

УДК 557.97

## О КЛАССЕ, КЛАССИФИКАЦИИ И СИСТЕМАТИЗАЦИИ

**С.В. Микони**

Санкт-Петербургский институт информатики и автоматизации Российской академии наук,  
Санкт-Петербург, Россия  
smikoni@mail.ru

### Аннотация

Рассмотрены определения понятия классификации в отечественных источниках. Обнаружены существенные различия в имеющихся трактовках, связанные с обозначением классификации, как процесса, так и его результата. Предложено именовать эти понятия разными терминами. Для устранения неточностей в определении классификации, её методов и связи с систематизацией знания предложено привлечь математические модели класса, классифицирования и системы. Рассматриваются модели формирования классов, используемые различными методами классифицирования объектов. Систематизированы свойства методов классифицирования и получаемых с их применением классификаций. Рассмотрен метод деления объёма понятий. На конкретном примере показана ошибка, допущенная в цитируемой работе за счёт неправильного выбора модели формирования класса. Рассмотрены индуктивные методы классифицирования объектов и метод архетипа. На основании анализа модели системы показывается соотношение классификации и систематизации, которое иллюстрируется на примере периодической системы химических элементов Д.И. Менделеева. Уточнены понятия классификации.

**Ключевые слова:** модель, класс, классифицирование, классификация, таблица Менделеева, система, систематизация, деление понятия, таксономия.

**Цитирование:** Микони, С.В. О классе, классификации и систематизации / С.В. Микони // Онтология проектирования. – 2016. – Т.6, №1(19). – С. 67-80. – DOI: 10.18287/2223-9537-2016-6-1-67-80.

### Введение

В примечании к делению индивидов на категории в работе [1] редколлегия журнала отметила относительность «любой классификации» и пригласила читателей высказаться на эту тему. Эта тема действительно актуальна для формирования онтологии любой предметной области (ПрО), ибо согласно Ж.Б. Ламарку «всякая наука начинается с классификации».

Понятие «классификация (классифицирование)» нашло отражение в работах по философии и формальной логике [2-5], причём определение этого понятия в отмеченных источниках претерпевало изменения. Приведём его варианты в хронологическом порядке.

- 1) *Классификация* – это «раскрытие внутренней необходимой связи между группами (классами, родами и т.д.), по которым распределены классифицируемые предметы» [2].
- 2) *Классификация* – это «система соподчинённых понятий (классов объектов) какой-либо области знания или деятельности человека, часто представляемая в виде различных по форме схем (таблиц) и используемая как средство для установления связей между этими понятиями или классами объектов, а также для ориентировки в многообразии понятий или соответствующих объектов» [3].
- 3) *Классификация* – это «общенаучное и общеметодологическое понятие, означающее такую форму систематизации знания, когда вся область изучаемых объектов представлена

в виде системы *классов*, или групп, по которым эти объекты распределены на основании их сходства в определённых свойствах» [4].

- 4) *Классификация* – это «многоступенчатое, разветвленное деление логического объёма понятия. Результатом классификации является система соподчиненных понятий (род, вид, подвид и т.д.)» [5].

Во всех определениях классификация представляется как *система взаимосвязанных классов*. Ключевыми словами системы классов являются «класс» и «связь». Первая часть четвёртого определения отражает процесс образования классов. Как процесс его правильнее назвать *классифицированием*, а его результат – *классификацией*, что соответствует паре аналогичных терминов: *систематизация* (процесс) и *система* (результат систематизации). Будем также называть классифицированием процесс отнесения объекта с известными свойствами к одному из классов.

Классификации во всех указанных источниках делятся на *естественные* и *искусственные*, причём в [2, 3] они различаются по степени *существенности* оснований деления. В [2] им ставятся в соответствие *содержательные* и *формальные* классификации. К первым отнесены классификации, отражающие объективные законы природы, а ко вторым – классификации, получаемые логическим делением понятий. Естественные классификации в [2, 3] в отличие от искусственных классификаций именуется также *научными* по причине отражения ими законов природы.

Во всех источниках в качестве примера естественной (научной) классификации приводится периодическая система химических элементов Д.И. Менделеева, а в качестве примера искусственной классификации – универсальные десятичные классификации (УДК). Наряду с названными классификациями в [2] приводится метод архетипа, применяемый в биологической систематике.

Помимо своей разнородности приведённые определения классификации и пояснения к ним вызывают ряд вопросов. Только в определении [5] процесс классификации (классифицирование) отделён от его результата. Определение [2] не отражает процесс образования классов. Они считаются заранее известными. Определение [5] сводит классифицирование только к делению логического объёма понятия. Определение [4] рассматривает классификацию как форму систематизации знания (значит, одну из форм?), но во всех источниках классификация отождествляется с систематизацией на примере периодической *системы* химических элементов Д.И. Менделеева, что вызывает вопрос о соотношении этих понятий.

Определения классификации и пояснения к ним не охватывают все её методы, что отражается на используемой терминологии. Существенность признаков, используемых для деления классификаций на естественные и искусственные, никак не определена.

Надёжным фундаментом для ответов на эти вопросы могут служить только математические модели. Тем более, что в [2] отмечается, что «формально-логические правила классифицирования разработаны слабо». С этой целью в качестве объектов исследования рассмотрим класс, классификацию и систематизацию с привлечением их математических моделей. Применение этих моделей позволит уточнить перечисленные понятия и уточнить связи между ними. Это поможет, в том числе, аргументированно ответить на вопрос, заданный редакцией журнала по поводу конкретного примера классификации.

