УДК 004.9+005.94

А. Е. Стрижак

Институт телекоммуникаций и глобального информационного пространства
НАН Украины, Украина, г. Киев

ТАКСОНОМИЧЕСКИЕ ХАРАКТЕРИСТИКИ ОНТОЛОГИЧЕСКИХ СИСТЕМ

Описываются структурные свойства таксономий, как системных компонентов онтологических систем. Рассматриваются взаимоотношения между таксономиями и таксономическими категориями. Для этого рассматриваются множественные бинарные отношениям упорядоченности и их свойства. Определяются типы бинарных отношений. Для множественных структур вводится категория гиперотношения. На основании типов множественных отношений между концептами предметных областей определяются условия устойчивости таксономий онтологических систем.

ОНТОЛОГИЧЕСКАЯ СИСТЕМА, КОНЦЕПТЫ, ТАКСОНОМИЯ, УПОРЯДОЧЕННОСТЬ, БИНАРНЫЕ ОТНОШЕНИЯ, КАТЕГОРИЯ

Введение

При решении широкого спектра сложных прикладных задач часто приходится использовать распределенные в сети информационные массивы. В большинстве случаев эта информация имеет слабоструктурированную организацию, политематическую интерпретацию и не всегда высокий уровень достоверности. Также следует отметить проблему интероперабельности, которая понижает уровень релевантности и требует разработки методов и средств корректной семантической интерпретации каждой информационной единицы. Одним из направлений развития программных средств корректного использования распределенной информации является создание и применение онтологических систем [1, 2].

Одной из главных особенностей онтологических систем является формирование предметных классификаторов и их агрегированное использование при решении прикладных задач. В статье исследуются свойства таксономических структур, рассматриваются проблемы, связанные с их использованием при создании классификаторов онтологических систем, и организации в их среде корректного описания объектов тематических предметных областей.

1. Упорядоченность таксономических категорий

Одной из системных компонентов онтологической системы является таксономия [3], отображающая определенную иерархию взаимодействия концептов. При этом собственно иерархия задается при помощи бинарных отношений, определяющих характер взаимодействия между концептами онтологии. На основании указанных отношений, определяющих взаимодействие между концептами, реализуется процедура разбиения множества концептов онтологии на классы [4]. Для построения онтологических классификаторов конкретных предметных областей (ПрО) на основе сформированных классов, используются таксономические категории. Список указанных категорий определяется в процессе решения прикладной задачи на основе свойств, объединяющих концепты по тематическим признакам. Сложность множества взаимоотношений между ними можно увидеть на примере часто используемого в естествознании следующего множества таксономических категорий $\bar{T}$ [5]:

тип (phylum) – <имя сложного концепта>
подтип (subphylum) – <имя сложного концепта>
класс (classis) – <имя сложного концепта>
подкласс (subclassis) – <имя сложного концепта>
отряд (у растений — порядок) (ordo) – <имя сложного концепта>
подотряд (subordo) – <имя сложного концепта>
семейство (familia) –<имя сложного концепта>
подсемейство (subfamilia) –<имя сложного концепта>
род (genus) –<имя сложного концепта>
подрод (subgenus) – <имя сложного концепта>
вид (species) – <имя сложного концепта>
подвид (subspecies) – <имя сложного концепта>
разновидность (varietas) – <имя сложного концепта>
форма (forma) – <имя сложного концепта>.

Для формирования любой таксономической категории может быть использовано бинарное отношение – «*быть элементом категории*». Понятие *категория* здесь играет роль переменной и может принимать любое значение из выше приведенного списка. Также каждая категория является сложным концептом, которому, как определенному понятию предметной области, соответствует конкретное имя. Данное отношение позволяет организовывать иерархию концептов на уровне определения их принадлежности к конкретному тематическому классу. Здесь играет существенную роль понятие сложности концепта. Будем говорить, что концепт имеет простую структуру (простой концепт), если структурно он не может быть разложен на другие концепты. Тогда понятие сложного концепта может быть рекурсивно определено как структурно включающее непустое

множество простых концептов. Т. е. простой концепт связан со сложным отношением – «*быть элементом категории*».

