МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
“ЛЭТИ” ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)

Кафедра алгоритмической математики

Демонстрация различных методов резолюций

Выполнили студенты группы 9308:
Соболев Матвей;
Степовик Виктор.

Распределение работы (команда 9):
Соболев М.С. -- Создание программной реализации и разбор алгоритмов
Степовик В.С. -- Разбор теории, примеров и создание презентации

Ссылки на работы:
Ссылка на документ с методами
Ссылка на презентацию
Ссылка на Git-репозиторий с программой, инструкцией и файлами тестов
Ссылка на книгу Ченя и Ли
Ссылка на учебник С.Н. Позднякова и С.В. Рыбина

Методы резолюций

В данном документе представлена теоретическая часть работы нашей команды. Я, Степовик Виктор, в данном документе разбираю теоретическую часть работы нашей команды: оттолкнувшись от основ (логики высказываний и предикатов) мы плавно окунемся в разбор понятия “метод резолюции” и углубимся в нюансы его применения в разделе “стратегии метода резолюции”.

В этом документе вы сможете ознакомиться с основами и теорией метода резолюции, а в другом документе -- с основной частью нашей работы -- стратегиями и методами резолюций.

Введение

Вторая половина 60-х годов в области искусственного интеллекта выделялась особым увеличением интереса к машинному доказательству теорем. Широкое распространение и интенсивность этого интереса вызываются не только растущим сознанием, что умение делать логические выводы есть неотъемлемая часть человеческого интеллекта, но, возможно, в большей степени являются следствием того статуса, который приобрела техника машинного доказательства теорем в конце 60-х годов. Основы машинного доказательства теорем были заложены Эрбраном в 1930 г. Его метод был неосуществим практически до изобретения электронных вычислительных машин. Только после основополагающей статьи Дж.А.Робинсона в 1965г. и развития метода резолюций были сделаны важные шаги к созданию программ,реализующих доказательство теорем. После 1965г. были предложены многочисленные усовершенствования метода резолюций.

Выходит, что метод резолюций - первый шаг человечества к искусственному интеллекту?

Утверждение, что формула логически следует из формул; мы будем называть теоремой. Рассуждение, устанавливающее, что некоторая теорема верна, т. е. что формула логически следует из других формул, будет называться доказательством этой теоремы. Проблема машинного доказательства теорем состоит в рассмотрении машинных методов для нахождения доказательств теорем.

Есть много задач, которые удобно преобразовать в задачи доказательства теорем. Мы перечислим некоторые из них.

1. В вопросно-ответных системах утверждения могут быть представлены логическими формулами. Тогда, чтобы ответить на вопрос, используя данные факты, мы доказываем, что формула, соответствующая ответу, выводима из формул, представляющих эти факты.

2. В задаче анализа программ мы можем описать выполнение программы формулой A, а условие, что программа закончит работу, другой формулой B. Тогда проверка того, что программа закончит работу, эквивалентно доказательству того, что формула B следует из формулы A.

3. В проблеме изоморфизма графов мы хотим знать, изоморфен ли граф подграфу другого графа. Эта проблема не только представляет математический интерес, эта проблема практическая. Например, структура органического соединения может быть описана графом. Следовательно, проверка того, является ли под структура структуры некоторого органического соединения структурой другого органического соединения, есть проблема изоморфизма. Для ее решения мы можем описать графы формулами. Таким образом, задача может быть сформулирована как задача доказательства того, что формула, представляющая граф, следует из формулы, представляющей другой граф.

4. В проблеме преобразования состояний имеется набор состояний и набор операторов над состояниями. Когда один из операторов применяется к состоянию, получается новое состояние.

Исходя из начального состояния, попытаемся найти последовательность операторов, которая преобразует начальное состояние в некоторое желаемое. В этом случае мы можем описать состояния и правила перехода логическими формулами. Следовательно, преобразование начального состояния в желаемое может рассматриваться как проверка того, что формула, представляющая желаемое состояние, следует из формулы, представляющей как состояния, так и правила перехода.

Из вышепредставленных утверждений можем сделать вывод, что проблема автоматизации доказательств является важной областью в искусственном интеллекте.

Основы метода резолюций

(!) -- В данном пункте даны сведения, необходимые для понимания метода резолюции. Вы можете перейти к следующему пункту, если темы “логика высказываний” и “логика предикатов” вам хорошо знакомы.

Математическая логика рассматривает языки, основная цель которых - обеспечить символизм (систему формальных обозначений) для рассуждений, встречающихся не только в математике, но и в повседневной жизни

Логика высказываний

В логике высказываний нас интересуют утвердительные предложения, которые могут быть истинными (“И”) или ложными (“Л”). каждое такое утвердительное предложение называется высказыванием.
Каждое высказывание будем обозначать символом (заглавной буквой), который принято называть атомарной формулой или атомом (P = “Снег белый”).
Из атомарных формул мы можем строить составные высказывания при помощи логических связок (в логике высказываний их всего пять: не $\sim$; и $\wedge$; или $\vee$; если...то $\longrightarrow$; тогда, и только тогда $\leftrightarrow$)

Определение. Правильно построенные формулы (или короче— формулы) в логике высказываний определяются рекурсивно следующим образом:
1. Атом есть формула.
2. Если G— формула, то (~G) — формула.
3 Если G и H—формулы, то (G/\H), (G\/H), (G—> H) и (G<->H)— формулы.
4. Никаких формул, кроме порожденных применением указанных выше правил, нет.

Отношения атомов в составных высказываниях можно описать следующей таблицей

G	H	$\sim G$	$(G \wedge H)$	$(G \vee H)$	$(G \rightarrow H)$	$(G \leftrightarrow H)$
И	И	Л	И	И	И	И
И	Л	Л	Л	И	Л	Л
Л	И	И	Л	И	И	Л
Л	Л	И	Л	Л	И	И

Данная таблица называется таблицей истинности. Она служит для интерпретации составных высказываний: т.к. каждый атом может быть “И” либо “Л”, то для определённой комбинации составное высказывание также может принимать как “И” так и “Л” значения. Количество комбинаций значений атомарных формул = 2^N, где N - количество атомов в составном высказывании.