## 1 Модели класса

Поскольку класс является ключевым понятием в классифицировании объектов, рассмотрим различные модели класса. Под классом понимается открытое множество [6]:

$$(1) \quad A = \{x \mid Pr(x)\} \text{ или } A = \{x : Pr(x)\}.$$

Принадлежность элемента  $x$  множеству  $A$  определяется относительно свойства  $Pr$ , присущего всем его элементам. Истинность одноместного предиката  $Pr(x)$  свидетельствует о том, что элемент  $x$  обладает заданным свойством  $Pr$ . Таким образом, предикат  $Pr(x)$  представляет собой *логическое правило* формирования множества  $A^1$ . Оно характеризует факт принадлежности элемента  $x$  множеству  $A$  и на этой основе называется *характеристической функцией* множества  $A$ :  $Pr: A \rightarrow \{0, 1\}$ . Значение 0 (ложь) означает, что  $x \notin A$ , а значение 1 (истина) означает, что  $x \in A$ .

*Открытым* множество  $A$  называют потому, что его мощность (количество элементов) не фиксировано. По способу формирования множества его можно назвать *разрешимым* или *рекурсивным*, поскольку его характеристическая функция позволяет определить, принадлежит ли анализируемый элемент  $x$  множеству  $A$  с заданным свойством элементов.

Логическое правило формирования множества  $A$  усложняется, если его элементы должны обладать *одновременно*  $n > 1$  свойствами:

$$(2) \quad A = \left\{ x : \bigwedge_{j=1}^n Pr_j(x) \right\}.$$

Фактически формула (2) предлагает выполнение  $n$  логических правил формирования множества  $A$ . Элементами класса «Отличник», например, являются только те ученики, которые получили оценку «Отлично» по *всем*  $n$  предметам. Класс «Троечник» предполагает наличие *хотя бы одной* удовлетворительной оценки:

$$(3) \quad A = \left\{ x : \bigvee_{j=1}^n Pr_j(x) \right\}.$$

Именованный класс представляет собой понятие. Мощность  $|A|$  множества  $A$  называется *объёмом* понятия, а совокупность признаков, по которым определяется объём – *содержанием* понятия [7].

Совокупность существенных признаков извлекается из *определения* понятия, выделяющего его из множества других понятий ПрО. Признаки являются *существенными* по отношению к той точке зрения, с которой определяется данное понятие. Например, по отношению к знанию, учёного можно охарактеризовать, как *знающего, познающего* и *порождающего* (новое знание). В силу фиксированного числа признаков содержание понятия «учёный» задаётся перечислительным способом:

$$Учёный(x) = \{ \text{Знающий}(x), \text{Познающий}(x), \text{Порождающий}(x) \}.$$

Человек  $x$ , удовлетворяющий этим требованиям, может считаться учёным. Объём понятия «учёный» не всегда удобно представлять перечислительным способом. Через разрешимое множество оно представляется подстановкой в формулу (2):

$$Учёный(x) = \{ x \mid \text{Знающий}(x) \wedge \text{Познающий}(x) \wedge \text{Порождающий}(x) \}.$$

Применение двоичных предикатов достаточно для выделения конкретного понятия из множества других понятий ПрО, но недостаточно для сопоставления порождённых из него видовых понятий. Например, как отличить молодого учёного от немолодого или зрелого от незрелого? Эта задача решается двухместным предикатом с целевым значением  $c_j$   $j$ -го свойства в роли базы сравнения. С учётом этого формула (2) принимает следующий вид:

$$(4) \quad A = \left\{ x : Pr_{>}(f_j(x), c_j) \right\}$$

Пусть учёный считается молодым в возрасте до сорока лет включительно. В соответствии с формулой (4) это условие представится следующим образом:

<sup>1</sup> Строго говоря, благодаря правилу отнесения к классу формула (1) представляет собой модель *формирования* класса.

**Молодой учёный**  $(x) = \{ x \mid Pr_{\leq}(\text{Возраст учёного } (x)), 40 \}$ .

Учёный, не удовлетворяющий этому правилу, считается немолодым. Здесь появляется искушение отождествить немолодого учёного со зрелым. Но зрелый учёный отличается от незрелого не возрастом, а своим вкладом в науку. А это уже другое основание деления. Человек, получивший учёную степень, может заняться другим видом деятельности и достигнуть зрелости в нём, а не в науке. Однако в общем случае зависимость зрелости учёного от его возраста исключать нельзя. Для её выявления используется статистика.

Формулу (4) так же, как и (1), можно обобщить на  $n$  свойств:

$$(5) \quad A = \left\{ x : \bigwedge_{j=1}^n Pr_{>}(f_j(x), c_j) \right\}.$$

Значение  $c_j$  на шкале признака, принятого за основание деления, является пороговым между двумя классами, означающими наличие и отсутствие  $j$ -го свойства. В теории принятия решений оно является основанием для *отбора* объектов в один из классов [8].

При наличии более одного класса шкала признака делится на  $m > 1$  диапазонов, а  $k$ -му классу,  $k = \overline{1, m}$ , ставится в соответствие диапазон  $[c_{kj, \min}, c_{kj, \max}]$  значений  $j$ -го признака,  $j = \overline{1, n}$ . Отсюда число диапазонов значений признака, участвующего в классификации объектов, должно соответствовать числу классов  $m$ . Граница между  $k$ -м и  $l$ -м смежными классами на шкале  $j$ -го признака называется *нечёткой*, если имеет место пересечение поставленных им в соответствие диапазонов значений:

$$[c_{kj, \min}, c_{kj, \max}] \cap [c_{lj, \min}, c_{lj, \max}] \neq \emptyset.$$

Принадлежность объекта  $x$  каждому из смежных классов в интервале  $[c_{lj, \min}, c_{kj, \max}]$  означает частичное обладание  $j$ -м свойством, что соответствует философии *частичной* истины [9]. Для её оценивания применяется функция принадлежности  $\mu_{jk}(x)$   $k$ -у классу по  $j$ -у свойству. Наличие общей шкалы у функций принадлежности разным классам позволяет классифицировать объект  $x$  на основании средневзвешенной аддитивной функции принадлежности  $k$ -у классу по  $n$  признакам [8]:

$$(6) \quad \mu_k(x) = \sum_{j=1}^n w_j \cdot \mu_{jk}(x), \quad k = \overline{1, m}.$$

В формуле (6)  $w_j$  означает вес (важность)  $j$ -го признака. Если все признаки имеют равную важность, их веса равны  $w_j = 1/n$ .

