

Естественная классификация как законы природы. Естественная классификация действительности в процессе восприятия.

Витяев Е.Е.

Институт математики им. С.Л.Соболева СО РАН, vityaev@math.nsc.ru

1. Естественная классификация как закон природы. Понятие «естественной» классификации развивалось в СССР и в России в рамках классификационного движения. За пределами России понятие «естественной» классификации не разрабатывалось. В рамках этого направления был систематизирован опыт естествоиспытателей по созданию естественных классификаций. В.Ю. Забродин систематизировал критерии «естественности» классификации, которые в разное время выдвигались естествоиспытателями [1]. Приведем некоторые из этих критериев.

1. Смирнов Е.С. “Таксономическая проблема заключается в “индикации”: от бесконечно большого числа признаков нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки”;
2. Уэвель В. “Чем больше общих утверждений об объектах дает возможность сделать классификация, тем она естественней”;
3. Любичев А.А. “Наиболее совершенной системой является такая, где все признаки объекта определяются положением его в системе”.

Определим понятие «естественной» классификации, объясняющее перечисленные выше свойства [2,3]. Будем предполагать, что дана некоторая онтология Предметной Области (ПО), включающая в себя систему понятий и величин, а также систему законов (аналитических и синтетических), связывающих между собой понятия и величины. Будем предполагать, что синтетические законы получены некоторым KDD (Knowledge Discovery in Data Bases) методом.

Объекты ПО являются целостными образованиями, соединяющими в себе понятия, величины и законы ПО. Если на систему законов ПО смотреть как на систему аксиом, сформулированную в системе понятий ПО, которой должны удовлетворять объекты ПО, то каждый объект как реально существующий экземпляр, на котором действуют эти понятия и законы удовлетворяет этой системе аксиом. Поэтому объекты являются объектами-моделями онтологии. Совокупность всех таких объектов-моделей определяет семантику онтологии ПО.

Определим закономерную модель $M_a = \langle \Omega_a, Z_a \rangle$ объекта **a**, где Ω_a - значения всех понятий и величин, принимающих на объекте **a** определенные значения (истинности, числовые и т.д.); Z_a - законы применимые к объекту и связывающие на нем значения этих понятий и величин. Рассмотрим некоторый класс \mathcal{C} объектов. Закономерную модель $M_{\mathcal{C}} = \langle \Omega_{\mathcal{C}}, Z_{\mathcal{C}} \rangle$ класса \mathcal{C} определим как пересечение всех закономерных моделей объектов класса \mathcal{C} .

По критерию Е.С. Смирнова проблема индикации состоит в том, что бы найти такие признаки, из которых предсказываются все остальные признаки объекта. Такими признаками в закономерной модели класса $M_{\mathcal{C}}$ являются порождающие признаки: по значениям порождающих признаков $\langle x_{i1}^{j1}, \dots, x_{ik}^{jk} \rangle$, где x_i^j - значение признака x_i , и закономерностям из $Z_{\mathcal{C}}$ предсказываются все остальные значения признаков $\Omega_{\mathcal{C}}$ объектов класса. Понятно, что набор значений порождающих признаков определяется неоднозначно. По критерию Уэвель, чем больше общих утверждений об объектах можно сделать и чем богаче множество закономерностей $Z_{\mathcal{C}}$, тем она естественней. Богатое множество закономерностей позволяет меньшему множеству порождающих признаков предсказать все остальные признаки и делает классификацию более «естественной».

Предположим, что все классы $\{\mathcal{C}_{i \in I}\}$ нам известны и мы знаем все закономерные модели этих классов $M_{\mathcal{C}_i}$. Естествоиспытатели для построения систематики должны найти такие порождаю-

щие признаки $\langle x_{i1}, \dots, x_{ik} \rangle$, значения $\langle x_{i1}^{j1}, \dots, x_{ik}^{jk} \rangle$ которых являются порождающими для каждого класса систематики $\{C_{i \in I}\}$. В этом случае каждый класс однозначно определяется своим набором значений порождающих признаков и, что и определяет место объекта в системе. Задача естествоиспытателя состоит в том, что бы найти наиболее компактный и информативный набор порождающих признаков. Это и является совершенной по А.А.Любичеву системой.

Нами разработана программная система NatClass, которая реализует приведенные определения классификации и систематики. Она использовалась для решения задач биоинформатики [5].

2. Естественная классификация действительности в процессе восприятия. Процесс восприятия направляется и организуется «образом мира» [4]. Систематика создаваемая естествоиспытателями является для них «образом мира» ПО. Из систематики как «образа мира», вытекают следующие свойства восприятия естествоиспытателем объектов систематики:

- (1) Для построения систематики нужно выделить небольшое количество порождающих признаков, значения которых определяют класс к которому принадлежит объект, а также все остальные признаки объекта. Когда человек учится распознавать объекты классов он неизбежно должен найти эти признаки, даже если они будут варьироваться от класса к классу ввиду неоднозначности выбора порождающих признаков, в противном случае он не сможет по этим признакам предвосхитить (антиципацией) значения всех остальных признаков объекта и, значит, воспринять объект.
- (2) Когда из порождающих признаков предсказываются все остальные признаки, то это предсказание непрерывно во времени сравнивает «образ мира» с наличной стимуляцией и является процессом активного движения от «образ мира» к внешнему миру – непрерывным во времени и множественным по признакам процессом проверки адекватности «образ мира» на реальных стимулах из внешнего мира. Только, если все многочисленные предсказания будут совпадать с реальными стимулами непрерывно во времени, только тогда есть восприятие и, в этом случае, «образ мира» работает адекватно, в противном случае он разрушается [4].
- (3) Из определения систематики следует, что признаки не одинаково информативны по отношению к предсказанию остальных признаков - некоторые признаки «вторичны» и хорошо предсказываются по другим признакам, другие же являются «источниками» информации и, либо вообще не предсказываются другими признаками, либо только такими же признаками, которые являются «источниками» информации.
- (4) В естественной классификации самыми элементарными классами, содержащими минимальное количество признаков, являются классы формирующие некоторый вторичный язык описания объектов, который формируется автоматически.
- (5) Как показано во многих наших работах [5] наиболее эффективным методом обнаружения синтетических закономерностей является семантический вероятностный вывод. Нами предложена модель нейрона [6], основанная на семантическом вероятностном выводе, а так же модель работы мозга [6] реализующая все основные функции Теории Функциональных Систем П.К.Анохина. В этой же работе описана формальная модель схем восприятия.

1. Забродин В.Ю. О критериях естественной классификации. – НТИ, сер.2, 1981, №8.
2. Витяев Е.Е. Классификация как выделение групп объектов, удовлетворяющих разным множествам согласованных закономерностей. // Анализ разнотипных данных (Выч. сист. 99), Новосибирск, 1983, с. 44-50.
3. Витяев Е.Е., Костин В.С. Естественная классификация как закон природы // Интеллектуальные системы и методология. (Материалы научно-практического симпозиума "Интеллектуальная поддержка деятельности в сложных предметных областях"), вып.4, Новосибирск, 1992, с. 107-115.
4. С.Д.Смирнов Психология образа: проблема активности психического отражения. МГУ, М., 1985, с.232.
5. Scientific Discovery web site <http://www.math.nsc.ru/AP/ScientificDiscovery>
6. Витяев Е.Е. Формальная модель работы мозга, основанная на принципе предсказания // Модели Когнитивных Процессов. (Выч. системы, 164), Новосибирск, 1998, стр. 3-61.
http://math.nsc.ru/AP/ScientificDiscovery/pages/Related_papers.html#Principles%20of%20Brain%20Activity