ISSN 2079-3316 ПРОГРАММНЫЕ СИСТЕМЫ: ТЕОРИЯ И ПРИЛОЖЕНИЯ т. 13, № 4(55), с. 163–179

научная статья МАТЕМАТИЧЕСКИЕ ОСНОВЫ ПРОГРАММИРОВАНИЯ

УДК 519.68:510.64

10.25209/2079-3316-2022-13-4-163-179

Фронтальный алгоритм решения SAT задачи

Юрий Михайлович **Сметанин**✉

Удмуртский государственный университет, Ижевск, Россия

✉*gms1234gms@rambler.ru* *(подробнее об авторе на с. 177)*

Аннотация. Алгоритм вычисления семантического значения конъюнктивных формул вида $U = F(X_1, X_2, ..., X_n)$ в неклассической пропозициональной логике L_{S_2} также вычисляет множество всех решений логического уравнения

$$F(x_1, x_2, ..., x_n) = 1,$$

где $F(X_1, X_2, ..., X_n)$ – формула булевой алгебры множеств, составляющих дискретную диаграмму Венна. Элементы этих множеств являются неотрицательными целыми числами. На основе этого алгоритма строится новый алгоритм для решения задачи SAT. Существенная разница между ним и семейством алгоритмов, основанных на $DPLL$, и $CDCL$, – замена булевых переменных множествами. Это позволяет эффективно проверить выполнимость не одного, а многих наборов значений логических переменных $x_1, x_2, ..., x_n$. (*see abstract in English on p. 178*)

Ключевые слова и фразы: неклассическая пропозициональная логика, основанная на модели с невырожденной булевой алгеброй, исчисление дискретных диаграмм Венна, задача SAT

Для цитирования: Сметанин Ю. М. *Фронтальный алгоритм решения SAT задачи* // Программные системы: теория и приложения. 2022. **Т. 13**. № 4(55). С. 163–179. http://psta.psiras.ru/read/psta2022_4_163-179.pdf

Введение

Задача верификации логического следования в семантическом смысле для классической логики высказываний

$$F_p(x_1, x_2, ..., x_n) \models F_s(x_1, x_2, ..., x_n) \equiv F_p(\widetilde{x}_n) \models F_s(\widetilde{x}_n)$$

сведена к необходимости вычисления множеств U_p и U_s, входящих в суждения неклассической многозначной логики L_{s_2} вида

$$U_p = F_p(X_1^0, X_2^0, ..., X_n^0), \quad U_s = F_s(X_1^0, X_2^0, ..., X_n^0),$$

получаемых из формул $F_p(\widetilde{x}_n), F_s(\widetilde{x}_n)$ заменой булевых переменных x_i на фиксированные для данного n множества целых неотрицательных чисел (конституентые множества) X_i^0. При этом доказано, что логическое следование

$$F_p(\widetilde{x}_n) \models F_s(\widetilde{x}_n)$$

имеет место при наличии отношений $U_p \subset U_s$ либо $U_p = U_s$.

Тип данных «множество» реализуется в языках программирования в универсуме $\{0, 1, ..., 2^n - 1\}$ для небольших значений $n \leq 8$, сложность вычисления операций булевой алгебры множеств для $n > 8$ имеет экспоненциальную оценку.

Показано [1], что множество U, вычисленное по формуле $U = F(X_1, X_2, ..., X_n)$, содержит числа, которые в двоичной системе счисления указывают на все выполняющие подстановки логического уравнения

$$F(x_1, x_2, ..., x_n) = 1.$$

В этом уравнении переменные $x_1, x_2, ..., x_n$ являются характеристическими функциями модельных множеств $X_1, X_2, ..., X_n$; $X_i = U \cdot X_i^0$.

Силлогистическая система L_{s_2}, описанная ниже, имеет областью интерпретации множество $U^0 = \{0, 1, ..., 2^n - 1\}$, n —это число модельных множеств, используемых в ее суждениях. Для интерпретации суждений Φ в этой силлогистике (вычисления их логического объема) конъюнктивные суждения приводятся к атомарной формуле NOB_S (1) $F_s(X_1, X_2, ..., X_n) = U$, ее значение $U = F_s(X_1, X_2, ..., X_n)$ принимается за смысловое содержание Φ. Для верификации логического следования нужно использовать исчисление конституентных множеств и логику L_{S_2} [1–3].

Побочным результатом алгоритмизации вычисления семантического значения формул в L_{S_2} является алгоритм поиска всех выполняющих подстановок ассоциированного с $U = F(X_1, X_2, ..., X_n)$ уравнения

$F(x_1, x_2, ..., x_n) = 1$. На базе этого алгоритма построен новый эффективный алгоритм решения задачи SAT, отличный от семейства алгоритмов, основанных на $DPLL$ и $CDCL$. Существенное сходство данных семейств алгоритмов в том, что логика их работы основана на анализе значений булевых переменных в вырожденной (degenerate) булевой алгебре с универсумом равным $\{1\}$. Выполняющая подстановка ищется путем последовательного присваивания значений булевых переменных.

