Метод резолюций

учебные материалы
для курса «Дискретная математика–2» ФКН ВШЭ

(осенний семестр 2017 г.)

Формула с булевыми переменными и пропозициональными связками задаёт булеву функцию. Если эта функция равна 1 на любом наборе значений переменных (*присваивании*), то такая формула называется *тавтологией*.

Вопрос, который обсуждается в этих записках: как по формуле проверить, является ли она тавтологией?

Есть очевидный способ. Если в формулу входит n булевых переменных, то различных присваиваний всего 2^n. Поэтому достаточно вычислить значение формулы на всех присваиваниях, после чего уже легко решить, является ли эта формула тавтологией.

Конечно, такой алгоритм медленный — 2^n очень быстро растёт с ростом n. Однако есть способы ускорить перебор возможных присваиваний. Эти способы настолько хороши, что современные программы способны проверять тавтологичность формул с многими тысячами переменных.

Метод резолюций, о котором пойдёт речь, лежит в основе всех таких способов.

1 О доказательствах тавтологичности

Мы будем рассматривать булевы формулы, составленные из пропозициональных связок $\to$, $\wedge$, $\vee$, $\neg$ и булевых переменных. Для однозначного выбора порядка применения связок в формулах также используются скобки.

Задача 1. Проверьте, что $x \to x$, $x \vee \neg x$ являются тавтологиями.

Если формула длинная, то проверка тавтологичности становится непростым делом. Сократить перебор можно, если присваивать значения переменным по очереди и упрощать формулу по мере возможности. Для этого можно использовать равенства из предыдущей задачи и такие же легко проверяемые тождества с булевыми

функциями. Например:

$$\begin{array}{llll} 0 \to x = 1, & 0 \vee x = x, & 0 \wedge x = 0, & \neg\neg x = x, \\ 1 \to x = x, & 1 \vee x = 1, & 1 \wedge x = x, & \\ x \to 0 = \neg x, & x \vee y = y \vee x, & x \wedge y = y \wedge x, & \\ x \to 1 = 1, & x \to y = \neg x \vee y & x \wedge y = \neg(x \to \neg y). & \end{array}$$

Пример 2. Докажем, что формула $(a \to (b \to c)) \to ((a \to b) \to (a \to c))$ является тавтологией.

Если $b = 1$, то эта формула упрощается так

$$(a \to (1 \to c)) \to ((a \to 1) \to (a \to c)) = (a \to c) \to (1 \to (a \to c)) = (a \to c) \to (a \to c) = 1.$$

А если $b = 0$, то иначе, но всё равно упрощается

$$(a{\to}(0{\to}c)){\to}((a{\to}0){\to}(a{\to}c)) = (a{\to}1){\to}(\neg a{\to}(a{\to}c)) = \neg\neg a{\vee}\neg a{\vee}c = (a{\vee}\neg a){\vee}c = 1.$$

Помимо полного или частичного перебора всех присваиваний есть другая идея: придумать *доказательство* тавтологичности.

Какими могут быть доказательства? Первая идея состоит в том, чтобы использовать обычный для математики аксиоматический метод. Начнём с того, что выберем какой-нибудь простой набор тавтологий в качестве аксиом. Из этих тавтологий можно получать другие, пользуясь правилом modus ponens, которое символически записывается как

$$\frac{A,\ A \to B}{B}.$$

Смысл у этой записи простой: если уже доказано, что формулы A и $A \to B$ являются тавтологиями, то и формула B также является тавтологией. (Неформально: если истинна посылка импликации, то истинно и заключение импликации.)

Задача 3. Проверьте, что из $a = 1$, $a \to b = 1$ следует $b = 1$.

Пользуясь правилом modus ponens, можно расширять список доказанных тавтологий. Если в этот список войдёт интересующая нас формула, то мы получили доказательство её тавтологичности.

Мы описали систему доказательств, которую сейчас называют *система Фреге*. Таких систем много, в зависимости от выбора исходного набора аксиом. Не любой набор тавтологий годится в качестве системы аксиом. Во-первых, он должен быть достаточно простым, чтобы по формуле легко было проверить, что она является аксиомой. Во-вторых, он должен быть *полным*.

Полнота набора аксиом по определению означает, что любая тавтология может быть доказана, исходя из этого набора. Если система аксиом неполна, для некоторых тавтологий не будет доказательств, что нас не устраивает.

Полные и достаточно простые наборы аксиом существуют, см., например, книгу Н. К. Верещагина и А. Шеня «Языки и исчисления».

Хотя системы Фреге дают принципиальный способ доказать любую тавтологию, у них есть существенный недостаток. Предположим, вы хотите доказать тавтологичность формулы в системе Фреге. С каких аксиом начинать доказательство и каков его вид? Общего ответа на этот вопрос нет и сомнительно, чтобы удовлетворительный ответ был. Система Фреге слишком мощная, поэтому доказательства в ней устроены слишком сложно.

Мы рассмотрим более простую систему доказательств. Более точно, это будет не система доказательств, а система *опровержений*. Если формула Φ — тавтология, то её отрицание $\neg\Phi$ — *невыполнимо*, то есть обращается в 0 на любом присваивании. Конечно, верно и обратное. Поэтому доказательство тавтологичности Φ — это то же самое, что доказательство невыполнимости $\neg\Phi$.

Одним из видов доказательства невыполнимости формулы является *опровержение* предположения о её выполнимости. Это обычный для математики способ доказательств от противного. Мы предполагаем, что Ψ выполнима (равна 1 на некотором присваивании) и приходим к противоречию. Такое рассуждение и будем называть опровержением формулы Ψ.

2 От общих формул к КНФ

Система опровержений, основанная на резолюциях, которую мы обсудим дальше, применима не к произвольным булевым формулам, а только к КНФ. Напомним терминологию: *литерал* — это переменная или её отрицание, *дизъюнкт* (clause) — дизъюнкция литералов, а *конъюнктивная нормальная форма* (КНФ) — это конъюнкция дизъюнктов.

Пример: $(x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$. Легко убедиться, что такая КНФ выполнима.

Напомним, что формулы называются *эквивалентными*, если они задают одну и ту же булеву функцию.

Теорема 1. *Любая булева функция представляется в виде КНФ.*

Другими словами, для любой формулы есть эквивалентная ей КНФ.

