К ВОПРОСУ О ПОСТРОЕНИИ ЭФФЕКТИВНОГО ВЫЧИСЛИМОГО ПРЕДСТАВЛЕНИЯ ДЛЯ ПРОСТОГО АБДУКТИВНОГО ВЫВОДА*

Н.В. Головко

Проблема построения эффективно вычислимых функций, а точнее, вычислимых функций с эффективным представлением во многом является классической проблемой [1], особенно если мы говорим о применимости понятия эффективного вычисления к проблеме соотношения дедуктивного и эмпирического способов познания. В частности, одна из наиболее актуальных задач, стоящих, например, в области исследования проблем искусственного интеллекта, связана с поиском эффективного вычислимого представления, описывающего динамику смены состояний того, что мы называем искусственным интеллектом. Другими словами, эффективного вычислимого представления функции, которая играет роль функции объяснения в эпистемологии, логике, философии науки и, таким образом, находит свое отражение в исследованиях по искусственному интеллекту. На наш взгляд, одним из представлений такой функции может выступать абдуктивный вывод.

В современной литературе можно найти достаточно много работ, посвященных анализу логической структуры абдуктивного вывода, тем не менее проблема построения его эффективного вычислимого представления все еще является одной из наиболее сложных [2]. Сложность связана, во-первых, с особенностями самого абдуктивного вывода (немонотонный амплиативный гипотетический вывод, допускающий много интерпретаций), а во-вторых, с особенностями современных языков логического программирования (в широком смысле), не позволяющих в полной мере

* Работа выполнена при поддержке гранта Междисциплинарный интеграционный проект СО РАН – 2006 (проект № 1: «Вычислимость и рациональность: исследование сферы применимости тезиса Черча – Тьюринга и понятия эффективного вычисления к проблеме соотношения дедуктивного и эмпирического способов познания когнитивных и физических процессов»).

формализовать его. Цель данной работы – продемонстрировать возможность построения эффективно вычислимого представления для простого (без дополнительной проверки на связность или объяснительную силу) абдуктивного вывода с использованием семантических таблиц Бета – Хинтикка. Задача, которая стоит перед нами, – построить эффективное представление (алгоритм) для простого абдуктивного вывода, адекватное задачам использования его в вычислительных процедурах, например в моделях искусственного интеллекта.

Что такое абдуктивный вывод?

Один из ведущих представителей современной аналитической философии Я. Хинтикка в своей работе «Что такое абдукция? Основная проблема современной эпистемологии» [3] указывает на то, что одной из наиболее интригующих особенностей абдуктивного вывода является способность приводить к объяснению имеющихся фактов, а также получать новое знание не только в контексте здравого смысла, но и в более строгом контексте теоретических научных построений. Поскольку термин «абдукция» используется в литературе для обозначения процесса построения объяснения в достаточно разнообразных ситуациях, приведем примеры абдуктивного рассуждения, ограничивающие предмет нашего исследования [4]:

- *на уровне здравого смысла*. Всем известно, что трава во дворе дома будет мокрой после дождя или после работы дождевальной системы. Проснувшись утром, мы обнаружили, что трава мокрая. Соответственно, мы предлагаем гипотезу объяснения: ночью шел дождь либо работала дождевальная система.
- *на уровне научного рассуждения*. Иоганн Кеплер пришел к заключению, что орбиты планет представляют собой эллипсы, на основании наблюдения того, что значения долготы орбиты Марса не соответствуют предполагаемой круговой орбите. Естественно, прежде чем предположить, что орбиты планет описываются эллиптически, он проанализировал несколько других гипотез, не прошедших последующих проверок. Более того, Кеплер был вынужден сделать ряд достаточно сильных допущений, прежде чем открыл свой знаменитый закон «заметания площадей». Будучи сторонником гелиоцентрической гипотезы, он предположил, что именно Солнце является причиной наблюдаемого движения планет, а

потом, опираясь на наблюдения Марса, обобщил свои представления на другие планеты, предположив «одинаковость» состояний всех планет [5].

В данной работе предметом анализа будет ситуация, аналогичная первому примеру, – объяснение простых фактов. Типология абдуктивных рассуждений достаточно разнообразна (на что косвенно указывает второй пример) – от приведенных примеров до построения объяснения причинной взаимосвязи между фактами или указания на противоречие либо «обнаружение неисправности». Более того, типология абдукции как вывода к объясняющим гипотезам включает в себя постановку вопросов такого рода: является ли абдукция результатом построения объяснения или процессом? является ли она истолкованием имеющихся посылок или выбором из имеющихся данных? в каком смысле можно рассматривать абдукцию как наилучшее объяснение? и др. [6]. На наш взгляд, означенный случай «простой абдукции», связывающий объяснение с «простыми» фактами, является более подходящим для целей данной работы. В последующих работах мы постараемся, с одной стороны, перейти к анализу более сложных случаев абдукции, а с другой – более четко проинтерпретировать понятие «абдуктивное объяснение» в том смысле, в каком оно используется в работах по искусственному интеллекту.