Поскольку в общем случае имеет место *частичная* принадлежность объекта  $x$   $k$ -му классу по  $n$  признакам ( $\mu_k(x) \in [0, 1]$ ),  $k = \overline{1, m}$ , класс  $h^*$ , которому объект  $x$  принадлежит в большей степени, определяется по максимальной величине функции принадлежности:

$$(7) \quad h^* = \arg(\max_k \mu_k(x)).$$

Формула (7) представляет правило отнесения объекта  $x$  к одному из классов. Поскольку функция принадлежности  $\mu_k(x)$  измеряется в шкале  $[0, 1]$ , для её отнесения к классу  $h^*$  может также использоваться пороговое значение в этой же шкале, например,  $\mu_k(x) > 0,5$ .

Правило (7) реализуется в модели выбора класса  $h^*$ :

$$(8) \quad h^* = \{x : \arg(\max_k \mu_k(x))\}$$

Рассмотренные модели классификации изложены в порядке их усложнения, связанного с получением большего объёма информации о свойствах классифицируемых объектов – от наличия  $j$ -го свойства до его количественного измерения, от одного до нескольких свойств.

Иными словами, переход к каждой следующей модели предполагает более глубокое изучение ПрО.

При геометрической трактовке объектов точками в  $n$ -мерном пространстве их свойств существуют два способа задания класса: его границами и компактностью размещения точек.

Если границы класса задаются гиперповерхностью  $F(x_1, \dots, x_n)$ , условие принадлежности объекта  $x$  классу  $A$  отражается нахождением его внутри границ класса:

$$(9) \quad A = \{x : Pr_{\leq}(F(x_1, \dots, x_n), 0)\}.$$

В случае кусочно-линейного представления гиперповерхности  $F(x_1, \dots, x_n)$  правило (9) преобразуется в конъюнкцию гиперплоскостей, представляющих границы класса  $A$ .

Условие компактности размещения точек в  $n$ -мерном пространстве свойств объекта представляется *таксономической* моделью класса:

$$(10) \quad A_k = \{x : Pr_{\leq}(d(x, c_k), d_{k, \max})\}.$$

В формуле (9)  $c_k$  – это точка, принятая за центр (представитель)  $k$ -го класса,  $d(x, c_k)$  – нормированное евклидово расстояние в  $n$ -мерном пространстве между объектом  $x$  и центром  $k$ -го класса,  $d_{k, \max}$  – максимальное расстояние до границы  $k$ -го класса. Модель (9) задаёт границы класса окружностью с центром  $c_k$  и радиусом  $d_{k, \max}$ .

В основу нахождения таксонов более сложной конфигурации (рисунок 1) кладётся гипотеза  $\lambda$ -компактности [10]. Она опирается на понятие  $\lambda$ -расстояния, которое учитывает нормированное расстояние  $d$  между элементами множества и локальную плотность  $\tau$  множества в окрестностях этих элементов.



Рисунок 1 – Два таксона сложной конфигурации в двумерном пространстве свойств

Вычисленные  $\lambda$ -расстояния всех отрезков, соединяющих точки множества  $A$ , отображаются из евклидова в  $\lambda$ -пространство<sup>2</sup>. С применением алгоритма Крускала или Прима [6] находится минимальная стягивающая сеть (остов) полного  $\lambda$ -графа. Исходя из гипотезы  $\lambda$ -компактности, утверждается, что точки, близкие в  $\lambda$ -пространстве, образуют  $\lambda$ -компактные сгустки, представимые таксонами.

## 2 Методы классифицирования объектов

Разнообразие моделей класса соответствует разнообразие методов классифицирования объектов [2]. Эти методы характеризуются следующими признаками:

- 1) способом создания классификации
  - a. дедуктивные;
  - b. индуктивные;

<sup>2</sup> В формуле (10)  $d$ -расстояния заменяются на  $\lambda$ -расстояния.

- 2) используемым математическим аппаратом
  - а. логические;
  - б. вычислительные;
- 3) наличием образца для классифицирования
  - а. методы, имеющие образец;
  - б. методы, не имеющие образца.

Как и любые методы проектирования, дедуктивный и индуктивный методы различаются направленностью создания классификации. Дедуктивный метод реализует *нисходящий* способ, а индуктивный метод – *восходящий* способ проектирования. Дедуктивный метод *конкретизирует* свойства классифицируемых объектов, а индуктивный метод *обобщает* их. Как и при решении других задач проектирования, эти методы, как правило, используются совместно, начиная с создания классификации от общего к частному или наоборот, и сменяя друг друга.

Как дедуктивный, так и индуктивный, методы могут использовать логические или вычислительные операции, либо и те, и другие, для разделения, либо обобщения классифицируемых объектов по их свойствам. Логическая операция оценивает наличие либо отсутствие у объекта рассматриваемого свойства. Вычислительные операции позволяют классифицировать объекты на основе количественного оценивания их свойств. Реализующие их методы называют *параметрическими*.

К методам, имеющим образец для создания классификации, относится разработанный в биологической систематике метод *архетипа*. Для каждого вида особей устанавливается типовой образец (архетип). При установлении места особи в систематике её признаки сравниваются с признаками архетипов различных видов, и делается вывод о принадлежности её к тому или иному виду.

Рассмотрим методы классифицирования объектов более подробно, уделив наибольшее внимание логическому методу деления объёма понятий, который имеет прямое отношение к вопросу, предложенному редколлегией журнала для обсуждения.