Доказательство. Составим такую КНФ из дизъюнктов, которые обращаются в 0 ровно на одном присваивании.

Заметим, что дизъюнкция равна 0 только тогда, когда каждый её член равен 0. Скажем, дизъюнкция $\neg x_1 \vee x_2 \vee x_3 \vee \neg x_4$ обращается в 0 только при значениях переменных $x_1 = 1$, $x_2 = 0$, $x_3 = 0$, $x_4 = 1$. Если в формулу входят ещё какие-то переменные, то присваиваний, на которых эта дизъюнкция обращается в 0, будет больше одного (остальные переменные могут принимать произвольные значения). Но если дизъюнкт содержит все переменные (или их отрицания), то он обращается в 0 ровно на одном присваивании, как и в разобранном примере.

Обозначим через D_a дизъюнкт, который обращается в 0 только на присваивании a. Правило построения такого дизъюнкта ясно из разобранного примера: если $x_i = 0$ в присваивании a, то включаем в D_a литерал x_i; если $x_i = 1$, то включаем литерал $\neg x_i$. Тогда $D_a(a) = 0$, $D_a(x) = 1$ для любого $x \neq a$.

Составим список $N_0 = \{a_1, \ldots, a_m\}$ всех присваиваний, на которых функция f обращается в 0. Искомая КНФ имеет вид

$$C_f(x) = \bigwedge_{i=1}^{m} D_{a_i}(x).$$

Действительно, если $f(x) = 0$ на некотором присваивании, то оно входит в список N_0, $x = a_j$. Тогда $C_f(x) = C_f(a_j) = 0$, так как один из членов конъюнкции, а именно, D_{a_j} обращается в 0.

Если же $f(x) = 1$, то все члены конъюнкции равны 1, так как $D_{a_i}(x) = 1$ при $x \neq a_i$, а x в список N_0 не входит. □

У КНФ, построенной в доказательстве теоремы 1, есть существенный недостаток. Её размер пропорционален количеству нулей функции f. Для той же самой функции могут найтись намного более короткие формулы, представляющие эту функцию.

Задача 4. Докажите, что **a)** формула

$$(x_1 \wedge x_2) \vee (x_3 \wedge x_4) \vee \cdots \vee (x_{2n-1} \wedge x_{2n})$$

задаёт булеву функцию с 3^n нулями.

b) Любая КНФ, эквивалентная этой формуле, содержит не менее $(3/2)^n$ дизъюнктов.

Формулы — это частный случай схем, в которых каждое промежуточное значение, не считая переменных, используется не больше одного раза. *Размером* формулы мы и будем считать размер соответствующей схемы, то есть количество присваиваний в схеме. Размер также равен количеству подформул в формуле, отличных от переменных. Если представлять формулу в виде дерева (см. рис. 1), то размер — это количество вершин дерева формулы, отличных от листьев. Он также совпадает с количеством пропозициональных связок в формуле.

Как видно из примера задачи 4, у некоторых формул все эквивалентные КНФ имеют экспоненциально большой размер. Чтобы свести проверку выполнимости формулы к проверке выполнимости КНФ, не слишком теряя в размере данных, нужно ослабить требование эквивалентности.

Если нас интересует проверка выполнимости, то сгодится любая КНФ, выполнимость которой равносильна выполнимости исходной формулы. Нужно только, чтобы эта КНФ получалась из исходной формулы эффективно, с помощью достаточно быстрого алгоритма.

Объясним, как получить такую КНФ. Этот способ годится и для произвольных схем, не только формул.

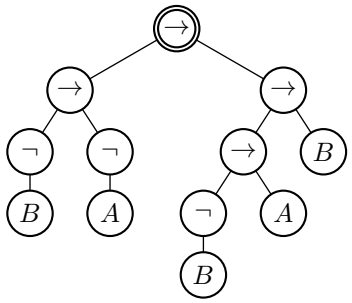


Рис. 1: Дерево формулы $(\neg B \to \neg A) \to ((\neg B \to A) \to B)$

Пусть функция $f(x_1, \ldots, x_n)$ вычисляется схемой Ф

$$g_1,\ g_2,\ \ldots,\ g_m = f,$$

где g_k — либо одна из переменных $x_1, \ldots, x_k$, либо функция, которая получается из предыдущих с помощью одной из связок: для каких-то $\ell, r < k$ выполняется одно из равенств

$$g_k = g_\ell \to g_r,\ \text{или}\ g_k = g_\ell \vee g_r,\ \text{или}\ g_k = g_\ell \wedge g_r,\ \text{или}\ g_k = \neg g_\ell.$$

Построим КНФ C_Φ по следующему правилу.

К переменным $x_1, \ldots, x_n$ добавим переменные z_j для каждого элемента j схемы Ф, который не является входной переменной $x_1, \ldots, x_n$.

Для каждого такого элемента j построим также КНФ E_j. Эта КНФ представляет одну из функций, в зависимости от вида элемента схемы (импликация, дизъюнкция, конъюнкция или отрицание):

$$z_k \equiv (z_\ell \to z_r),\ \text{или}\ z_k \equiv (z_\ell \vee z_r),\ \text{или}\ z_k \equiv (z_\ell \wedge z_r),\ \text{или}\ z_k \equiv \neg z_\ell.$$

Здесь использована связка равносильности, задаваемая таблицей истинности

x	y	$x \equiv y$
0	0	1
0	1	0
1	0	0
1	1	1

(то есть $x \equiv y = 1$ тогда и только тогда, когда значения аргументов равны).

Из теоремы 1 заключаем, что в каждой из построенных КНФ не более 8 дизъюнктов.

Соберём КНФ C_Φ из этих КНФ, добавив к ним дизъюнкт z_m:

$$C_\Phi = z_m \wedge \bigwedge_{j=1}^{m} E_j.$$

В этой КНФ не более $n+m$ переменных и не более $1+8m$ дизъюнктов, каждый из которых содержит не более трёх литералов.

Построение такой КНФ выполняется эффективно, как видно из её описания.