Что касается общего логического представления абдуктивного вывода, то в данной работе мы будем придерживаться следующего:

$$\theta,\alpha \Rightarrow \varphi,$$

где θ – исходная базовая теория;φ – «новый» факт, требующий объяснения; α – результат абдукции (решение). Вместе с тем мы потребуем выполнения ряда условий, отражающих дополнительные характеристики и ограничения. Во-первых, выводящий параметр должен устанавливать подходящую логическую взаимосвязь между базовой теорией, эксплanансом и экспланандумом в удобной форме, допускающей формализацию, например для переноса модели на ЭВМ. Во-вторых, необходимо заранее оговаривать, какой тип абдуктивного вывода рассматривается: например, φ – новое, неожиданно обнаруженное явление или явление, вступающее в противоречие с θ, и т.д. В-третьих, следует заранее оговаривать тип предполагаемого решения: факт, правило, закон и др.

Мы будем исходить из того, что в самом общем случае абдукция представляет собой процесс поиска объяснения, результатом которого является объяснение, имеющее конкретную «структуру вывода». В частности, при выводе объяснения мы обязаны делать выбор между всеми возможными объяснениями и «наилучшими» из них, что находит отражение в логической форме объяснения (атомарное, конъюнктивное, дизъюнктивное), а также в конкретном процессе его конструирования. Поскольку, как правило, эффективное вычислимое представление абдукции используется в рамках построения моделей искусственного интеллекта, мы будем рассматривать абдуктивный вывод именно так, как принято в данной области исследований, – как *обратную дедукцию, снабженную дополнительными условиями* [7]. Например, для заданных и искомое действительно будет объяснением, если будут выполнены условия:

1) θ∪α‘φ’;

2) α удовлетворяет условию связности с θ;

3) α является «минимальным» (наилучшим, удовлетворяющим дополнительным условиям и др.) объяснением;

4) α имеет четко заданную синтаксическую форму, например в виде атомарной формулы или их конъюнкции и т.д.

Не вдаваясь в детали, определим абдукцию как «усиленную» форму классического вывода, укладывающегося как в схему рассуждений тех, кто стремится использовать его вычислимое представление в моделях искусственного интеллекта, так и в стандартную гемпелевскую схему построения объяснения [8].

Рассмотрим приведенный нами пример, связывающий объяснение простых фактов (дождь – мокрая трава), в рамках классической пропозициональной логики. Пусть θ: $r \rightarrow w$, $s \rightarrow w$ и $\varphi = w$. Прежде чем принять формулу α в качестве объяснения для φ, мы должны потребовать выполнения *условия выводимости* θ, α‘φ. Очевидно, что этому условию будет удовлетворять большое число формул. Вместе с уже упомянутыми возможными объяснениями (r – дождь, s – дождевальная система) мы можем рассматривать, например, их конъюнкции с другими формулами, которые тоже будут объяснениями ($r \wedge \neg w$). Кроме того, ничто не мешает нам привлекать к объяснению новые факты или правила, например предположить, что дети баловались с поливальным шлангом, в результате чего вся трава перед домом оказалась мокрой. Поэтому, например, мы можем выписать следующие формулы, удовлетворяющие условию выводимости:

$r, s, r \wedge s, r \wedge z, r \wedge \neg w, s \wedge \neg w, w, c, c \rightarrow w, \theta \rightarrow w$ и др.

Некоторые из этих «объяснений» будут естественным образом отсеяны. Однако не лишним будет провести дополнительный отбор объяснений, потребовав от них удовлетворения *условию связности* $\theta\neg\alpha$:

$$r,s,r\wedge s,r\wedge z,r\wedge\neg w,w,c,c\rightarrow w,\theta\rightarrow w.$$

Далее, логично было бы ожидать, что объяснение α является необходимым, т.е. $\theta–\varphi$ – *условие необходимости*). На первый взгляд, само по себе это условие не может отразиться на содержании списка предполагаемых объяснений, поскольку α вообще не участвует в аргументе (за исключением вырожденныхслучаев $\alpha=\varphi$). В то же самое время данное условие может быть полезным как для того, чтобы избежать так называемого «внешнего объяснения» (привлечения к объяснению дополнительной информации: вмешательство детей или соседа), так и для того, чтобы подчеркнуть достаточность α, – самого α не должно быть достаточно, чтобы объяснить φ, т.е. должно выполняться условие $\alpha–\varphi$ – *условие достаточности*. Соответственно придем к следующему списку:

$$r,s,r\wedge s,r\wedge z,\theta\rightarrow w.$$

Отметим, что в данном списке все еще остаются объяснения, которые, на наш взгляд, не удовлетворяют достаточной степени общности предполагаемого объяснения, – это $r\wedge z$ и $\theta\rightarrow w$. В связи с этим можно предположить наличие более тонких критериев достаточности объяснения, или объяснительной силы [9].

Как уже отмечалось выше, предметом настоящего исследования является именно простой абдуктивный вывод, не связанный с дополнительными проверками на связность, минимальность, объяснительную силу и прочие свойства получаемых решений. На наш взгляд, приведенные ниже построения без труда получают обобщение на любое из возможных ограничений, налагаемых на предполагаемое объяснение. Однако прежде чем перейти к анализу эффективного представления простого абдуктивного вывода и построению соответствующего ему алгоритма, дадим краткую характеристику методу семантических таблиц, который, по нашему мнению, является одним из наиболее адекватных инструментов для решения поставленной перед нами задачи.