### 3 Деление понятий

В основе деления объёма понятий по свойствам лежит отношение *общее–частное*. Оно устанавливает родо-видовую связь между общим (родовым) и частным (видовым) понятиями за счёт расширения содержания общего понятия  $A$  *видовым отличием*  $B_1$  по основанию деления<sup>3</sup>  $B$  [11]:

$$(11) \quad B_1A = A \cup B_1$$

Получаемое таким образом понятие  $B_1A$  называется *видовым*. Поскольку все существенные признаки исходного понятия  $A$  при его объединении с видовым отличием  $B_1$  сохраняются, формула (11) отражает *наследование признаков* при порождении нового понятия.

Объём видового понятия сужается на основании формулы, двойственной формуле (11):

$$(12) \quad B_1A = A \cap B_1$$

Если основание деления  $B$  имеет два значения  $B_1$  и  $B_2$ , члены деления представляют собой несовместимые понятия, объёмы которых не пересекаются:

$$(13) \quad B_1A \cap B_2A = \emptyset,$$

а объединение объёмов даёт объём исходного понятия  $A$  (правило соразмерности деления):

$$(14) \quad B_1A \cup B_2A = A.$$

<sup>3</sup> В системном анализе основание деления называют системообразующим признаком.

Например, если за основание деления понятия *контроль* принять его место в технологическом процессе изготовления изделий (на его входе и выходе), то объёмы полученных видовых понятий *входной* и *выходной контроль* не совпадают (пересечение их является пустым). Данный пример характеризует дихотомическое деление понятия *контроль* по выбранному основанию деления, так как последнее не порождает других членов деления, кроме приведённых выше:  $VX K \cup BIX K = K$ . При этом видовые понятия  $VX K$  и  $BIX K$  наследуют все признаки родового понятия *контроль* ( $K$ ).

Объект, оцененный правилом (2), включается в класс  $A$ , если он обладает всеми  $n$  свойствами этого класса. В противном случае, при дихотомическом делении понятия он включается в противоположный класс  $\neg A$ .

Деление понятия по объёму не обязательно оказывается дихотомическим. Причиной тому является использование обобщённого основания деления. Если в приведённом выше примере место контроля не ограничивать входом и выходом технологического процесса, то результат деления понятия *контроль* будет включать три члена деления:  $VX K \cup ОП K \cup BIX K = K$ , где  $ОП K$  означает пооперационный контроль.

Если обобщённое основание деления «место контроля в технологическом процессе» разделить на «оконечный контроль технологического процесса» и «вид окончного контроля», то получим двухступенчатое дихотомическое деление. На первой ступени по ответу *Да* контроль разделится на *оконечный* и *промежуточный*, а по ответу *Нет* на второй ступени – на *входной* и *выходной*. Таким образом, расщепляя основание деления, можно  $N > 2$  членов деления свести к двум на каждом шаге за  $k = \lceil \log_2 N \rceil$  шагов<sup>4</sup> [8]. При этом каждое последующее основание деления зависит от предыдущего, конкретизируя его.

Приведённый пример деления понятия *контроль* отражает последовательный способ деления, когда второе основание деления зависит от первого. В том случае, когда основания деления одного понятия независимы, реализуется *параллельный* способ деления объёма понятия. Получаемые по независимым основаниям дихотомического деления видовые понятия называют *координатными* [11].

Координатные понятия позволяют формировать *межвидовые* и *смешанные* понятия. Межвидовое понятие создаётся объединением координатных понятий, полученных по *разным* основаниям деления. В качестве примера приведём понятие *входной производственный контроль*. Видовое понятие *производственный* получено относительно деления технологического процесса на *производственный* и *торговый*. Смешанное понятие создаётся совмещением противоположных видовых понятий. Актуальным примером является появление в современных анкетах графы *Другой* при определении пола лица, принимаемого на работу. Человек, характеризующийся этим понятием, совмещает свойства мужчины и женщины.

Выбор признаков, принимаемых за основания деления, зависит от цели классификации. Здесь уместно провести аналогию с деревом целей в модели принятия решений [8]. Состав оснований деления, как и частных целей в дереве целей, определяется поставленной целью. Соответствие общей и частных целей регламентируется закономерностью целостности системного анализа [12].

Процессы конкретизации и обобщения связаны с поиском подходящих терминов для обозначения новых понятий. Здесь следует руководствоваться правилами терминологии [13, 14]. Установление взаимосвязи между известными понятиями выполняется на основе их определений. Существенные признаки, извлечённые из определений сопоставляемых понятий, позволяют установить, находятся ли они в отношении «общее-частное». Для решения этой задачи в [15] был предложен язык определения понятий (ЯОП).

<sup>4</sup> Символы  $\lceil \cdot \rceil$  означают ближайшее большее целое число.

Относительность классификаций, отмеченная редколлегией журнала, объясняется неоднозначным подбором оснований деления понятий и их значений под заданную цель. Помимо относительности (условности) любой классификации следует оценивать её *правильность*. Она определяется выбором модели, соответствующей свойствам формируемого класса. Именно несоблюдение этого условия в [1] вызвало замечание редколлегии журнала.

