Дискретная математика

том 3 выпуск 3 * 1991

УДК 519.

ОБ ОРГАНИЗАЦИИ БАЗ ЗНАНИЙ, ОРИЕНТИРОВАННЫХ НА АВТОМАТИЧЕСКОЕ РЕШЕНИЕ ЗАДАЧ

А.С. Подколзин

В работе описывается схема организации и использования баз знаний, хранящих многообразия приемов автоматического решения задач. На основе этой схемы была разработана и отлажена процедура автоматического решения задач из некоторых разделов элементарной математики. Процессы решения задач данной процедурой носят, в основном, градиентный характер, и лишь изредка в них возникает ветвление, приводящее к появлению элементов перебора. Последовательность предпринимаемых здесь действий, после определенной регулировки управляющей компоненты рассматриваемых приемов, почти не отличается от естественной последовательности действий при решении задач человеком. Предлагаемая общая схема упрощает согласование друг с другом процедур решения задач из различных предметных областей и разработку энциклопедических процедур решения задач, охватывающих сразу многие предметные области.

§ 1. Общая схема функционирования автоматического решателя задач

Основным для рассматриваемой общей схемы является понятие задачи, которое складывается из четырех типов задач: на доказательство, на описание, на преобразование и на исследование. Задание класса допустимых задач начинается с указания некоторого логического языка L. Алфавит этого языка состоит из конечного множества логических символов, конечного множества отличных от них символов переменных и скобок. Термы языка L суть логические символы, переменные, а также всевозможные слова вида $\varphi(t_1 \ldots t_n)$, где φ – логический символ; $t_1, \ldots, t_n$ – термы; $n \geqslant 1$. Выделяются специальные подмножества термов, называемых соответственно утверждениями и выражениями, и фиксируется интерпретация языка L на "универсуме" U, при которой каждое утверждение определяет предикат, а каждое выражение – функцию со значениями из U. При задании утверждений и выражений используются конструкции, характерные для естественного математического языка и существенно расширяющие рамки языка классической логики предикатов.

Задачи определяются как некоторые конструктивные объекты, в состав которых входят:

а) список утверждений $f_1, \ldots, f_n$, называемых посылками задачи (в процессе решения задачи значения свободных переменных утверждений $f_1, \ldots, f_n$ предполагаются зафиксированными, а сами эти утверждения – истинными);

б) набор $(a_1, \ldots, a_n)$ целых неотрицательных чисел, называемых весами посылок (a_i – вес посылки f_i) и используемых, как указывается ниже, для управления процессом "просмотра" задачи при поиске очередного приема (в исходной ситуации веса посылок задачи равны 0);

в) набор $(K_1, ..., K_n, K_0)$, элементы K_i которого суть наборы объектов, называемых при $i \neq 0$ комментариями к i-й посылке задачи, а при $i = 0$ – общими комментариями к посылкам задачи (в комментариях регистрируется различная информация, накопленная в процессе решения задачи и используемая при оценке целесообразности применения тех или иных приемов, например, информация о ранее предпринимавшихся неудачных шагах, блокирующая их повторение; в исходной ситуации все K_i пусты).

Компоненты а), б), в) являются общими для задач всех перечисленных выше типов. Прочие компоненты определяются в зависимости от типа задачи следующим образом.

1. Задача на доказательство:

г) утверждение f', называемое условием задачи (если из истинности посылок $f_1, ..., f_n$ вытекает истинность условия f', то ответом на задачу служит логический символ "истина", иначе – ответ не определен);

д) целое неотрицательное число a' – вес условия задачи;

е) набор K' объектов, называемых комментариями к задаче.

Компоненты д), е) здесь аналогичны компонентам б), в), связанным с посылками.

2. Задача на описание:

г) список утверждений $f'_1, ..., f'_m$, называемых условиями задачи;

д) набор $(a'_1, ..., a'_m)$ весов условий;

е) набор $(K'_1, ..., K'_m, K'_0)$, элементы K'_i которого суть наборы комментариев: при $i \neq 0$ – к условию f'_i, а при $i = 0$ – общих комментариев к задаче;

ж) набор $(c_1, ..., c_p)$ объектов, называемых целями задачи (класс утверждений, являющихся ответами на задачу, определяется в зависимости от набора ее целей; при этом некоторые из целей могут не накладывать каких-либо ограничений на ответ задачи, а лишь указывать те или иные требования оптимизации ответа и определять соответствующие тенденции в процессе решения задачи; как правило, задача на описание имеет цель, указывающую множество ее неизвестных (вхождение неизвестных в посылки не допускается), а условия должны являться следствиями ответа и посылок задачи; при этом ответ можно рассматривать как частичное либо полное описание множества наборов значений неизвестных, удовлетворяющих условиям задачи; вид этого описания регламентируется списком целей; варьируя список целей задачи, можно включать либо выключать целые группы приемов, участвующих в ее решении, и тем самым обеспечивать необходимую гибкость обращений к вспомогательным задачам);

з) (может отсутствовать) – задача на исследование Z', называемая блоком анализа данной задачи на описание Z (задача Z' первоначально имеет своими посылками все посылки и условия задачи Z; она вводится для извлечения таких следствий условий задачи Z, которые позволили бы однозначно определить значения ее неизвестных, либо выразить их через другие неизвестные, либо указать целесообразность применения тех или иных приемов для решения задачи Z).

3. Задача на преобразование:

г) выражение t, называемое условием задачи;

д) вес a' условия задачи;

е) набор K' комментариев к задаче;

ж) набор целей задачи.

Ответ θ задачи на преобразование представляет собой такое выражение, что из истинности посылок вытекает равенство значений выражений t и θ, причем некоторые из целей накладывают ограничения на вид ответа, а другие – те или иные тенденции оптимизации в процессе приводящих к ответу преобразований.

4. Задача на исследование:

г) набор целей задачи (эти цели управляют процессом вывода представляющих интерес следствий из посылок задачи, упрощения посылок и удаления избыточных посылок; в результате таких действий исходная задача на исследование Z преобразуется в некоторую новую задачу Z', и при исчерпании средств, отведенных для решения задачи Z, в качестве ответа на нее выдается задача Z').

Текущая ситуация, возникающая в процессе решения задачи Z, описывается последовательностью задач $Z = Z_1, Z_2, ..., Z_N$, где Z_{i+1} – вспомогательная задача, введенная при решении задачи Z_i $(i = 1, ..., N - 1)$; Z_N – текущая решаемая задача; $N \geqslant 1$. Каждая задача Z_i $(i = 1, ..., N)$ характеризуется натуральным числом M_i, называемым ее максимальным уровнем и определяющим уровень средств, отведенных для ее решения (по исчерпании этих средств выдается отказ), целым неотрицательным числом m_i, называемым текущим уровнем этой задачи и определяющим уровень средств, среди которых ведется поиск очередного преобразования текущей ситуации, $m_i \leqslant M_i$, а также вспомогательной информацией, необходимой для возобновления прерванной процедуры решения задачи Z_i по окончании решения задачи Z_{i+1}.

Изменение текущей ситуации происходит в результате срабатывания приемов решения задач, организованных в специальную базу знаний. Каждый такой прием представляет собой алгоритмическую процедуру, осуществляющую следующие действия.

1) Усмотрение в текущей ситуации возможности применить некоторое ее допустимое (в смысле логической корректности) преобразование.

2) Проверка условий, определяющих эвристическую целесообразность применения данного преобразования (как правило, происходит одновременно с действиями из п. 1).

3) Реализация преобразования.

Для сравнения эвристической целесообразности преобразований вводится дискретная шкала оценок (целые неотрицательные числа), причем управляющая компонента приемов (п. 2) устроена так, что на каждом шаге выполняется поиск преобразования с наименьшей возможной оценкой. В качестве оценки целесообразности преобразования здесь выступает то наименьшее значение текущего уровня решаемой задачи, при котором управляющая компонента приема допускает его применение.

