

УДК 004.08

ВОПРОСНО-ОТВЕТНАЯ СИСТЕМА, УЧИТЫВАЮЩАЯ СЕМАНТИКУ ПРЕДМЕТНОЙ ОБЛАСТИ

Попов С.В.

ООО «Научно-внедренческая фирма БП+», Москва, e-mail: s-v-popov@yandex.ru

Информация обладает собственной ценностью, иногда существенно превышающей ценность материальных объектов. Быстрый релевантный поиск дает возможность не только приобретать новые знания, но и обогащаться материально. Поэтому поисковые системы (поисковики) информации в интернете представлены в изобилии. Как правило, в них используется разновидность поиска по ключевым словам, что связано с недостаточной развитостью в настоящее время семантического анализа и автоматического логического вывода, позволяющих получать следствия из полученных данных. Основная сложность поиска необходимой информации в большом объеме данных состоит в том, что, как правило, при поиске определенного объекта имеется еще большая совокупность объектов, которые отличаются от искомого «немного», «чуть-чуть», «самую малость» и пр. Поэтому с содержательной точки зрения они отличаются мало, и человеку с его образным мышлением трудно подобрать существенные признаки, выделяющие нужный ему объект. В таком случае реально полезна подсказывающая система, указывающая те признаки, которые в данной предметной области представляются наиболее существенными. Тем самым человеку предоставляется возможность не вспомнить необходимый признак, что сопряжено с определенными ресурсными затратами, а выбрать из представленного набора. В работе предпринята попытка, используя оригинальный формализм, построить теорию вопросно-ответных систем, которая базируется на семантике предметной области. Эта семантика выражается в терминах совместности / несовместности, что представляется достаточно универсальным механизмом для описания предметных областей и построения систем поиска информации.

Ключевые слова: поиск информации, предметная область, несовместимость элементов, семантика, граф ортогональности, единичный n -мерный куб

QUESTION-ANSWERING SYSTEM THAT TAKES INTO ACCOUNT THE SEMANTICS OF THE SUBJECT AREA

Popov S.V.

LLC «Nauchno-vnedrencheskaya firma BP+», Moscow, e-mail: s-v-popov@yandex.ru

Information has its own value, sometimes greatly exceeding the value of material objects. Fast, relevant search provides the opportunity not only to acquire new knowledge, but also to enrich themselves financially. Therefore, search engines (search engines) information on the Internet abound. Typically, they use a kind of keyword search, which is associated with underdevelopment at the present time semantic analysis and automatic logical inference, allowing the investigation of the obtained data. The main difficulty of finding the necessary information in big data is that, as a rule, when searching for a particular object, there is still a large set of objects, which differ from the desired «slightly», «slightly», «very little» and so forth. Therefore, from a substantive point of view, they differ little, and the man with his imaginative thinking is hard to find significant signs that produce the desired object. In this case, it is really useful prompting system that indicates those features that are in a given subject area appear to be the most significant. Thus, the man given the opportunity to remember the essential sign, which is associated with a defined resource costs, and to choose from the presented set. In the work the attempt is made, using the original formalism, to construct a theory of question-answering systems, which is based on the semantics of the subject area. This semantics is expressed in terms of a compatibility / not compatibility, that is quite a universal mechanism for the description of the subject areas and building systems for information retrieval.

Keywords: information search, the subject area, the incompatibility of the elements, semantics, orthogonally graph, the unit n -dimensional cube

В связи с глобальной информатизацией общества информация (хотя этот термин и не имеет четкого определения) ныне обладает собственной ценностью, иногда существенно превышающей ценность материальных объектов. Быстрый релевантный поиск дает возможность не только приобретать новые знания, но и обогащаться материально. Поэтому поисковые системы (поисковики) информации в интернете представлены в изобилии [1, 2]. Как правило, в них используется разновидность поиска по ключевым словам, что связано с недостаточной развитостью в настоящее время семантического анализа и автомати-

ческого логического вывода, позволяющих получать следствия из полученных данных. Последний механизм может как сократить множество выдаваемых статей, полученных только по ключевым словам, так и ускорить поиск необходимой информации.