Однако следует учитывать, что собственно концепты представлены только конкретными именами. Их активное использование в онтологической системе требует определения их смыслов, которые дают возможность интерпретировать конкретные концепты при выборе действий, применимых к ним в процессе решения прикладных задач. Такие смыслы могут быть заданы как бинарными отношениями между активными концептами, так и унарными свойствами самих концептов, характеризующих их прикладные признаки.

В общем виде бинарные отношения таксономических категорий могут быть представлены простой моделью бинарных отношений «*объект – объект*» [6]. Простую модель бинарных отношений можно обобщить до уровня множественных отношений типа «*группа объектов – объект*». Тогда отношения между таксономическими категориями можно рассматривать, как некие гиперотношения Gr вида – $YGrx$, где Y – множество всех возможных множеств концептов X таксономической категории $\tilde{T}$, а x один из концептов этого множества. Понятно, что если рассматривать множественное бинарное отношение типа «*группа объектов – группа объектов*», то его можно представить в виде некоего множества гиперотношений $\{Gr\}$, которое может быть представимо через множество бинарных отношений типа «*группа объектов – объект*». Поэтому достаточно корректно в дальнейшем рассматривать именно гиперотношения Gr, которые всегда можно доопределить до множества групповых бинарных взаимоотношений, приняв во внимание их декартовое пространство.

Указанная модель бинарных отношений задает отношение упорядоченности для всех концептов и их возможных множеств в структуре таксономических категорий онтологической системы. Такая упорядоченность может быть представлена следующими типами бинарных отношений, которые могут быть заданы над множеством всех концептов X онтологической системы O:

ацикличность – γ,

частичная упорядоченность – $\tilde{p}$,

линейная упорядоченность – p.

Отношение линейной упорядоченности p выводимо из отношения ацикличности:

$$x_i^j \gamma x_l^k \xrightarrow{\alpha} x_i^j p x_l^k \quad (1)$$

и также может быть выведено из отношения частичного порядка:

$$x_i^j \tilde{p} x_l^k \xrightarrow{\alpha} x_i^j p x_l^k \quad (2)$$

Бинарные отношения ацикличности и частичного порядка позволяют сформировать из общих концептов онтологической системы множества таксономических категорий $\tilde{T}$, на основании которых могут быть сформулированы предикативные выражения в виде логических функции и уравнений. Т. е. из концептов таксономических категорий мы можем строить утверждения, описывающие состояния онтологической системы, интерпретационные функции которой используются при решении прикладных задач. Бинарное отношение линейной упорядоченности позволяет нам определять последовательность действий над концептами и их множествами в среде онтологической системы.

Рассмотрение множественных отношений вида (1) – (2), образуемых между концептами, включая множества всех множеств концептов вида $\{X\}$ онтологической системы O и самими множествами, позволяет их типизировать следующим образом [6]. Согласно (2) бинарное отношение частичной упорядоченности может быть доопределено до бинарного отношения строгой частичной упорядоченности, которая позволяет задать над концептами онтологической системы бинарное отношение линейного порядка. Т. е. можно всегда для бинарного отношения частичного порядка вида (2) найти монотонно возрастающую последовательность из множества натуральных чисел N, определяющих n - й порядок бинарного взаимодействия концептов онтологической системы. Такая бинарная гиперупорядоченность может быть определена в виде множества бинарных упорядоченностей – P и представлена в виде следующего выражения – $P = \{p_n\}, n \in N$:

$$x \gamma y \Leftrightarrow \forall n \in N : x p_n y . \quad (3)$$

Рассматривая множества гиперотношений Gr, можно выделить следующие свойства, характеризующие их как бинарные [6]:

1. Гиперцикличность – если для Gr не существует гиперциклического множества $X \subseteq U$ такого, что –

$$\forall x \in X \exists Y \subseteq X : YGrx . \quad (4)$$

2. Иррефлексивность

$$YGrx \Rightarrow (Y \setminus \{x\})Grx . \quad (5)$$

3. Гипертранзитивность

$$YGrx, x \in X, XGrz \Rightarrow ((Y \cup X) \setminus \{x\})Grz . \quad (6)$$

4. Регулярность

$$YGrx, Y' \supseteq Y \Rightarrow Y'Grx . \quad (7)$$

Свойство множественной упорядоченности таксономических категорий позволяет формировать и/или выделять из множества концептов онтологической системы подмножество концептов, связанных между собой конкретным бинарным отношением упорядоченности. Таким образом, если множество гиперотношений Gr обладает набором свойств (3) – (7), обеспечивающих отображение различных активных состояний онтологической