Метод семантических таблиц и построение эффективного представления

Метод семантических таблиц как метод поиска опровержения формул логики исчисления предикатов был предложен в середине 1950-х годов независимо Э. Бетом [10] и Я. Хинтиккой [11]. Фактически метод семантических таблиц сам по себе является мощным средством построения эффективного представления, а основная проблема заключается в том, чтобы подобрать соответствующую интерпретацию семантической таблице и операциям над ней. В данной работе мы будем использовать более общее представление о семантических таблицах, как оно сформулировано у Р. Смальяна или М. Фиттинга [12].

Основная идея такова. Для того чтобы проверить, следует ли формула φ из множества посылок θ, необходимо убедиться в том, что для предложения $\theta\cup\{\neg\varphi\}$ можно построить *открытую* таблицу. Таблица представляет собой бинарное дерево, которое строится для начального множества предложения с использованием соответствующих правил для каждой из логических связок (встречающихся в исходном наборе предложений), что и задает структуру ветвей дерева. Таблица является *закрытой* в том случае, если исходное множество предложений невыполнимо, т.е. выполняется θ‘φ. Другими словами, если результирующая таблица имеет открытые ветви, то формула φ не является следствием θ; таблица является закрытой, если каждая ветвь содержит атомарную формулу и ее отрицание.

Правила конструирования таблиц достаточно просты, поскольку без труда допускают переинтерпретацию всех основных связок: двойного отрицания, конъюнкции, дизъюнкции, импликации и их отрицаний [13]. Приведем пример простейшей таблицы. Пусть $\theta = \{r\rightarrow w\}$и $\varphi = w$. Соответственно наша цель – убедиться в том, что θ‘φ, т.е. построить таблицу $T(\theta\cup\{\neg\varphi\})$ (рис. 1).

Результирующая таблица является открытой, поэтому θ–φ. Открытая ветвь указывает нам на «контрпример», т.е. на случай, когда θ истинно, в то время как φ ложно (r,w ложны). В самом общем случае построение таблицы продолжается до тех пор, пока формулы в вершинах таблицы (графа) не «превратятся» в атомарные высказывания или их отрицания. Более того, такой принцип построения является гарантией того, что в вершинах таблицы могут встречаться только подформулы θ или $\neg\varphi$.

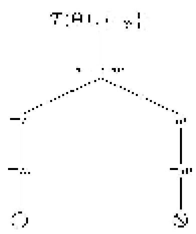

Рис. 1. Семантическая таблица $T(\theta \cup \{\neg\varphi\}$, $\theta = \{r \rightarrow w\}$, $\varphi = w$. Здесь и в последующих таблицах знак о обозначает, что ветвь открыта, $\otimes$ – что ветвь закрыта

Если $T(\theta)$ имеет открытые ветви, то θ непротиворечиво и совместно, – каждая открытая ветвь соответствует возможности проверки (верификации) модели. Если у $T(\theta)$ все ветви закрыты, то множество формул θ несовместно. В целом можно привести следующие несколько правил построения таблиц.

1. Закрытая ветвь. Ветвь является закрытой, если содержит некоторую формулу или ее отрицание.

2. Открытая ветвь. Ветвь является открытой, если она не закрытая.

3. Доказательство формулы X. Доказательством формулы X является закрытая таблица для $–X$.

4. Доказательство θ‘φ. Доказать, что θ‘φ, – значит построить закрытую таблицу для $\theta \cup \{\neg\varphi\}$.

Основная задача, которая стоит перед нами, – это построить эффективное представление для абдуктивного вывода, используя метод семантических таблиц. На наш взгляд, абдуктивный вывод можно представить как своеобразное «расширение» семантической таблицы.

Рассмотрим более сложный пример. Пусть $\theta = \{b, c \rightarrow r, r \rightarrow w, s \rightarrow w\}$. Построим таблицу $T(\theta)$, для того чтобы наметить возможные пути верификации самого множества формул θ (рис. 2). В результате мы получили открытую таблицу. Следовательно, исходное множество формул является непротиворечивым, а каждая открытая ветвь соответствует возможности проверки модели. Например, вторая слева ветвь показывает, что можно задать модель для θ, принимая c и r ложными, а b и w – истинными (в данном случае можно говорить о существовании двух моделей в зависимости от истинностного значения s). Построение таблицы дает возможность провести своеобразную оценку формул, включенных в таблицу, в соответствии с «поведением» ветвей.