Пример:

Истинностная таблица формулы $(P \wedge Q) \to (R \leftrightarrow (\sim S))$

P	Q	R	S	$(\sim S)$	$(P \wedge Q)$	$(R \leftrightarrow (\sim S))$	$(P \wedge Q) \to (R \leftrightarrow (\sim S$
И	И	И	И	Л	И	Л	Л
И	И	И	Л	И	И	И	И
И	И	Л	И	Л	И	И	И
И	И	Л	Л	И	И	Л	Л
И	Л	И	И	Л	Л	Л	И
И	Л	И	Л	И	Л	И	И
И	Л	Л	И	Л	Л	И	И
И	Л	Л	Л	И	Л	Л	И
Л	И	И	И	Л	Л	Л	И
Л	И	И	Л	И	Л	И	И
Л	И	Л	И	Л	Л	И	И
Л	И	Л	Л	И	Л	Л	И
Л	Л	И	И	Л	Л	Л	И
Л	Л	И	Л	И	Л	И	И
Л	Л	Л	И	Л	Л	И	И
Л	Л	Л	Л	И	Л	Л	И

Если составная формула истинна для **любой комбинации значений атомов**, то её называют тавтологией или общезначимой формулой, а если формула ложна - это противоречие.

Методом истинностных таблиц можно доказать, что:

а) $(P \wedge \sim P)$ - противоречива,а значит не общезначима

б) $(P \vee \sim P)$ - общезначима, следовательно не противоречива

в) $(P \longrightarrow \sim P)$ - не общезначима, но и не противоречива

Определение эквивалентности

Говорят, что две формулы F и G эквивалентны или что F эквивалентна G (обозначается F = G), тогда и только тогда, когда истинностные значения F и G совпадают при каждой интерпретации F и G.

Для логики высказываний справедливы следующий набор эквивалентных формул, с помощью которых можно выполнять преобразования.

$F \leftrightarrow G = (F \rightarrow G) \wedge (G \rightarrow F)$

$F \rightarrow G = \sim F \vee G$

(a) $F \vee G = G \vee F$; (b) $F \wedge G = G \wedge F$

(a) $(F \vee G) \vee H = F \vee (G \vee H)$; (b) $(F \wedge G) \wedge H = F \wedge (G \wedge H)$

(a) $F \vee (G \wedge H) = (F \vee G) \wedge (F \vee H)$; (b) $F \wedge (G \vee H) = (F \wedge G) \vee (F \wedge H)$

(a) $F \vee \square = F$; (b) $F \wedge \blacksquare = F$

(a) $F \vee \blacksquare = \blacksquare$; (b) $F \wedge \square = \square$

(a) $F \vee \sim F = \blacksquare$; (b) $F \wedge \sim F = \square$

$\sim (\sim F) = F$

(a) $\sim (F \vee G) = \sim F \wedge \sim G$; (b) $\sim (F \wedge G) = \sim F \vee \sim G$

где $\blacksquare$ - атом истинности (белева 1), $\square$ - атом противоречия (булев 0).

Определение. *Литера* есть атом или отрицание атома.

Определение. Говорят, что формула F находится в *конъюнктивной форме*, тогда и только тогда, когда F имеет вид $F \rightleftarrows F_1 \wedge \ldots \wedge F_n$, $n \geqslant 1$, где каждая из $F_1, F_2, \ldots, F_n$ есть дизъюнкция литер.

Пример 2.6. Пусть P, Q и R — атомы. Тогда $F \rightleftarrows (P \vee \vee \sim Q \vee R) \wedge (\sim P \vee Q)$ есть формула в конъюнктивной нормальной форме. Для этой формулы $F_1 = (P \vee \sim Q \vee R)$ и $F_2 = = (\sim P \vee Q)$. Ясно, что F_1 есть дизъюнкция литер P, $\sim Q$ и R и F_2 есть дизъюнкция литер $\sim P$ и Q.

Определение. Говорят, что формула F находится в *дизъюнктивной нормальной форме*, тогда и только тогда, когда F имеет вид $F \rightleftarrows F_1 \vee \ldots \vee F_n$, $n \geqslant 1$, где каждая из $F_1, F_2, \ldots, F_n$ есть конъюнкция литер.

Логические следствия

Как в математике, так и в обычной жизни нам часто нужно решить, следует ли одно утверждение из другого или нескольких других утверждений. Данная задача приводит к понятию “логического следствия”.

Логическое следствие заключается в доказательстве следования предположения “G” из логической конъюнкции имеющихся утверждений F1,F2...Fn
т.е. составное высказывание (F1/\F2/\.../\Fn -> G) должно быть общезначимо.

Пример:

В этом примере четыре утверждения.

(1) Если предварительные процентные ставки растут, то курс акций падает.

(2) Если курс акций падает, то большинство людей несчастны.

(3) Предварительные процентные ставки растут.

(4) Большинство людей несчастны.

Эти утверждения сначала выразим в следующей символической форме:

(1′) $P \to S$,

(2′) $S \to U$,

(3′) P,

(4′) U.

Мы покажем, что (4′) истинно, как только (1′) $\wedge$ (2′) $\wedge$ (3′) истинно.

Сперва преобразуем формулу $((P \to S) \wedge (S \to U) \wedge P)$ (представляющую (1′) $\wedge$ (2′) $\wedge$ (3′)) в нормальную форму:

$$\begin{aligned}((P \to S) \wedge (S \to U) \wedge P) &= ((\sim P \vee S) \wedge (\sim S \vee U) \wedge P)\\ &= (P \wedge (\sim P \vee S) \wedge (\sim S \vee U))\\ &= (((P \wedge \sim P) \vee (P \wedge S)) \wedge (\sim S \vee U))\\ &= ((\square \vee (P \wedge S)) \wedge (\sim S \vee U))\\ &= (P \wedge S) \wedge (\sim S \vee U)\\ &= (P \wedge S \wedge \sim S) \vee (P \wedge S \wedge U)\\ &= (P \wedge \square) \vee (P \wedge S \wedge U)\\ &= \square \vee (P \wedge S \wedge U)\\ &= P \wedge S \wedge U.\end{aligned}$$

Следовательно, если $((P \to S) \wedge (S \to U) \wedge P)$ истинна, то $(P \wedge S \wedge U)$ истинна. Так как $(P \wedge S \wedge U)$ истинна, только если P, S и U все истинны, мы заключаем, что U истинна. Так как U истинна, как только $(P \to S)$, $(S \to U)$ и P истинны, то U называется в логике *логическим следствием* $(P \to S)$, $(S \to U)$ и P. Более формально мы определяем логическое следствие следующим образом.

