La *estrella* L_1^* consiste de todas aquellas palabras que pueden ser escritas de la forma $W_1 W_2 \dots W_n$ donde todo W_i se encuentra en L_1 y $n \geq 0$. (Nótese que esta definición incluye a ϵ en cualquier L^*)
- La *intercalación* $L_1^* L_2$ consiste de todas aquellas palabras que pueden ser escritas de la forma $v_1 w_1 v_2 w_2 \dots v_n w_n$; son palabras tales que la concatenación $v_1 \dots v_n$ está en L_1 , y la concatenación $w_1 \dots w_n$ está en L_2

Una pregunta que se hace típicamente sobre un determinado lenguaje formal L es cuán difícil es decidir si incluye o no una determinada palabra v . Este tema es del dominio de la teoría de la computabilidad y la teoría de la complejidad computacional.

Por contraposición al lenguaje propio de los seres vivos y en especial el lenguaje humano, considerados lenguajes naturales, se denomina lenguaje formal a los lenguajes «artificiales» propios de las matemáticas o la informática, los lenguajes artificiales son llamados lenguajes formales (incluyendo lenguajes de programación). Sin embargo, el lenguaje humano tiene una característica que no se encuentra en los lenguajes de programación: la diversidad.

En 1956, Noam Chomsky creó la jerarquía de Chomsky para organizar los distintos tipos de lenguaje formal.

Verdades concernientes a los lenguajes formales

Teorema 1

El conjunto de lenguajes en general (incluyendo los no-formales) es incontable.

Lema 1

El conjunto de lenguajes en un alfabeto no vacío dado es incontable.

Afirmar que un alfabeto es no-vacío equivale a que ese alfabeto contenga al menos un símbolo, ergo, basta demostrar que el conjunto de lenguajes en el alfabeto $\{a\}$ es incontable. Como sabemos, un lenguaje L en $\{a\}$ es un subconjunto de $\{a\}^*$, esto nos lleva a la conclusión de que, el conjunto de todos los lenguajes en $\{a\}$ es justamente $2^{\{a\}^*}$ (el conjunto de todos los subconjuntos o conjunto potencia de $\{a\}^*$) y es evidente que $\{a\}^*$ es infinito (de hecho; contable), también ha sido demostrado que si A es un conjunto infinito (contable o incontable), entonces 2^A es mayor que A porque 2^A pasa a ser un conjunto infinito de órdenes del infinito, al ser mayor, no existirá biyección entre A y 2^A , lo que hace a 2^A un conjunto infinito incontable, la prueba ha finalizado.

Demostración del Teorema 1

Puede derivarse fácilmente que la aseveración delineada en el Teorema 1 es verdadera, porque el conjunto de lenguajes en general es justamente una unión infinita de conjuntos del tipo 2^A , donde A es un conjunto infinito contable.

Teorema 2

Los lenguajes son conjuntos contables.

Se sabe que un lenguaje L en un alfabeto Σ es un subconjunto de Σ^* y como ya se hizo mención, Σ^* es infinito contable, por ende, L es como mucho un conjunto infinito contable (del mismo tamaño que Σ^*), la prueba ha culminado.

Teorema 3

El conjunto de lenguajes formales es contable.

Como sabemos un lenguaje formal puede ser generado por una gramática formal (o de estructura de frase), lo cual implica que todo lenguaje formal puede ser aceptado por una MT, lo que a su vez implica que se puede definir una biyección entre el conjunto de lenguajes formales y el conjunto de las MT's (debido a la propiedad transitiva de la relación "existe biyección entre A y B "). Para demostrar el teorema se utilizará el concepto de codificación de MT's que se introduce en el estudio de las MT's universales, generalmente se codifica una MT con una función que tiene precisamente como dominio al conjunto de las MT's (lo llamaremos X) y como codominio $\{0, 1\}^*$, esa función puede ser una biyección si el codominio pasa a ser Y (un subconjunto de $\{0, 1\}^*$) y como $\{0, 1\}^*$ es contable, ese subconjunto también será contable y como existe dicha biyección (entre X e Y), la aserción ha sido demostrada, prueba concluida.

Notas y referencias

- [1] Mellema, Gregory, « formal language (<http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t116.e929>) » (en inglés), *The Oxford Companion to Philosophy*, Oxford University Press, , consultado el 13 de octubre de 2009.
- [2] Shapiro, Stewart, « Classical Logic (<http://plato.stanford.edu/archives/win2009/entries/logic-classical/>) », en Edward N. Zalta (en inglés), *Stanford Encyclopedia of Philosophy* (Winter 2009 Edition), , «Again, a formal language is a recursively defined set of strings on a fixed alphabet.».

Enlaces externos

- Autómatas y lenguajes (<http://homepages.mty.itesm.mx/rbrena/AyL.html>) - Libro electrónico gratuito por Ramon F. Brena Pinero.

Fuentes y contribuyentes del artículo

Lenguaje formal *Fuente:* <http://es.wikipedia.org/w/index.php?oldid=60188648> *Contribuyentes:* AchedDamiman, Acratta, Airunp, AlfonsoERomero, Allforrous, Amadís, Ascánder, Bienchido, Cfau, Cobalttempest, Danakil, Davius, Echani, Egaida, Ensada, Farisori, Greek, Interwiki, Ivan.Romero, JMCC1, Javier Carro, Jerowiki, Jjmaster, Jkbw, Kenrych, Luis Felipe Schenone, MONIMINO, Maldoror, Mansoncc, ManuelGR, Matdrides, Netito777, Paco79, Rbrena, RoyFocker, SpeedyGonzalez, Tano4595, UA31, Unf, Waka Waka, 74 ediciones anónimas

Fuentes de imagen, Licencias y contribuyentes

Archivo:Entidades sintácticas 2.svg *Fuente:* http://es.wikipedia.org/w/index.php?title=Archivo:Entidades_sintácticas_2.svg *Licencia:* Creative Commons Attribution-Sharealike 3.0
Contribuyentes: Luis Felipe Schenone

Licencia

Creative Commons Attribution-Share Alike 3.0 Unported
[//creativecommons.org/licenses/by-sa/3.0/](https://creativecommons.org/licenses/by-sa/3.0/)