системы O на основе выделения бинарного отношения частичной упорядоченности $\tilde{p}$ для множеств концептов $\{X\}$, то всегда можно определить таксономию и/или таксономическую категорию $T \in \tilde{T}$, для концептов которой могут быть сформулированы предикативные выражения, принимающие значения истинности. Т. е. мы всегда можем из множества концептов X выделить непустое подмножество концептов $\{x\}$, обладающих свойством бинарного отношения частичной упорядоченности $\tilde{p}$, которое и связывает их между собой.

$$YGrx \Rightarrow \exists y \in Y : y\tilde{p}x \qquad (8)$$

При этом частичная упорядоченность $\tilde{p}$ является элементом множества бинарных упорядоченностей – P, которое может быть расширено до гиперотношения Gr, и тем самым определять множественную упорядоченность над подмножествами концептов $\{x\}$ в виде множественного отношения – *«группа объектов – объект»*, имеющего свойства (4) – (8):

$$YGrx \Leftrightarrow n \in N \exists y \in Y : yp_n x : p_n \in P \subseteq Gr\ . \qquad (9)$$

Предикативные выражения, формулируемые на основе концептов таксономической категории с заданным множественным отношением упорядоченности вида (1) – (9), принимают только значения истинности. Это позволяет формировать на основе терминов-концептов онтологической системы лингвистические выражения, осмысленно отображающие состояния онтологической системы.

Так для множества таксономических категорий $\tilde{T} = \{$*тип (phylum); подтип (subphylum); класс (classis); подкласс (subclassis); отряд (у растений — порядок) (ordo); подотряд (subordo); семейство (familia); подсемейство (subfamilia); род (genus); подрод (subgenus); вид (species); подвид (subspecies); разновидность (varietas); форма (forma)*$\}$, множественное бинарное отношение частичной упорядоченности обеспечивает сохранение всех типов взаимодействия между ними. И также позволяет формировать все множества таксономий из концептов формируемой онтологической системы. Однако задав над указанными категориями отношение линейной упорядоченности, мы сужаем их перечень, т. к. ряд категорий, таких как, например: класс (classis); форма (forma); вид (species) могут занимать в выражении (1), представляющего бинарное отношение – «быть элементом категории», как левую, так и правую часть. Тем самым становятся невыполнимыми свойства отношения (5) и (7) – иррефлексивности и регулярности.

2. Структурирование на основе множественной упорядоченности

Любая онтологическая система формируется на основе структурированного представления предметной области ее применения. Обычно основу структурирования составляют классы объектов [4], на которые условно разбивается множество всех понятий, свойства которых определяют семантику предметной области. Непосредственно свойства концептов позволяют определить множество отношения между ними. Множественные бинарные отношения упорядоченности составляют тип гиперотношений, которые обладают свойствами (1) – (9) и обеспечивают формирование из классов концептов различные таксономические структуры.

Следует отметить, что таксономическая категория формируется на основе выделения определенного подмножества концептов имеющих общее свойство, которое их всех характеризует. Такое свойство может быть унарным для каждого концепта, однако при определении указанного класса как сложного концепта, такое свойство позволяет определить над всеми концептами класса множественное бинарное отношение *«быть элементом класса»*. При этом следует отметить, что унарное свойство, являющееся общим для множества концептов предметной области, может интерпретироваться как признак этих концептов или как критерий для выбора концептов данного класса.

Так при выборе концептов кандидатов для формирования определенной категории – *класс (classis),* первоначально определяется свойство-признак, на основании которого данная категория может быть представлена как сложный концепт. При этом включение в категорию-класс конкретного простого концепта или понятия с менее сложной структурой основывается на характеризующем его общем для всех концептов этой категории-класса унарном свойстве-критерии. Однако указанное общее свойство позволяет определить для данной категории бинарное отношение *«часть-целое»,* которое также задает частичную упорядоченность вида (2) над всеми концептами формируемой категории.