Определение. Пусть даны формулы $F_1, F_2, \ldots, F_n$ и формула G. Говорят, что G есть *логическое следствие* формул $F_1, \ldots, F_n$ (или G *логически следует из* $F_1, \ldots, F_n$), тогда и только тогда, когда для всякой интерпретации I, в которой $F_1 \wedge F_2 \wedge \ldots \wedge F_n$ истинна, G также истинна. $F_1, F_2, \ldots, F_n$ называются *аксиомами* (или *постулатами*, или *посылками*) G.

Если G есть логическое следствие формул $F_1, \ldots, F_n$, то формула $((F_1 \ \ldots \ F_n) \to G)$ называется *теоремой*, а G называется также *заключением теоремы*.

Чтобы показать, что логическое следствие истинно для каждой модели формул можно использовать несколько методов:

1)метод таблиц истинности

Истинностная таблица для $((P \to Q) \wedge \sim Q) \to \sim P$

P	Q	$P \to Q$	$\sim Q$	$(P \to Q) \wedge \sim Q$	$\sim P$	$((P \to Q) \wedge \sim Q) \to \sim P$
И	И	И	Л	Л	Л	И
И	Л	Л	И	Л	Л	И
Л	И	И	Л	Л	И	И
Л	Л	И	И	И	И	И

2) Преобразование в конъюнктивную нормальную форму

$$
\begin{aligned}
((P \rightarrow Q) \wedge \sim Q) \rightarrow \sim P &= \sim ((P \rightarrow Q) \wedge \sim Q) \vee \sim P \\
&= \sim ((\sim P \vee Q) \wedge \sim Q) \vee \sim P \\
&= \sim ((\sim P \wedge \sim Q) \vee (Q \wedge \sim Q)) \vee \\
&\vee \sim P \\
&= \sim ((\sim P \wedge \sim Q) \vee \square) \vee \sim P \\
&= \sim ((\sim P \wedge \sim Q)) \vee \sim P \\
&= (P \vee Q) \vee \sim P \\
&= (Q \vee P) \vee \sim P \\
&= Q \vee (P \vee \sim P) \\
&= Q \vee \blacksquare \\
&= \blacksquare
\end{aligned}
$$

Логика предикатов (логика первого порядка)

Исходные элементы в логике высказываний — это атомы. Из атомов мы строим формулы. Затем мы используем формулы, чтобы выразить различные сложные мысли. В этой простой логике атом представляет повествовательное предложение, которое может быть или истинно или ложно, но не то и другое вместе. Атом рассматривается как единое целое. Его структура и состав не анализируются. Однако есть много мыслей, которые не могут быть рассмотрены таким простым способом.

Чем логика предикатов отличается от логики высказываний?

Допустим, что мы желаем представить утверждение «x больше 3». Сначала мы определим предикат БОЛЬШЕ (x, y), который означает «x больше y». (Отметим, что предикат есть отношение[1].) Тогда выражение «x больше 3» представляется выражением БОЛЬШЕ $(x, 3)$.

Вообще говоря, для построения атомов нам разрешается использовать следующие четыре типа символов:

(1) Индивидные символы или константы. Это обычно имена объектов такие, как Мэри, Джон и 3.

(2) Символы предметных переменных. Это обычно строчные буквы x, y, 2,..., возможно, с индексами.

(3) Функциональные символы. Это обычно строчные буквы f,g h,... или осмысленные слова из строчных букв такие, как ‘отец’ и ‘плюс’.

(4) Предикатные символы. Это обычно прописные буквы P, Q, R,... или осмысленные слова из прописных букв такие, как БОЛЬШЕ или ЛЮБИТ.

Определение. *Термы* определяются рекурсивно следующим образом:

(i) Константа есть терм.

(ii) Переменная есть терм.

(iii) Если f есть n-местный функциональный символ и $t_1, \ldots, t_n$ — термы, то $f(t_1, \ldots, t_n)$ — терм.

(iv) Никаких термов, кроме порожденных применением указанных выше правил, нет.

Пример 3.1. Так как x и 1 — термы и *плюс* — двухместный функциональный символ, то *плюс* $(x, 1)$ есть терм согласно приведенному определению.

Определение. Если P — n-местный предикатный символ и $t_1, \ldots, t_n$ — термы, то $P(t_1, \ldots, t_n)$ — *атом*.

Символы $\forall$ и $\exists$ называются соответственно *кванторами* (*все*) *общности* и *существования*. Если x — переменная, то $(\forall x)$ читается как *«для всех x»*, *«для каждого x»* или *«для всякого x»*, тогда как $(\exists x)$ читается *«существует x»*, *«для некоторых x»* или *«по крайней мере для одного x»* [1]).

Определение. Вхождение переменной x в формулу называется *связанным* тогда и только тогда, когда оно совпадает с вхождением в кванторный комплекс $(\forall x)$ или $(\exists x)$ или находится в области действия такого комплекса. Вхождение переменной в формулу *свободно* тогда и только тогда, когда оно не является связанным.

Определение. Переменная *свободна* в формуле, если хотя бы одно ее вхождение в эту формулу свободно. Переменная *связана* в формуле, если хотя бы одно ее вхождение в эту формулу связано.

Переменная может быть и связана и свободна одновременно:

$P(x,y) \wedge \exists x T(x,y)$ - “x” свободна и связана одновременно

Определение. *Правильно построенные формулы* или, короче, *формулы* логики первого порядка рекурсивно определяются следующим образом:

(i) Атом есть формула. (Отметим, что «атом» — это сокращение для атомарной формулы.)

