

## FORMALIZACIÓN Y EXTENSIÓN DE LOS PROBLEMAS TAXONÓMICOS

Yosu Yurramendi Mendizabal  
Universidad del País Vasco/ EMU

La taxonomía es un sistema o método de clasificación de los seres, cuyas bases se establecieron en las ciencias naturales durante los siglos XVII y XVIII.

Su objetivo es la determinación natural entre los animales y plantas, de un conjunto de relaciones que representa sus "afinidades".

A pesar de las dificultades que conlleva alcanzar este objetivo, durante los siglos XVIII y comienzos del XIX se trazó un sistema de relaciones que, en sus líneas generales, ha sido mantenido hasta nuestros días. Sin embargo, sus imperfecciones son a menudo patentes.

El planteamiento de esta tarea puede ser englobado en otro más amplio que es propio a todo ser humano, y que H. Daudin en De Linné a Lamarck. Méthodes de classification et idée de série en botanique et zoologie (1740-1790, PUF, 1983 (1ª ed. en 1926-1927), en su página 2 describe de esta manera:

"Designer d'un même nom tous ceux (objets) qui se manifestent assez pareils pour que leurs différences, spécifiques ou individuelles, puissent être, dans la plupart des cas, négligées sans inconvénient - rapprocher ensuite, sous une désignation commune, les genres qui, bien que très discernables et distingués par des vocables différents, intéressent cependant la conduite humaine par des propriétés ou des relations communes - ces deux actes ne sont-ils impliqués nécessairement dans toute démarche de l'intelligence à l'égard du donné?"

Pourtant, si le premier semble bien se confondre, en quelque sorte, avec l'existence même du langage et ne pouvoir être expliqué que du jour ou la g n se de ce fait fondamental aurait  t  rendue accessible   l'analyse, peut- tre le second, qui proc de d'un analyse d j  constitu , releve-t-il de facteurs plus complexes, plus sp ciaux et, par l  m me, moins  loign s de l'observation."

Estas denominaciones comunes tales como "p jaro", "pez", etc. se basan a menudo en afinidades emp ricas tales como el ser "alado", "b pedo", "acu tico", etc. que pueden ser subordinadas unas a otras.

Citamos de nuevo a H. Daudin:

"Or tout l'effort scientifique de la classification botanique et zoologique,   l' poque moderne, va  tre de tirer de l'observation exacte, et d'elle seule, des motifs suffisantes (el subrayado es nuestro) de grouper les plantes et les animaux selon leurs propri t s intrins ques, ind pendamment de la dignit , de la valeur o  de l'int ret pratique qui peut leur  tre imput  par des repr sentations traditionnelles, (...)  
Mais, en m me temps, le travail de la classification scientifique, avec plus ou moins d'assurance, tend   constituer, au-dessus de toutes les liaisons diverses qui r sultent de la pluralit  de leurs caract res, un tableau unique et complet de leurs "rapports" r els.

Antecedentes de este trabajo, podemos encontrarlos ya en la obra de Arist teles, que intenta utilizar el m todo de la "divisi n" de su maestro Plat n por el cual un objeto es desdoblado en dos opuestos tantas veces como sea necesario hasta llegar a una perfecta determinaci n de la especie de ese objeto (por ejemplo, la del pescador de ca a en El Sofista) y que, sin embargo, decide rechazar en su obra naturalista, pues llega forzosamente a desdoblar "g neros" por la oposici n de caracteres. Por ejemplo, hay "p jaros terrestres" y "p jaros acu ticos".

La clasificaci n aristot lica de los animales en cuadr pedos viv paros, cuadr pedos ov paros, p jaros, peces, serpientes, crustaceos, etc. en

su "Historia de los animales", se justifica por encontrarse en cada uno de los géneros pertenecientes a una misma clase, los mismos órganos. En la introducción de esta obra se explicita que el mundo zoológico se presta a ser dividido según varias consideraciones distintas e independientes, como por ejemplo, los tipos de vida o las partes de los cuerpos de los animales. Aún así, enfatiza las consideraciones referentes a los aparatos de nutrición y generación, pues, en particular, distingue los animales "con sangre" de los que "no tienen sangre".

En todo caso, a decir de los especialistas, a Aristóteles le interesa más la relación lógica entre las nociones de género y especie, que la construcción de un sistema natural que dé cuenta de los géneros de los animales.

Es a partir de Linneo (1700-1778) que se considera el logro de un sistema que explique, en parte, junto a una lexicología, los diversos géneros de plantas y animales.

En botánica, algunos precursores tales como Bahuin, Ray y Tournefort, marcaron el camino para la consecución de un sistema natural, señalando su gran importancia teórica y práctica, y utilizando para ello los caracteres orgánicos de las plantas, incluidos los del aparato reproductor (la flor y el fruto) que para Linneo serían los determinantes de su sistema, aunque dichos caracteres observables estaban todavía imperfectamente discernidos.