Согласно обычной схеме организации многообразия приемов в базу знаний, они объединяются в древовидную процедуру $\mathcal{D}$, узлы которой определяют выделение тех или иных элементов задачи и проверку связанных с ними условий, а концевые вершины соответствуют различным приемам. При перемещении от корня процедуры $\mathcal{D}$ к концевой вершине происходит выделение в задаче ряда объектов, определяющих конкретный вариант применения приема, и проверка условий его применимости. Поиск приема осуществляется здесь в процессе перебора всевозможных допустимых (в смысле истинности проверяемых условий) путей, ведущих от корня к концевым вершинам, и его можно рассматривать как "сканирование" многообразия приемов. Недостатком такой схемы является то, что для различных приемов приходится осуществлять независимые повторные просмотры задачи; чтобы сделать один просмотр задачи общим для целой группы приемов, была применена другая схема, при которой поиск приема основан на "сканировании" задачи. Оказалось, что для подавляющего большинства приемов локализация в задаче ситуаций, определяющих варианты их применения, начинается с обнаружения вхождения в посылку либо в условие некоторого фиксированного для рассматриваемого приема логического символа. Это сделало целесообразным объединение всех приемов, соответствующих одному и тому же логическому символу φ, в одну древовидную процедуру $\mathcal{D}_\varphi$ указанного выше типа, а приемов, применение которых не связано с выделением в задаче вхождения какого-либо конкретного логического символа — в четыре процедуры $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$, $\mathcal{D}_4$ соответственно для решения задач на доказательство, описание, преобразование и исследование. В результате вся база знаний оказывается организованной по принципу энциклопедии, где за каждым понятием (логическим символом φ либо типом задачи) закреплена соответствующая ему информация (процедура $\mathcal{D}_\varphi$ либо одна из процедур $\mathcal{D}_1$ – $\mathcal{D}_4$). Для поиска приема предпринимается последовательный просмотр всех подряд логических символов условий и посылок задачи ("сканирование" задачи) и обращение для каждого такого вхождения символа φ к процедуре $\mathcal{D}_\varphi$ (перед началом очередного цикла сканирования реализуется соответствующая типу задачи процедура $\mathcal{D}_i$, $i \in \{1, 2, 3, 4\}$). Более подробно функционирование решателя задач при рассмотрении текущей ситуации $Z_1, Z_2, ..., Z_N$ осуществляется по следующей схеме.

1) Происходит обращение к процедуре $\mathcal{D}_i$, $i \in \{1, 2, 3, 4\}$, соответствующей типу задачи Z_N. Если при этом не срабатывает ни один из приемов, либо сработавшие приемы изменили лишь комментарии задач и ни один из них не указал явно на необходимость повторного рассмотрения задачи (в таких случаях говорим, что процедура не внесла существенных изменений в текущую ситуацию), то переход к п. 3, иначе – к п. 2.

2) Если в результате срабатывания приема определился ответ на задачу Z_N либо был выдан отказ на нее, то возобновляется прерванный ранее прием решения задачи Z_{N-1}, в процессе реализации которого возникла задача Z_N (при $N = 1$ выдается ответ либо отказ на исходную задачу Z_1). В остальных случаях текущий уровень задачи Z_N заменяется на 0, и переход к п. 1.

3) Осуществляется последовательный просмотр всех условий и посылок F задачи Z_N, вес v которых равен текущему уровню m_N задачи Z_N либо равен $m_N + 1$ (заметим, что в исходной ситуации текущий уровень задачи Z_1 полагается равным 0). Сначала просматриваются условия, затем посылки; порядок просмотра условий (посылок) определяется соответствующими списками задачи Z_N. Если $v = m_N$, то происходит однократный просмотр всех вхождений логических символов φ в F (слева направо); если же $v = m_N + 1$, то – серия таких просмотров, в процессе которых значение текущего уровня задачи Z_N полагается последовательно равным $0, 1, ..., m_N$. Для каждого рассматриваемого вхождения логического символа φ в F осуществляется обращение к процедуре $\mathcal{D}_\varphi$ (исходные данные этой процедуры содержат полную информацию о "координатах" данного вхождения символа φ в задачу Z_N). Если процедура не внесла существенных изменений в текущую ситуацию, то переход к рассмотрению очередного вхождения логического символа, иначе – переход к п. 2. Если просмотр терма F закончился безрезультатно, то вес его увеличивается на 1 (новые либо видоизмененные посылки и условия получают вес 0). Если просмотр всех условий и посылок задачи Z_N закончился безрезультатно, то текущий уровень задачи Z_N увеличивается на 1. Если в результате он становится больше, чем максимальный уровень M_N, то на задачу Z_N выдается отказ, иначе – переход к п. 1.

Использование весов посылок и условий позволяет сузить область просмотра при поиске очередного приема, исключая из нее посылки и условия, имеющие большой вес (они уже были достаточно хорошо рассмотрены ранее, и срабатывание связанного с ними приема маловероятно). В результате происходит локализация рассмотрения задачи, направляемого в первую очередь на новые либо видоизмененные посылки и условия; рассмотрение же всей задачи в целом имеет место, как правило, лишь на начальном этапе ее решения. По мере повышения текущего уровня m_N задачи Z_N в просмотр вовлекаются ранее отложенные посылки и условия F, вес v которых больше m_N; это происходит при $v = m_N + 1$ (см. п. 3), причем предварительно осуществляется поиск приема, срабатывающего при рассмотрении F для меньших, чем m_N, значений текущего уровня. Переключение "внимания" при рассмотрении задачи может происходить также в результате срабатывания приемов, уменьшающих веса тех или иных условий и посылок.

Предлагаемая общая схема ориентирована главным образом на разработку "энциклопедических" процедур решения задач, охватывающих сразу многие предметные области, так как при переходе к новой предметной области происходит формирование групп приемов, связанных с новыми понятиями, лишь в незначительной степени затрагивающее группы приемов, соответствующих ранее рассмотренным понятиям, и практически не увеличивающее время их поиска. Далее мы опишем множество приемов, которое организовано в соответствии с указанной схемой, возникло в результате анализа процессов решения задач человеком и охватывает несколько различных предметных областей. В настоящей статье ограничимся приемами, входящими в процедуры $\mathcal{D}_1 - \mathcal{D}_4$, некоторыми "логическими" приемами общего характера и приемами, относящимися к алгебре множеств.

§ 2. Логические приемы решения задач

Начнем с описания основных приемов, входящих в процедуры $\mathcal{D}_1 - \mathcal{D}_4$ (будем называть их общими приемами решения задач соответствующих типов). Простейшие общие приемы решения задач на доказательство устраняют дублирование в посылках и выдают ответ при появлении посылки, являющейся отрицанием другой посылки либо при совпадении условия с одной из посылок (одна из наиболее типичных заключительных ситуаций). Чтобы своевременно усматривать появление условия среди посылок задачи, используются два различных приема, один из которых связан с рассмотрением нового условия, а другой – с рассмотрением новой посылки. Такое применение различных приемов для усмотрения одной и той же ситуации, отправляясь от различных ее элементов, характерно для ситуаций, в описании которых участвуют несколько "равноправных" посылок либо условий задачи. Иногда здесь выделяется один основной прием, а прочие приемы лишь уменьшают вес той посылки (либо условия), с рассмотрением которой связан основной прием.

Для сокращения времени решения задачи на доказательство в тех случаях, когда условие g не является следствием посылок, используется прием, устанавливающий на основании ряда признаков (наличие свободной переменной условия, не встречающейся в посылках и т. п.) "сомнительность" утверждения g и предпринимающий попытку опровергнуть его. Приемы, осуществляющие упрощение выражений, встречающихся в задаче, делятся на две группы. К первой из них относятся сравнительно простые, малотрудоемкие приемы, срабатывающие в результате просмотра входящих в ту же задачу логических символов, ко второй – более трудоемкие приемы, включаемые только при решении вспомогательных задач на преобразование специально выделенных в первоначальной задаче выражений. Для такого выделения подлежащих упрощению выражений служит специальный прием решения задач на доказательство. Этот прием просматривает все входящие в посылки и в условие задачи выражения, не расположенные внутри большего выражения, и выбирает те из них, которые обладают рядом признаков. Наряду с признаками общего характера (большая длина выражения, наличие повторяющихся переменных и т. п.) здесь используется множество признаков, связанных с конкретными логическими символами, встречающимися в рассматриваемом выражении. Это множество признаков, как и множество приемов, организовано "по принципу энциклопедии", причем проверка признаков происходит в процессе просмотра логических символов, встречающихся в выражении. Аналогичный подход, основанный на организации в специальную базу знаний ("базу признаков") множества признаков, связанных с выделением в рассматриваемой ситуации вхождения конкретного логического символа, позволяет при управлении многими приемами общего характера учитывать специфические особенности конкретной ситуации и значительно повышать целесообразность их применения. Помимо перечисленных основных, имеется ряд вспомогательных общих приемов решения задач на доказательство, управляющих весами посылок и максимальным уровнем задачи, выдающих отказ при "зацикливании", а также ряд приемов, сопровождающих приемы отдельных логических символов.