(ii) Если F и G — формулы, то $\sim(F)$, $(F \vee G)$, $(F \wedge G)$, $(F \rightarrow G)$ и $(F \leftrightarrow G)$ — формулы.

(iii) Если F — формула, а x — свободная переменная в F, то $(\forall x) F$ и $(\exists x) F$ — формулы.

(iv) Формулы порождаются только конечным числом применений правил (i), (ii) и (iii).

Определение. *Интерпретация* формулы F логики первого порядка состоит из непустой (предметной) области D и указания «оценки» (значения) всех констант, функциональных символов и предикатных символов, встречающихся в F.

1. Каждой константе мы ставим в соответствие некоторый элемент из D.

2. Каждому n-местному функциональному символу мы ставим в соответствие отображение из D^n в D. (Заметим, что $D^n = \{(x_1, \ldots, x_n) \mid x_1 \in D, \ldots, x_n \in D\}$.)

3. Каждому n-местному предикатному символу мы ставим в соответствие отображение D^n в $\{И, Л\}$.

Иногда, чтобы акцентировать внимание на области D, мы говорим об интерпретации формулы на D. Когда мы ищем «оценку», т. е. определяем истинностное значение формулы в интерпретации на области D, $(\forall x)$ будет интерпретироваться как «для всех элементов x из D» и $(\exists x)$ — как «существует элемент x из D».

Для каждой интерпретации формулы на области D формула может получить истинностное значение $И$ или $Л$ согласно следующим правилам:

1. Если заданы значения формул G и H, то истинностные значения формул $\sim G$, $(G \wedge H)$, $(G \vee H)$, $(G \to H)$ и $(G \leftrightarrow H)$ получаются с помощью табл. 2.1 в главе 2.

2. $(\forall x)\, G$ получает значение $И$, если G получает значение $И$ для каждого x из D; в противном случае она получает значение $Л$.

3. $(\exists x)\, G$ получает значение $И$, если G получает значение $И$ хотя бы для одного x из D; в противном случае она получает значение $Л$.

$$G: (\forall x)(P(x) \to Q(f(x), a)).$$

В G имеется одна константа a, один одноместный функциональный символ f, один одноместный предикатный символ P и один двухместный предикатный символ Q. Ниже приводится интерпретация I формулы G.

Область: $D=\{1, 2\}$.

Оценки для a: $\frac{a}{1}$.

Оценки для f:

$f(1)$	$f(2)$
2	1

Оценки для P и Q:

$P(1)$	$P(2)$	$Q(1, 1)$	$Q(1, 2)$	$Q(2, 1)$	$Q(2, 2)$
Л	*И*	*И*	*И*	*Л*	*И*

Если $x=1$, то

$$\begin{aligned} P(x) \to Q(f(x), a) &= P(1) \to Q(f(1), a) \\ &= P(1) \to Q(2, 1) \\ &= Л \to Л = И. \end{aligned}$$

Если $x=2$, то

$$\begin{aligned} P(x) \to Q(f(x), a) &= P(2) \to Q(f(2), a) \\ &= P(2) \to Q(1, 1) \\ &= И \to И = И. \end{aligned}$$

Так как $P(x) \to Q(f(x), a)$ истинно для всех элементов x из области D, то формула $(\forall x)(P(x) \to Q(f(x), a))$ истинна в интерпретации I.

Определение. Формула G *непротиворечива (выполнима)* тогда и только тогда, когда существует такая интерпретация I, что G имеет значение *И* в I. Если формула G есть *И* в интерпретации I, то мы говорим, что I есть *модель* формулы G и I *удовлетворяет* G.

Определение. Формула G *противоречива (невыполнима)* тогда и только тогда, когда не существует интерпретации, которая удовлетворяет G.

Определение. Формула G *общезначима* тогда и только тогда, когда не существует никакой интерпретации, которая удовлетворяет формуле G.

Определение. Говорят, что формула F в логике первого порядка находится в *предваренной нормальной форме*, тогда и только тогда, когда формула F имеет вид

$$(\mathrm{Q}_1 x_1) \ldots (\mathrm{Q}_n x_n)(M),$$

где каждое $(\mathrm{Q}_i x_i)$, $i = 1, \ldots, n$, есть или $(\forall x_i)$ или $(\exists x_i)$, и M есть формула, не содержащая кванторов. $(\mathrm{Q}_1 x_1) \ldots (\mathrm{Q}_n x_n)$ называется *префиксом*, а M — *матрицей* формулы F.

Приведем несколько формул, находящихся в предваренной нормальной форме:

$$(\forall x)(\forall y)(P(x, y) \wedge Q(y)),$$
$$(\forall x)(\forall y)(\sim P(x, y) \rightarrow Q(y)),$$
$$(\forall x)(\forall y)(\exists z)(Q(x, y) \rightarrow R(z)).$$

Приведённые ниже эквиваленции справедливы для логики предикатов справедливы для

$$(\mathrm{Q}x) F[x] \vee G = (\mathrm{Q}x)(F[x] \vee G);$$
$$(\mathrm{Q}x) F[x] \wedge G = (\mathrm{Q}x)(F[x] \wedge G);$$
$$\sim ((\forall x) F[x]) = (\exists x)(\sim F[x]);$$
$$\sim ((\exists x) F[x]) = (\forall x)(\sim F[x]).$$

Имеются два других закона:

(3.3a) $(\forall x) F[x] \wedge (\forall x) H[x] = (\forall x)(F[x] \wedge H[x])$,

(3.3b) $(\exists x) F[x] \vee (\exists x) H[x] = (\exists x)(F[x] \vee H[x])$,

т. е. квантор всеобщности $\forall$ и квантор существования $\exists$ можно распределять по $\wedge$ и $\vee$ соответственно.

Доказательства законов (3.3a) и (3.3b) нетрудны. Мы предоставим их читателю. Однако квантор всеобщности $\forall$ и квантор существования $\exists$ *нельзя* распределять по $\vee$ и $\wedge$ соответственно, т. е.

$$(\forall x) F[x] \vee (\forall x) H[x] \neq (\forall x)(F[x] \vee H[x])$$

и

$$(\exists x) F[x] \wedge (\exists x) H[x] \neq (\exists x)(F[x] \wedge H[x]).$$

Преобразование формул в предваренную нормальную форму.