При формировании понятия «трансцендентный человек» в работе [1] допущена ошибка в выборе модели класса, а именно, выбрана простейшая модель класса (1), не отвечающая поставленной задаче. Поясним эту мысль, отступив в терминах от высокого слога философии. Речь, на наш взгляд, идёт о целеустремлённом человеке (ЦУЧ). Ему недостаточно быть человеком, способным проектировать своё будущее (ПрЧ). Здесь нельзя не согласиться с мнением редколлегии журнала в том, что практически каждый человек, в той или иной мере, задумывается (философствует) о будущем. А поскольку измерить меру философствования каждого человека не представляется возможным, необходимо привлечь, как минимум, ещё один признак для выделения целеустремлённого человека из числа других людей. К таким признакам логично отнести настойчивость в достижении цели (настойчивый человек – НЧ). Тогда целеустремлённый человек характеризуется моделью (2), содержащей более одного признака:

$$\text{ЦУЧ}(x) = \{ x \mid \text{ПрЧ}(x) \wedge \text{НЧ}(x) \}.$$

Этой формулой снимается возражение редколлегии в том, что практически все люди философствуют (проектируют будущее). Мало поставить цель (создать проект будущего). Нужно ещё иметь волю для достижения поставленной цели. В России человека, не имеющего этого качества, называют прожектёром. Такие люди могут, например, много говорить об импортозамещении, но не предпринимать конкретных шагов в этом направлении. Не претендуя на точный смысл терминов, выбранных для приведённого примера, хочется надеяться на то, что он доступно иллюстрирует идею правильного подбора модели класса для формирования нового понятия.

Более точным определением классифицирования по объёму понятий, чем в [5], на наш взгляд, является «многократное параллельное или/и последовательное деление объёма понятий по выбранным основаниям деления на виды в отношении *общее–частное*».

Принцип формирования видовых понятий путём деления объёма исходного понятия соответствует нисходящему способу проектирования – от общего к частному. По этому способу метод создания классификации понятий является *дедуктивным*, а по использованию логического аппарата – *логическим*.

#### 4 Другие методы классифицирования объектов

Альтернативой дедуктивному логическому методу деления объёма понятий являются *индуктивные* методы создания классификаций, использующие модели класса (9) и (10) и статистические, либо экспериментальные данные.

Границы между сгустками точек (таксонами) в  $n$ -мерном пространстве признаков устанавливаются либо анализом расстояний между точками, либо обучением модели классифицирования на обучающей выборке.

Аналитический подход к формированию классов (таксонов) реализуется *индуктивными вычислительными* методами на основе вычисления расстояний между точками в  $n$ -мерном пространстве признаков [10]. Обучение модели классифицирования реализуется *индуктивными логическими* методами, формирующими границы класса на основе выборки с положительными и отрицательными примерами принадлежности классу. Для решения этой задачи успешно применяются нейросетевые модели.



В классификационной системе с установленными опытным или экспертным способом нечёткими границами классов выбор класса осуществляется *параметрическим* методом по максимальной обобщённой принадлежности классу на основе модели (8). Здесь выбор класса осуществляется не только по совокупности отношения признаков к целевым значениям как на основе модели (5), но и по их обобщённым оценкам, например, с применением формулы (6), использующей нечёткие оценки свойств.

В биологической систематике успешно применяется метод классифицирования особей на основе составленного классификатором архетипа. Архетип отражает характерные свойства группируемых объектов. Например, архетипом птицы является объект живой природы, обладающий туловищем, головой, крыльями, хвостовым оперением и ногами. В силу разнообразия этих частей организма в живой природе при сопоставлении особей они анализируются по отдельности, чему соответствует членение системы в отношении «целое-часть».

Определение принадлежности рассматриваемой особи заданному архетипу осуществляется методами мерономии и таксономии [16]. Методы мерономии основаны на умозрительном анализе классифицируемых объектов, а методы таксономии – на экспериментальном подходе. Таксономия и мерономия находятся в отношении двойственности. Чем больше классификатор знает о реальной структуре таксонов, тем он лучше может определить архетип. И, наоборот, более точное определение архетипов позволяет гораздо точнее определить таксономическую структуру.

Части архетипа принимаются за мероны, играющие роль оснований деления в классифицировании понятий. В приведённом примере за мероны птицы принимаются все выделенные части тела. Относительно только одного мерона, такого как крыло, к птицам можно отнести и летучую мышь. Поэтому для определения принадлежности архетипу особи сопоставляются *по всем* меронам. Если же их не хватает для выявления сходства / различия особей, добавляются дополнительные мероны, либо детализируются уже выбранные мероны.

При отвлечении архетипа от живой природы он может использоваться и для классифицирования объектов искусственной природы. Например, архетип птицы можно обобщить в архетип «летательные объект», если голову птицы представить как орган управления, туловище, как корпус, ноги, как орган контактирования с земной (или водной) поверхностью. К этому архетипу относится и такой искусственный объект, как самолёт, имеющий наряду с крыльями и хвостовым оперением орган управления, конкретизированный кабиной пилота, корпус – фюзеляжем и орган контактирования с поверхностью – шасси. Такая аналогия с птицей и послужила идеей для изобретателей первых летательных аппаратов.

Идея архетипа, как обобщённого представителя класса, успешно используется в таких задачах распознавания образов, как распознавание рукописных символов (букв, цифр и пр.) и многих других.

## 5 Характеристика классификаций

Классификация, как результат классифицирования объектов ПрО, характеризуется некоторой совокупностью свойств, систематизированных следующим образом:

- 1) относительно подхода к целостности системы
  - a. *внутренняя* система;
  - b. *внешняя* система;
- 2) относительно природы классифицируемых объектов
  - a. *естественная* классификация;
  - b. *искусственная* классификация;
- 3) по числу классов

- а. *известное;*
- б. *неизвестное;*
- 4) по форме представления
  - а. *табличная;*
  - б. *графовая;*
- 5) по структуре связей между классами
  - а. *древообразная;*
  - б. *сетевая;*
- б) по степени обоснованности
  - а. *научная;*
  - б. *эмпирическая.*

Внутренний подход состоит в том, что исходная целостность мыслится как нерасчленённая, а присущая ей организация позволяет выделять в ней естественные членения на компоненты, которые сами могут рассматриваться как системы [12]. Примером внутренней системы является система понятий ПрО, полученная многократным делением выбранного родового понятия.