Существенным отличием предлагаемого фронтального алгоритма является замена булевых переменных модельными множествами. Это позволяет эффективно организовать проверку выполнимости сразу некоторого подмножества наборов значений булевых переменных $x_1, x_2, ..., x_n$, используя исчисление конституентных множеств и логику L_{S_2} [**1**–**3**].

1. Логика (силлогистическая система) L_{S_2}

Атомарные суждения логики (1) выражают объемные отношения множеств в универсуме U. Универсум зависит от числа модельных множеств $U^0 = \{0, 1, ..., 2^n - 1\}$. Семантика дана равносильностями (2)–(4).

(1) $NOB_S = \langle A(X,Y), Eq(X,Y), IO(X,Y), X \subset U, X = U \rangle$,

(2) $A(X,Y) \equiv (X \subset Y) \cdot (X \subset U) \cdot (X' \subset U) \cdot (Y \subset U) \cdot (Y' \subset U)$

(3) $Eq(X,Y) \equiv (X = Y) \cdot (X \subset U) \cdot (X' \subset U) \cdot (Y \subset U) \cdot (Y' \subset U)$

(4) $IO(X,Y) \equiv (X \cdot Y \neq \varnothing) \cdot (X \cdot Y' \neq \varnothing) \cdot (X' \cdot Y \neq \varnothing) \cdot (X' \cdot Y' \neq \varnothing)$

(5) $A(X,Y) \equiv G_{13}(X,Y); \quad q(X,Y) \equiv G_9(X,Y); \quad IO(X,Y) \equiv G_{15}(X,Y).$

На рисунке 1 показаны жергонновы отношения между двумя модельными множествами и универсумом и их связь с атомами (1).

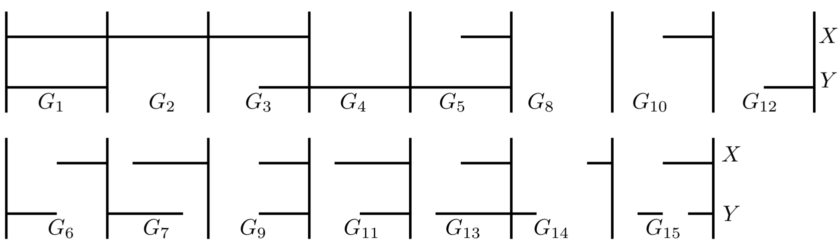

Рисунок 1. Модельные схемы, отражающие жергонновы отношения между объемами терминов X, Y

Определено понятие закона как ППФ, семантическое значение которой есть универсум сопоставленной ей алгебраической онтологии (11), смотри определение 1, противоречием является ППФ, у которой семантическим значением является пустое множество. Определено непарадоксальное (релевантное) логическое следование в семантическом смысле ($\vDash_N$), позволяющее построить логику на основе атомарных высказываний (1). Семантическим значением формулы в L_{S_2} является непустое семейство множеств из неотрицательных целых чисел. NOB_S является альтернативой категорическим суждениям базиса Аристотеля

$$A_S = \langle AXY, EXY, IXY, OXY \rangle,$$

неоднозначную интерпретируемых в 15-ти модельных схемах силлогистики (смотри рисунок 1). Например, в NOB_S общеутвердительное суждение $AXY \equiv$«все X есть Y» для традиционной силлогистики имеет 2 смысла. $AXY \equiv \underbrace{Eq(X,Y)}_{G_9} + \underbrace{A(X,Y)}_{G_{13}}$. Из квадрата Пселла следует, что это же суждение в пятнадцати модельных схемах имеет семь смыслов, задаваемых дизъюнкцией попарно несовместных конъюнкций атомов из (1):

$$\begin{aligned} AXY \equiv\ & \underbrace{(X=Y)\cdot(X=U)}_{G_1} + \underbrace{(X'=U)\cdot(Y=U)}_{G_4} \\ & + \underbrace{(X\subset U)\cdot(X'\subset U)\cdot(Y=U)}_{G_5} + \underbrace{(X=Y)\cdot(X'=U)}_{G_8} \\ & + \underbrace{Eq(X,Y)}_{G_9} + \underbrace{(X'=U)\cdot(Y\subset U)\cdot(Y'\subset U)}_{G_{12}} + \underbrace{A(X,Y)}_{G_{13}} \end{aligned}$$

Здесь множество $X\cdot Y'$ —пересечение X и дополнения Y' до универсума. Вместо X и Y можно подставить любые правильно построенные формулы (ППФ) $F_1(\widetilde{X}_n), F_2(\widetilde{X}_n)$, $\widetilde{X}_n = \langle X_1, X_2, \ldots, X_n\rangle$ алгебры множеств. В логике L_{S_2} используются три логических операции: отрицание, дизъюнкция и конъюнкция. Из атомарных суждений (1) можно составить конъюнкции —конъюнктивные формулы (КФ). Остальные правильно построенные формулы L_{S_2} являются неконъюнктивными формулами (НКФ). Показано, что любая НКФ может быть представлена как дизъюнкция КФ, являющихся попарно противоречивыми. Семантическим значением КФ является множество натуральных чисел либо пустое множество. Семантическим значением НКФ является семейство множеств из натуральных чисел либо пустое множество.