Основная сложность поиска необходимой информации в большом объеме данных состоит в том, что, как правило, при поиске определенного объекта имеется еще большая совокупность объектов, которые отличаются от искомого «немного», «чуть-чуть», «самую малость» и пр. Поэтому с содержательной точки зрения они отличаются мало, и человеку с его образным мышлением

трудно подобрать существенные признаки, выделяющие нужный ему объект. В таком случае должна существовать подсказывающая система, указывающая те признаки, которые в данной предметной области представляются наиболее существенными. Тем самым человеку предоставляется возможность не вспоминать необходимый признак, что сопряжено с определенными ресурсными затратами, а выбирать из представленного набора. Однако такая система подсказок предполагает значительную проработку предметной области, так как иначе подсказки будут носить нерелевантный характер.

Основными типами запросов на поиск информации служат запрос на нахождение конкретного объекта, существование которого не вызывает сомнения, и объекта, обладающего определенными свойствами, наличие которого в информационной системе под вопросом. Например, такими объектами могут быть вычислительное устройство определенной конструкции, для которого указываются его конструктивные особенности, или авторского свидетельства на интересующую тему, в котором дается описание устройства с неточно заданными параметрами. Если в первом случае имеется гарантия, что такое устройство имеется и при достаточном терпении мы его найдем, то вторая ситуация характеризуется неопределенностью, устройство может не присутствовать в информационной системе. В качестве примера также можно упомянуть построение объекта, обладающего определенными свойствами, когда известно, какие составные элементы приносят то или иное свойство. Вопросам семантики посвящено очень много работ, с основополагающими идеями можно ознакомиться в [3, 4].

В этой работе предпринята попытка, используя оригинальный формализм, построить теорию вопросно-ответных систем, которая базируется на семантике предметной области. Эта семантика выражается в терминах совместимости / несовместимости, что представляется достаточно универсальным механизмом для описания предметных областей и построения систем поиска информации по семантическим признакам. В настоящей работе не затрагиваются вопросы нечеткого описания объектов, которые подразумевают градации присутствия признаков наподобие: присутствует всегда, присутствует иногда, встречается редко и т.п. Этот вопрос будет поднят в следующих публикациях.

Содержательное описание предметной области

Под *предметной областью* будем понимать пару $OR = (Elements, Objects)$, где

Elements есть конечное множество *базисных элементов*, а *Objects* – совокупность *объектов*, которые суть подмножества множества базисных элементов, т.е. $Objects \subseteq 2^{Elements}$. Один базисный элемент может входить в несколько объектов в качестве составной части. Если два базисных элемента одновременно не принадлежат никакому объекту, то они называются *несовместимыми*, в противном случае – *совместимыми*. Содержательно несовместимость базисных элементов может интерпретироваться по-разному: это может быть несовместимость двух признаков при описании понятий, двух химических элементов при попытке создать новое соединение или двух разных команд в одном управляющем устройстве и т.п. Например, две команды ускорение и торможение, выполняемые одновременно, несовместимы. Два химических элемента, вступающие в реакцию, несовместимы при попытке создать устойчивое соединение, в которое они входят. Два признака молока – кислое и парное – также несовместимы. Два пользователя распределенной информационной системы несовместимы, если они не имеют доступа к одинаковой информации. Более подробно семантика предметной области, основанная на несовместимости, описана в [5].

Тем самым на множестве *Elements* задано бинарное коммутативное отношение $Ort \subseteq Elements \times Elements$ – описывающее отношение *несовместимости*.