El sistema propuesto por Linneo para definir las características de las plantas, se apoya en la consideración exclusiva del aparato sexual (según el número, la proporción, la situación de los estambres en relación a los pistilos) y en la rigidez del número de unidades jerarquizadas: clase, orden, género, especie y variedad, de las cuales las tres primeras comportan la "disposición teórica" y las dos últimas la "disposición práctica". H. Daudin en una nota a pie de página (op.cit. p. 38) dice:

"Recourant aux analogies comme il (Linneo) aime à le faire, il la motive par des séries pareilles de cinq termes fondamentaux qu'il extrait de la géographie, de l'art militaire et, enfin, de la "philosophie"

## Los problemas taxonómicos

scolastique (genus summum, intermedium, proximum, species, individuum). Cette dernière série, quoique fort étrangère en elle même, aux besoins de la cause, a bien pu lui être réellement un motif de limiter à cinq termes la série de ses unités taxinomiques."

Las pretensiones del llamado "método natural" pueden ser adivinadas en el siguiente párrafo de la "Philosopphia Botanica" (1751) de Linneo (p.290) al que hemos accedido por la obra de H. Daudin:

"Les genres et les especes une fois établis, déterminés par leur différences essentielles, nous avons atteint le point culminant dans l'étude des végétaux. Si les Botanistes pouvaient en arriver enfin à être en mesure de déterminer toutes les espèces par leur nom essentiel, il n'y aurait pas à vouloir aller plus loin ("non datur ultra").

Sin embargo, los problemas surgen de inmediato. Citamos todavía a H. Daudin (op.cit. p.41):

"il arrive, en fait, fréquemment que des plantes qui concordent parfaitement par telle de ses "parties", ou par telle "dimension" de l'une d'elles, se distinguent plus ou moins fortement par telle autre. Auquel des caracteres devra-t-on dans de pareils cas, attribuer une signification prépondérante? Pas d'autre moyen de le savoir que d'examiner quel est celui qui se montre le plus constant dans les ensembles d'espèces dont l'unité générique est acquise."

En zoología, el método lineano supuso la división natural del reino animal según la estructura interna, es decir, según el valor funcional de los aparatos vitales (circulatorio, respiratorio, masticador, reproductor etc.) El cambio, por tanto, fue conceptual, ya que las clases formadas hasta entonces (según la clasificación aristotélica) por afinidades visibles, se conservaron en su mayoría.

No obstante, había contemporáneos como Adanson y Buffon, que veían la necesidad de apoyarse en varios criterios para constituir las diversas fami-

lias de animales y plantas. Así, Buffon en su obra "De la maniere de traiter et d'etudier l'histoire naturelle" (1749), dice:

"... que le seul moyen de faire une méthode instructive et naturelle, c'est de mettre ensemble les choses qui se ressemblent, et de séparer celles qui diffèrent les unes de les autres. Si les individus ont une ressemblance parfaite, ou les différences si petites..., ces individus seront de la même espece; si les différences commencent à être sensibles,... les individus seront d'une autre espèce mais du même genre que les premiers; si ces différences sont encore plus marquées,... non seulement d'une autre espèce mais même d'un autre genre..., et cependant ils seront de la même classe, parce qu'ils se ressemblent plus qu'ils ne diffèrent; mais si au contraire le nombre des différences excède celui de ressemblances, alors les individus ne sont pas de la même classe."

Observamos, además, en este párrafo, que Buffon se dispone a ordenar el sistema natural según una jerarquía de tres grados, es decir, clase, género y especie, en contraste con los cinco a los que Linneo se ciñe.

A este respecto, hay que añadir que otro contemporáneo, Donati, en su obra "Della Storia Naturale marina dell'Adriatico" (1750), concebía hasta siete grados.

En el lenguaje de la Teoría de Conjuntos, estas nociones de clasificación pueden ser fácilmente definidas.

Dado un conjunto de objetos,  $X$ , una jerarquía de clases (correspondientes a las clases, órdenes, géneros, especies y variedades lineanas) es un subconjunto del conjunto de las partes de  $X$ ,  $\mathcal{C} \subset \mathcal{P}(X)$ , definida de la siguiente manera:

\* La clase compuesta por todos los objetos pertenece a la jerarquía:  $X \in \mathcal{C}$  pero no así la clase que no tiene ningún objeto (en el caso de que se haya postulado su existencia):  $\emptyset \notin \mathcal{C}$ .

\* Cada una de las clases elementales, es decir, aquellas compuestas por un sólo objeto, pertenece a la jerarquía:  $\forall x \in X, \{x\} \in \mathcal{C}$ .