Шаг 1. Используем законы

$$F \leftrightarrow G = (F \to G) \wedge (G \to F), \tag{2.1}$$

$$F \to G = \sim F \vee G, \tag{2.2}$$

чтобы исключить логические связки $\leftrightarrow$ и $\to$.

Шаг 2. Повторно используем закон

$$\sim (\sim F) = F, \tag{2.9}$$

законы де Моргана

$$\sim (F \vee G) = \sim F \wedge \sim G, \tag{2.10a}$$

$$\sim (F \wedge G) = \sim F \vee \sim G \tag{2.10b}$$

и законы

$$\sim ((\forall x)\, F[x]) = (\exists x)\,(\sim F[x]), \tag{3.2a}$$

$$\sim ((\exists x)\, F[x]) = (\forall x)\,(\sim F[x]), \tag{3.2b}$$

чтобы пронести знак отрицания внутрь формулы.

Шаг 3. Переименовываем связанные переменные, если это необходимо.

Шаг 4. Используем законы

$$(Qx)\, F[x] \vee G = (Qx)\,(F[x] \vee G), \tag{3.1a}$$

$$(Qx)\, F[x] \wedge G = (Qx)\,(F[x] \wedge G), \tag{3.1b}$$

$$(\forall x)\, F[x] \wedge (\forall x)\, H[x] = (\forall x)\,(F[x] \wedge H[x]), \tag{3.3a}$$

$$(\exists x)\, F[x] \vee (\exists x)\, H[x] = (\exists x)\,(F[x] \vee H[x]), \tag{3.3b}$$

$$(Q_1 x)\, F[x] \vee (Q_2 x)\, H[x] = (Q_1 x)\,(Q_2 z)\,(F[x] \vee H[z]), \tag{3.4a}$$

$$(Q_3 x)\, F[x] \wedge (Q_4 x)\, H[x] = (Q_3 x)\,(Q_4 z)\,(F[x] \wedge H[z]), \tag{3.4b}$$

чтобы вынести кванторы в самое начало формулы для получения формулы, находящейся в предваренной нормальной форме.

В случаях, подобных этим, мы должны поступать специальным образом. Так как каждая связанная переменная в формуле может рассматриваться лишь как место для подстановки какой угодно переменной, то каждую связанную переменную x можно переименовать в z и формула $(\forall x) H[x]$ перейдет в $(\forall z) H[z]$, т. е. $(\forall x) H[x] = (\forall z) H[z]$. Предположим, что мы выбираем переменную z, которая не встречается в $F[x]$. Тогда

$$(\forall x) F[x] \vee (\forall x) H[x] = (\forall x) F[x] \vee (\forall z) H[z]$$

(путем замены всех x, входящих в $(\forall x) H[x]$, на z)

$$= (\forall x)(\forall z)(F[x] \vee H[z]) \quad \text{(по 3.1a).}$$

Аналогично, мы имеем

$$(\exists x) F[x] \wedge (\exists x) H[x] = (\exists x) F[x] \wedge (\exists z) H[z]$$

(путем замены всех x, входящих в $(\exists x) H[x]$, на z)

$$= (\exists x)(\exists z)(F[x] \wedge H[z]) \quad \text{(по 3.1b).}$$

Следовательно, для этих двух случаев мы все еще можем вынести все кванторы в этой формуле влево. В общем случае мы имеем

(3.4a) $(Q_1 x) F[x] \vee (Q_2 x) H[x] = (Q_1 x)(Q_2 z)(F[x] \vee H[z])$,
(3.4b) $(Q_3 x) F[x] \wedge (Q_4 x) H[x] = (Q_3 x)(Q_4 z)(F[x] \wedge H[z])$,

Сколемовская нормальная форма

Пусть формула F находится в предваренной нормальной форме $(Q_1 x_1) \ldots (Q_n x_n) M$, где M есть *конъюнктивная* нормальная форма. Положим, Q_r есть квантор существования в префиксе $(Q_1 x_1) \ldots (Q_n x_n)$, $1 \leqslant r \leqslant n$. Если никакой квантор всеобщности не стоит в префиксе левее Q_r, мы выберем новую константу c, отличную от других констант, входящих в M, заменим все x_r, встречающиеся в M, на c и вычеркнем $(Q_r x_r)$ из префикса. Если $Q_{s_1}, \ldots, Q_{s_m}$ — список всех кванторов всеобщности, встречающихся левее Q_r, $1 \leqslant s_1 < s_2 < \ldots < s_m < r$, мы выберем новый m-местный функциональный символ f, отличный от других функциональных символов, заменим все x_r в M на $f(x_{s_1}, x_{s_2}, \ldots, x_{s_m})$ и вычеркнем $(Q_r x_r)$ из префикса. Затем весь этот процесс применим для всех кванторов существования в префиксе; последняя из полученных формул есть *скулемовская стандартная форма* — для краткости, *стандартная форма* формулы F. Константы и функции, используемые для замены переменных квантора существования, называются *скулемовскими* функциями.

Для сколемовской формы справедливо высказывание: формула противоречива в том, и только том случае когда её нормальная форма противоречива.

Теорема Эрбрана

Множество дизъюнктов S невыполнимо тогда и только тогда, когда существует конечное невыполнимое множество S’ основных примеров дизъюнктов S.

Доказательство Теоремы Эрбрана

Предположим, что существует конечное невыполнимое множество S’ основных примеров дизъюнктов в S. Так как каждая интерпретация I для S содержит интерпретацию S’ множества S’ и I’ опровергает S’, то I должна также опровергать S'. Однако S’ опровергается в каждой интерпретации I’. Следовательно, S’ опровергается в каждой интерпретации I множества S. Поэтому S опровергается в каждой интерпретации множества S’; значит, S невыполнимо.

Пример 4.19. Рассмотрим

$$
\begin{aligned}
S &= \{P(a),\ \sim P(x) \vee Q(f(x)),\ \sim Q(f(a))\};\\
H_0 &= \{a\};\\
S_0' &= P(a) \wedge (\sim P(a) \vee Q(f(a))) \wedge \sim Q(f(a))\\
&= (P(a) \wedge \sim P(a) \wedge \sim Q(f(a))) \vee (P(a) \wedge Q(f(a)) \wedge \sim Q(f(a))\\
&= \square \vee \square = \square.
\end{aligned}
$$

Что такое методы резолюции?