При внешнем подходе к выявлению целостности она мыслится не как возможность естественного членения на компоненты, но как возможность естественного *объединения* в классы заранее имеющих объектов. Общность этих объектов состоит в наличии у них единой природы, позволяющей естественным образом сопоставлять между собой эти объекты и образовывать из них естественные классы. Примером внешней системы является биологическая систематика птиц.

Естественная классификация отражает объективные законы природы. Она может быть как внутренней (систематика химических элементов), так и внешней (систематика птиц). Искусственная классификация отражает системность объектов искусственной природы, создаваемых человеком. Искусственным происхождением объясняется и фиксированное (известное) число классов. Оно неизвестно при экспериментальном подходе к созданию классификации. Пример табличной классификации представлен в таблице 1.

Таблица 1 - Табличная классификации по двум основаниям деления

Ф1 \ Ф2	Внутренняя	Внешняя
Естественная	Е, Вн	Е, Вш
Искусственная	И, Вн	И, Вш

В качестве оснований деления для формирования двух фасет с координатными понятиями выбраны: природа классифицируемых объектов (Естественная (Е)/ Искусственная (И)) и отношение к целостности ПрО (Внутренняя (Вн) / Внешняя (Вш)). В клетках таблиц приведены межвидовые понятия, сформированные объединением координатных понятий из разных фасет.

Табличная форма классификации может быть преобразована в ранжированный (иерархический) граф с корневой вершиной *классификация*, если принять межвидовые понятия за терминальные вершины графа, а координатные понятия – за вершины промежуточного уровня графа. Поскольку полустепени захода терминальных вершин равны двум, граф имеет сетевую структуру связей между понятиями. В графе типа *дерево* все промежуточные и терминальные (висячие) вершины имеют полустепень захода, равную единице.

К научным следует относить научно обоснованные классификации, отражающие закономерности ПрО. Не обладающие этими свойствами классификации имеют эмпирический характер. Надо отметить, что любая классификация, имеющая опытное или теоретическое начало, подвергается многократной доработке для того, чтобы отразить закономерности

ПрО. Именно так обстояло дело и с периодической системой химических элементов, предложенной Д.И. Менделеевым и усовершенствованной его последователями.

## 6 Классификация как разновидность системы

Определённый парадокс заключается в различии терминов *классификация* и *система* при ссылке в качестве примера *классификации* на периодическую *систему* химических элементов Д.И. Менделеева, приводимой во всех источниках.

Для нахождения соотношения между классификацией и системой рассмотрим модель *системы*. В [16] система определена как «целостность, определяемая некоторой организующей общностью». Система как совокупность элементов, находящихся в определенных *отношениях* друг с другом и со средой [17], может быть представлена реляционной системой [6]:

$$(15) \quad M_{str} = \langle A, R \rangle.$$

Здесь  $A$  – множество компонентов (элементов) системы  $A = A_s \cup A_e$ ,

$A_s$  – множество компонентов (элементов) системы, выделенных из их общей совокупности  $A$ ,

$A_e$  – множество объектов, отнесённых к среде,

$R$  – множество отношений, отражающих внутренние  $R_s$  и внешние  $R_e$  связи системы:

$$R = R_s \cup R_e.$$

Согласно формуле (15) модель  $M_{str}$  отражает *многообразие* связей между элементами системы, отражая её структурный аспект. Многообразие связей является одним из признаков, характеризующих сложность системы.

Применительно к онтологии ПрО множество  $A$  представляет собой совокупность используемых в ней понятий. Множество отношений  $R$  охватывает как общесистемные связи между понятиями (общее-частное (наследования), часть-целое (партитивную), причинно-следственную (каузальную), функциональную, алфавитную и др.), так и связи, специфические для конкретной ПрО.

Поскольку для деления объёмов понятий используется отношение *общее-частное* между понятиями ПрО, классифицирование на основе одного отношения решает частную задачу систематизации понятий ПрО, и в этом смысле находится к ней в отношении «частное-общее» [18]. В этом смысле *классификационная* система (классификация) представляет собой хотя и важнейшую, но только часть онтологии ПрО.

Почему же таблица Менделеева носит название *системы*, а не классификации? Да потому, что на множестве свойств химических элементов устанавливается не только отношение *общее-частное*, но и отношение *порядка* на атомных массах (весах) элементов и периодах таблицы.

Номер столбца в таблице определяет квантовую структуру верхней оболочки атома, благодаря чему элементы этого столбца и обязаны *сходством* химических свойств, образуя относительно них класс. Номер строки в таблице определяет число электронных оболочек атома. В строке таблицы устанавливается разнонаправленное отношение *порядка* по атомному радиусу, энергии ионизации и электроотрицательности. Именно комплекс отношений на свойствах химических элементов обеспечил огромную роль периодической *системы* химических элементов Д.И. Менделеева в дальнейших исследованиях о строении материи.

## Заключение

Во всех цитированных работах по классификации отмечается их важная роль в познавательной и предсказательной деятельности. По мнению Ж.Б. Ламарка «всякая наука начина-

ется с классификации», а по мнению Д.И. Менделеева «наука начинается там, где начинают измерять». Кажущееся противоречие этих высказываний о начале науки исключается, если обобщить измерения на номинальную шкалу. Именно, в номинальной шкале, т.е. именами, и измеряются классифицируемые понятия. Это указывает на важнейшую роль терминологии в познавательной деятельности. В этом смысле плохую роль в понимании классификаций сыграла неоднозначность этого термина. Он трактуется и как процесс, и как результат процесса, т.е. порождает сам себя.

Предложенное в работе именование процесса создания классификации классифицированием позволило рассмотреть отдельно и систематизировать свойства методов классифицирования и самих классификаций.

Для правильного классифицирования объектов важно выбрать адекватную методу создания классификации модель формирования класса, что показано на конкретном примере. Предложенные в работе модели должны облегчить такой выбор.