• Si dos clases de la jerarquía tienen algún objeto en común, entonces una de ellas está forzosamente incluida en la otra:

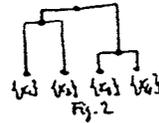
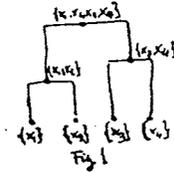
$$C, C' \in \mathcal{C} : C \cap C' \neq \emptyset \Rightarrow C \subset C' \vee C' \subset C.$$

Un ejemplo de jerarquía sería el siguiente:

$$X = \{x_1, x_2, x_3, x_4\}$$

$$\mathcal{C} = \{\{x_1\} \{x_2\} \{x_3\} \{x_4\} \{x_1, x_2\} \{x_3, x_4\} \{x_1, x_2, x_3, x_4\}\}$$

Al representar por medio de un grafo (Fig. 1) llamado dendrograma, las condiciones requeridas por esa jerarquía, se puede optar entre múltiples soluciones (Fig. 2). Dicha multiplicidad surge de las distintas elecciones que se ofrecen para emplazar las clases superiores de la jerarquía, en relación a la altura que les separa de la inferiores, es decir, de las clases elementales.



Este mismo problema se encuentra en el sistema linneano: conociendo las semejanzas entre dos unidades jerárquicas dadas, ¿se puede fijar a partir de ellas el grado que ocupará en la escala su reunión?

La altura que les separa debe reflejar de alguna manera el parecido de los objetos reunidos. Así, si los objetos de una clase no elemental son globalmente más parecidos que los de la otra, la primera será emplazada más cerca de las clases elementales que la segunda, de tal modo que su distanciamiento refleje esa disparidad.

Esta faceta debe ser tomada en cuenta a la hora de construir una jerarquía de un conjunto de objetos dados, pues ayuda a la inteligibilidad del sistema.

Otra faceta a tomar en cuenta es la del procedimiento de construcción. Aunque ésta ya ha sido señalada anteriormente de un modo superficial al mencionar las disposiciones teórica y práctica del sistema leneano, conviene prestarle un poco más de atención. La construcción de la jerarquía puede

hacerse ascendente (formando primero las clases de los objetos más parecidos y luego aplicando sucesivamente esta misma regla sobre el conjunto en que se toman las nuevas clases por nuevos objetos), o bien descendente (escindiendo gradualmente el conjunto inicial de objetos en clases distintas, según la desemejanza global de sus objetos, hasta obtener las clases elementales), o bien como lo concebía Linneo, ascendente y descendente.

Hay que hacer notar que en cada paso se construya una partición del conjunto de objetos. Dichas particiones están encajadas unas en otras.

Los algoritmos de construcción de una jerarquía de clases, llamados "básicos" o "generales", se definen de la siguiente manera:

a) Algoritmo ascendente (o agregativo, o aglomerativo):

1. Tomar la partición de las clases elementales.
2. Reunir dos clases de la partición en curso, de manera que maximice un criterio de semejanza definido a priori; por tanto, se construye una nueva partición.

3. Si la partición en curso está formada únicamente por el conjunto de todos los objetos, terminar; si no, volver al paso 2.

b) Algoritmo descendente (o disgregativo, o divisivo):

1. Tomar la partición formada únicamente por el conjunto de los objetos.

2. Dividir o disgregar una de las clases de la partición en curso, de manera que minimice un criterio de semejanza definido a priori; por tanto, se construye una nueva partición.

3. Si la partición en curso está formada por las clases elementales, terminar; si no, volver al paso 2.

Se observa que por la aplicación de cualquiera de los dos algoritmos, se construye una jerarquía binaria, es decir, una jerarquía en la que cualquier clase no elemental se obtiene, bien por agregación de dos clases, bien por la escisión en dos clases.

Para que los algoritmos sean operativos no falta más que precisar lo que se entiende por "parecido" o por "desemejanza" de clases. Es decir, estamos en el corazón mismo del problema que se plantearon los botanistas y zoólogos del siglo XVIII, pues para tal definición se debe estar en consonancia con los caracteres tomados en cuenta a los objetos de estudio.

Lo primero que es preciso dilucidar es, si el criterio de semejanza se atiene a un carácter (Linneo) o a varios (Adanson, Buffon), es decir, si los agrupamientos (o escisiones en su caso) son monotéticos o politéticos.

Otro punto que es preciso tomar en cuenta, es el referente al modo de considerar "globalmente" las clases.