Лемма 2. *Функция $f = g_m$ принимает на каком-то присваивании значение* 1 *тогда и только тогда, когда выполнима КНФ C_Φ.*

Доказательство. Пусть $f(a) = 1$ для какого-то присваивания a:

$$x_1 = a_1,\ x_2 = a_2,\ \ldots,\ x_n = a_n.$$

Добавим к этому присваиванию значения остальных переменных z_j из КНФ C_Φ так, чтобы КНФ стала истинной. Если уже найдены значения z_j для $j < k$, то очередное значение z_k присваиваем в соответствии с типом связки

$$z_k = z_\ell \to z_r,\ \text{или}\ z_k = z_\ell \vee z_r,\ \text{или}\ z_k = z_\ell \wedge z_r,\ \text{или}\ z_k = \neg z_\ell.$$

Индукцией по j легко проверить, что на построенном присваивании все КНФ E_j обращаются в 1 и $z_m = 1$ (так как это и есть значение функции $f(a)$).

В обратную сторону: пусть $C_\Phi(a, z) = 1$. Это означает, что для каждой подформулы выполняются равенства указанного выше вида (в зависимости от вида элемента) и $z_m = 1$. Но тогда значения переменных z_j как раз и есть значения функций g_j, входящих в схему, на присваивании a исходных переменных $x_1, \ldots, x_n$. Поэтому $g_m(a) = f(a) = 1$. □

3 Правило резолюции. Опровержения резолюциями

Итак, мы выяснили, что достаточно проверять выполнимость КНФ, сводя остальные формулы к этому случаю.

Удобство КНФ заключается, в частности, в том, что возможно ограничить количество используемых дизъюнктов. Действительно, с помощью легко проверяемых тавтологий

$$x \vee x \equiv x; \quad x \vee \neg x \equiv 1$$

приводим каждый дизъюнкт к стандартному виду: дизъюнкция литералов, в которой каждой переменной отвечает не более одного литерала. Дизъюнкт, содержащий одновременно x и $\neg x$, тождественно равен 1. Поэтому такой дизъюнкт не влияет на выполнимость КНФ. Мы будем отбрасывать такие дизъюнкты и не включаем их в дизъюнкты стандартного вида.

Далее мы всегда предполагаем, что дизъюнкты приведены к стандартному виду. В таком случае можно понимать под дизъюнктом множество литералов (с дополнительным условием, что в множество не входят одновременно x и $\neg x$).

Какой дизъюнкт отвечает пустому множеству литералов? Буквально такой формулы нет. Однако удобно считать, что это особый, тождественно ложный, дизъюнкт. Мы будем обозначать его $\bot$ и включать в стандартные дизъюнкты.

Задача 5. Найдите количество стандартных дизъюнктов от n переменных.

Сформулируем теперь **правило резолюции**. Формально оно записывается как

$$\frac{A \vee x,\ B \vee \neg x}{A \vee B}.$$

Содержательно это правило означает очень простое утверждение: если $A \vee x = 1$ и $B \vee \neg x = 1$, то $A \vee B = 1$. Действительно, если $x = 0$, то из первого равенства получаем $A = 1$; если $x = 1$, то из второго равенства получаем $B = 1$. В любом случае $A \vee B = 1$.

При использовании правила резолюции мы предполагаем, что дизъюнкты приводятся к стандартному виду.

Например, из пары дизъюнктов $x_1 \vee x_2 \vee x_3$ и $x_2 \vee \neg x_3 \vee x_4$ по правилу резолюции выводится дизъюнкт $x_1 \vee x_2 \vee x_4$ (правило резолюций по переменной x_3).

Для пары дизъюнктов $x_1 \vee x_2$ и $\neg x_1 \vee \neg x_2$ правило резолюции применимо и по переменной x_1, и по переменной x_2. В обоих случаях получается дизъюнкт, тождественно равный 1, так что мы его отбрасываем, как договаривались. Другими словами, резолюции ничего не дают из этой пары.

А что будет, если применить резолюции к паре дизъюнктов x_1 и $\neg x_1$? Формально можно записать $x_1 = \bot \vee x_1$; $\neg x_1 = \bot \vee \neg x_1$ и по правилу резолюций выводится нулевой дизъюнкт $\bot$.

Содержательно это можно пересказать так: предположим, что КНФ $x_1 \wedge \neg x_1$ выполнима. Тогда для какого-то присваивания значений переменным выполняется $x_1 = 1$, $\neg x_1 = 1$, а из сделанного выше наблюдения следует, что $\bot = 1$. Но $\bot$ ложен для всех присваиваний. Получили противоречие.

Конечно, невыполнимость $x_1 \wedge \neg x_1$ ясна и без этих слов. Однако этот пример подсказывает, как определить *опровержение* КНФ C методом резолюций. Это такая последовательность дизъюнктов $D_1, \ldots, D_T$, что её начало состоит из дизъюнктов КНФ C, каждый из остальных дизъюнктов последовательности получается из предыдущих применением правила резолюции, а последний член равен $\bot$.

Пример 6. Для КНФ $C = x_1 \wedge (x_2 \vee \neg x_1) \wedge \neg x_2$ одно из опровержений методом резолюций имеет вид

$$(\neg x_2),\ (x_1),\ (x_2 \vee \neg x_1),\ (x_2),\ \bot$$

(для наглядности мы окружили каждый дизъюнкт скобками).

Заметим, что построение опровержения использует только синтаксические свойства КНФ, то есть только вид формул. Оказывается, существование опровержения методом резолюций равносильно невыполнимости КНФ, то есть семантическому свойству (зависящему от интерпретации формул).

Теорема 3. *КНФ невыполнима тогда и только тогда, когда для неё существует опровержение методом резолюций.*

Доказательство. В одну сторону (корректность) мы уже фактически всё проверили на примерах, перескажем те же рассуждения в общем виде.

Пусть $D_1, \ldots, D_T = \bot$ — опровержение КНФ C. Если на каком-то присваивании все дизъюнкты КНФ истинны, то на том же присваивании истинны и все остальные дизъюнкты в опровержении. Формально это нужно доказывать по индукции, применяя утверждение о правиле резолюций, которое мы проверили с самого начала.

Однако $\bot$ ложен на любом присваивании. Отсюда следует, что КНФ невыполнима.