Таким образом, множественное бинарное отношение частичной упорядоченности, может быть задано на основе бинарного отношения *«часть-целое»,* которое может быть переписано в виде множественного отношения – «*группа объектов – объект*». Здесь фраза «*группа объектов*» определяет имя сложного концепта-класса, который является таксономической категорией для рассматриваемой предметной области. Выбор концептов для их включения в формируемую таксономическую категорию по общему семантическому свойству мы можем рассматривать как объединение одноэлементных множеств, где каждое такое множество определено конкретным концептом предметной области с выбранным общим для них свойством-критерием. Конструктивность применения процедуры выбора здесь проявляется возможностью точного определения необходимого элемента для

формирования таксономии. Предпочтительность критериального выбора в данном случае не зависит от метода ранжирования. Таксономия задается концептами с общим свойством, над которыми определяется множественное бинарное отношение частичной упорядоченности вида (2).

Таксономия может быть определена для любого сложного концепта, т. к. по умолчанию она может быть образована бинарным отношением линейной упорядоченности *«часть-целое»*. Отношение *«часть-целое»* может быть расширено до множественного отношения «*группа объектов – объект*» и далее до «*быть элементом класса*» и/или «*быть элементом категории*». Различие между понятиями *«класс»* и *«категория»* заключается в полноте отображения предметной области. Понятие *«категория»* достаточно полно отображает семантические возможности предметной области, в то же время понятие *«класс»* определяется просто отобранным множеством концептов с общими семантическими свойствами. Поэтому понятие *«таксономия»* не является эквивалентом понятия *«таксономическая категория»*.

Гиперотношение Gr, заданное над множественными бинарными отношениями упорядочивания, определяющих таксономию, позволяет представить предикативные выражения, формулируемые на основе ее концептов с заданным множественным отношением упорядоченности вида (1) – (9) в виде рекурсивного предиката [8]:

$$Pr(x_1,...,x_n)=1,$$
$$\text{если } \neg Pr(x_1,...,x_n) \wedge Pr(x_1,...,x_n)=0,$$
$$\text{если } Pr(x_1,...,x_n), \qquad (10)$$

где $x_i \in X; 1 \le i \le n$.

Фактически предикат вида (10) позволяет определить, какие концепты онтологической системы обладают выделенным свойством. Сформировав множества классов концептов онтологической системы при помощи предикативного выражения вида (10), мы получаем таксономию, над концептами которой задано множественное бинарное отношение *«часть–целое»*. Такая таксономия может иметь сложную структуру нисходящих иерархий. Каждая составная иерархия представляет собой класс концептов, имеющих как минимум одно общее семантическое свойство:

$$Pr(x_1,...,x_n)=0 \Rightarrow \exists T \subseteq \breve{T}: \forall x \in X \exists Y \subseteq X: T = YGrx \ . \quad (11)$$

Гиперотношения Gr вида (1) – (9) определяют взаимодействия между концептами каждой таксономии, выделенной из разных классов концептов предметной области. Соответственно рекурсивный предикат вида (10) позволяет определить новые виды таксономии онтологической системы. Особую роль в формировании новых таксономий играет, помимо рекурсивных предикатов, функция выбора. В терминах таксономических категорий она имеет следующий вид:

$$\forall T \left[\varnothing \notin T \Rightarrow \exists f: T \to \cup T, \forall T \in \breve{T}(f(T) \in \breve{T} \right] \quad (12)$$

Тогда характеристическая функция рекурсивного предиката (10), может определять условия применимости функции выбора при формировании множества таксономий онтологической системы. Если значение характеристической функции равно нулю, то мы выделили множество концептов, из которых невозможно сформировать истинные высказывания, и значит, они не обладают ни одним из бинарных гиперотношений Gr упорядоченности. В противном случае, характеристическая функция принимает значение 1, а из множества выбранных концептов формулируются истинные высказывания. И соответственно над ними могут быть определены гиперотношения Gr. Т. е. всегда можно найти непустое множество концептов онтологической системы, между которыми существует хотя бы один из типов отношения упорядоченности вида (1) – (7) и концепты, связанные таким отношением, могут образовывать истинные высказывания. Практически это условие задает индуктивность процесса формирования множеств концептов, между которыми устанавливаются гиперотношения Gr, и фактически конструируется таксономия и/или таксономическая категория. В формальном виде, на основании утверждений представленных в виде (10) – (12) это условие представимо в следующем виде:

$$\exists P \subseteq Gr, \exists p \in P, \forall x \in X \exists Y \subseteq X: YGrx \Rightarrow \exists T = YGrx \ . \quad (13)$$

Индуктивность выражения (13) задается последовательностью применения предикативного выражения вида (10) и (11) к множеству концептов онтологической системы O.