Para responder a estas dos cuestiones, el camino usual ha sido el de representar los objetos en un espacio métrico. La información es recogida en una tabla de números reales en la que o bien se expresa la semejanza existente entre cada par de objetos, o bien se expresan para cada objeto, las coordenadas referentes a los caracteres observados, en un espacio vectorial sobre el que posteriormente se define una métrica (distancia) adecuada. En cualquier caso, se tienen expresadas las semejanzas de los objetos por medio de números reales, tanto sean los criterios utilizados monotéticos como politéticos.

La semejanza existente entre dos clases pertenecientes a la jerarquía, puede ser expresada utilizando criterios locales (por ejemplo, dos clases son tan parecidas como lo son los dos objetos más parecidos, uno de cada clase), o criterios globales (por ejemplo, dos clases son tan parecidas como lo son sus baricentros, en el caso de tratarse de un espacio vectorial euclídeano).

Un problema que puede plantearse a la hora de operativizar los algoritmos, es el del número de cálculos requeridos a cada paso: si el conjunto considerado en un paso determinado del algoritmo se compone de  $n$  elementos, el número de cálculos requeridos es  $n(n-1)/2$  (por ejemplo, si  $n = 200$  hay que hacer 19900 comparaciones). Uno de los procedimientos accesorios más ha-

## Los problemas taxonómicos

bituales para aligerar el problema, consiste en mantener constante el espacio métrico en todos los pasos del algoritmo utilizado, de manera que la mayoría de los cálculos efectuados en los pasos anteriores a uno dado permanecen también constantes, y por tanto, el número de cálculos a efectuar en la construcción de la jerarquía disminuye considerablemente. Aún así, si los caracteres tomados en cuenta son muchos, el cálculo suele ser todavía muy pesado.

La capacidad de memoria y la rapidez de cálculo de los ordenadores actuales ha facilitado la resolución, en parte, de este problema. En realidad habría que decir que la formalización y el desarrollo de los procedimientos taxonómicos ha sido favorecido por el desarrollo de los ordenadores, ya que sin ellos, la imposibilidad práctica los habría hecho inconcebibles.

La automatización del proceso de construcción de una jerarquía, en el caso de utilizar criterios politéticos, cuestiona otro problema muy importante, cual es el de la inteligibilidad de la jerarquía obtenida: ¿Cuáles son las características comunes o semejanzas de los objetos pertenecientes a una misma unidad jerárquica? Por otra parte, podría también plantearse la siguiente cuestión: ¿cuáles son las características que separan dos unidades jerárquicas dadas?

El modo de responder a este tipo de cuestiones es conceptualmente simple, pero operativamente, en el caso de tratarse de una jerarquía con muchas unidades, no lo es tanto. Las respuestas se basan en el criterio de semejanza usado en la construcción de la jerarquía, pero su obtención requiere de algoritmos nada sencillos para la organización de los cálculos.

Otro problema que se suele plantear en el caso de la construcción de una jerarquía sobre un conjunto de objetos grande, es el de la elección del nivel a partir del cual la jerarquía es "interesante". Esta cuestión la podríamos plantear de la siguiente manera: en la parte inferior de la representación de una jerarquía, las clases se reúnen de acuerdo a una gran semejanza en relación a aquella que reúne las clases en la parte superior, es

decir, la información que se desecha cuando se reúnen dos clases en la parte inferior del dendrograma es menor que la que se desecha cuando se hace lo mismo en la parte superior ¿A qué nivel habrá que considerar el dendrograma, de manera que guarde la mayor parte de las desemejanzas (información) en un número relativamente pequeño de unidades jerárquicas?

Una vez establecido lo que se entiende por cantidad de información (desemejanza), se determinan arbitrariamente la porción que se quiere retener, o alternativamente, se retiene la parte superior del dendrograma a partir del paso en el que la pérdida de información es considerable (desde un punto de vista también arbitrario) en relación a los pasos precedentes.

El método taxonómico, aparte de ofrecer un análisis-síntesis del problema tratado, es también útil desde el punto de vista decisional: si un nuevo elemento es añadido al sistema, las unidades jerárquicas anteriores pueden ser utilizadas para emplazarlo sucesivamente a través del recorrido trazado por el dendrograma, desde su vértice superior (la clase de todos los objetos) hasta su base (las clases elementales), utilizando para ello el criterio de semejanza anterior.

En fin, en la somera exposición que hemos hecho de los problemas taxonómicos (unos, planteados desde los inicios del método, y otros propios de los procedimientos actuales), hemos pretendido mostrar la importancia que ha tenido para su ulterior desarrollo la formalización de las nociones utilizadas.

Este procedimiento formalizado es considerado hoy día, en las distintas ramas de la investigación como son la antropología, la sociología, la lingüística, la medicina, la economía, etc., como un instrumento potente para analizar y sintetizar los múltiples datos ofrecidos por la realidad tratada.