В обратную сторону (полнота) требуются новые идеи. Начнём со списка дизъюнктов КНФ, обозначим его через L. Будем его расширять, добавляя те дизъюнкты, которые выводятся из уже существующих по правилу резолюции. Процесс рано или поздно остановится, поскольку возможных членов последовательности — дизъюнктов стандартного вида — лишь конечное число. Обозначим полученное множество дизъюнктов через D: из дизъюнктов из множества D уже невозможно вывести нового дизъюнкта с помощью правила резолюции. Если в D есть $\bot$, то получено опровержение и КНФ невыполнима. Осталось доказать, что если в D нет $\bot$, то D совместно и, следовательно, исходная КНФ выполнима.

Построим присваивание $a = (\alpha_1, \ldots, \alpha_n)$, на котором истинны все дизъюнкты из D, присваивая значения переменным по очереди. Выбирая значение очередной переменной, мы следим за тем, чтобы ни один из дизъюнктов списка не обратился в 0. Более точно, мы поддерживаем следующий инвариант: после присваивания значений первым i переменным все дизъюнкты из D, содержащие только первые i переменных, истинны для этого присваивания. В начале процесса $i = 0$, никакие значения еще не присвоены и инвариант выполнен, поскольку пустого дизъюнкта в D нет.

Пусть уже присвоены значения первым $(i-1)$ переменным $x_1 = \alpha_1$, …, $x_{i-1} = \alpha_{i-1}$ (при $i = 1$ это означает, что ещё ни одной переменной значение не присвоено) так, что инвариант истинен. Переменной x_i присваиваем значение 0, если те дизъюнкты из D, в которые входят только первые i переменных, истинны на присваивании $x_1 = \alpha_1$, …, $x_{i-1} = \alpha_{i-1}$, $x_i = 0$. В противном случае присваиваем значение 1, если те дизъюнкты из списка, в которые входят только первые i переменных, истинны на присваивании $x_1 = \alpha_1$, …, $x_{i-1} = \alpha_{i-1}$, $x_i = 1$. Если же в обоих случаях нашлись дизъюнкты D_0 и D_1, ложные на присваиваниях $x_1 = \alpha_1$, …, $x_{i-1} = \alpha_{i-1}$, $x_i = 0$ и $x_1 = \alpha_1$, …, $x_{i-1} = \alpha_{i-1}$, $x_i = 1$ соответственно, то процесс останавливается.

Заметим, что если в результате этого процесса присвоены значения всем переменным, то все дизъюнкты из списка истинны на построенном присваивании (так мы сформулировали правило выбора значений переменных).

Осталось доказать, что таким образом мы сможем присвоить значения всем n переменным с сохранением инварианта. Рассуждая от противного, допустим процесс остановился, не присвоив значение переменной x_k. Это означает, как уже сказано выше, что какой-то дизъюнкт D_0 из списка ложен на присваивании $x_1 = \alpha_1$, …, $x_{k-1} = \alpha_{k-1}$, $x_k = 0$, а другой дизъюнкт D_1 ложен на присваивании $x_1 = \alpha_1$, …, $x_{k-1} = \alpha_{k-1}$, $x_k = 1$. Из истинности инварианта на предыдущем шаге следует, что в

D_0 и D_1 входит переменная x_k (а остальные переменные имеют номера меньше k). При этом в D_0 входит литерал x_k, а в D_1 — литерал $\neg x_k$, то есть

$$D_0 = D'_0 \vee x_k, \qquad D_1 = D'_1 \vee \neg x_k.$$

Более того, оба дизъюнкта D'_0, D'_1 (содержащие только переменные $x_1, \ldots, x_{k-1}$) ложны на присваивании $x_1 = \alpha_1$, …, $x_{k-1} = \alpha_{k-1}$.

Применение правила резолюции по переменной x_k к паре D_0, D_1 даёт дизъюнкт $D'_0 \vee D'_1$, который ложен при присваивании $x_1 = \alpha_1$, …, $x_{k-1} = \alpha_{k-1}$. Это противоречит истинности инварианта на предыдущем шаге, поскольку дизъюнкт $D'_0 \vee D'_1$ принадлежит D. Это противоречие доказывает, что преждевременная остановка процесса невозможна и D совместно. □

Контрольный вопрос 1. Почему рассуждение из доказательства неприменимо к списку, содержащему нулевой дизъюнкт $\bot$? На каком шаге остановится процесс последовательного присваивания значений переменным?

Заметим, что рассуждения, примененные в доказательство Теоремы 3 годятся и для бесконечных множеств дизъюнктов.

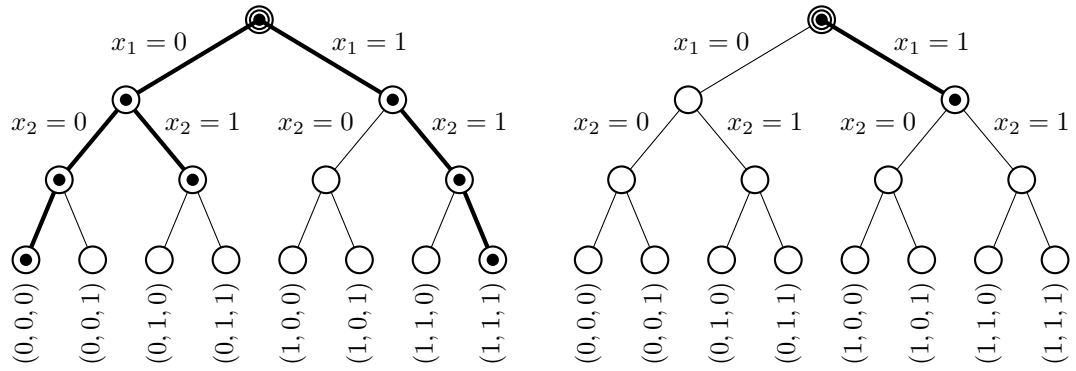
Теорема 4. *Множество дизъюнктов невыполнимо тогда и только тогда, когда из него в ИР выводится пустой дизъюнкт* $\bot$.

Доказательство. Невыводимость пустого дизъюнкта из выполнимых множеств доказывается точно так же, как в конечном случае.