В случае задач на описание прежде всего проверяется наличие условий, не зависящих от неизвестных. Для каждого такого условия f предпринимается попытка доказать, что оно является следствием посылок. Если эта попытка удается, то условие f удаляется (при единственном условии – выдается ответ "истина"). Иначе – рассматривается список целей задачи, и в зависимости от него либо выдается отказ (если требуется найти полный пример, т. е. для любых удовлетворяющих посылкам значений известных величин указать набор удовлетворяющих условиям значений неизвестных), либо предпринимается попытка доказать отрицание f (в случае ее удачи выдается ответ "ложь"), либо, наконец, решается новая задача на описание, полученная перенесением f из списка условий в список посылок. При получении ответа g на нее к конъюнкции утверждений f и g (а если f было единственным условием, то – к f) применяется специальная процедура редактирования ответа, формирующая окончательный ответ решавшейся задачи на описание. Необходимость дополнительного редактирования ответа в данном и ряде других приемов объясняется тем, что заготовка ответа возникает путем объединения в рамках одного терма нескольких независимо полученных и не согласованных между собой с учетом целей задачи подтермов (например, g может иметь не зависящий от неизвестных конъюнктивный член f', следствием которого является f, и в этом случае f должно быть удалено). Основная часть процедуры редактирования заключается в сканировании заготовки ответа и выполнении преобразований, приводящих ее "градиентным" образом к виду, соответствующему целям задачи; это сканирование реализовано как процесс решения вспомогательной задачи на описание со специальным списком целей. Как правило, процедура редактирования ответа применяется лишь для исходной задачи Z_1, так как ответы вспомогательных задач редактируются "автоматически" в рамках тех задач, где они используются.

Приемы, устраняющие дублирование в посылках задачи на описание и выдающие ответ "истина" при появлении посылки, являющейся отрицанием другой посылки, – такие же, как у задач на доказательство.

Если в задаче на описание требуется найти пример (полный либо частичный, когда набор значений неизвестных указывается лишь для части удовлетворяющих посылкам значений известных величин), то предпринимается попытка подобрать такие ее элементарные (не имеющие связанных переменных и логических связок, отличных от отрицания) условия $f_1, ..., f_s$, содержащие все неизвестные $x_1, ..., x_n$, и такие выражения $t_1, ..., t_n$, что все утверждения $S^{x_1 \dots x_n}_{t_1 \dots t_n} f_i$ $(i = 1, ..., s)$ являются посылками задачи. Подбор начинается с условия f_1, содержащего хотя бы одну не являющуюся неизвестной переменную; после того как $t_1, ..., t_n$ определены, для каждого отличного от $f_1, ..., f_s$ условия f предпринимается попытка быстрого усмотрения из посылок утверждения $S^{x_1 \dots x_n}_{t_1 \dots t_n} f$ и при успешном завершении всех таких попыток в качестве ответа на задачу выдается утверждение вида $x_1 = t_1 \& ... \& x_n = t_n$. Возможность применения данного приема на заключительном этапе решения задачи обеспечивается множеством других приемов, выполняющих необходимые переформулировки условий и вывод следствий в посылках задачи. Аналогичный прием используется для подбора значений $t_1, ..., t_n$ неизвестных $x_1, ..., x_n$ из рассмотрения содержащих неизвестные посылок $f_1, ..., f_s$ блока анализа задачи на описание.

Если в задаче требуется найти не пример, а явное описание допустимых значений неизвестных, полное либо частичное, то заключительный этап ее решения обеспечивается различными приемами логических символов, усматривающими, что множество условий задачи (либо часть его) уже имеет вид явного описания, и выдающими ответ задачи (соответственно, осуществляющими исключение части неизвестных, для которых найдено явное описание через остальные неизвестные). В тех случаях, когда совокупность условий задачи на описание распадается на независимые (не имеющие общих неизвестных) группы, происходит сведение ее к системе независимо решаемых "подзадач". Аналогично тому как это делалось для задач на доказательство, осуществляется упрощение выражений, входящих в условия и посылки задачи на описание. Периодические обращения к рассмотрению блока анализа Z^* задачи на описание Z предпринимаются при фиксированных значениях $\mathcal{U}_1, \mathcal{U}_2, ..., \mathcal{U}_n$ текущего уровня задачи Z ($\mathcal{U}_1 < \mathcal{U}_2 < ... < \mathcal{U}_n$), причем каждый раз к посылкам задачи Z^* добавляются все новые (ранее не заносившиеся в блок анализа) условия задачи Z. Завершение очередного рассмотрения блока анализа происходит либо при исчерпании средств, определяемых уровнем $\mathcal{U}_i$, либо в результате срабатывания приема, усматривающего целесообразность возвращения к решению задачи Z (например, после нахождения значения одной из неизвестных).

В случае задач на преобразование ответ определяется при исчерпании средств, отведенных для решения (когда текущий уровень задачи оказывается равен ее максимальному уровню). Ответом здесь становится условие θ решаемой задачи, если только оно удовлетворяет всем требованиям, накладываемым на ответ целями этой задачи. Если задача на преобразование имеет "вычислительный" характер, т. е. сводится к последовательному выполнению операций над выражениями специального подкласса (например, десятичными записями чисел), то сначала преобразуются операнды θ_i условия $\varphi(\theta_1 \; ... \; \theta_n)$, не входящие в указанный подкласс, а затем уже преобразуется непосредственно условие задачи. Аналогичные приемы, но уже связанные с конкретными логическими символами, используются для выделения и независимого преобразования – с различными целями – тех или иных фрагментов условия решаемой задачи на преобразование. Как и для задач на доказательство, применяются приемы, устраняющие дублирование в посылках задачи на преобразование и выполняющие упрощение посылок. Такие же приемы применяются и при решении задач на исследование.

При решении задачи на описание часто оказывается полезным выделение зависящих от неизвестных выражений θ, имеющих более одного вхождения (явного либо "неявного", возникающего после некоторых простых преобразований) в условия этой задачи. Подстановка вместо вхождений таких выражений θ новых неизвестных может уменьшить число неизвестных задачи либо существенно упростить ее вид, сделать решаемой "стандартными" средствами. Для обнаружения многократных вхождений, зависящих от неизвестных выражений θ, используется блок анализа Z задачи на описание Z^*, в котором реализована тенденция поиска указанных выражений θ и обозначения их новыми переменными. Возникающие при этом переменные x присоединяются к списку неизвестных задачи Z; все вхождения θ в посылки задачи Z (как явные, так и найденные "неявные") заменяются на их обозначения x и вводятся новые посылки "равно (θx)" задачи Z.

Перейдем к описанию приемов, связанных с логическими константами, логическими связками, кванторами и равенством. Для обозначения логических символов используются далее выделенные кавычками слова либо вспомогательные словообразования (записываемые без пробелов). Логическая константа "истина" удаляется из списков посылок и условий (если условие "истина" – единственное, то выдается ответ). Ряд приемов осуществляет упрощение утверждений, в которых логическая константа "истина" либо "ложь" оказывается операндом логической связки либо квантора. Если константа "ложь" является посылкой задачи на доказательство либо на описание, то выдается ответ "истина"; если же – "ложь" – посылка блока анализа Z задачи на описание Z^*, то она присоединяется к списку условий задачи Z^*, после чего возобновляется решение последней. При появлении условия "ложь" задачи Z на доказательство либо на описание ответ на эту задачу определяется в зависимости от того, противоречива ли совокупность ее посылок. Для установления непротиворечивости множества посылок предпринимается попытка решения задачи на описание, подбирающей пример таких значений свободных переменных посылок, при которых они истинны. Если это удается сделать, то вводится общий комментарий к посылкам задачи Z и некоторых внешних задач, списки посылок которых те же, что у Z, позволяющий впоследствии усматривать непротиворечивость посылок без решения вспомогательных задач, и выдается ответ либо отказ на Z. Если же непротиворечивость посылок задачи Z установить не удается, то предпринимается попытка доказать их противоречивость, после чего также выдается ответ либо отказ на Z. Посылки возникающей при этом вспомогательной задачи на доказательство снабжаются общим комментарием, блокирующим повторение описанной процедуры в случае, если ее условие окажется преобразованным к виду "ложь". В некоторых случаях при обращении к задаче на описание вводятся специальные цели, отменяющие указанный анализ непротиворечивости ее посылок и ускоряющие выдачу отказа либо ответа "ложь" при обнаружении условия "ложь".