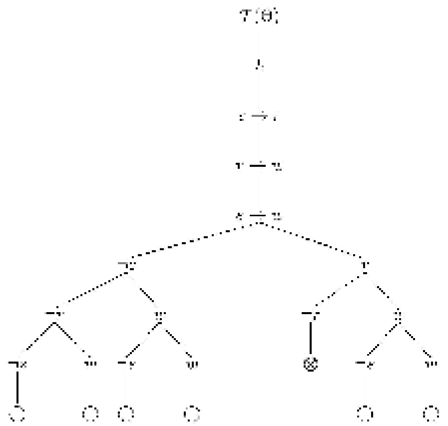

Рис. 2. Семантическая таблица $T(\theta)$, $\theta = \{b, c\to r, r\to w, s\to w\}$

Отметим, что как только мы добавим формулу (и тем самым расширим таблицу), ряд возможных моделей будут признаны «неподходящими», – соответствующие им ветви окажутся закрытыми. Пусть $\varphi = w$. Построим таблицу $T\,\theta \cup \{\neg\varphi\}$ и тем самым проведем «проверку» θ‘w (рис. 3). Несмотря на то что исходное множество формул остается непротиворечивым, все ветви, за исключением одной, являются закрытыми. В частности, большая часть моделей, адекватных $T(\theta)$, становятся неадекватными при ложном w. Оставшаяся открытая ветвь указывает на модель, удовлетворяющую множеству формул $\theta \cup \{\neg\varphi\}$ (c, r, s, w ложны, b истинно), которая указывает на то, что θ – w.

Пожалуй, наиболее важное для нас следствие из приведенных примеров, состоит в следующем. В том случае, если φ не является подходящим следствием θ, мы очевидным образом получаем все случаи, для которых верно, что θ – φ, и которые графически представляются открытыми ветвями. Таблица на рис. 3 демонстрирует возможность того, что «контрпримеры» можно «подкорректировать», изменив исходное множество формул, например добавив посылки. Таким образом, мы можем вполне легально в конце концов добиться того, чтобы стало следствием некоторого «расширенного представления». На наш взгляд, данное обстоятельство удачно соответствует тому представлению об абдукции, о котором говорилось выше. Мы можем представить абдукцию как *процесс «расширения» таблицы с помощью подходящих формул, которое приведет к тому, что открытые ветви станут закрытыми*.

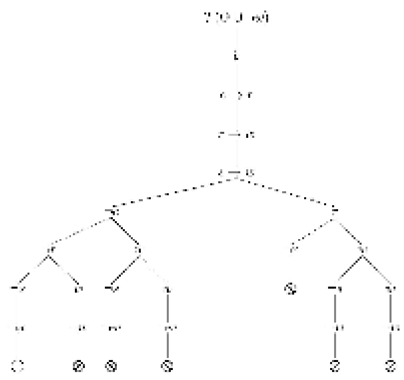

Рис. 3. Семантическая таблица $T\theta \cup \{\neg\varphi\}$, q = $\{b, c \to r, r \to w, s \to w\}$

В приведенном выше примере (см. рис. 3) оставшаяся открытой ветвь имеет вид

Таким образом, уже можно представить себе ряд формул, добавление которых в таблицу закроет эту ветвь (и тем самым всю таблицу), например:

$\neg b, c, r, s, w, c \wedge r, r \wedge w, s \wedge w, s \wedge \neg w, c \vee w.$

Некоторые из приведенных формул можно рассматривать как абдуктивное объяснение для посылок θ.

Отметим, что добавление формулы в таблицу может изменить таблицу достаточно сильно. Многое зависит от формы добавляемой формулы, а также от формы тех формул, которые уже содержались в θ. Если мы добавляем атомарную формулу, то расширенная таблица будет содержать лишь добавление этой формулы «снизу» ко всем открытым ветвям (ср. рис. 2 и рис. 3). Если добавляемая формула является сложной, то расширенная таблица может принять совершенно другую форму, например дизъюнкция разобьет каждую открытую ветвь на две. Здесь мы не будем вдаваться в подробности относительно обоснования полноты и способов расширения таблиц, анализа вариантов, к которым может приводить добавление формулы (в частности, она может не закрывать открытую ветвь, закрывать все открытые ветви или закрывать часть открытых ветвей, а остальные оставлять открытыми), и возникающих в связи с этим проблем [14]. Сказанного уже вполне достаточно, для того чтобы построить эффективное представление для простого абдуктивного вывода.

Пусть θ – множество формул, φ – предложение и α – абдуктивное объяснение. Тогда *мы сможем говорить об эффективном представлении простого абдуктивного вывода в том случае, если нам удастся построить закрытую таблицу* $T((\theta \cup \{\neg\varphi\} \cup \{\alpha\})$, что будет означать $(\theta, \alpha`\varphi)$ [15]. Другими словами, простым абдуктивным объяснением будут те формулы, которые будут закрывать открытые ветви таблицы $T(\theta \cup \{\neg\varphi\})$. В целях увеличения эвристического потенциала предлагаемого представления потребуем, чтобы анализируемые абдуктивные объяснения выбирались среди тех формул, которые закрывают некоторые (не обязательно все) открытые ветви $T(\theta)$.