Если в “основах” мы обсуждали как решать задачи путём доказательства теорем, то теперь узнаем процедуры поиска этих доказательств.

Тьюринг и Чёрчь, в свое время, доказали что не существует никакой общей процедуры, проверяющей общезначимость формул в логике предикатов.
Однако,существуют алгоритмы, способные доказать общезначимость формулы за конечное число шагов,а в случае не общезначимых формул - не завершают свою работу вовсе (работают бесконечно).
Важным подходом к авто доказательству теорем был сформулирован Эрбраном в 1930г, и заключался в доказательстве противоречивости отрицания формулы (если ф-ла не ложна, следовательно, ф-ла истинна).

На этом подходе базируется метод резолюций, введённый Робинсоном, являющийся наиболее эффективным в выполнении поставленной задачи

Основная идея метода резолюций состоит в том, чтобы проверить, содержит ли S пустой дизъюнкт □. Если S содержит □, то S невыполнимо. Если S не содержит □, то проверяется следующий факт: может ли □ быть получен из S.

Правило резолюции

1)В дизъюнктах D_i D_j (i != j) множества дизъюнктов S находим контрарную пару: L и ~L
2)Добавляем в S новый дизъюнкт D_n (его называют резольвентой D_i D_j), образованный дизъюнкцией D_1 и D_2 с исключенными из них L и ~L

Пример 0

S = {D_1 ; D_2 }, D_1 = (P\/~Q \/ R) ; D_2 = (~P \/ M)
По правилу резолюции:

1) нашли контрарную пару: P из D_1 и ~P из D_2
2) После исключения P из D_1 и ~P из D_2, соединяем оставшиеся части дизъюнкцией: D_3 = (~Q\/R) \/ (M)
3) S = {D_1 ; D_2 ; D_3 }, D_1 = (P\/~Q\/R) ; D_2 = (~P \/ M) ; D_3 = (~Q \/ R\/ M)

Определение резолютивного вывода

Пусть S — множество дизъюнктов. Резолютивный вывод C из S есть такая конечная последовательность C_1,C_2, ..., C_k дизъюнктов, что каждый C_i или принадлежит S или является резольвентой дизъюнктов, предшествующих C_i, и C_k = C. Вывод □ из S называется опровержением (или доказательством невыполнимости) S.

Пример 5.5. Для множества

$$\left.\begin{array}{ll} (1) & P \vee Q, \\ (2) & \sim P \vee Q, \\ (3) & P \vee \sim Q, \\ (4) & \sim P \vee \sim Q \end{array}\right\} S$$

мы породим следующие резольвенты:

(5) Q из (1) и (2),
(6) $\sim Q$ из (3) и (4),
(7) □ из (5) и (6).

Так как получен □, то S невыполнимо.

Подстановка и унификация

Если в логике предикатов перед использованием правила резолюции необходимо унифицировать термы в контрарной паре, т.е. привести к одинаковой сигнатуре.

Имеем : $C_1: \quad P(x) \vee Q(x),$
$C_2: \quad \sim P(f(x)) \vee R(x).$

Для приведения P и ~P к одинаковой сигнатуре можно подставить в C_2 вместо “x” - “a”, а в C_2 вместо “x” подставить f(a), тогда получим:

$C_1': \quad P(f(a)) \vee Q(f(a)),$
$C_2': \quad \sim P(f(a)) \vee R(a).$, отсюда резольвента: $C_3': \; Q(f(a)) \vee R(a).$

Выражение полученное из выражения подстановкой терма называется “примером” выражения

В процедуре доказательства по методу резолюций, для того чтобы отождествлять контрарные пары литер, мы очень часто должны унифицировать (склеивать) два или более выражения, т. е. мы должны находить подстановку, которая может сделать несколько выражений тождественными. Итак, мы сейчас рассмотрим унификацию выражений.

Определение подстановки

Подстановка θ называется унификатором множества {E_1;E_2;...;E_k}, если E_1(θ)= E_2(θ)=...= E_k(θ).

Определение: Множество рассогласований σ непустого множества выражений W получается выявлением первой (слева) позиции, на которой не для всех выражений из W стоит один и тот же символ, и затем выписыванием из каждого выражения в W подвыражения, которое начинается с символа, занимающего эту позицию. Множество этих подвыражений и есть множество рассогласований в W

Для W= $\{P(x, \underline{f(y, z)}), \; P(x, \underline{a}), P(x, \underline{g(h(k(x)))})\}$ получаем σ = $\{f(y, z), \; a, \; g(h(k(x)))\}$

Алгоритм унификации

σ - множество всех рассогласований (изначально пустое)
D_i - текущие рассогласования (изначально пустое)
W - множество выражений
x/z - условное рассогласование
ПОКА W - унифицируемо ВЫПОЛНИТЬ:

ПОКА не найдено рассогласование ИЛИ не конец выражения ВЫПОЛНИТЬ:

Посимвольное сравнение выражений начиная слева;
Выписываем из каждого выражения в D_i первое встреченное ...
…. рассогласование x/z;
КЦ
ЕСЛИ рассогласование найдено ВЫПОЛНИТЬ:
Дополняем σ найденным рассогласованием;
Выполняем подстановку σ в W
Освобождаем D_i (теперь = пустому множеству)
ИНАЧЕ W - неунифицируемо -> остановка алгоритма
КЦ

Пример 1

Пример 5.11. Найти наиболее общий унификатор для

$$W=\{P(a,\ x,\ f(g(y))),\ \ P(z,\ f(z),\ f(u))\}.$$

1. $\sigma_0=\varepsilon$ и $W_0=W$. Так как W_0 — не единичный дизъюнкт, то σ_0 не является наиболее общим унификатором для W.

2. Множество рассогласований $D_0=\{a,\ z\}$. В D_0 существует переменная $v_0=z$, которая не встречается в $t_0=a$.

