ESTRUCTURA ALGEBRAICA DEL LENGUAJE

Por Pedro Suárez, S.J.

ESQUEMA : 1) Introducción

2) Definiciones previas

3) Verificación

4) Observaciones finales

5) Bibliografía

1) INTRODUCCION

Las palabras que forman un idioma, por ejemplo el Español, parecen a primera vista rechazar todo intento de "matematización". Sin embargo, si se analiza el lenguaje vemos que éste tiene piezas o unidades (las palabras) que a su vez están constituidas por letras. Los lógicos, desde Aristóteles hasta Bertrand Russell y nuestros contemporáneos, han estudiado la estructura del pensamiento en el aspecto de su validez y lo han esquematizado -hasta cierto punto- en reglas que determinan la corrección y validez de los razonamientos. Sin embargo, la misma riqueza del pensamiento humano y su múltiple expresión hacen imposible una esquematización total, desde el punto de vista de la lógica matemática. Hoy se habla de una lógica que incluya el factor <u>tiempo</u> y esto da como resultado un sistema lógico muchísimo más rico y complejo que la lógica simbólica de Leibnitz, Boole, etc., pero que al mismo tiempo confiesa la imposibilidad de enmarcar <u>toda</u> expresión del pensamiento humano dentro de sus moldes[1].

En este modesto trabajo no nos fijaremos en la estructura lógica de los razonamientos. Solamente nos limitaremos a descubrir y describir la estructura algebraica que manifiesta el <u>lenguaje en cuanto formado por palabras, y éstas por</u>

letras. Nos concretaremos a considerar un lenguaje particular, el Español, pero lo que se afirma en el Teorema central de este trabajo puede aplicarse (mutatis mutandis) a cualquier otro idioma occidental de estructura similar al nuestro (por ejemplo: Inglés, Francés, Alemán, Portugués, etc.).

Aplicando las definiciones que daremos en 2) demostraremos el siguiente

Teorema: Sea B el conjunto de 27 letras del alfabeto a, b, c, d, e, f, g, h, i, j, k, l, m, n, ñ, o, p, q, r, s, t, u, v, w, x, y, z. Sea X el conjunto de todas las palabras de la lengua española. Denótese por ( ) la letra o palabra "vacía". Defínase una operación asociativa "yuxtaposición" sobre B, tal que si x_1, x_2 son dos elementos cualesquiera de B, la yuxtaposición de x_1 y x_2 es sencillamente x_1x_2. Entonces se afirma que X presenta una estructura algebraica de monoide no conmutativo bajo la operación de yuxtaposición generado no libremente, sino con relaciones, por B.

Antes de pasar a dar las definiciones pertinentes hay que hacer algunas observaciones:

i) Se han excluido de la lista de "generadores" (elementos en B) las letras ch, ll, rr, que resultan de la yuxtaposición de dos letras.

ii) Se ha incluido la letra no castellana w para completar la lista de posibles palabras en el castellano moderno con palabras que provienen de lenguas extranjeras pero que se han hecho de uso común en nuestro idioma. Los puristas del idioma la podrán eliminar, dejando a B con 26 elementos sin que el espíritu del teorema sufra menoscabo.

iii) No se han considerado otros signos del lenguaje como el acento o la diéresis. En este primer intento de descubrir la estructura algebraica del lenguaje preferimos dejar para más adelante considerar este problema, o mejor aún, dejar como "ejercicio para el lector" el hallar la manera de corregir o amplificar este estudio con la consideración de estos signos.

2) DEFINICIONES PREVIAS

Sea C un conjunto cualquiera. Sea C x C el producto cartesiano de C consigo mismo, i.e. $C \times C = \{(x_1,x_2) / x_1, x_2 \in C\}$.

i) Una operación binaria s sobre C es una función $s: C \times C \to C$. Es decir, si x_1,x_2 son dos elementos cualesquiera de C, se tiene $s(x_1,x_2)$ es un elemento de C.