Приемы логического символа "и" осуществляют замену посылок, а в случае задачи на описание – также условий, имеющих вид "и $(f_1 \ldots f_n)$", на совокупность утверждений $f_1, \ldots, f_n$. Если условие задачи на доказательство имеет вид "и $(f_1 \ldots f_n)$", то она сводится к последовательности задач, имеющих условия $f_1, \ldots$ $\ldots, f_n$. Во избежание повторного вывода следствий из одних и тех же посылок здесь предусмотрено перенесение в список посылок задачи с условием f_i наиболее интересных посылок, полученных при решении задачи с условием f_{i-1} (в частности, быть может, самого условия f_{i-1}). Осуществляется приведение конъюнкций к стандартному виду, основанное на преобразованиях

"и $(f_1 \ldots f_{i-1}$ и $(g_1 \ldots g_m) f_{i+1} \ldots f_n)$" $\to$ "и $(f_1 \ldots f_{i-1} g_1 \ldots g_m f_{i+1} \ldots f_n)$";

"и $(\ldots f \ldots$ не $(f) \ldots)$" $\to$ "ложь",

удалении повторяющихся конъюнктивных членов и лексикографическом упорядочении операндов.

При стандартизации утверждений вида "не (f)" применяются преобразования, заносящие отрицание под знак логической связки либо квантора. Если условие задачи имеет вид "не (f)", где f – посылка, либо посылка имеет вид "не (f)", а f – условие, то условие заменяется на логическую константу "ложь". В тех случаях, когда задача на доказательство имеет условие вида "не (f)" и посылку вида "не (g)", предусматривается замена указанной посылки на f, а условия – на g.

Аналогично конъюнкциям, дизъюнкции приводятся к стандартному виду при помощи преобразований

"или $(f_1 \ldots f_{i-1}$ или $(g_1 \ldots g_m) f_{i+1} \ldots f_n)$" $\to$ "или $(f_1 \ldots f_{i-1} g_1 \ldots g_m f_{i+1} \ldots$ $\ldots f_n)$";

"или $(\ldots f \ldots$ не $(f) \ldots)$" $\to$ "истина"

(последнее преобразование не применяется для посылок, введенных с целью разбора случаев по f), удаления повторяющихся дизъюнктивных членов и лексикографического упорядочения операндов. Если задача на доказательство имеет условие "или $(f_1 \ldots f_n)$", то оно заменяется на f_n, а отрицания утверждений $f_1, \ldots,$ $\ldots, f_{n-1}$ заносятся в список посылок.

С каждым вхождением α в задачу утверждения либо выражения A естественным образом связываются вхождения в задачу утверждений f, в предположении истинности которых рассматривается A. Множество таких утверждений f будем далее называть контекстом вхождения α. Если в контексте некоторого вхождения утверждения f вида "или $(f_1 \ldots f_n)$" встречается утверждение "или $(f_{i_1} \ldots f_{i_m})$"; $\{ i_1, \ldots, i_m \} \subseteq \{ 1, \ldots, n \}$; $m \geqslant 1$, то f заменяется на "истина"; если же в этом контексте имеется отрицание некоторого f_i, то – на "или $(f_1 \ldots f_{i-1} f_{i+1} \ldots f_n)$". В том случае, когда два различных дизъюнктивных члена имеют общий конъюнктивный член, предпринимается вынесение его за скобку.

Посылка "или $(f_1 \ldots f_n)$" задачи Z на доказательство, описание либо преобразование может быть использована для решения этой задачи путем разбора "случаев" $f_1, \ldots, f_n$. Число рассматриваемых здесь случаев f_j иногда удается сократить, замечая, что одни из них, в силу симметричности условий и посылок задачи, сводятся к другим. Специальный прием, усматривая сводимость случая f_j к случаю f_i при некотором переобозначении переменных, регистрирует эту информацию в виде комментария к данной посылке; приемы, реализующие разбор случаев и использующие эту информацию, включаются лишь после анализа симметрии. Если при решении задачи на описание не удается найти ответ для всех случаев f_i $(i = 1, \ldots, n)$, причем требуется получить полный ответ, то выдается отказ, иначе – формируется частичный ответ на основе тех случаев f_i, в которых задача была решена.

Для упрощения дизъюнкций "или $(f_1 \ldots f_n)$", не используемых при разборе случаев, а также до разбора случаев, если он происходит при достаточно большом значении текущего уровня задачи, применяются приемы, один из которых пытается доказать отрицание утверждения f_i как следствие отрицаний $f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n$ и затем удалить f_i, а другой – доказать f_i как следствие этих отрицаний и заменить дизъюнкцию на константу "истина".

Если задача на описание Z имеет условие f вида "или $(f_1 \ldots f_n)$", то предпринимается попытка получения ответа на нее путем объединения частичных ответов, возникающих при решении вспомогательных задач Z'_i, получаемых из Z заменой условия f на f_i $(i = 1, \ldots, n)$. В том случае, когда требуется найти полный пример, накопление частичных ответов происходит до тех пор, пока дизъюнкция накладываемых ими ограничений на известные величины не оказывается следствием посылок задачи Z; при этом решается, вообще говоря, лишь часть задач $Z'_1, \ldots, Z'_n$. Так как в этой ситуации могут быть достаточны даже частичные ответы на задачи $Z'_1, \ldots, Z'_n$, они решаются с целью получения частичных примеров. Прием включается сначала при сравнительно небольшом значении текущего уровня задачи, затем – при "среднем" значении, и наконец – при большом. Максимальный уровень задачи Z'_i полагается здесь равным текущему уровню задачи Z. Такое постепенное увеличение уровня средств, привлекаемых для рассмотрения отдельных случаев f_i, предотвращает неоправданно долгое рассмотрение одного случая в тех ситуациях, когда ответ легко усматривается из другого.

Если при выводе следствий в блоке анализа Z задачи на описание Z^* появляется дизъюнктивная посылка, содержащая неизвестные, то она переносится в список условий задачи Z^* и инициирует соответствующий разбор случаев.

Квантор общности обозначается логическим символом "для любого" и служит для образования утверждений вида

"для любого $(x_1 \ldots x_m f_1)$",

"для любого $(x_1 \ldots x_m$ если $f_1 \ldots f_{n-1}$ то $f_n)$",

где $x_1, \ldots, x_m$ – различные переменные; $f_1, \ldots, f_n$ – утверждения. Логические символы "если", "то" в других ситуациях не используются, так как импликации, не связанные с квантором общности, удобнее представлять в дизъюнктивной форме. Для приведения к стандартному виду кванторной импликации "для любого $(x_1 \dots x_m$ если $f_1 \dots f_{n-1}$ то $f_n)$",

где $n \geqslant 1$, применяются следующие преобразования:

1) из кванторной приставки исключаются переменные x_i, не имеющие свободных вхождений в $f_1, ..., f_n$;

2) если один из антецедентов $f_1, ..., f_{n-1}$ является отрицанием другого либо совпадает с консеквентом f_n, то импликация заменяется на константу "истина";

3) если отрицание антецедента совпадает с консеквентом, то этот антецедент удаляется;

4) если антецедент f_i $(1 \leqslant i \leqslant n - 1)$ имеет вид "и $(g_1 \dots g_k)$", то он заменяется на группу новых антецедентов $g_1, ..., g_k$;

5) если антецедент f_i имеет вид "или $(g_1 \dots g_k)$", то импликация заменяется на конъюнкцию импликаций

"для любого $(x_1 \dots x_m$ если $f_1 \dots f_{i-1} g_j f_{i+1} \dots f_{n-1}$ то $f_n)$"; $j = 1, ..., k$;

6) если консеквент f_n имеет вид "и $(g_1 \dots g_k)$", то импликация заменяется на конъюнкцию импликаций

"для любого $(x_1 \dots x_m$ если $f_1 \dots f_{n-1}$ то $g_j)$"; $j = 1, ..., k$;