Осталось доказать совместность бесконечных множеств дизъюнктов, из которых невыводим пустой дизъюнкт. Пусть дано бесконечное множество дизъюнктов L, из которого невыводим пустой дизъюнкт. Сначала расширим L добавив в него все выводимые из него дизъюнкты (с помощью любого конечного числа применения правила резолюции). Обозначим через D полученное множество дизъюнктов. По предположению оно не содержит пустого дизъюнкта.

Если L счетно, то и количество переменных в L и D счетно, и мы можем использовать ту же конструкцию выполняющего набора, что и в конечном случае. А именно, мы индукцией по n находим такие значения первых n переменных, при которой все дизъюнкты из D, содержащие только их, истинны. Только теперь надо сделать счетное число шагов, определив значения всех переменных. Поскольку после n-шага значение n-ой переменной не меняется, после счетного количества шагов мы придадим некоторые значения всем переменным. При этих значениях все дизъюнкты из множества D (а значит и из L) истинны. В самом деле, каждый дизъюнкт содержит лишь конечное число переменных. Если дизъюнкт содержит только первые n переменных, то его истинность обеспечена уже на n-ом шаге, а после этого шага значения входящих в него переменных уже не меняются.

Для несчетных списков нужна теорема Цермело. По этой теореме можно вполне упорядочить множество всех переменных, входящих в D. Затем нужное присваивание значений переменным строится рекурсивно так же, как и в счетном случае. Мы не излагаем это рассуждение подробно, поскольку оно требует трансфинитной рекурсии, выходящей за рамки этих лекций. □



(а) $(\neg x_1 \vee x_2) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_3 \vee x_1)$ (б) $x_1 \wedge (\neg x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$

Рис. 2: Перебор по дереву

4 Перебор по дереву и метод резолюций

Итак, применение резолюций к проверке выполнимости КНФ состоит в том, чтобы увеличивать количество резолюций и пытаться вывести противоречие. Как только противоречие получено, можно останавливаться: появилось доказательство, что КНФ невыполнима. В противном случае нужно строить все возможные выводы резолюциями, пока не получится нерасширяемый набор дизъюнктов.

Разумеется, реализовать эту идею в виде алгоритма можно по-разному. Самой удачной оказалась идея соединить выводы резолюциями с направленным поиском выполняющего присваивания.

Эта идея уже появилась в неявном виде в доказательстве полноты опровержений резолюциями. Там мы поочередно присваивали значения переменным так, чтобы при этом ни один из дизъюнктов не обратился в ложь. Если этому условию невозможно удовлетворить, ветка перебора обрывается: на всех продолжениях текущего частичного присваивания хотя бы один из дизъюнктов будет ложным. Наглядно этот процесс изображён на рис. 2. В вершинах дерева, не отмеченных точками, обращается в ложь один из дизъюнктов. Например, на третьем уровне дерева рис. 2(а) частичное присваивание $x_1 = 1$, $x_2 = 0$ обращает в ложь дизъюнкт $\neg x_1 \vee x_2$. На рис. 2(б) допустимых вершин (т.е. тех, в которых ни один дизъюнкт ещё не обратился в ложь) всего две.

Для поиска выполняющего присваивания можно использовать более общий вид перебора по дереву. Не обязательно ветвится по переменным в порядке их номеров. Вместо этого можно каждый раз выбирать наиболее перспективную переменную. Показатели перспективности могут быть разными, от их выбора зависит качество алгоритма перебора. Мы не обсуждаем здесь возможные стратегии выбора переменных. Вместо этого отметим, что такой перебор по дереву очень хорошо согласуется с построением опровержения методом резолюций. Каждая оборванная ветка добавляет новый дизъюнкт в список, как мы уже видели в доказательстве теоремы 3.

Скажем, на рис. 2(а) присваивание $x_1 = 0$, $x_2 = 1$ невозможно продолжить.

Другими словами это можно пересказать так: в выполняющем наборе дизъюнкт $x_1 \vee \neg x_2$ обязан быть истинным. Поэтому этот дизъюнкт можно добавить в список.

Обратите внимание, что этот дизъюнкт выводится по правилу резолюции из дизъюнктов $\neg x_2 \vee x_3$ и $\neg x_3 \vee x_1$.

Комбинируя метод резолюций и перебор по дереву, удаётся построить эффективные на практике алгоритмы проверки выполнимости КНФ, о которых шла речь во введении.

5 Исчисление резолюций для формул первого порядка

Сначала дадим несколько определений. Говорят, что формула находится в *предваренной нормальной форме*, если она имеет вид

$$Q_1 x_1 \ldots Q_n x_n \, A,$$

где $Q_1, \ldots, Q_n$ — кванторы (всеобщности или существования), а формула A не содержит кванторов. Вот пример такой формулы:

$$\forall y \exists z \forall u \exists v \, (P(x, y) \to Q(z, u, v)).$$

Здесь P, Q — предикатные символы. *Универсальной формулой* называется любая формула в предварённой нормальной форме, содержащая только кванторы всеобщности. Пример:

$$\forall x \forall y \forall z \, (P(x, y) \to Q(z, u, v)).$$

Теперь мы обобщим понятие дизъюнкта на формулы первого порядка. Будем называть *дизъюнктом* формулу первого порядка, являющуюся дизъюнкцией атомных формул и отрицаний атомных формул. Пример дизъюнкта:

$$P(f(x), c) \vee \neg Q(y) \vee R.$$

Здесь P, Q, R — двухместный, одноместный и нульместный предикатные символы, а f, c одноместный и нульместный функциональные символы. Мы считает, что порядок формул в дизъюнкте не имеет значения, и кратность вхождения тоже не имеет значения. Например, мы отождествляем слеюдующие три дизъюнкта: $Q(y) \vee R$, $R \vee Q(y)$ и $R \vee Q(y) \vee R$.

Будем называть *универсальными дизъюнктами* универсальные формулы, которые можно получить, навешивая кванторы всеобщности на дизъюнкты. Пример универсального дизъюнкта:

$$\forall y \forall z \, (P(f(x), z) \vee \neg Q(y) \vee R).$$

Обычные дизъюнкты также считаются универсальными дизъюнктами.

5.1 Исчисление резолюций для доказательства несовместности множества универсальных дизъюнктов

Напомним, что множество замкнутых формул называется несовместным, если не существует модели, в которой все формулы множества истинны. Если в множестве имеются незамкнутые формулы, то оно называется несовместным, если не существует модели и значений свободных переменных, для которых в этой модели все формулы множества истинны.