Si la función s cumple además con la regla $s(s(x_1,x_2),x_3) = s(x_1, s(x_2,x_3))$ para cualesquiera x_1,x_2,x_3 elementos de C, entonces s es una operación asociativa.

Por ejemplo, la suma (+) de dos números naturales es una operación binaria asociativa $+ : N \times N \to N$, donde N es el conjunto de los números naturales.

ii) Un semigrupo asociativo (o simplemente semigrupo) es un par (S,+) donde S es un conjunto cualquiera y + es una operación binaria asociativa sobre S. Por ejemplo (N, +) es un semigrupo. A menudo se elimina el signo + de la notación y se dice solamente "S es un semigrupo", con la operación sobreentendida.

iii) Un monoide es un semigrupo (asociativo) M en el cual existe un "elemento neutro" 0 con respecto a la operación + del monoide M, que verifica la condición $0 + x = x = x + 0$ para todo $x \in M$. Ejemplo: $N \cup \{0\}$ es un monoide.

iv) Sea C un conjunto cualquiera. Defínase una operación binaria asociativa + sobre C tal que existe un subconjunto $B = \{C_1, C_2, \ldots, C_n\} \subset C$ que verifica la condición siguiente: si x es un elemento arbitrario de C, x puede ser expresado unívocamente como la suma de un número finito de elementos de B, es decir $x = \sum_{c_j \in B} c_j$. Se dice entonces que c_i $(1 \leq i \leq n)$ es un generador de C y que

B es un conjunto base de C o una base de generadores de C (como semigrupo). Si existe además un elemento neutro de C bajo la operación + entonces se dice que C es un monoide generado por B o que B es una base de generadores de C como monoide.

v) Antes de pasar adelante demos un ejemplo. Supongamos un conjunto B = {a, b, c} sobre el cual se define la operación asociativa "yuxtaposición", es decir si x, y, z ε B, entonces "yuxtaposición" de x e y es simplemente xy. Además (xy)z = x(yz) y entonces se escribe simplemente xyz para expresar la yuxtaposición de tres elementos. Los elementos "generados" por B serán entonces de la forma

a	acc	bbbbbbabb
aa	cca	bacccba
aba	ccbac	aaccaa
abbca	bca	etc...

Sea M el conjunto de todos estos y otros posibles elementos generados por B. Obviamente $B \subset M$. M será entonces un <u>semigrupo libre</u> generado por (o con base) B. Los "elementos" de M serán todas las palabras finitas formadas por las letras a, b y c. Si se añade a M un elemento "identidad" o "neutro" (la palabra vacía) con respecto a la operación de yuxtaposición, entonces M será un monoide libre con base B.

Sin embargo, supongamos que queremos excluir de la lista de elementos de M (palabras posibles) alguna palabra o alguna combinación de generadores. Por ejemplo, eliminemos en M palabras en que aparezca una misma letra repetida dos veces seguidas por yuxtaposición (aa, bb, cc). Entonces, el elemento o palabra abacb permanece, pero las palabras aaba, cbb, baccba, etc. desaparecerían de nuestra lista de palabras posibles o elementos de M. Matemáticamente tal exclusión es un ejemplo de lo que se llama una <u>relación algebraica</u> entre los generadores de una estructura algebraica. En este ejemplo se dice que M es generado por B como semigrupo (o monoide) pero no lo es "libremente" sino con las relaciones aa, bb, cc.

Si a la lista de propiedades de un semigrupo o monoide S se añade la ley conmutativa ($x + y = y + x$ para todo $x,y \in S$) se obtiene entonces un semigrupo o monoide conmutativo.

vi) Definamos la longitud de una palabra como el número de letras (no necesariamente distintas) que la forman. Denótese la longitud de una palabra p por L(p). Ejemplo: si p = aacb entonces L(p) = 4. Se verifica L(p) = 0 si y sólo si p = palabra vacía ( ).

vii) Sea M un monoide generado (libremente o no) por un número finito de generadores. Para toda palabra $p \in M$ se tendrá $L(p) < \infty$. Esta condición (que llamaremos Condición F) está implícita en la definición dada en iv), pero la recalcamos para enfatizar que se excluyen las palabras de longitud "infinita" de la lista de palabras (elementos de M) generados por una base B.