7) если консеквент f_n имеет вид "или $(g_1 \dots g_k)$", то он заменяется на g_k, а отрицания утверждений $g_1, ..., g_{k-1}$ присоединяются к списку антецедентов;

8) если утверждения $f_1, ..., f_n$ допускают такое разбиение на непересекающиеся подмножества, при котором никакая переменная x_i, $i = 1, ..., m$, не встречается в утверждениях из различных подмножеств, то осуществляется преобразование исходной импликации в дизъюнкцию импликаций (возможно, вырожденных), соответствующих отдельным подмножествам разбиения;

9) если консеквент f_n имеет вид "не (g_1)", а некоторый антецедент f_i $(1 \leqslant i \leqslant \leqslant n - 1)$ – вид "не (g_2)", то f_i заменяется на g_1, а f_n – на g_2;

10) если некоторый антецедент имеет вид "равно $(x_i \theta)$" либо консеквент – вид "не (равно $(x_i \theta)$)", где $1 \leqslant i \leqslant m$ и θ не содержит x_i, то применяется преобразование, исключающее связанную переменную x_i путем "фиксации" ее значения θ;

11) исключаются повторные вхождения антецедентов и происходит их лексикографическое упорядочение.

Для того чтобы рассмотрение утверждений, принадлежащих контексту некоторого вхождения в задачу, не требовало постоянного переобозначения входящих в них переменных, связанных внешними по отношению к данному утверждению кванторами либо другими "связывающими" символами φ, такое переобозначение связанных переменных делается один раз, сразу же при появлении в задаче нового вхождения связывающего символа φ (в том числе символа "для любого"). Целью переобозначения связанных переменных $x_1, ..., x_k$, входящих в подтерм $\varphi(x_1 \dots x_k \theta_1 \dots \dots \theta_n)$ посылки либо условия f задачи Z, является при этом удовлетворение следующих требований:

а) переменные $x_1, ..., x_k$ не имеют свободных вхождений в посылки (а если f – условие, то и в условия) задачи и связанных вхождений в термы $\theta_1, ..., \theta_n$;

б) переменные $x_1, ..., x_k$ отличны от неизвестных;

в) переменные $x_1, ..., x_k$ не имеют связанных вхождений в левые части равенств, содержащихся в посылках (а если f – условие, то и в условиях) задачи;

г) если рассматриваемое вхождение подтерма $\varphi(x_1 \dots x_k \theta_1 \dots \theta_n)$ расположено внутри левой части входящего в f равенства, то $x_1, ..., x_k$ не входят в отличные от f условия и посылки задачи.

Требования в), г) нужны здесь для того, чтобы в процессе реализации различных приемов можно было рассматривать без дополнительного переобозначения связанных переменных контексты вхождений в условия и посылки, модифицированные (только на период выполнения приема) путем применения замен $t_2 \to t_1$, где "равно $(t_1\ t_2)$" – посылка либо условие задачи.

Если "для любого $(x_1 \ldots x_k$ если $f_1 \ldots f_{n-1}$ то $f_n)$" – условие задачи на доказательство, то происходит замена его на f_n и присоединение утверждений $f_1, \ldots, f_{n-1}$ к посылкам задачи.

Во многих случаях существенного упрощения условий и посылок задачи удается добиться путем расшифровки по определениям ("развертки") входящих в них понятий, логического упрощения результата такой развертки и последующей "свертки" по определениям. Такая процедура (для распознавания целесообразности ее применения существует специальная база признаков) повторяется в процессе решения задачи несколько раз, с постепенным увеличением уровня привлекаемых средств, причем для управления переключением фаз "развертки" и "свертки" используются специальные комментарии к соответствующим посылкам и условиям. В процессе "развертки" кванторная импликация

"для любого $(x_1 \ldots x_k$ если $g_1 \ldots g_{i-1}$ существует $(y_1 \ldots y_m g_i')g_{i+1} \ldots g_{n-1}$ то $g_n)$"

преобразуется к виду

"для любого $(x_1 \ldots x_k y_1 \ldots y_m$ если $g_1 \ldots g_{i-1} g_i' g_{i+1} \ldots g_{n-1}$ то $g_n)$",

а кванторная импликация

"для любого $(x_1 \ldots x_k$ если $g_1 \ldots g_{n-1}$ то для любого $(y_1 \ldots y_m$ если $f_1 \ldots f_{p-1}$, то $f_p))$"

– к виду

"для любого $(x_1 \ldots x_k y_1 \ldots y_m$, если $g_1 \ldots g_{n-1} f_1 \ldots f_{p-1}$, то $f_p)$"

(при $n = 1$ последнее преобразование применяется безотносительно к тому, происходит ли "развертка"); в процессе "свертки" реализуются обратные преобразования.

Если кванторная импликация f вида

"для любого $(x_1 \ldots x_k$ если $f_1 \ldots f_{n-1}$ то $f_n)$", $n \geqslant 2$,

является посылкой задачи, то она используется для вывода следствий в списке посылок этой задачи. Сначала (для текущего уровня $\mathcal{U}_0$) предпринимается попытка найти в посылках задачи утверждения вида $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_i$ $(i = 1, \ldots, n-1)$ и вывести из них утверждение $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_n$. Затем (для текущего уровня $\mathcal{U}_1 > \mathcal{U}_0$) аналогичным образом находятся следствия вида $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_j'$ посылок

$$S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_1, \ldots, S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_{j-1}, S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_n', S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_{j+1}, \ldots, S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_{n-1}$$

(f_j', f_n' – отрицания утверждений f_j, f_n). Наконец, для текущего уровня $\mathcal{U}_2 > \mathcal{U}_1$ при поиске среди посылок задачи утверждений, представимых в виде $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_i$ ($S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_n'$), осуществляются вспомогательные преобразования посылок, приводящие их к такому виду. В список посылок заносятся не все извлеченные указанным образом при помощи утверждения f следствия, а лишь те из них, которые по ряду признаков оцениваются как интересные для решения задачи. Если f находится в фазе "свертки", причем число извлеченных при помощи f следствий не очень велико, а текущий уровень достаточно высок, то для извлечения новых следствий из f решаются вспомогательные задачи на описание, позволяющие определить такие выражения $\theta_1, \ldots, \theta_k$, что $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_1, \ldots, S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_{n-1}$ суть следствия посылок задачи, а $S_{\theta_1 \ldots \theta_k}^{x_1 \ldots x_k} f_n$ отличается от ранее полученных следствий из f.

Если посылка задачи на доказательство представляет собой кванторную импликацию и ее не удается путем преобразований "свертки", привлекая средства достаточно большого уровня, перевести в элементарное утверждение, то осуществляется "перестановка" этой посылки f и условия задачи g: отрицание g присоединяется к посылкам, а утверждение f удаляется из посылок, и его отрицание становится условием задачи.

При решении задачи на описание применяется прием, усматривающий в посылках задачи кванторную импликацию f, консеквент которой допускает унификацию с условием g_0 (либо консеквентом условия g_0) этой задачи, определяющий на основе f совокупность утверждений $g_1, ..., g_n$, следствием которых является g_0, и предпринимающий попытку решить задачу путем замены g_0 на $g_1, ..., g_n$. Этот прием срабатывает только для достаточно больших значений текущего уровня и при выполнении ряда дополнительных требований, связанных с распознаванием реализуемости новых условий $g_1, ..., g_n$.

Для приведения к стандартному виду утверждений

"существует $(x_1 ... x_k f)$"

применяются преобразования, аналогичные преобразованиям 1, 8, 10, 11 кванторных импликаций (в случаях 8, 10 здесь f представляется как "и $(f_1 ... f_n)$", а также переобозначение связанных переменных. Кроме того, если f имеет вид "или $(f_1 ... f_n)$", то указанное утверждение преобразуется к виду дизъюнкции утверждений

"существует $(x_1 ... x_k f_j)$"; $j = 1, ..., n$.

Аналогичное преобразование выполняется, если дизъюнкцией является конъюнктивный член утверждения f, причем имеет место фаза "развертки". При отсутствии "развертки" применяется обратное преобразование частичной склейки дизъюнктивных членов A_i, A_j утверждения "или $(A_1 ... A_r)$", имеющих, соответственно, такие конъюнктивные члены

"существует $(x_1 ... x_k$ и $(g_1 ... g_m))$",

"существует $(y_1 ... y_s$ и $(g'_1 ... g'_p)$",

что некоторое g'_{i_1} получается из g_{i_0} переименованием без отождествлений переменных $x_1, ..., x_k$ на переменные списка $y_1, ..., y_s$.