Простой абдуктивный вывод и его эффективное вычислимое представление

Прежде чем окончательно перейти к алгоритмическому представлению простой абдукции приведем несколько определений, характеризующих степень закрытости ветвей семантических таблиц [16]. Рассмотрим множество $(\Gamma_1,...,\Gamma_n)$ – множество всех открытых ветвей $T(\theta)$. Обозначим множество формул, закрывающих открытую ветвь Γ_i, как $BTC(\Gamma_i) = \{x \mid \neg x \in \Gamma_i\}$. Тогда множество формул, закрывающих все ветви таблицы, будет пересечением $TTC(\theta) = \bigcap_{i=1}^{i=n} BTC(\Gamma_i)$. Обозначим множество формул, которые закрывают данную ветвь, но оставляют некоторые из открытых ветвей открытыми, как $BPC(\Gamma_i) = BTC(\Gamma_i) - TTC(\theta)$. Тогда множество всех формул, частично закрывающих таблицу, будет объединением $TPC(\theta) = \bigcup_{i=1}^{i=n} BPC(\Gamma_i)$. Мы будем использовать таблицы как простой тест на адекватность предполагаемого объяснения, вместе с тем контролируя множество потенциальных формул, из которых оно «выбирается». Наше представление будет соответствовать различным формам, которые может принимать формула, претендующая на то, чтобы быть объяснением (атомарная формула, конъюнкция или дизъюнкция формул). Начнем построение эффективного представления для объяснения в виде атомарной формулы.

Идея достаточно проста: нам необходимо найти те атомарные формулы, которые закрывают каждую открытую ветвь таблицы $T(\theta\cup\{\neg\varphi\})$ и соответствуют множеству $TTC(\theta\cup\{\neg\varphi\})$. Рассмотрим пример $\theta = \{\neg a \vee b\}$, $\varphi = b$ (рис. 4).

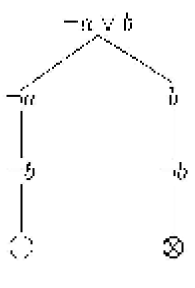

Рис. 4. Семантическая таблица $T(\theta \cup\{\neg\varphi\})$, $\theta = \{\neg a \vee b\}$, $\varphi = b$

Соответственно возможны два решения в виде атомарных формул, – это α и b. Перейдем к эффективному представлению для конъюнкции и дизъюнкции формул.

Поскольку объяснение в виде атомарной формулы может не существовать вовсе или не удовлетворять принятым ограничениям, постольку определенный интерес может представлять анализ более сложных форм объяснения: конъюнктивной и дизъюнктивной. Эффективное представление объяснения, принимающего форму конъюнкции $\alpha = \alpha_1 \wedge ... \wedge \alpha_n$, очевидным образом близко к построению объяснения, имеющего атомарную форму. Мы ищем формулы, которые закрывают открытые ветви, но в данном случае нам интересны формулы, которые закрывают некоторые, но не все ветви, т.е. формулы, входящие в $TPC(\theta \cup \{\neg\varphi\})$. Именно они и явятся составляющими конъюнктами объяснения. Каждое из этих «частичных» решений (конъюнктов) будет делать факт φ «менее удивительным» – в том смысле, что будет закрывать свою ветвь, а вместе они будут представлять собой общее решение.

Не уменьшая общности рассуждений, можно утверждать, что объяснение в виде дизъюнкции формул будет сконструировано из найденных ранее атомарных и «частичных» решений, т.е. решений в виде конъюнкции. К сожалению, объем статьи не позволяет нам более полно проанализировать построение объяснения в виде дизъюнкции формул, однако можно привести пример, который, на наш взгляд, удачно проиллюстрирует основную идею. Пусть $\theta = \{\alpha\}$ и $\varphi = b$. Построим таблицу $T(\theta \cup \{\neg b\})$ (рис. 5).

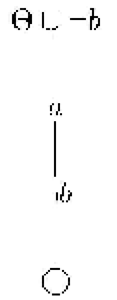

Рис. 5. Семантическая таблица $T(\theta \cup \{\neg b\})$, $\theta = \{\alpha\}$, $\varphi = b$

Очевидно, что возможно два атомарных решения: $\neg\alpha$ и b. Однако первое из них не удовлетворяет условию совместности, а второе является тривиальным. Более того, в данном случае отсутствуют «частичные» решения, поскольку вся таблица состоит только из одной ветви.

Один из возможных выходов из сложившейся ситуации связан с построением комбинации из различных атомарных решений. Например, можно представить решение в виде дизъюнкции $\neg\alpha \vee b$. Для того чтобы убедиться в том, что это действительно решение, построим таблицу $T((\theta\cup\{\neg b\})\cup(\neg\alpha\vee b)$ (рис. 6).

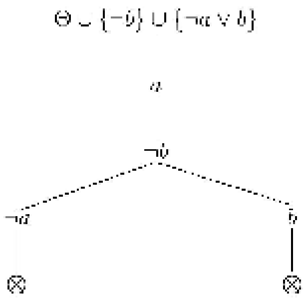

Рис. 6. Семантическая таблица $T((\theta\cup\{\neg\varphi\})\cup\{\alpha\})$, $\theta = \{\alpha\}$, $\varphi = b$, $\alpha = \{\neg a\vee b\}$

На наш взгляд, приведенный пример как нельзя лучше отражает эффективное представление решения в виде дизъюнкции формул, которое, например, ищется в виде импликации [17].