Исчисление резолюций позволяет доказать несовместность любого несовместного множества формул. Сначала мы объясним, как это сделать для множеств, состоящих только из универсальных дизъюнктов.

Исчисление резолюций (ИР) имеет два правила: *правило резолюций* и *правило подстановки*. Первое аналогично правилу резолюций для пропозициональных формул и позволяет вывести $B \vee C$ из $\neg A \vee B$ и $A \vee C$, где A — любая атомная формула, а B, C — любые дизъюнкты. Второе правило позволяет вывести из любого универсального дизъюнкта $\forall x A(x)$ формулу $A(t)$, где t — любой терм. Заметим, что терм t не обязан быть корректным для подстановки вместо x в $A(x)$, то есть, в формуле $A(x)$ могут быть кванторы всеобщности по переменным, входящим в t. Графически, эти два правила можно изобразить так:

$$\frac{\neg A \vee B, \quad A \vee C}{B \vee C} \qquad \frac{\forall x\, A(x)}{A(t)}.$$

Как и раньше, целью является получение пустого дизъюнкта из исходных универсальных дизъюнктов с помощью многократного применения этих двух правил.

Вот пример опровержения в исчислении резолюций. Пусть исходный набор состоит из универсальных дизъюнктов $\forall x\, P(x)$ и $\neg P(f(y))$. Применив правило подстановки к первому универсальному дизъюнкту, мы можем вывести $P(f(y))$. После этого можно применить правило резолюции к полученному дизъюнкту и второму дизъюнкту и вывести пустой дизъюнкт.

Теорема 5 (теорема корректности исчисления резолюций для формул первого порядка)**.** *Если из множества универсальных дизъюнктов можно вывести в ИР пустой дизъюнкт, то это множество несовместно.*

Доказательство. Это почти очевидно. Пусть дано множество универсальных дизъюнктов, из которого можно вывести в ИР пустой дизъюнкт. Рассуждая от противного, допустим, что есть модель и такие значения свободных переменных, что все дизъюнкты из этого множества истинны в этой модели. Как и раньше, правило резолюций сохраняет истинность (в этой модели): если $\neg A \vee B$ и $A \vee C$ истинны, то и $B \vee C$ истинно.

Это же верно и для правила подстановки: если $\forall x A(x)$ истинный универсальный дизъюнкт, то таков и $A(t)$. Здесь важно, что формула $A(x)$ универсальная (для формул общего вида нужно еще потребовать, чтобы замена x на t была корректной). В самом деле, истинность универсальной формулы $\forall x A(x)$ означает, что при подстановке любых элементов носителя вместо всех переменных, связанных кванторами,

полученная формула истинна в нашей модели. Аналогично, истинность универсальной формулы $A(t)$ означает, что при подстановке любых элементов носителя вместо всех переменных, связанных кванторами, полученная формула истинна в нашей модели. Каждая формула, полученная этим способом из $A(t)$, может быть получена и из формулы $\forall x A(x)$ (вместо переменной x надо подставить значение терма t), поэтому истинна.

Таким образом, все дизъюнкты, которые можно вывести из исходного множества, истинны в данной модели при данных значениях свободных переменных (если при применении правила подстановки появились новые свободные переменные, им можно присвоить любые значения). Получаем противоречие, поскольку пустой дизъюнкт ложен. □

Теорема 6 (теорема полноты исчисления резолюций для формул первого порядка)**.** *Из любого несовместного множества универсальных дизъюнктов можно вывести в ИР пустой дизъюнкт.*

Доказательство. Пусть дано несовместное множество универсальных дизъюнктов S. Сначала рассмотрим случай, когда все универсальные дизъюнкты из S замкнуты (не имеют свободных переменных).

Выберем любую переменную x и рассмотрим множество S', состоящее из (обычных) дизъюнктов вида $D(t_1, \ldots, t_n)$, где $t_1, \ldots, t_n$ — некоторые термы, составленные из функциональных символов, входящих в сигнатуру (в частности, и констант, входящих в сигнатуру) и переменной x, а универсальный дизъюнкт $\forall x_1 \ldots \forall x_n\, D(x_1, \ldots, x_n)$ входит в S.

Каждый дизъюнкт из S' можно вывести из исходных универсальных дизъюнктов применением правила подстановки. Поэтому достаточно доказать, что из S' можно получить пустой дизъюнкт с помощью правила резолюции.

Заменим каждую атомную формулу, входящую в формулы из S', на новую пропозициональную переменную (одинаковые атомные формулы заменяем на одну и ту же переменную). Полученное множество пропозициональных дизъюнктов обозначим через S''. Докажем, что S'' несовместно.

Рассуждая от противного, допустим, что можно подобрать значения подставленных пропозициональных переменных, при которых все формулы из S'' истинны. Докажем, что тогда S совместно. В качестве носителя модели рассмотрим множество всех термов, составленные из функциональных символов, входящих в сигнатуру (в частности, и констант, входящих в сигнатуру) и переменной x. Носитель не пуст, поскольку есть хотя бы один такой терм (например, сама переменная x). (Если в сигнатуре есть хотя бы одна константа, то можно было бы в качестве носителя взять множество всех термов без переменных — оно тогда было бы непустым.) Функциональные символы проинтерпретируем естественным образом: значение функции f на термах $t_1, \ldots, t_n$ равно терму $f(t_1, \ldots, t_n)$. Предикатные символы проинтерпретируем в соответствии со значениями пропозициональных переменных: значение предиката P на термах $t_1, \ldots, t_n$ равно значению пропозициональной переменной,

на которую мы заменили атомную формулу $P(t_1, \ldots, t_n)$. Если эта атомная формула не входила в S'', то это значение можно определить любым образом.

Теперь выберем значение переменной x: её значение положим равным терму, состоящий из этой переменной. В построенной модели при таком значении x атомная формула $P(t_1, \ldots, t_n)$ принимает в точности то же значение, что и пропозициональная переменная, на которую она была заменена. Поэтому все дизъюнкты из S' истинны в этой модели при этих значениях переменных.