Если посылка f вида

"существует $(x_1 ... x_k$ и $(g_1 ... g_m))$"

задачи Z такова, что $g_1, ..., g_m$ — элементарные утверждения, $m \geqslant 1$, причем из других элементарных посылок непосредственно усматривается истинность утверждений $S^{x_1 ... x_k}_{\theta_1 ... \theta_k} g_1, ..., S^{x_1 ... x_k}_{\theta_1 ... \theta_k} g_m$ при некоторых $\theta_1, ..., \theta_k$, то f исключается из списка посылок задачи Z. Если же последнее требование не выполняется, причем Z — либо задача на доказательство, либо задача на описание, цели которой допускают появление новых переменных в посылках и ответе, то выбираются не используемые в задаче Z переменные $y_1, ..., y_k$, и посылка f заменяется на группу утверждений $S^{x_1 ... x_k}_{y_1 ... y_k} g_j; j = 1, ..., m$.

Задача на доказательство, имеющая условие f вида

"существует $(x_1 ... x_k g)$",

сводится к задаче на описание, имеющей единственное условие g, неизвестные $x_1, ..., x_k$ и решаемой с целью получения полного примера. Аналогичным образом происходит сведение задачи Z на описание, имеющей условие f указанного вида, к другой задаче на описание, у которой условие f заменено на g, а переменные $x_1, ...$ $..., x_k$ присоединены к списку неизвестных. Если на последнюю задачу получен ответ h, то ответ на задачу Z получается применением к утверждению

"существует $(x_1 ... x_k h)$"

процедуры редактирования ответа. Неизвестные $x_1, ..., x_k$ задачи на описание, вводившиеся в двух последних приемах, являются вспомогательными в том смысле, что при решении задачи нет необходимости явно определять их допустимые значения, а достаточно лишь установить существование таких значений. Если в процессе решения удается усмотреть, что группа содержащих некоторую такую переменную x_i условий представляет собой "стандартное" описание непустого множества ее значений, то все эти условия можно удалить, а переменную x_i исключить из списка неизвестных задачи. Информация о том, что переменные $x_1, \ldots, x_k$ являются вспомогательными неизвестными, регистрируется в списке целей задачи и учитывается при ее решении.

Полученная при выводе следствий посылка

"существует $(x_1 \ldots x_k g)$"

блока анализа Z задачи на описание Z', содержащая невырожденную информацию о неизвестных задачи Z', переносится в список условий задачи Z', где инициирует указанное выше удаление квантора существования.

В режиме "развертки" происходит замена утверждений вида

"существует $(x_1 \ldots x_k$ и $(g_1 \ldots g_{i-1}$ существует $(y_1 \ldots y_n$ и $(h_1 \ldots h_s))g_{i+1} \ldots$ $\ldots g_m))$"

на

"существует $(x_1 \ldots x_k y_1 \ldots y_n$ и $(g_1 \ldots g_{i-1} h_1 \ldots h_s g_{i+1} \ldots g_m))$"

(при $m = 1$ – безотносительно к наличию режима "развертки"); в режиме "свертки" реализуется обратное преобразование.

Если условие f задачи Z на описание имеет вид "для любого $(x_1 \ldots x_k$ если $f_1 \ldots f_{n-1}$ то существует $(y_1 \ldots y_m f_n))$", причем не требуется получить полное описание множества допустимых наборов значений неизвестных, то предпринимается попытка подобрать для произвольных, удовлетворяющих утверждениям $f_1, \ldots, f_{n-1}$ значений переменных $x_1, \ldots, x_k$ конкретные, удовлетворяющие утверждению f_n значения переменных $y_1, \ldots, y_m$ (решаются вспомогательные задачи на описание с условием f_n, неизвестными $y_1, \ldots, y_m$ и посылками $f_1, \ldots, f_{n-1}$, объединенными с посылками задачи Z, отличными от f ее условиями и некоторыми другими утверждениями; целью решения этих задач является получение частичных примеров), и далее решается задача, полученная из Z заменой f на утверждение h, описывающее условия, при которых такой подбор оказался возможным. Попытки подбора значений переменных $y_1, \ldots, y_m$ предпринимаются, вообще говоря, неоднократно (каждый раз вводится дополнительная посылка, указывающая, что подбор происходит в той ситуации, для которой ранее значения $y_1, \ldots, y_m$ найти не удавалось), в результате чего утверждение h "накапливается" в виде дизъюнкции утверждений $h_1, \ldots, h_s$, полученных при отдельных попытках.

Перейдем к описанию основных приемов, связанных с равенством. Некоторые из имеющихся здесь приемов возникли в связи с процедурами решения задач по элементарной алгебре и элементарной геометрии (выделение подсистем из n уравнений с n неизвестными и т. п.); их рассмотрение мы отложим до описания соответствующих процедур.

Простейшие приемы символа "равно" заменяют на логическую константу "истина" утверждение вида "равно (tt)"; выдают ответ задачи на описание, имеющей единственное условие "не (равно (xt))", где x – неизвестная, не входящая в t, и выдают отказ при появлении такого условия "равно (xt)" либо "не (равно (xt))" задачи на доказательство, что переменная x не входит в t и в посылки задачи. Для произвольной посылки f задачи Z будем называть зоной этой посылки совокупность всех отличных от f посылок задачи Z и всех ее условий. Аналогично, зоной условия f задачи Z называем совокупность всех отличных от f условий этой задачи. Если утверждение f вида "равно $(\theta_1 \theta_2)$" оказывается посылкой задачи либо условием задачи на описание, то в рамках этой задачи возникают различные обозначения θ_1, θ_2 одного и того же объекта, и становится целесообразной унификация этих обозначений. Такая унификация осуществляется путем замены вхождений выражения θ_1 в принадлежащие зоне равенства f утверждения на выражение θ_2. Чтобы направление замен, осуществляемых в процессе унификации обозначений, согласовывалось с рядом соображений эвристического характера (например, во многих случаях целесообразна замена выражения, содержащего неизвестные, на выражение, не содержащее их; более длинного выражения – на более короткое и т. п.), выполняется предварительная "нормализация" равенства f, заключающаяся в возможной перестановке его левой и правой частей. Для блокировки унификации обозначений в тех случаях, когда она нежелательна, используются специальные комментарии посылок и условий.

Если утверждение f вида "равно (xt)" либо "равно (tx)" является условием задачи Z на описание, причем x – неизвестная этой задачи, не входящая в t, то осуществляется сведение задачи Z к задаче Z', получающейся из Z удалением условия f, подстановкой t вместо x во все оставшиеся условия и удалением x из списка неизвестных (в случае, когда f – единственное условие, сразу же выдается ответ). При появлении в посылках блока анализа Z задачи на описание Z' равенства, выражающего неизвестную задачи Z' через прочие ее переменные, происходит перенесение этого равенства в список условий задачи Z'.

Непосредственное усмотрение истинности либо ложности утверждения "равно $(t_1 t_2)$" происходит в тех случаях, когда выражения t_1, t_2 суть стандартные обозначения (имена) некоторых объектов, а также на основе сравнения типов (множество, функция, число и т. п.) значений этих выражений. Для некоторых таких типов предусмотрены процедуры, устанавливающие истинность либо ложность равенства путем просмотра утверждений, относящихся к контексту его вхождения в задачу. Расшифровка по определению равенства осуществляется в процессе "развертки", в зависимости от типа объектов, являющихся значениями его левой и правой частей.

Если условие f некоторой задачи имеет вид "не (равно $(t_1 t_2)$)", то может быть предпринята попытка решить уравнение "равно $(t_1 t_2)$" относительно некоторой переменной x, входящей в него, и заменить f на отрицание найденного ответа. Аналогично, если посылка f имеет вид "равно $(t_1 t_2)$" либо является отрицанием утверждения такого вида, то предпринимается попытка (при наличии ряда признаков, устанавливающих ее целесообразность) решения данной посылки как уравнения относительно некоторого входящего в f выражения θ, рассматриваемого в качестве неизвестной.