Итак, описав возможные эффективные представления для поиска решения простого абдуктивного вывода, мы можем преступить к построению алгоритма, соответствующего эффективному представлению для вычисления простой абдукции. Что будет являться входами и выходами нашего алгоритма, какова его рабочая процедура (процесс)?

- ***Входы***

. Исходное множество формул θ, представляющих принятую теорию.

. Формула φ – формула, представляющая «факт, который необходимо объяснить».

. Предварительные условия, накладываемые на θ и φ, например $\theta–\varphi$ или $\theta\neg\varphi$ и др.

- ***Выходы***

. Множество абдуктивных объяснений $\{\alpha_1,...,\alpha_n\}$, таких что:

1) $T((\theta\cup\{\neg\varphi\})\cup\{\alpha_i\})$ закрыта;

2) формулы α_i соответствуют принятым ранее ограничениям на словарь теории и форму объяснения.

- ***Рабочая процедура (процесс)*:**

. Поиск множества $(\Gamma_1,...,\Gamma_n)$ – открытых ветвей исходной таблицы.

. *Поиск объяснения в виде атомарной формулы:*

1) *вы*числение $TTC(\Gamma)$.

2) проверка формул, входящих в $TTC(\Gamma)$, на связность с θ и на тривиальность решения.

. *Поиск объяснения в виде конъюнкции формул:*

1) для каждой открытой ветви Γ_i вычисление $BPC(\Gamma_i)$;

2) проверка всех ветвей Γ_i на предмет того, чтобы не все ветви оказались закрытыми. В противном случае алгоритм останавливается – решения в виде конъюнкции формул нет;

3) каждое $BPC(\Gamma_i)$ содержит формулу, которая частично закрывает таблицу. Решение конструируется как конъюнкция всех формул, выбираемых отдельно из каждого $BPC(\Gamma_i)$. Например, пусть $a_1 \in BPC(\Gamma_1)$, $b_1 \in BPC(\Gamma_2)$, ..., $x_1 \in BPC(\Gamma_n)$тогда решение $a_1 = a_1 \wedge b_1 \wedge ... \wedge x_1$ и т.д.;

4) проверка всех решений на связность и единственность.

. *Поиск объяснения в виде дизъюнкции формул:*

1) составление решений путем комбинирования найденных объяснений в виде атомарных формул и конюънкций формул друг с другом, а также путем комбинирования их с φ. Например, пусть α_1 и α_m – две атомарные формулы-объяснения, тогда одно из возможных решений будет $\alpha_1 \vee \alpha_m$; также решениями будут $\alpha_1 \vee \varphi$ и $\alpha_m \vee \varphi$;

2) проверка всех решений на связность.

. Проверка всех решений (в виде атомарных формул, конъюнкции и дизъюнкции формул) на «состоятельность» объяснения, удовлетворение условию выводимости θ,α‘φ. Например, полученные решения α должны отвечать следующим требованиям: во-первых, теория не должна сама объяснять φ (это одно из условий на θ и φ выше); во-вторых, объяснение не должно быть тривиальным α–φ. В-третьих, решение должно удовлетворять условию связности θ–$\neg\alpha$ и т.д.

На наш взгляд, представленный алгоритм полностью отвечает задаче построения эффективного вычислимого представления для простого абдуктивного вывода. Полученный алгоритм позволяет достаточно просто представить решения в той или иной логической форме и подобрать множество формул, анализ которых можно продолжить с целью выбора наиболее оптимальной из них. Более того, полученный алгоритм позволяет без труда обобщить представления о «входах», «выходах» и «процессе» как на известные языки логического программирования (Лисп, Пролог), так и на стандартные машинные языки, что является еще одной положительной эвристикой предлагаемого подхода.

Примечания

1. «Исходным понятием в данном случае служит понятие “алгоритм” в его самом широком значении “настолько точного и детального предписания, что следовать ему могут даже механические устройства, если они подходящим образом сконструированы и налажены”. Алгоритм *P реализует* функцию *f*, если для любого заданного на входе *x* алгоритм *P* производит на выходе *f(x)*, когда *f(x)* определено, и алгоритм *P* не имеет результата, когда *f(x)* не определено. Функция *f вычислима*, если существует алгоритм, который реализует *f*. … Представление функции – это синтаксическое выражение, для которого существуют его интерпретации как функции. Поскольку у одной и той же функции может быть несколько представлений, то не всякое представление функции может определять алгоритм ее вычисления… Представление функции φ *эффективно*, если φ определяет алгоритм, который реализует функцию, описываемую φ. Иными словам, нам необходимо построить формальный язык, в котором определяются представления функций, а также интерпретацию этого формального языка, при которой представления функций интерпретируются как эти функции. Еслиφ эффективно, то, интерпретируя φ, мы будем знать алгоритм, который реализует функцию, обозначаемую φ, – мы приобретаем это знание без использования какой-либо информации, которая не являлась бы частью той, что необходима для понимания самого представления φ.» (*Гончаров С.С., Ершов Ю.Л., Самохвалов К.Ф.* Введение в логику и методологию науки. – М.: Интерпракс, 1994. – С. 32–33).