3) VERIFICACION

La demostración del teorema es un simple ejercicio: verificar que B = {letras del alfabeto} genera las palabras del idioma castellano, no libremente, con la condición F y otras que añadiremos más adelante.

Que las palabras del idioma son generadas por las letras del alfabeto por yuxtaposición es trivial. Sin embargo, no toda palabra obtenida por yuxtaposición de letras está permitida en nuestra lengua. La siguiente lista de relaciones expresa este hecho:

Relaciones de tipo A: no se permiten repeticiones de la misma letra de modo que no exista palabra en castellano con esa combinación de letras. Obviamente la lista sería demasiado grande, pero a modo de ejemplo, se excluirán las combinaciones aaa, ii, ff, ppp, rrr, etc. y toda palabra que las contenga.

Relaciones de tipo B: se excluyen palabras o combinaciones de letras que no se puedan pronunciar en el idioma. Por ejemplo, fxhhw, mstspgt, etc.

Relaciones de tipo C: se excluyen palabras que, aunque pronunciables, no tengan sentido en Español. Por ejemplo: pejtelnu, nentefki, etc. El lector puede añadir otras relaciones a su arbitrio de modo que, eliminadas estas palabras "prohibidas" quede el idioma que hablamos, leemos y escribimos.

Obviamente, la ley de conmutatividad no se aplica en el monoide del lenguaje, pues salvo raras excepciones el orden de las letras altera el sentido de las palabras y, más aún, en ejemplos como "casa", y "asac" la primera es una palabra permitida mientras que la segunda no lo es, por una relación de tipo C.

4) OBSERVACIONES FINALES

a) El lector puede observar que las palabras del idioma constituyen un sistema algebraico inferior al grupo, en el sentido de que "algo le falta" al idioma para llegar a constituir un grupo. Este "algo" es el inverso con respecto a la operación de yuxtaposición. Es evidente que dada una palabra cualquiera p no existe otra palabra p' tal que pp' = ( ) = p'p.

b) ¿Qué hacer con las palabras formadas por iniciales, como son UASD, UCMM, IBM, etc., o las palabras como INTEC, RAHINTEL, UCAMAIMA, etc., que están formadas por sílabas sacadas de otras palabras con sentido completo? El purista del idioma quizás prefiera no admitirlas en el diccionario, pero el algebrista no tendrá dificultad en incluirlas entre las palabras permitidas y no entre las relaciones. De una u otra manera la estructura algebraica de monoide generado con relaciones no queda afectada.

NOTAS

[1] En la Universidad Nacional de Colombia (Bogotá) el Dr. Carlos E. Vasco está realizando investigaciones sobre una "Lógica" que tenga una estructura temporal interna. Al Dr. Vasco le agradezco algunas observaciones interesantes sobre el presente trabajo durante el tiempo que coincidimos en el Institut des Hautes Etudes Scientifiques (París) en Julio de 1978.

BIBLIOGRAFIA

Gentile, Enzo R. Estructuras Algebraicas. O.E.A., Washington, D.C.: 1967.

______________. Estructuras Algebraicas II. O.E.A., Washington, D.C.:1971.

Luna, Eduardo; Núñez, Amarilis; Alba, Orlando; Núñez, Apolinar. Reglas del acento en Español expresadas en Lenguaje Matemático. Magister, UCMM, Santiago, 1977.

O'Brien, Horacio Hernán. Estructuras Algebraicas III. O.E.A., Washington, D.C.: 1973.