3. Свойства упорядоченностей таксономий

Формирование таксономических структур на основе процедуры выбора и множества бинарных гиперотношений упорядоченности позволяет определить ряд свойств, обеспечивающих конструктивность применения таксономий для реализации различных состояний онтологической системы. Особенностью нашего подхода является алгебраическое представление выбора соответствующего предпочтения при формировании множеств классов из концептов предметной области. Индуктивность предикативной формы представления процесса формирования таксономических структур (9) – (10), расширенная условием существования таксономии (12), позволяет определить ряд конструктивных свойств таксономий как упорядоченных бинарных структур. Указанная индуктивность накладывает ряд условий на входо-выходные состояния [9] онтологической системы при выборе концептов, на основе которых

формируется таксономия. Т. к. входо-выходные состояния конструируемой онтологической системы O определяют устойчивость формирования таксономических структур, то регулировать их взаимодействие на подсистемном уровне возможно при соблюдении ряда условий. Эти условия определяются типом отношений бинарной упорядоченности (1) – (7) и обеспечивают независимость значений характеристической функции рекурсивного предиката (10) от последовательности контекстов индуктивного выбора. В сложных системах, какими являются онтологии, выделяют следующие условия устойчивости [6, 7, 10]:

Условие наследования:

$$\exists T' \subseteq T \Rightarrow \exists f(T') \supseteq f(T) \cap T' . \qquad (14)$$

Условие независимости от не включенных в таксономию концептов:

$$\exists f(T') \subseteq T' \subseteq T \Rightarrow f(T') = f(T)' . \qquad (15)$$

Условие согласия:

$$\cap_n f(\mathrm{T}_n) \subseteq f(\cup_n T_n) . \qquad (16)$$

Свойства таксономических структур, представленные выражениями (14) – (16), позволяют в определенной степени проинтерпретировать интуитивные рассуждения при выборе конкретных свойств концептов, на основании которых реализуется выбор предпочтений по их включению в таксономию. Так условие наследования обеспечивает включение концептов, имеющих эквивалентные свойства-критерии. Условие независимости позволяет задать набор аксиом для онтологии, исключающие возникновение противоречий при выборе. Условие согласия позволяет обеспечить корректность разбиения множества концептов предметной области онтологической системы на классы. При этом согласно исследованиям [6, 7, 10] обеспечивается формирование класса функций, определяющих достаточно полно множественность бинарных отношений упорядоченности. Более того, можно утверждать, что условия (14) – (16) достаточно точно определяют саму систему выбора при условии, что свойства-критерии, на основе анализа которых реализуется формирование таксономии, заданы.

Представление свойств концептов таксономии может быть реализовано в виде предикативного выражения (11). Тогда применение условий (14) – (16) к концептам предметной области онтологии позволяет формировать таксономические структуры – классы и/или категории в виде деревьев признаков [11]. Это обеспечивает возможность представления таксономий, концепты которых взаимодействуют друг с другом на основе множества гиперотношений $\{Gr\}$, в виде графа без циклов. Т. е. справедливо следующее утверждение: любая иерархическая структура, для которой справедливы условия наследования, независимости и согласия, может быть представлена в виде бинарного дерева.

Дополнительно отметим следующее: условие согласия (16) определяет свойство пластичности множества таксономий $\breve{T}$ онтологической системы. Свойство пластичности позволяет утверждать, что для произвольной онтологической системы всегда существует непустое множество тавтологий, которое может быть расширено новейшей тавтологией. Таким образом, при использовании определенных концептов онтологии возможно выполнять подстановку множеств высказываний, которые задаются на множестве концептов онтологии на основании их тематических свойств и бинарных отношений упорядоченности вида (1) – (7). При этом все указанные концепты должны быть представлены в виде предикативного выражения (11). Это утверждение может быть перефразировано следующим образом: если определенная онтология O формируется на основе таксономии $\breve{T}$ множества концептов X, то всегда есть определенное непустое множество высказываний $\{Pr(x_1,...,x_n)\}$, которое образуются концептами x этой онтологии, при условии, что эти концепты связаны между собой бинарными отношениями вида:

$$r^m(x_i^j, x_l^k \mid x_i^j \in X_i; r^m \in R \neq \varnothing) . \qquad (17)$$

где $R = \prod_i^n X_i$.