3. Пусть

$$\begin{aligned}\sigma_1&=\sigma_0\circ\{t_0/v_0\}=\varepsilon\circ\{a/z\}=\{a/z\},\\ W_1&=W_0\{t_0/v_0\}\\ &=\{P(a,\ x,\ f(g(y))),\ \ P(z,\ f(z),\ f(u))\}\{a/z\}\\ &=\{P(a,\ x,\ f(g(y))),\ \ P(a,\ f(a),\ f(u))\}.\end{aligned}$$

4. W_1 — не единичный дизъюнкт, так как нашли множество рассогласований D_1 для W_1:

$$D_1=\{x,\ f(a)\}.$$

5. Из D_1 мы найдем, что $v_1=x$ и $t_1=f(a)$.

6. Пусть

$$\begin{aligned}\sigma_2&=\sigma_1\circ\{t_1/v_1\}=\{a/z\}\circ\{f(a)/x\}=\{a/z,\ f(a)/x\},\\ W_2&=W_1\{t_1/v_1\}\\ &=\{P(a,\ x,\ f(g(y))),\ \ P(a,\ f(a),\ f(u))\}\{f(a)/x\}\\ &=\{P(a,\ f(a),\ f(g(y))),\ \ P(a,\ f(a),\ f(u))\}.\end{aligned}$$

7. W_2 — не единичный дизъюнкт, так как мы нашли множество рассогласований D_2 для W_2: $D_2=\{g(y),\ u\}$. Из D_2 мы найдем, что $v_2=u$ и $t_2=g(y)$.

8. Пусть

$$\begin{aligned}\sigma_3&=\sigma_2\circ\{t_2/v_2\}=\{a/z,\ f(a)/x\}\circ\{g(y)/u\}=\\ &=\{a/z,\ f(a)/x,\ g(y)/u\},\\ W_3&=W_2\{t_2/v_2\}\\ &=\{P(a,\ f(a),\ f(g(y))),\ \ P(a,\ f(a),\ f(u))\}\{g(y)/u\}\\ &=\{P(a,\ f(a),\ f(g(y))),\ \ P(a,\ f(a),\ f(g(y)))\}\\ &=\{P(a,\ f(a),\ f(g(y)))\}.\end{aligned}$$

9. Так как W_3 — единичный дизъюнкт, то $\sigma_3=\{a/z,\ f(a)/x,\ g(y)/u\}$ есть наиболее общий унификатор для W.

Важно также понимать что литеры с разным количеством термов (“разноместные”) не могут быть унифицированны.

Пример 2

Пример 5.12. Определим, унифицируемо ли множество

$$W = \{Q(f(a),\ g(x)),\ Q(y,\ y)\}.$$

1. Пусть $\sigma_0 = \varepsilon$ и $W_0 = W$.
2. W_0 — не единичный дизъюнкт, так как мы нашли множество рассогласований D_0 для W_0: $D_0 = \{f(a),\ y\}$. Из D_0 мы знаем, что $v_0 = y$ и $t_0 = f(a)$.
3. Пусть

$$\sigma_1 = \sigma_0 \circ \{t_0/v_0\} = \varepsilon \circ \{f(a)/y\} = \{f(a)/y\},$$
$$W_1 = W_0\{t_0/v_0\} = \{Q(f(a),\ g(x)),\ Q(y,\ y)\}\{f(a)/y\}$$
$$= \{Q(f(a),\ g(x)),\ Q(f(a),\ f(a))\}.$$

4. W_1 — не единичный дизъюнкт, так как мы нашли множество рассогласований D_1 для W_1: $D_1 = \{g(x),\ f(a)\}$..
5. Тем не менее нет элемента, который был бы переменной. Следовательно, алгоритм унификации завершается, и мы заключаем, что W не унифицируемо.

Отмечу, что вышеуказанный алгоритм унификации всегда кончает работу для любого конечного непустого множества выражений, так как иначе породилась бы бесконечная последовательность Wσ_1,Wσ_2,Wσ_3 , ... конечных непустых множеств, обладающая тем свойством, что каждое последующее множество содержит на одну переменную меньше, чем предшествующее (а именно, Wσ_k, содержит переменную “v_i”, но Wσ_(k+1) её уже не содержит). Но это невозможно, так как W содержит только конечное число различных переменных. Введем новое понятие.

Определение склейки

Если две литеры (одинаковые буквы, разные термы) имеют наиболее общий унификатор, то их можно унификацией привести к одной литере, и такую операцию называю склейкой:

{P(x); P(f(x))} = {P(f(a))}

Пусть C_1 и C_2 - два дизъюнкта, а L и ~L принадлежат им соответственно, тогда если L и ~L имеют наиболее общий унификатор σ, то РЕЗОЛЬВЕНТОЙ называют выражение типа ((C_1)σ - Lσ)\/((C_2)σ - (~L)σ), а L и ~L - отрезаемыми литерами.

Пример 5.15. Пусть $C_1 = P(x) \vee P(f(y)) \vee R(g(y))$ и $C_2 = {\sim} P(f(g(a))) \vee Q(b)$. Склейка C_1 есть $C_1' = P(f(y)) \vee R(g(y))$. Бинарная резольвента C_1' и C_2 есть $R(g(g(a))) \vee Q(b)$. Следовательно, $R(g(g(a))) \vee Q(b)$ есть резольвента C_1 и C_2.

Пример (доказательство теоремы из планиметрии)

Пример 5.16. Показать, что внутренние накрест лежащие углы, образованные диагональю трапеции, равны.

Чтобы доказать эту теорему, мы сперва аксиоматизируем ее. Пусть $T(x, y, u, v)$ обозначает, что $xyuv$ — трапеция с верхней левой вершиной x, верхней правой вершиной y, нижней правой вершиной u и нижней левой вершиной v. Пусть $P(x, y, u, v)$ обозначает, что отрезок xy параллелен отрезку uv, и пусть

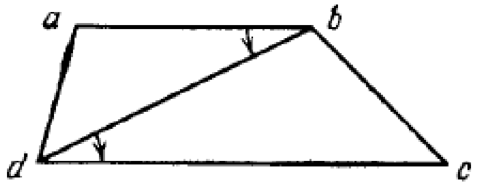

Рис. 7.