Базисные элементы можно представлять обладающими определенными значениями ряда элементарных признаков. Признаки суть элементы конечного алфавита, значения каждого признака также принадлежат конечному множеству. Поэтому базисный элемент можно мыслить как вектор, каждый компонент которого указывает на наличие / отсутствие определенного признака. Тем самым признак однозначно соответствует компоненту вектора, разным компонентам соответствуют разные признаки и наоборот. Значения признаков, в свою очередь, бывают *совместимыми* и *несовместимыми*. Разные значения одного признака – *несовместимы*, они не встречаются в описании одного элемента. Это вполне естественно, так как в предметной области желательно максимально разделять базисные элементы, используя один признак. Предполагаем, что признаки, участвующие в описании базисных элементов, независимые, поэтому значения разных признаков совместимые.

В данном случае используется некоторая условность, так как несовместимые значения признаков при описании одной предметной области могут быть совместимыми

при описании другой предметной области. Но такая условность оправдана тем, что мы априорно фиксируем последнюю. Таким образом, исследование предметной области сводится к изучению базисных элементов, с определенным на них отношением совместности / несовместности.

Построим по множеству *Elements* и отношению *Ort* граф G_{Ort} несовместности: его узлы суть элементы множества *Elements*, и два узла b_1, b_2 смежные тогда и только тогда, когда выполняется отношение $Ort(b_1, b_2)$. Мы несколько усилим понятие объектов ПО: множество *Objects* – совокупность объектов, которые суть максимальные подмножества базисных элементов, то есть к условию $Objects \subseteq 2^{Elements}$ добавляется требование, чтобы каждый объект нельзя было расширить за счет добавления какого-либо элемента из *Elements*. Такое усиление объяснимо с содержательной точки зрения, если подмножество множества *Elements* расширяется, то такой объект можно считать лишь промежуточным, не несущей всей информацией. С другой стороны, совокупность максимальных подмножеств является полной характеристикой предметной области в терминах отношения несовместности.

Введенное определение предметной области будем называть *содержательным*, так как отношение несовместности может носить неформальный характер и определяться, исходя из содержательных соображений. Сейчас мы не рассматриваем разновидности определения понятия несовместности. Так несовместность может быть: всегда, в большинстве случаев, как правило, иногда, редко и т.п. Мы вводим предикат *Ort*, чтобы сделать изложение достаточно прозрачным.

Нетрудно увидеть, что каждому объекту $B \in Objects$ в графе G_{Ort} соответствует максимальный пустой подграф G_B и наоборот, каждому максимальному пустому подграфу графа G_{Ort} соответствует единственный объект из *Objects*. Таким образом, мы можем говорить не об объектах предметной области, а о пустых подграфах графа G_{Ort} .

Логическая семантика предметной области

Введем понятие *логической матрицы*, строки которых соответствуют переменным, а столбцы – дизъюнктам или конъюнктам в зависимости от того, представляет матрица КНФ или ДНФ логической функции. Элементы логической матрицы суть 0, 1, $_$ (обозначая соответственно отрицательную литеру, положительную или

отсутствие ее). Поэтому каждый столбец матрицы с n строками можно рассматривать как интервал в единичном n -мерном кубе E^n , а произвольная совокупность M^* столбцов определяет интервал $I^* \subseteq E^n$, который есть пересечение интервалов, определяемых этой совокупностью. Будем называть этот интервал I^* *определяемым* подматрицей M^* . Это пересечение может быть как пустым, так и не пустым, что определяется видом M^* .

В [5] показано, что по графу G_{Ort} строится логическая матрица с n строками и m столбцами, где $m = |Elements|$, и столбцы которой помечены разными элементами множества *Elements*, которая обладает следующими свойствами.

1. Два разных столбца b_1 и b_2 помечены интервалами σ_1 и σ_2 единичного n -мерного куба, пересечение которых $\sigma_1 \cap \sigma_2$ пусто тогда и только тогда, когда $Ort(b_1, b_2)$. Будем называть такие столбцы матрицы – *ортогональными*.

2. Каждый пустой подграф $G_0 \subseteq G_{Ort}$, $G_0 = \{b_1, b_2, \dots, b_q\}$ определяет подматрицу $M(b_1, b_2, \dots, b_q)$ такую, что определяемый ею интервал n -мерного единичного куба не пуст. Для непустых графов определяемые ими интервалы пусты.