Если утверждение g имеет вид "или $(g_1 \ldots g_n)$", причем при некоторых i, j из $\{1, \ldots, n\}$, $i \neq j$, g_i и g_j имеют конъюнктивные члены вида

"существует $(x_1 \ldots x_k A)$", "существует $(y_1 \ldots y_m B)$" $(k \geqslant 0, m \geqslant 0)$

соответственно, и некоторые конъюнктивные члены утверждений A, B суть равенства

"равно $(\theta_1 \theta_2)$", "равно $(\theta_2 \theta_3)$",

где θ_1 не зависит от x_1, ..., x_k (а тогда, если текущий уровень достаточно велик и уже произошли необходимые применения процедуры переобозначения связанных переменных, не зависит также и от y_1, ..., y_m, то предпринимается попытка выделить в θ_2 и θ_3 такую не зависящую от x_1, ..., x_k "общую часть" θ_4, что для некоторых τ_1, τ_2 и новой переменной x, входящей в θ_4, выражения $S^x_{\tau_1}\theta_4$ и $S^x_{\tau_2}\theta_4$ получаются из θ_2, θ_3 эквивалентными преобразованиями. После этого (при отсутствии фазы "развертки") дизъюнкция утверждений g_i, g_j заменяется на

"существует (x и (равно $(\theta_1 \theta_4)$ или $(g'_i g'_j)$))",

где g'_i, g'_j получены из g_i, g_j заменой рассматриваемых равенств на

"равно $(x\tau_1)$", "равно $(x\tau_2)$".

В тех случаях, когда число неизвестных задачи на описание более одной, предпринимаются попытки использовать уравнения, входящие в ее список условий либо в список посылок ее блока анализа, для получения явного выражения одной из неизвестных задачи через другие.

§ 3. Приемы решения задач по алгебре множеств

Приведем краткое описание приемов, осуществляющих решение некоторого класса задач по алгебре множеств. Здесь будут рассматриваться атомарные утверждения вида "множество (t)", "принадлежит $(t_1 t_2)$", "содержится $(t_1 t_2)$" $(t_1 \subseteq t_2)$, а также выражения вида "пусто" (константа ϕ), "объединение $(t_1 \ldots t_n)$", "пересечение $(t_1 \ldots t_n)$" $(n \geqslant 2)$, "разность $(t_1 t_2)$", "булеан (t)" (множество всех подмножеств t), "объединение всех $(x_1 \ldots x_n f t)$" (объединение всех множеств, являющихся значениями выражения t для таких наборов значений переменных $x_1, \ldots, x_n$, что истинно утверждение f), "пересечение всех $(x_1 \ldots x_n f t)$" (пересечение всех множеств, являющихся значениями выражения t для удовлетворяющих f значений $x_1, \ldots \ldots x_n$), "класс $(x f)$" (множество всех значений переменной x, для которых истинно утверждение f), "семейство $(x_1 \ldots x_n f t)$" (множество всех значений выражения t для удовлетворяющих f значений $x_1, \ldots, x_n$), "перечень $(t_1 \ldots t_n)$" (конечное множество, элементы которого суть значения выражений $t_1, \ldots, t_n$), "набор $(t_1 \ldots t_n)$" (упорядоченный набор значений выражений $t_1, \ldots, t_n$), "прямое произведение $(t_1 \ldots \ldots t_n)$" $(t_1 \times \ldots \times t_n)$, "элемент (t)" (некоторый элемент непустого множества t), "внешний элемент (t)" (некоторый элемент, не принадлежащий множеству t).

Если в процессе решения задачи на описание множество всех условий, содержащих некоторую неизвестную x, оказывается стандартным описанием значений этой неизвестной, то осуществляется исключение неизвестной x (а если она была единственной, то – выдача ответа). Стандартными описаниями при этом считаются следующие множества утверждений:

$\{$"множество (x)"$\}$, $\{$"равно (пересечение $(t_1 \ldots t_{i-1} x t_{i+1} \ldots t_n)$ пусто)"$\} \cup A$,

$\{$"содержится $(x t_1)$"$\} \cup A$, $\{$"содержится $(t_0 x)$"$\} \cup B$,

$\{$"принадлежит $(x t_0)$"$\}$.

Здесь

$A \subseteq \{$"множество (x)", "содержится $(t_0 x)$"$\}$;

$B \subseteq \{$"множество (x)"$\}$; $t_0, t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n$

– выражения, не содержащие x. При необходимости в ответ включается условие на значения известных величин, выражающее "совместность" соответствующего стандартного описания. Если задача на описание имеет более одной неизвестной и "множество (x)" – ее условие, а x – неизвестная, то предпринимается попытка исключения этой неизвестной, основанная на выделении совокупности всех условий, содержащих x, и извлечения из нее явного описания для x через значения прочих неизвестных.

Ряд приемов осуществляет упрощение выражений на основе известных тождеств, а также непосредственное усмотрение истинности либо ложности утверждений вида

"множество (t)", "принадлежит $(t_1 t_2)$", "содержится $(t_1 t_2)$"

из контекста их вхождения в задачу. Для приведения утверждений к стандартному виду используются замены:

"содержится (t пусто)" $\to$ "равно (t пусто)";

"равно (объединение $(t_1 \ldots t_n)$ пусто)" $\to$ "и (равно (t_1 пусто) $\ldots$ равно (t_n пусто))";

"равно (разность $(t_1 t_2)$ пусто)" $\to$ "содержится $(t_1 t_2)$"

и тому подобные упрощающие преобразования.

Несколько приемов связаны с приведением к стандартному виду кванторных импликаций f:

"для любого $(x_1 \ldots x_k$ если $f_1 \ldots f_n$ то $f_0)$".

Так, если f_0 имеет вид "не (принадлежит $(x_i t)$)", $i \in \{1, \ldots, k\}$, а некоторое f_j, $j \in \{1, \ldots, n\}$, не имеет вида "принадлежит $(x_s \theta)$", где $s \in \{1, \ldots, k\}$, то f_0 заменяется на отрицание f_j, а f_j – на "принадлежит $(x_i t)$". Если из истинности утверждений

$f_1, ..., f_n$ усматривается, что значением некоторой переменной x_i, $i \in \{1, ..., k\}$, является множество, причем истинность утверждения f_0 при произвольном значении $x_i = A$ влечет истинность этого утверждения при всех значениях x_i, включающих A, то предпринимается попытка определить наименьшее по включению множество A_0, удовлетворяющее условиям $f_1, ..., f_n$ (при $x_i = A_0$), и f заменяется на утверждение вида

"для любого ($x_1 ... x_k$ если $f_{j_1} ... f_{j_m}$ равно ($x_i \widetilde{A_0}$) то f_0)".

Здесь $\widetilde{A_0}$ – выражение, значением которого (при истинных $f_1, ..., f_n$) является указанное множество A_0; список $f_{j_1}, ..., f_{j_m}$ получен из $f_1, ..., f_n$ удалением ряда утверждений, оказывающихся избыточными при добавлении утверждения "равно ($x_i \widetilde{A_0}$)". Аналогичные действия предпринимаются, если из истинности f_0 при $x_i = A$ вытекает истинность этого утверждения при всех x_i, являющихся подмножествами A, но вместо "наименьшего" множества A_0 берется "наибольшее" множество.