Цитированные определения классификации отражают, по существу, только нисходящий подход к их созданию, не затрагивая восходящее проектирование классификаций на основе экспериментальных данных. Между тем, эти подходы равноправны и дополняют друг друга. Для восходящего проектирования классификаций предложены таксономические модели формирования класса.

Ключевой задачей в создании классификации является способ нахождения свойств классифицируемых объектов. Они могут выявляться умозрительным путём на основе умозаключений и доказательств их истинности, либо опытным путём на основе проведения экспериментов и их теоретических обобщений. По существу, эти подходы различаются лишь отправной точкой исследования – от теории или от практики. Использование архетипа оказалось эффективным средством выявления сходства / различия сложных объектов.

В цитированных источниках только естественные классификации рассматриваются как научные. Однако помимо законов природы существуют законы человеческого мышления, изучаемые философией, психологией и реализуемые системным анализом. Поэтому, на наш взгляд, к научным можно относить и те искусственные классификации, которые согласуются с этими законами. Непременными условиями научности следует также считать соответствие частных целей классификации её общей цели и подтверждение практикой. В отличие от научных эмпирические классификации не отвечают этим требованиям.

Классификации как системы, отражающие только отношение «общее-частное» на множестве свойств объектов, являются важным начальным этапом создания онтологии ПрО. Однако изучение функциональных, порядковых, причинно-следственных и других закономерностей ПрО требует применения соответствующих видов отношений. Применение дополнительных отношений и обеспечило успех периодической системы химических элементов Д.И. Менделеева в познании свойств материи.

## Благодарности

Работа поддержана Программой фундаментальных исследований Отделения нанотехнологий и информационных технологий РАН (проект № 0073-2015-0007) в рамках бюджетной темы № 0073–2014–0009.

Автор благодарен редколлегии журнала за должное внимание к указанной проблеме и приглашение к дискуссии (том 5, № 4, 2015, с. 456).

## СПИСОК ИСТОЧНИКОВ

- [1] **Резник, Ю.М.** К феноменологии возможных миров человека: экзистенциальная онтология проектирования / Ю.М. Резник // Онтология проектирования. – 2015. – Т. 5, №4(18). – С. 450-462. doi: 10.18287/2223-9537-2015-5-4-450-462.
- [2] Философская энциклопедия. Под редакцией Ф.В. Константинова. В 5-ти томах. – М.: Советская энциклопедия. 1960-1970.
- [3] Философский энциклопедический словарь. Гл. редакция: Л.Ф. Ильичёв, П.Н. Федосеев, С.М. Ковалёв, В.Г. Панов. – М.: Советская энциклопедия. 1983.
- [4] Новая философская энциклопедия. Под редакцией В. С. Стёпина. В 4-х томах. – М.: Мысль. 2001.
- [5] Философия: Энциклопедический словарь. Под редакцией А.А. Ивина. – М.: Гардарики. 2004.
- [6] **Микони, С.В.** Дискретная математика для бакалавра: множества, отношения, функции, графы. – СПб.: Лань, 2012. 192 с.
- [7] Статья «Понятие». Большая Советская Энциклопедия. 3-е изд. т. 25. – М.: Изд-во БСЭ, 1976. – С. 473-474.
- [8] **Микони, С.В.** Теория принятия управленческих решений: Учебное пособие. – СПб.: Лань, 2015. – 448 с.
- [9] **Zadeh, L.A.** Toward perception-based theory of probabilistic reasoning with imprecise probabilities // Сборник докладов Междунар. конф. по мягким вычислениям и измерениям SCM'2003. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2003, С. 69–75.
- [10] **Загоруйко, Н.Г.** Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во института математики, 1999. – 270 с.
- [11] **Микони, С.В.** Модели и базы знаний. Учебное пособие. – СПб.: ПГУПС, 2000. – 154 с.
- [12] **Микони, С.В., Ходаковский, В.А.** Основы системного анализа: Учебное пособие. – СПб.: ПГУПС, 2011. – 142 с.
- [13] **Лотте, Д.С.** Основы построения научно-технической терминологии. – М.: Изд-во АН СССР, 1961. – 157 с.
- [14] **Сложеникина, Ю. В.** Основы терминологии. – М.: Книжный дом «Либроком», 2013. 120 с.
- [15] **Микони С.В., Чахирева А.Л.** Формализованный язык для определения понятий // Научно-техническая информация. Сер.2. 1987. № 1. – С. 23-27.
- [16] **Шрейдер, Ю.А., Шаров А.А.** Системы и модели. – М.: Радио и связь, 1982. – 151 с.
- [17] **Берталанфи, Л.** Общая теория систем: Критический обзор // Исследования по общей теории систем. – М.: Прогресс, 1969. – С. 23–82.
- [18] **Микони, С.В.** Роль и место классификаций в системном анализе // Труды IV-й Международной конференции «Системный анализ и информационные технологии». – Челябинск: Изд-во ЧГУ, 2011, Том 2, – С.39-42.

## ABOUT THE CLASS, CLASSIFICATION AND SYSTEMATIZATION

S.V. Mikoni

*St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia*  
[smikoni@mail.ru](mailto:smikoni@mail.ru)

**Abstract**

We consider a few classification definitions and explanations in domestic sources. It detected the ambiguity of the classification term. He refers to both the process and its result. It is proposed to call these concepts are different terms. In order to eliminate a number of uncertainties in the interpretation of classification, its methods and systematization of knowledge offered to attract mathematical models of class and system. The models of the formation of classes, used by different methods of classifying objects are considered. Classifying methods properties and classifications obtained on the basis of classifying methods are systematized. Mathematical model of dividing the volume of the concepts is discussed. Wrong choice of the class model is shown on a concrete example. Inductive methods of classifying objects and archetype method are considered. The system model is used to distinguish the processes of classification and systematization. It is illustrated by the example of the Mendeleev periodic system of chemical elements. Classification concepts are refined.