2. См., например: *Aliseda-Llera A.* Seeking explanations: Abduction in logic, philosophy of science and artificial intelligence. – Palo Alto: Stanford Univ. Press, 1997; *Mayer M., Pirri F.*

First order abduction via tableau and sequent Calculi // Bulletin of the IGPL. – 1993. – V. 1. – P. 99–117.

3. См.: *Hintikka J.* What is abduction? The fundamental problem of contemporary epistemology / Inquiry as Inquiry: A Logic of Scientific Discovery / Hintikka J. Selected Papers. – N.Y.: Kluwer Academic Publishers, 1999. – V. 5. – P. 91–113.

4. Наиболее типичный пример практики абдуктивного рассуждения – постановка диагноза. Фактами, которыми обладает доктор, являются симптомы заболевания (температура, отек и др.). Доктор гипотетически приходит к постановке диагноза (определению заболевания, его стадии или тяжести) основываясь на знании причинной взаимосвязи симптомов и конкретного заболевания. В дальнейшем именно эта гипотеза служит основанием для медикаментозного или хирургического лечения.

5. См.: *Hanson N.* Patterns of scientific discovery. – Cambridge: Cambridge Univ. Press, 1961.

6. См., например: *Harman G.* Change in view: Principles of reasoning. – Cambridge: MIT Press, 1986; *Josephson J., Josephson S.* Abductive inference: computation, philosophy, technology. – Cambridge: Cambridge Univ. Press, 1996; *Lipton P.* Inference to the best explanation. – London: Routledge, 2004; *Magnani L.* Abduction, reason and science: Processes of discovery and explanation. – N.Y.: Kluwer Academic Publishers, 2001; *Nagel E.* The Structure of science: Problems of the logic of scientific explanation. – Cambridge: Cambridge Univ. Press, 1979; *Ruben D.* Explaning explanation. – London: Routledge, 1990; *Salmon W.* Four decades of scientific explanation. – Minneapolis: Univ. of Minnesota Press, 1990; *Thagard P.* Computational philosophy of science. – Cambridge: MIT Press, 1988.

7. См., например: *Konolige K.* A general theory of abduction // Automated Abduction: Working Notes / Ed. by P. O'Rorke. – Palo Alto: Stanford Univ. Press, 1996. – P. 62–66; *Konolige K.* Abductive theories in artificial intelligence // Principles of Knowledge Representation / Ed. by G. Brewka. – Palo Alto: Stanford Univ. Press, 1996. – P. 129–152; *Kakas A., Kowalski R., Toni F.* Abductive logic programming // Journal of Logic and Computation. – 1993. – V. 6.– P. 719–770.

8. Стандартное представление о логическом выводе имеет вид: «заключение *C* следует из множества посылок *P*». Интуитивно ясно, что абдуктивный вывод является «обратным», в каком-то смысле заключение нам уже дано, а в результате мы должны обнаружить подходящие посылки. На самом деле направление вывода не играет большой роли, – намного важнее возможность расширить представление о посылках вывода. Уже у Гемпеля (См.: *Hempel C.* Aspects of scientific explanation // Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. By C. Hempel. – N.Y.: Free Press, 1965) можно встретить представление о том, что посылки вывода могут включать в себя достаточно разнообразные базовые предположения (background assumptions), которые

могут иметь «различный способ доступа или влияния(access)» по отношению к ним. В любом случае мы ищем соответствующие посылки, которые будут подкреплять заключение.

Стандартная интерпретация достоверности вывода предполагает по крайне мере два пути: «через семантику», когда мы опираемся на понятия «модель» и «интерпретация» (на каждой модели, на которой *P* истинно, *C* тоже истинно), и «через синтаксис» (теорию доказательства), когда мы опираемся на понятия «доказательство» и «выводимость» (доказательство достоверно, если проводится посредством принятых правил вывода). В данном случае мы вынуждены от нее отказаться, поскольку наша цель – вполне конкретное эффективное представление, которое можно будет использовать в качестве основания для построения моделей искусственного интеллекта. Накопленный опыт программирования (использование в программах внелогических операторов и проч.) и современное состояние дел в данной области заставляют нас изменить представление о достоверности заключения (см., например: *Bethem J., van.* Logic as programming // Fundamenta Informaticae. – 1992. – V. 17. – P. 285–318; *Kalsbeek M.* Meta logics for logic programming. – Amsterdam: Univ. of Amsterdam Press, 1995; *Kowalski R.* A Metalogic programming approach to multiagent knowledge and belief // Artificial Intelligence and Mathematical Theory of Computation / Ed. by V. Lifschitz. – N.Y.: Academic Press, 1991. – P. 231–246).