$E(x, y, z, u, v, w)$ означает, что угол xyz равен углу uvw. Тогда мы имеем следующие аксиомы:

A_1: $(\forall x)(\forall y)(\forall u)(\forall v)[T(x, y, u, v) \rightarrow P(x, y, u, v)]$	определение трапеции,
A_2: $(\forall x)(\forall y)(\forall u)(\forall v)[P(x, y, u, v) \rightarrow E(x, y, v, u, v, y)]$	внутренние накрест лежащие углы для параллельных линий равны,
A_3: $T(a, b, c, d)$	дана на рис. 7.

Из этих аксиом мы должны уметь вывести, что $E(a, b, d, c, d, b)$ истинно, т. е.

$$A_1 \wedge A_2 \wedge A_3 \rightarrow E(a, b, d, c, d, b)$$

есть истинная формула. Так как мы хотим доказать это путем опровержения выполнимости, мы отрицаем заключение и докажем, что

$$A_1 \wedge A_2 \wedge A_3 \wedge \sim E(a,\ b,\ d,\ c,\ d,\ b)$$

невыполнимо. Чтобы это сделать, преобразуем выражение в стандартную форму следующим образом:

$$S = \{\sim T(x,\ y,\ u,\ v) \vee P(x,\ y,\ u,\ v),\ \sim P(x,\ y,\ u,\ v) \vee \vee E(x,\ y,\ v,\ u,\ v,\ y),\ T(a,\ b,\ c,\ d),\ \sim E(a,\ b,\ d,\ c,\ d,\ b)\}.$$

Эта стандартная форма S есть множество из четырех дизъюнктов. Теперь методом резолюций докажем, что множество S невыполнимо:

(1)	$\sim T(x,\ y,\ u,\ v) \vee P(x,\ y,\ u,\ v)$	дизъюнкт в S,
(2)	$\sim P(x,\ y,\ u,\ v) \vee E(x,\ y,\ v,\ u,\ v,\ y)$	дизъюнкт в S,
(3)	$T(a,\ b,\ c,\ d)$	дизъюнкт в S,
(4)	$\sim E(a,\ b,\ d,\ c,\ d,\ b)$	дизъюнкт в S,
(5)	$\sim P(a,\ b,\ c,\ d)$	резольвента (2) и (4),
(6)	$\sim T(a,\ b,\ c,\ d)$	резольвента (1) и (5),
(7)	$\square$	резольвента (3) и (6).

Так как последний выведенный из S дизъюнкт пустой, то мы заключаем, что S невыполнимо. Шаги доказательства легко можно представить деревом на рис. 8. Это дерево называется деревом вывода, т. е. *дерево вывода* из множества S дизъюнктов есть растущее (вверх) дерево, каждому начальному узлу которого приписывается дизъюнкт из S и каждому следующему узлу приписывается резольвента дизъюнктов, приписанных к его непосредственным предшественникам. Дерево вывода мы называем *деревом вывода дизъюнкта* R, если R приписан корневому узлу дерева вывода. Так как дерево вывода есть просто дерево, представляющее вывод, впоследствии мы будем использовать термины «вывод» и «дерево вывода» как взаимозаменяемые.

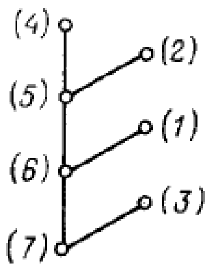

Рис. 8.

Доказательство полноты метода резолюций

Множество дизъюнктов S невыполнимо тогда и только тогда, когда существует вывод пустого дизъюнкта из S

Предположим, что существует <u>вывод</u> пустого дизъюнкта из S. Пусть есть некоторое количество резольвент X <u>в этом выводе.</u> Предположим, что S выполнимо на модели M. Эта модель должна удовлетворять дизъюнктам C_1 и C_2, принадлежащих S, и любым резольвентам данных дизъюнктов (путем унификации можно добиться примеров C_1 и C_2, дающих резольвенту). Следовательно модель M удовлетворяет дизъюнктам X, что невозможно,т.к. среди этих дизъюнктов есть пустой дизъюнкт, опровергающий выполнимость S. Поэтому S должно быть невыполнимо.

Для метода резолюций существуют стратегии, повышающие его эффективность: насыщения уровня, семантическая, лок-резолюция, линейная, поддержки, предпочтения единичных, вычёркивания...

Безусловно, метод резолюции - это наиболее удобный и универсальный способ авто доказательств.

Но осторожно: метод резолюций не экономит память ЭВМ! Неограниченное применение правила резолюций может вызвать порождение большого количества не относящихся к делу и излишних дизъюнктов! Так например, если использовать метод резолюции, выписывая сначала имеющиеся дизъюнкты, вычислять все возможные резольвенты, добавляя их в конец списка, до тех пор пока не будет найден пустой дизъюнкт (к слову, такая стратегия называется методом насыщения уровня), мы можем получить тавтологию (повторение) в списке дизъюнктов. Чтобы не транжирить свое время на поиск тавтологий мы можем применить стратегию вычеркивания.

Ссылки на использованные источники

1. Ч.Чень, Р.Ли “Математическая логика и автоматическое доказательство теорем” [книга] Ссылка на облако: https://cloud.mail.ru/public/DQAm/zSQynqanp
2. Статья Литературного клуба “Метод резолюции” [сайт]. URL: http://ipo.spb.ru/journal/content/931/Метод%20резолюций.pdf.
3. Корпоративный портал Томского Политехнического университета [сайт]. URL: https://portal.tpu.ru/SHARED/s/SHEFER/Study/Tab2/Tab3/Razdel_N4.pdf.
4. Открытые учебные программы Массачусетского Технологического Института (Massachusetts Institute of Technology) [сайт]. URL: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-825-techniques-in-artificial-intelligence-sma-5504-fall-2002/lecture-notes/Lecture7FinalPart1.pdf.
5. Департамент компьютерных наук и инженерии Индийского Технологического Института “Харагпер” (Indian Institute of Technology “Kharagpur”) [сайт]. URL: https://www.youtube.com/watch?v=cx5b0kyu-jU.