**Key words:** *model, class, classification, periodic table, system, systematization, concept division, taxonomy.*

**Citation:** *Mikoni SV. About the class, classification and systematization. Ontology of designing. 2016; 6(1): 67-80. DOI: 10.18287/2223-9537-2016-6-1-67-80.*

## References

- [1] **Reznik YM.** To a phenomenology of possible human worlds. *Ontology of designing*. 2015; 5(4): 450-462. doi: 10.18287/2223-9537-2015-5-4-450-462. (In Russian).
- [2] *Filosofskaya Entsiklopediya Pod redaktsiyey F. V. Konstantinova.* [Philosophical Encyclopedia. Edited by F.V. Konstantinov] V 5-ti tomakh. –M.: Sovetskaya entsiklopediya. 1960-1970. (In Russian).
- [3] *Filosofskiy entsiklopedicheskiy slovar'. Gl. redaktsiya: L. F. Il'ichov, P. N. Fedoseyev, S. M. Kovalev, V. G. Panov* [Philosophical Encyclopedic Dictionary. Chief Editorial: L.F. Ilichev, P.N. Fedoseyev, S. Kovalev, V.G. Panov] – M.: Sovetskaya entsiklopediya. 1983. (In Russian).
- [4] *Novaya filosofskaya entsiklopediya. Pod redaktsiyey V. S. Stepina* [The New Encyclopedia of Philosophy. Edited by V.S. Stepin] V 4-kh tomakh. –M.: Mysl'. 2001. (In Russian).
- [5] *Filosofiya: Entsiklopedicheskiy slovar'. Pod redaktsiyey A.A. Ivina* [Philosophy Encyclopedic Dictionary. Edited by A.A. Ivin] –M.: Gardariki. 2004. (In Russian).
- [6] **Mikoni SV.** Diskretnaya matematika dlya bakalavra: mnozhestva, otnosheniya, funktsii, grafy. [Discrete Mathematics for Bachelor: sets, relations, functions, graphs]. SPb.: Lan, 2012. 192 p. (In Russian).
- [7] Stat'ya «Ponyatie» [Article "Concept"]. *Great Soviet Encyclopedia*. 3rd ed. m. 25 - M.: Publishing House of the TSB, 1976: 473-474. (In Russian).
- [8] **Mikoni SV.** Teoriya prinyatiya upravlencheskih reshenij: Uchebnoe posobie [Theory of administrative decision making: A Tutorial] –SPb.: Lan', 2015. – 448 p. (In Russian).
- [9] **Zadeh LA.** Toward perception-based theory of probabilistic reasoning with imprecise probabilities // *Cbornik dokladov Mezhdunarodnoj konferencii po mjagkim vychislenijam i izmerenijam SCM'2003.* – SPb.: Izd-vo SPb-GEhTU «LEhTI», 2003: 69–75.
- [10] **Zagoruyko NG.** Prikladnyye metody analiza dannykh i znaniy [Applied methods of data mining and knowledge analysis] –Novosibirsk: Izd-vo instituta matematiki, 1999. – 270 p. (In Russian).
- [11] **Mikoni SV.** Modeli i bazy znaniy. Uchebnoe posobie. [Model and Knowledge Bases. Tutorial]. – SPb.: PGUPS, 2000. –154 p. (In Russian).
- [12] **Mikoni SV, Khodakovskij VA.** Osnovy sistemnogo analiza. Uchebnoe posobie. [Bases of the system analysis. Tutorial] – SPb.: PGUPS, 2011. –142 p.
- [13] **Lotte DS.** Osnovy postroeniya nauchno-tehnicheskoy terminologii [Bases of creation of scientific and technical terminology] – M.: Izd-vo AN SSSR, 1961. – 157 p. (In Russian).
- [14] **Slozhenikina YV.** Fundamentals of terminology. M.: Book House "LIBROKOM", 2013. 120 p. (In Russian).
- [15] **Mikoni SV, Chahireva AL.** Formalizovannyi yazyk dlya opredeleniya ponyatii [The formalized language for definition of concepts] // *Nauchno-tehnicheskaya informatsiya. Ser. 2.* 1987. № 1: 23-27. (In Russian).
- [16] **Shrejlder JuA, Sharov AA.** Sistemy i modeli. [Systems and models]. –M.: Radio i svjaz', 1982. – 151 p. (In Russian).
- [17] **Bertalanffy L.** Obshhaja teoriya sistem: Kriticheskij obzor [General Systems Theory: A Critical Review // Research on general systems theory.] // *Issledovaniya po obshhejj teorii sistem.* – M.: Progress, 1969: 23–82. (In Russian).
- [18] **Mikoni SV.** Rol' i mesto klassifikacij v sistemnom analize [The role and place of classifications in the system analysis] // *Trudy IV-j Mezhdunarodnoj konferencii «Sistemnyj analiz i informacionnye tekhnologii»,* – Cheljabinsk: Izd-vo ChGU, 2011. Vol 2: 39-42. (In Russian).

## Сведения об авторах



**Микони Станислав Витальевич**, 1936 г. рождения. Окончил Ленинградский институт инженеров железнодорожного транспорта им. Образцова в 1963 г., д.т.н. (1992), профессор (1994), ведущий научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН. Член Российской ассоциации искусственного интеллекта. В списке публикаций более 270 работ, из них 2 монографии и 7 учебных пособий в области технической диагностики, дискретной математики, системного анализа, теории принятия решений, искусственного интеллекта.

**Mikoni Stanislav Vitalievich** (b. 1936) graduated from the Obraztsov Institute of Engineers of Railway Transport (Leningrad) in 1963, D. Sc. Eng. (1992). Professor (1994). He is Russian Association of Artificial Intelligence member. He is author and co-author of more than 270 publications in the field of technical diagnostic, discrete mathematic, system analyses, artificial

intelligence, decision making theory.