Осталось проверить, что любой универсальный дизъюнкт из S истинен в этой модели при этих значениях переменных. В самом деле, рассмотрим любой такой дизъюнкт $\forall x_1 \ldots \forall x_n\, D(x_1, \ldots, x_n)$ и подставим вместо переменных $x_1, \ldots, x_n$ термы $t_1, \ldots, t_n$ Нам надо доказать, что полученный дизъюнкт $D(t_1, \ldots, t_n)$ истинен. По построению, дизъюнкт $D(t_1, \ldots, t_n)$ принадлежит S', а как мы только что доказали, все дизъюнкты из S' истинны.

Итак, S'' несовместно. По теореме о полноте пропозиционального ИР из S'' с помощью правила резолюции можно вывести пустой дизъюнкт. Повторяя этот вывод, можно вывести пустой дизъюнкт и из S' (вспомним, что у нас имеется взаимно однозначное соответствие между пропозициональными переменными из S'' и атомными формулами из S').

Заметим, что если в сигнатуре есть хотя бы одна константа, то мы могли не вводить переменную x, а рассмотреть множество всех замкнутых термов (термов, не содержащих переменных). Наличие хотя бы одной константы гарантирует непустоту носителя построенной модели.

Перейдём к случаю, когда в S есть незамкнутые формулы. Обозначим через V множество переменных, имеющих свободное вхождение в формулы из S. Повторим все рассуждения, но вместо термов от новой переменной x рассмотрим термы от переменных из V.

А именно, опять рассмотрим множество S', состоящее из (обычных) дизъюнктов вида $D(t_1, \ldots, t_n)$, где $t_1, \ldots, t_n$ — некоторые термы от переменных из V и универсальный дизъюнкт $\forall x_1 \ldots \forall x_n\, D(x_1, \ldots, x_n)$ входит в S. Опять достаточно доказать, что из S' можно получить пустой дизъюнкт с помощью правила резолюции. Для этого, снова построим множество S'', заменив каждую атомную формулу, входящую в формулы из S', на новую пропозициональную переменную.

Рассуждая от противного, докажем несовместность S''. Допустим, что можно подобрать значения подставленных пропозициональных переменных, при которых все формулы из S'' истинны. Докажем, что тогда S совместно (получив противоречие). В качестве носителя модели рассмотрим множество всех термов от переменных V. Поскольку V непусто, носитель тоже не пуст. Функциональные и предикатные символы проинтерпретируем, как и раньше. Теперь выберем значения переменных: каждой переменной сопоставим терм, состоящий из этой переменной. Опять в построенной модели при таких значениях переменных атомная формула $P(t_1, \ldots, t_n)$ принимает в точности то же значение, что и пропозициональная переменная, на которую она была заменена. Поэтому все дизъюнкты из S' истинны в этой модели при этих значениях переменных. Как и раньше, легко доказать, что любой универ-

сальный дизъюнкт из S истинен в этой модели при этих значениях переменных, что доказывает совместность S.

Полученное противоречие доказывает несовместность S''. Опять же, по теореме о полноте пропозиционального ИР из S'' с помощью правила резолюции можно вывести пустой дизъюнкт и, повторяя этот вывод, можно вывести пустой дизъюнкт и из S'. □

Аналогичными рассуждениями можно доказать *теорему Эрбрана*, утверждающую, из любой невыполнимой универсальной формулы правилом подстановки можно получить невыполнимое конечное множество бесквантворных формул. Мы могли действовать и немного другим образом: сначала доказать теорему Эрбрана и затем из неё и теоремы о полноте пропозиционального ИР вывести Теорему 6.

Задача 7. Докажите теорему Эрбрана.

5.2 Доказательства методом резолюций несовместности множеств формул общего вида

Несовместность множества формул общего вида можно доказать путем сведения к несовместности множества универсальных дизъюнктов с помощью следующих трех шагов, которые мы сначала опишем кратко.

Шаг 1. Каждую из формул исходного множества над преобразовать к предваренной нормальной форме, то есть, построить равносильную формулу, находящуюся в предваренной нормальной форме.

Шаг 2. После первого шага у нас появится множество формул, находящиеся в предваренной нормальной форме. Теперь каждую из них заменим на некоторую универсальную формулу (эта замена называется *сколемизацией* и будет определена ниже). Замененная формула вообще говоря не будет равносильна исходной. Однако несовместность сохранится: новое множество будет совместным тогда и только тогда, когда исходное множество совместно.

Шаг 3. На этом шаге мы заменим каждую универсальную формулу на множество универсальных дизъюнктов. И опять, новое множество будет совместным тогда и только тогда, когда старое множество совместно.

Таким образом, если к исходному множеству применить последовательно три шага, то полученное множество будет совместным тогда и только тогда, когда исходное множество совместно. Поэтому, если мы докажем несовместность полученного множества, выведя в ИР из него пустой дизъюнкт, то докажем и несовместность исходного множества (теорема 5). И обратно, если исходное множество несовместно, то из полученного множества можно вывести пустой дизъюнкт (теорема 6). Теперь подробно опишем все три шага.

Шаг 1. Преобразование к предваренной нормальной форме основано на следую-

щих равносильностях:

$$(A \to B) \sim (\neg A \vee B), \tag{1}$$
$$\neg\exists x\, A \sim \forall x\, \neg A, \quad \neg\forall x\, A \sim \exists x\, \neg A, \tag{2}$$
$$\neg(A \vee B) \sim (\neg A \wedge \neg B), \quad \neg(A \wedge B) \sim (\neg A \vee \neg B), \tag{3}$$
$$\exists x\, A(x) \sim \exists y\, A(y), \quad \forall x\, A(x) \sim \forall y\, A(y), \tag{4}$$
$$(\exists x\, A(x) \wedge B) \sim \exists x\, (A(x) \wedge B), \quad (\exists x\, A(x) \vee B) \sim \exists x\, (A(x) \vee B), \tag{5}$$
$$(\forall x\, A(x) \wedge B) \sim \forall x\, (A(x) \wedge B), \quad (\forall x\, A(x) \vee B) \sim \forall x\, (A(x) \vee B). \tag{6}$$