Выводы

Применение онтологических систем в большей степени связано с решением сложных тематических прикладных задач. От качества структуризации предметной области при построении онтологии зависит эффективность ее применения. Поэтому вопросы связанные с упорядочиванием множественности концептов онтологий, используемых при решении тематических задач, определяют конструктивность онтологической системы. Определение множественных отношений бинарной упорядоченности над тематическими концептами онтологической системы позволяет достичь высокого уровня корректности при формировании таксономических категорий и тематических классификаторов. При этом обеспечивается множественное взаимодействие между таксономическими структурами онтологии. Корректное формирование множественных бинарных отношений упорядоченности позволяет определять, при формировании таксономических структур онтологии, свойства выбора предпочтений, на основе которых принимается решение по включению определенных концептов.

Список литературы: 1. *Guarino N.* The Ontological Level / N. Guarino, R. Casati, N. Smith, G. White // Philosophy and the Cognitive Sciences. – Vienna : Holder-Pichler-Tempsky, 1994. – p. 443-456. **2.** *Гаврилова Т. А.* Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф.

Хорошевский. – СПб.: Питер, 2001. – 384 с. **3.** Шаталкин А.И. Таксономия. Основания, принципы и правила / А.И. Шаталкин. – М. : Товарищество научных изданий КМК, 2012. – 600 с. **4.** *Буч Г.* Объектно-ориентированное проектирование с примерами применения: Пер. с англ. – М.: Конкорд, 1992. – 519 с. **5.** *Фолта Я., Новы Л.* История естествознания в датах: Хронологический обзор = Jaroslav Folta, Luboš Nový. Dejinyprirod nýchvied v dátach. Chronologický prehlad. – Bratislava, Smena, 1981 / Пер. со словац. к.х.н. З. Е. Гельмана; Предисл. и общ. ред. д.х.н. А. Н. Шамина. — М.: Прогресс, 1987. — 496 с. **6.** *Малишевский А. В.* Качественные модели в теории сложных систем / А. В. Малишевський. – М. : Наука. Физматлит, 1998. – 528 с. **7.** Миркин Б.Г. Проблема группового выбора. – М.: Наука, 1974. – 256 с. **8.** Справочная книга по математической логике: В 4-х частях / Под ред. Дж.Барвайса. – Ч.III. Теория рекурсии: Пер. с англ. – М.: Наука. Главная редакция физико-математической литературы, 1982. – 360 с. **9.** *Макаров И.М., Лохин В.М., Манько С.В., Романов М.П.* Искусственный интеллект и интеллектуальные системы управления. — М.: Наука, 2006. — 333 с. **10.** *Саати Т.* Принятие решений. Метод анализ иерархий: Пер. с англ. / Т, Саати. – М.: Радио и связь, 1989. — 316 с. **11.** *Микони С. Д.* Теория и практика рационального выбора : монография / С. Д. Микони. – М. : Маршрут, 2014. – 463 с.

Поступила в редколлегию 9.06.2014

УДК 004.9+005.94

Таксономічні характеристики онтологічних систем / О. Є. Стрижак // Біоніка інтелекту: наук.-техн. журнал. – 2014. – № 2 (83). – С. 24–29.

Таксономії є множинною категорією. Упорядкованість таких категорії забезпечує їх застосування при формуванні онтологічних систем. Тому характеристиками їх ефективного застосування є множинні бінарні упорядкованості. Це дозволяє досить коректно визначити тематичні класи концептів онтології та визначати різні стани онтологічної системи.

Бібліогр.: 11 найм.

UDC 004.9+005.94

Taxonomic Characteristics of Ontological Systems / O. Ye. Stryzhak // Bionics of Intelligence: Sci. Mag. – 2014. – № 2 (83). – P. 24–29.

Taxonomies are multiple categories. The ordering of these categories provides used in the formation of ontological systems. Because the characteristics of effective use are multiple binary ordering. This allows one to correctly identify thematic classes ontology concepts and identify different states ontological system.

Ref.: 11 items.