Отметим, что при разных способах построения матрицы по графу G_{Ort} количество строк может варьироваться, тем самым поразному определяются интервалы единичного куба, соответствующие столбцам матрицы.

Получаем, что семантика несовместности элементов предметной области формально описывается в логических терминах, и отношению несовместности однозначно соответствует отношение ортогональности столбцов матрицы или пустоты пересечения интервалов.

Фильтрация столбцов логической матрицы

Введем операцию умножения векторов одинаковой длины в базисе 0, 1, $_$.

1. Значения 0 и 1 назовем *ортогональными*.

2. Если векторы в одном компоненте содержат ортогональные элементы (т.е. векторы ортогональные), то их произведение пусто.

3. В противном случае осуществляется компонентное умножение, на основе следующих базисных операций: $1 \times 1 = 1$, $1 \times _ = _ \times 1 = _$, $0 \times 0 = 0$, $_ \times _ = 0$, $_ \times 0 = 0$, $0 \times _ = _$.

Так как векторы в базисе 0, 1, $_$ определяют интервалы единичного куба, то очевидно, что операция умножения векторов однозначно определяет пересечение соответствующих интервалов.

Пусть M есть логическая матрица с n строками и m столбцами, столбцы поименованы b_1, b_2, \dots, b_m , и каждый из них определяет интервал $I(b_1), I(b_2), \dots, I(b_m)$ единичного n -мерного куба E^n . Введем операцию *фильтрация матрицы M произвольным вектором b длины n* . Ее результатом $\Phi(M, b)$ является матрица, которая получается из M заменой каждого ее столбца b_i на произведение $b \times b_i, i = 1, 2, \dots, m$. Пустые произведения в результирующую матрицу $\Phi(M, b)$ не включаются. Легко обобщить операцию фильтрации на множество $\mathbf{b} = \{b^1, b^2, \dots, b^q\}$ фильтрующих векторов: $\Phi(M, \mathbf{b}) = \Phi(\dots \Phi(\Phi(M, b^1), b^2, \dots, b^q)$.

Построение объекта предметной области по заданным характеристикам

Пусть из информационной системы надо выделить объект, обладающий заданными характеристиками, причем практически невозможно перечислить в точности все его свойства в силу того, что они составляют представительное множество и не всегда можно указать их все сразу. Однако в процессе уточнения в диалоге с вопросно-ответной системой такая задача вполне решается и в конечном итоге удастся получить искомый объект. Для определенности будем считать, что на начальном этапе имеется перечисление некоторого подмножества свойств, которые позволяют выделять подмножество объектов, обладающих этими свойствами, и задача вопросно-ответной системы состоит в том, чтобы подсказать, какие еще свойства необходимо указать, чтобы получить искомый ответ.

Сформулируем задачу в терминах предметной области, для которой задано отношение Ort на множестве элементов. Содержательно ее элементами могут трактоваться как конкретные сущности, так и признаки, позволяющие однозначно идентифицировать искомый объект. Тем самым для элементов предметной области задана содержательная семантика. Рассматриваемая предметная область характеризуется графом совместимости G_{Ort} , по которому в соответствии с описанной выше процедурой строится логическая матрица M . Теперь задача поиска искомого объекта состоит в построении максимального пустого подграфа, узлы которого суть элементы предметной области, имеющие содержательную интерпретацию и удовлетворяющие содержательным критериям принадлежности искомому объекту.

Полагаем, что исходно пользователь знает некоторые свойства искомого объекта. Пусть им соответствуют элементы из $Elements$, образующие множество $\mathbf{b} = \{b^1, b^2, \dots, b^q\}$,

в соответствии с которым вычисляется матрица $\Phi(M, \mathbf{b})$. В графе G_{Ort} элементы b^1, b^2, \dots, b^q выделяют пустой подграф, в общем случае он не максимальный, так как искомый объект еще не построен, и может расширяться разными способами. Однако, такое расширение происходит только за счет элементов, которым соответствуют столбцы матрицы $\Phi(M, \mathbf{b})$.