В процессе решения задач на описание выполняются эквивалентные преобразования, приближающие вхождения неизвестных к корню условия. Таковы, например, преобразования:

"равно (пересечение (t_1 ... t_{i-1} объединение ($\theta_1 ... \theta_k$) t_{i+1} ... t_n) пусто)" → → "и (равно (пересечение (t_1 ... $t_{i-1}\, \theta_1\, t_{i+1}$ ... t_n) пусто) ... равно (пересечение (t_1 ... $t_{i-1}\, \theta_k\, t_{i+1}$ ... t_n) пусто))";

"равно (пересечение (t_1 ... t_{i-1} разность ($\theta_1 \theta_2$) t_{i+1} ... t_n) пусто)" → "содержится (пересечение (t_1 ... $t_{i-1}\, \theta_1\, t_{i+1}$ ... t_n) θ_2)";

"содержится (t пересечение ($\theta_1 ... \theta_k$))" → "и (содержится ($t\theta_1$) ... содержится ($t\theta_k$))";

"содержится (пересечение (t_1 ... t_{i-1} объединение ($\theta_1 ... \theta_k$) t_{i+1} ... t_n) t_0)" → → "и (содержится (пересечение (t_1 ... $t_{i-1}\, \theta_1\, t_{i+1}$ ... t_n) t_0) ... содержится (пересечение (t_1 ... $t_{i-1}\, \theta_k\, t_{i+1}$ ... t_n) t_0))",

и ряд других. В каждом из указанных преобразований предполагается, что хотя бы одно из выражений θ_i содержит неизвестную задачи. При выполнении этих и некоторых других преобразований допускается предварительное применение к рассматриваемому условию таких замен $\tau_1 \to \tau_2$, что утверждение "равно ($\tau_2 \tau_1$)" принадлежит контексту вхождения данного условия в задачу. Для того чтобы отделить выражения, содержащие неизвестные, от выражений, не содержащих неизвестные, используются преобразования:

"содержится (t объединение ($\theta_1 ... \theta_k$))" → "содержится (разность (t объединение ($\theta_{i_1} ... \theta_{i_s}$)) объединение ($\theta_{j_1} ... \theta_{j_{k-s}}$))";

"содержится (пересечение ($\theta_1 ... \theta_k$) t)" → "равно (пересечение ($\theta_{j_1} ... \theta_{j_{k-s}}$ разность (пересечение ($\theta_{i_1} ... \theta_{i_s}$) t)) пусто)"

и т. п. Здесь $\theta_{i_1}, ..., \theta_{i_s}$ – все не содержащие неизвестные выражения списка $\theta_1, \ldots, \theta_k$; $s < k$; $\{j_1, \ldots, j_{k-s}\} = \{1, \ldots, k\} \setminus \{i_1, \ldots, i_s\}$; t не содержит неизвестные.

Если условие задачи на описание имеет вид "принадлежит ($t_1 t_2$)" либо "не (принадлежит ($t_1 t_2$))", где t_1 не содержит неизвестные, а t_2 – содержит, то оно переформулируется в терминах включения в t_2 одноэлементного множества "перечень (t_1)".

Несколько различных условий, имеющих вид явного описания значений одной и той же неизвестной x, могут быть объединены в одно условие такого вида. Для этого используются преобразования:

"содержится ($x t_1$)", "содержится ($x t_2$)" → "содержится (x пересечение ($t_1 t_2$))";

"равно (пересечение (t_1 ... $t_{i-1}\, x\, t_{i+1}$ ... t_n) пусто)", "содержится ($x t_0$)" → → "содержится (x разность (t_0 пересечение (t_1 ... $t_{i-1}\, t_{i+1}$ ... t_n)))";

"принадлежит ($x t_1$)", "принадлежит ($x t_2$)" → "принадлежит (x пересечение ($t_1 t_2$))"

и другие. В перечисленных преобразованиях выражения t_0, t_1, ..., t_n не зависят от неизвестных, причем последнее преобразование не применяется для задач, решаемых с целью получения примера (здесь используется обратное преобразование, так как усмотрение из посылок значения неизвестной удобнее выполнять для расчлененных на простейшие составляющие условий).

Если "простые" средства оказываются недостаточны для решения задачи на описание, в которой требуется найти пример, то применяются замены условий вида "содержится (tx)" либо "содержится (xt)", где x – неизвестная, не входящая в t, на "существует $(y$ и (равно $(x$ объединение $(ty))$ множество $(y)))$", "существует $(y$ и (равно $(x$ пересечение $(ty))$ множество $(y)))$" соответственно, т. е для x вводится параметрическое описание.

Ряд приемов осуществляет преобразования расшифровки по определениям – "развертки", а также обратные к ним преобразования "свертки". Целесообразность применения этих преобразований определяется на основе специальной базы признаков; как правило, расшифровка по определениям выполняется лишь по исчерпании средств, предусматривающих работу с соответствующими понятиями без такой расшифровки. Если значением условия t задачи на преобразование, решаемой с целью упрощения выражения t, является множество, то при достаточно большом текущем уровне этой задачи предпринимается попытка расшифровки по определениям утверждения "принадлежит (xt)" (x – новая переменная), применения к результату тех или иных упрощающих приемов и последующей "свертки" по определениям. В тех случаях, когда получается утверждение вида "принадлежит (xt')" и выражение t' по ряду признаков предпочтительнее, чем t, происходит замена условия t на t'.

В заключение упомянем прием, который при наличии условия вида "не (принадлежит $(t_1 t_2)$)" либо "для любого $(x_1$ ... x_k если f_1 ... f_n то не (принадлежит $(t_1 t_2)$))" задачи на описание Z предпринимает попытку решить эту задачу путем замены утверждения "не (принадлежит $(t_1 t_2)$)" на "равно (t_2 пусто)". Данный прием применяется лишь в тех случаях, когда не требуется получить полный ответ, t_2 зависит от неизвестных и не усматривается опровержение утверждения "равно (t_2 пусто)".

Описанная база знаний позволяет решать целый ряд стандартных задач по алгебре множеств [1]. Сюда относятся задачи на доказательство теоретико-множественных тождеств и включений (например, доказать, что если $A \cap B \cap C = \phi$ и $A \cup$ $\cup\ C \subseteq B$, то $A \cap C = \phi$); задачи на решение теоретико-множественных уравнений (например, найти все множества x, удовлетворяющие системе уравнений $A \setminus x = B$; $A \cup x = C$, или найти все такие множества A, B, C, D, что $A \times B \cup C \times D = (A \cup C) \times$ $\times\ (B \cup D)$, и т. п.), а также задачи на упрощение теоретико-множественных выражений. Присоединение к этой базе знаний простейших приемов логических символов "функция", "значение", "область" (т. е. область определения функции), "образ", "прообраз", "биекция" позволяет решать задачи, связанные с семействами множеств (например, доказать, что для непустого семейства множеств A_{ij}, $i \in I$, $j \in J$

$$\bigcup_{i \in I} \bigcap_{j \in J} A_{ij} \subset \bigcap_{j \in J} \bigcup_{i \in I} A_{ij}),$$

а также задачи на доказательство простых свойств отображений (например, доказать, что функция f является взаимно однозначной тогда и только тогда, когда для любых подмножеств A, B ее области определения выполняется $f(A \cap B) = f(A) \cap$ $\cap\ f(B)$, и т. п.). Пополняя при необходимости базу знаний новыми приемами, можно существенно расширить класс решаемых с ее помощью задач по теории множеств.

Анализ многообразия приемов, выявленных при имитации поведения человека, позволяет сделать ряд общих выводов относительно организации процедур автоматического решения задач. Прежде всего, для эффективного функционирования такие процедуры должны опираться на базу знаний не декларативного (теоремы, постулаты, определения и т. п.), как обычно делается в логическом программировании, а алгоритмического характера. Каждый прием, имеющийся в этой базе знаний, представляет собой небольшую алгоритмическую процедуру, хотя и основанную на одном или нескольких "теоретических" фактах, но возникающую из них как результат более или менее трудоемкого процесса программирования, при котором, в свою очередь, могут решаться различные задачи описания синтаксических структур, и лишь после этого пригодную для "практического" использования. Как правило, большая часть такой процедуры связана с усмотрением неявного вхождения в задачу ситуации, рассматриваемой в теореме, а собственно реализация содержащейся в теореме информации занимает лишь малую ее часть. Во-вторых, разработка сильных в эвристическом отношении решателей задач должна основываться не на тех или иных методах сокращения "перебора", а на методах организации в процессе решения различных градиентных тенденций. При этом процедуры "переборного" характера оказываются необходимы на предварительном этапе, для поиска отдельных преобразований, обеспечивающих градиентный процесс, а при решении задачи должно допускаться лишь сравнительно небольшое ветвление. В-третьих, применение логических преобразований как при наличии ветвления, так и при его отсутствии должно сопровождаться предварительной оценкой их целесообразности на основе признаков, учитывающих целевую направленность процесса и представляющих собой, по существу, предикатные аналоги тестов, используемых в комбинаторно-логических методах распознавания образов.

В заключение автор выражает глубокую благодарность В.Б. Кудрявцеву за поддержку и большую помощь, оказанную при работе.

СПИСОК ЛИТЕРАТУРЫ

1. Лавров И.А., Максимова Л.Л. Задачи по теории множеств, математической логике и теории алгоритмов. – М.: Наука, 1975.

Статья поступила 16.11.90