Выше мы упоминали выводящий параметр , его можно определять как семантически, так и синтаксически. Гораздо важнее так называемая «стратегия поиска» или «стратегия контроля». Поиск абдуктивного объяснения всегда будет связан с конкретной частной практической реализацией абдуктивного рассуждения. Соответственно можно предположить, что «стратегия поиска» должна быть связана с конкретным механизмом, приводящим к заключению, – нам не нужны процедуры, которые порождают *все* возможные абдуктивные объяснения, нам всегда будут интересны только те, которые мы сможем оценить «на практике». Цель последней – не соответствие определенной семантике, а просто механизм, порождающий и контролирующий отбор объяснений. Именно данное обстоятельство, заключающееся в том, что «стратегия контроля» намного «важнее» логики вывода, обусловило обращение к методу семантических таблиц Бета – Хинтикка. С одной стороны, Э. Бет пришел к идее семантических таблиц, проводя систематическое исследование понятий «прямой вывод» и «обратный вывод» (будучи прочитанными «в одну сторону», они показывают на возможные контрпримеры, «в другую» – указывают «направление» вывода). С другой стороны, Я. Хинтикка постоянно подчеркивает, что содержание абдуктивного вывода определяется дефинициональными и стратегическими правилами (логикой вывода и стратегией поиска достоверного заключения). По сути, мы включили в представление о посылках вывода информацию, касающуюся отбора адекватных объяснений, сохранив тем самым дух гемпелевской дедуктивно-номологической схемы объяснения.

9. В частности, в литературе можно встретить ряд более «тонких» ограничений, нацеленных на то, чтобы внимательнее проанализировать предполагаемое объяснение и отразить необходимые предпочтения. Например, стремясь получить «более простое» объяснение, как правило, накладывают следующее ограничение: если θ,α‘φ, то для всех β, таких что θ,β‘φ, необходимо, чтобы выполнялось условие ‘β→α.

10. См.: *Beth E.* Semantic entailment and formal derivability // The Philosophy of Mathematics / Ed. by J. Hintikka. – N.Y.: Oxford Univ. Press, 1969. – P. 9–41.

11. См. *Hintikka J.* Two papers on symbolic logic: Forms and quantification theory and reductions in the theory of types // Acta Philosophica Fenica. – 1955. – V. 8.– P. 1–115.

12. См.: *Smullyan R.* First order logic. – N.Y.: Springer-Verlag, 1968; *Fitting M.* First order logic and automated theorem proving. – N.Y.: Springer-Verlag, 1970.

13. См., например: *Aliseda-Llera A.* Seeking explanations…; *Mayer M., Pirri F.* First order abduction via tableau and sequent Calculi.

14. Об этом см.: *Smullyan R.* First order logic.

15. В данном случае, говоря о том, что абдуктивный вывод является *простым*, мы указываем на то, что не накладываем никаких дополнительных условий на объяснение α. В частности, говоря о более сложных видах абдукции, можно устанавливать *ограничения связности* объяснения. Оно включает в себя кроме простой абдукции еще и требование того, чтобы таблица $T((\theta \cup \{\alpha\})$ была открытой, т.е. (θ¬α). В этом случае алгоритм приведет нас ко всем тем объяснениям, которые несовместны с θ. Другое «популярное» ограничение – *ограничение на проверку объяснительной силы*. Оно включает дополнительные требования: 1) $T((\theta \cup \{\neg\varphi\})$ открыта, т.е. (θ–φ); 2) $\mathrm{T}(\alpha \cup \{\neg\varphi\})$ открыта, т.е.(α–φ) . Можно встретить и другие ограничения, например *улучшенное условие связности*, которое начинается не с простой абдукции, а с поиска всех формул α, закрывающих некоторые (но не все) открытые ветви $T(\theta)$, а затем уже для найденных на предыдущем шаге α проверяется, что $T((\theta \cup \{\neg\varphi\} \cup \{\alpha\})$ закрыта. В любом случае эти и другие ограничения играют роль ограничений на наши представления о типе абдуктивного вывода (простое объяснение фактов, причинная взаимосвязь между фактами, указание на противоречие и др.), о характере объяснения (динамическое или вероятностное, наилучшее и др.) или о характере поступающих данных (новые данные, ожидаемые предсказания, ретросказания и др.).

16. См.: *Smullyan R.* First order logic; *Fitting M.* First order logic and automated theorem proving.

17. Рассмотрим пример абдуктивного заключения. Нам известно, что появление облаков определенного типа, как правило, предшествуют дождю. Вечером вы замечаете, что все небо затянуто облаками, которые предвещают грозу. Утром вы обнаруживаете,

что трава перед домом мокрая. Для человека, остающегося в рамках здравого смысла, нетрудно проследить причинную взаимосвязь между тучами на небе вечером и сырой травой утром: ночью был дождь. Собственно, таблица на рис. 6 служит проверкой именно такой обнаруженной причинной взаимосвязи: $\{a\}$ – облака, $\{b\}$ – дождь, $a = \{a \to b\}$– обнаруженная причинная взаимосвязь.

Институт философии и права СО РАН,
г. Новосибирск

Golovko, N.V. **On the construction of effective calculable presentation for prime abductive inference**

The paper offers an algorithm which according to the author corresponds in full to meets a clear correlation with the problem of construction of effective calculable presentation for prime abductive inference. This algorithm makes possible to present solutions in one or another logical form in a rather simple way and to select a great number of formulas, their analysis may be continued in order to choose the best among them. In addition, this algorithm makes possible easily generalize the ideas of “inputs”, “outputs” and “process” both for logical programming languages and standard computer ones; it is another positive heuristics of the offered approach.