Сначала с помощью равносильности (1) мы избавляемся от импликации. Затем с помощью равносильностей (2) и (3) мы проносим отрицание к атомным формулам. С помощью (4) мы добиваемся, чтобы все переменные под кванторами были различными (вводя новые переменные). Наконец, с помощью равносильностей (5) и (6) мы выносим все кванторы наружу. Вот пример приведения к предваренной нормальной форме:

$$\begin{aligned}
&\neg(\exists x\, P(x) \to \exists x\, Q(x)) \sim \\
\sim&\neg(\neg\exists x\, P(x) \vee \exists x\, Q(x)) \sim \\
\sim&\exists x\, P(x) \wedge \neg\exists x\, Q(x) \sim \\
\sim&\exists x\, P(x) \wedge \forall x \neg Q(x) \sim \\
\sim&\exists x\, P(x) \wedge \forall y \neg Q(y) \sim \\
\sim&\exists x\, (P(x) \wedge \forall y \neg Q(y)) \sim \\
\sim&\exists x \forall y\, (P(x) \wedge \neg Q(y)).
\end{aligned}$$

Шаг 2. На этом шаге мы применяем *сколемизацию* (по имени норвежского логика Торалфа Сколема). Сколемизация сопоставляет каждой формуле в предварённой нормальной форме некоторую универсальную формулу. Сколемизация сохраняет выполнимость: исходная формула выполнима тогда и только тогда, когда её сколемизация выполнима.

Объясним, что это такое, на примере. Пусть, скажем, исходная формула имеет вид

$$\exists x \forall y \exists z \forall u \exists\, v A(x, y, z, u, v),$$

где $A(x, y, z, u)$ — произвольная бескванторная формула. Её сколемизацией будет формула

$$\forall y \forall u\, A(c, y, f(y), u, g(y, u)).$$

Здесь c — новая константа, а f, g — новые функциональные символы. Новая формула является формулой в расширенной сигнатуре (добавлением символов c, f, g).

Заметим, что новая формула может быть не равносильна старой. Однако если одна из них совместна, то совместна и другая. В самом деле, если новая формула истинна в некоторой модели M, то исходная формула истинна в той же модели

(надо только забыть интерпретации новых символов c, f, g). Действительно, нужные значения x, y, u, которые должны существовать, доставляются константой c и функциями f и g.

В обратную сторону рассуждение немного сложнее. Пусть исходная формула истинна в некоторой модели M. Тогда новая формула истинна в следующей модели, получаемой из M интерпретацией новых символов c, f, g. Сопоставим символу c то значение x, для которого в M истинна формула

$$\forall y \exists z \forall u \exists v\, A(x, y, z, u, v).$$

Такое значение существует, поскольку исходная формула истинна в M. Таким образом, формула

$$\forall y \exists z \forall u \exists v\, A(c, y, z, u, v)$$

истинна в полученной модели. Значит для каждого y можно указать z, для которого в ней истинна формула

$$\forall u \exists v\, A(c, y, z, u, v).$$

Сопоставим символу f любую функцию, которая сопоставляет каждому y некоторое такое z. Значит в полученной модели истинна формула

$$\forall y \forall u \exists v\, A(c, y, f(y), u, v).$$

Наконец, сопоставим символу g любую функцию, которая сопоставляет каждой паре y, u некоторое v, для которого в M истинно

$$A(c, y, f(y), u, g(y, u)).$$

В полученной модели истинна формула

$$\forall y \forall u\, A(c, y, f(y), u, g(y, u)),$$

что и требовалось.

Функции c, f, g, определённые описанным выше образом, называются *сколемовскими функциями*. В общем случае преобразование произвольной формулы в предварённой нормальной форме к равновыполнимой универсальной формуле аналогично: мы сначала заменяем каждую переменную x_i, по которой имеется квантор существования на терм $f(x_1, \ldots, x_{i-1})$, где f новый функциональный символ, а $x_1, \ldots, x_{i-1}$ — все предшествующие x_i переменные в кванторной приставке, связанные квантором всеобщности, а затем стираем все кванторы существования и связанные ими переменные. Доказательство того, что построенная универсальная формула равновыполнима с исходной также совершенно аналогично.

Пусть теперь дана не одна формула, а некоторое множество формул. Применим к каждой из них сколемизацию. При этом для каждой формулы будем использовать свои уникальные символы для сколемовских функций. Очевидно, что построенное множество формул равновыполнимо с исходным. При этом все формулы в нем универсальны.

Шаг 3. Теперь преобразуем каждую универсальную формулу

$$\forall x_1 \ldots \forall x_n \, A(x_1, \ldots, x_n)$$

из полученного множества следующим образом. Приведем формулу A к КНФ, то есть, построим формулу вида $B_1 \wedge \cdots \wedge B_k$, равносильную исходной и такую, что все формулы $B_1 \ldots B_k$ являются дизъюнктами. Поскольку формула $\forall y \, (U \wedge V)$ равносильна формуле $\forall y \, U \wedge \forall y \, V$, исходная универсальная формула равносильна формуле

$$\forall x_1 \ldots \forall x_n \, B_1 \wedge \cdots \wedge \forall x_1 \ldots \forall x_n \, B_k.$$

Следовательно, исходная универсальная формула выполнима тогда и только тогда, когда совместно множество универсальных дизъюнктов

$$\{\forall x_1 \ldots \forall x_n \, B_1, \ldots, \forall x_1 \ldots \forall x_n \, B_k\}.$$

Применим это преобразование ко всем формулам из полученного множества универсальных формул и объединим полученные множества универсальных дизъюнктов. Полученное множество равновыполнимо с исходным.

5.2.1 Пример доказательства несовместности для формул общего вида

Пусть дана (очевидно невыполнимая) формула

$$\neg(\exists x \, P(f(x)) \to \exists x \, P(x)).$$

Мы уже видели, что предваренной нормальной формой этой формулы будет, например, формула

$$\exists x \forall y \, (P(f(x)) \wedge \neg P(y)).$$

После сколемизации получаем формулу

$$\forall y \, (P(f(c)) \wedge \neg P(y)),$$

которая даёт множество из двух универсальных дизъюнктов

$$\{\forall y \, P(f(c)), \forall y \, \neg P(y)\}.$$

С помощью правила подстановки из них выводятся дизъюнкты $P(f(c))$ и $\neg P(f(c))$, из которых по правилу резолюции выводится пустой дизъюнкт.