Поэтому задача поиска состоит в том, чтобы путем поставленных вопросно-ответной системой вопросов построить максимальный пустой подграф, который и описывает искомый объект. То есть обладает содержательными характеристиками, соответствующими элементам предметной области, которые суть узлы пустого графа. Поэтому диалог пользователя и вопросно-ответной системы сводится к формулировке последних вопросов, на которые пользователь отвечает ДА/НЕТ. Вопросы касаются включения в описание искомого объекта тех элементов или признаков, которые представлены столбцами в матрице $\Phi(M, \mathbf{b})$. Поэтому вопросно-ответная система может задать любой из таких вопросов. Выбор признака, по которому задается вопрос о включении его в искомый объект, может происходить в соответствии с определенной стратегией, однако здесь мы не останавливаемся на ее представлении. Считаем, что априорно все признаки, представленные матрицей $\Phi(M, \mathbf{b})$, равноправны и вопрос может касаться любого из них.

Тем самым поиск искомого объекта в заданной предметной области состоит в последовательном включении новых элементов и фильтрации матрицы с учетом вхождения этих элементов в объект.

Построение объекта с определенными свойствами

Рассмотрим еще одну задачу, которая также успешно решается в рамках вопросно-ответной системы, базирующейся на представлении предметной области в виде множества элементов и отношения ортогональности. Пусть имеется несколько свойств, образующих множество $S = \{S_1, S_2, \dots, S_r\}$, которыми должен обладать конструируемый объект. Каждое свойство S_i достигается включением в объект хотя бы одного элемента из определенного множества $b_p, i = 1, 2, \dots, r$. Как и прежде, обозначим множество всех элементов через $Elements$, на котором задано отношение несовместимости Ort . Напомним, что пара несовместимых элементов не может входить в один объект, они взаимоисключающие. Соответственно, имеется граф G_{Ort} и соответствующая ему логическая матрица M .

Построим по паре $(\mathbf{S}, Elements)$ двудольный граф G_{SE} , одна его доля состоит из свойств S_1, S_2, \dots, S_r , а вторая – из элементов множества $Elements$. Узлы S_i и b_j смежные в точности в том случае, если $b_j \in S_i$, $i = 1, 2, \dots, r, j = 1, 2, \dots, m$. Тогда искомым объектом будет существовать в случае, если в графе G_{SE} существует двудольный подграф, включающий все узлы множества \mathbf{S} и некоторое множество узлов E^* , $E^* \subseteq Elements$, обладающих следующим свойством: для каждого $i = 1, 2, \dots, r$, пересечение $S_i \cap E^*$ не пусто. Тем самым можно констатировать, что искомым объектом обладает свойством \mathbf{S} и конструктивно состоит из элементов E^* . Описание объекта в виде совокупности E^* элементов предполагает, что множество E^*

выделяет в графе G_{Ort} пустой подграф, узлами которого служат элементы E^* .

Список литературы

1. Ашманов Н., Иванов А. Оптимизация и продвижение сайтов в поисковых системах. – СПб.: Питер, 2008. – 400 с.
2. Константинова Н.С., Митрофан О.А. Онтологии, как системы хранения знаний. – СПб.: СПбГУ, 2006. – 125 с.
3. Карнап Р. Значение и необходимость. Исследование по семантике и модальной логике. – М.: ЛКИ, 2007. – 215 с.
4. Сокулер Е.А. Семантика и онтология: к интерпретации некоторых моментов концепций Р. Карнапа и Л. Витгенштейна // Труды научно-исследовательского семинара Логического центра Института философии РАН. – М., 1999. – С. 49–59.
5. Попов С.В. Синтез предметных областей. Решение одного класса переборных задач. – LAP LAMBERT Academic Publishing, 2017. – 96 с.