Для интерпретации суждений в силлогистике традиционно используется модельная схема (6), выражающая объемные соотношения между модельными множествами в виде диаграммы Венна (смотри рисунок 2).

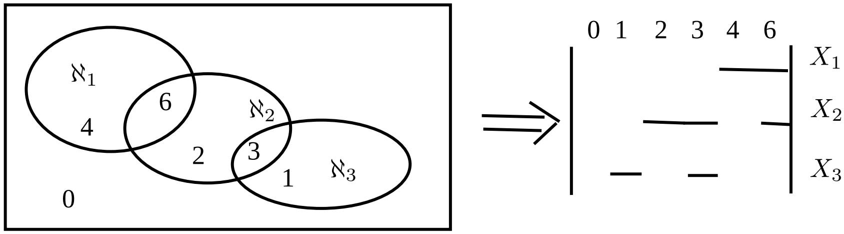

Рисунок 2. Преобразование модельной схемы в А-онтологию

При этом логические соотношения между терминами рассуждений, построенных в базисе Аристотеля, выявляются в алгебраической системе (8) с одним бинарным отношением — нестрогим включением множеств.

$$(6)\qquad M_S = \langle \Omega, \aleph_1, \aleph_2, ..., \aleph_n \rangle\,, \quad \widetilde{\aleph}_i = \langle \aleph_1, \aleph_2, ..., \aleph_n \rangle$$

где Ω–универсум, $\aleph_i \subseteq \Omega$ —модельные множества. Число таких схем не более $2^{(2^n)}$. Оно определяется семейством непустых конституент (7) составляемых из модельных множеств.

$$(7)\qquad \aleph_1^{\sigma_1} \cdot \aleph_2^{\sigma_2} \cdot \dots \cdot \aleph_n^{\sigma_n}, \quad \aleph_i^{\sigma_i} = \begin{cases} \aleph_i, & \sigma_i = 1 \\ \aleph_i', & \sigma_i = 0. \end{cases}$$

Модельные множества $\widetilde{\aleph}_i$ схемы (6) являются элементами носителя алгебраической системы (АС) (8)

$$(8)\qquad \Lambda = \langle \Sigma(\Omega, \widetilde{\aleph}_n),\ \{\ \cdot,\ +,\ \prime\ \},\ \{\subseteq\} \rangle.$$

Носителем $\Sigma(\Omega, \widetilde{\aleph}_n)$ —является семейство подмножеств универсума Ω, которые можно построить из модельных множеств $\aleph_i \subseteq \Omega$ в алгебре множеств. Удобство использования этой АС, в частности, в том, что по теореме Стоуна имеет место изоморфизм с моделью (9), лежащей в основе традиционной логики высказываний с вырожденной булевой алгеброй

$$(9)\quad \lambda = \langle \sigma(\omega, \tilde{x}_n),\ \{\cdot, +, '\},\ \{\leq\} \rangle, \quad \omega = \{1\}, \quad x_i \in\ \omega, \quad i = \overline{1, n}.$$

Недостатком является многосмысловость интерпретации суждений. В данной работе этот недостаток устранен за счет использования модели (10) на основе из конституентных множеств с двумя отношениями $\subset$ и $=$. Морфизм в форме гомоморфизма с вырожденной моделью (9) устанавливается также с помощью характеристической функции множества $\aleph \subseteq \Omega$

$$\forall e \in\ \Omega \left(\chi(\aleph) = \begin{cases} 1, & e \in\ \aleph \\ 0, & e \notin\ \aleph \end{cases}, \quad x_i = \chi(\aleph_i) \right), \quad i = \overline{1, n}.$$

2. Исчисление конституентных множеств

Область интерпретации ППФ L_{S_2} построена из образов n-арных модельных схем вида (6).

Набор $\mathbf{M}$ непустых конституент схемы (6) называется ее *характеристическим множеством*. Модельное множество равно объединению некоторых конституент из $\mathbf{M}$. Это множество можно задать множеством номеров конституент (конституентным множеством).

Таким образом, универсум Ω и модельные множества можно рассматривать как множества из номеров конституент, составляющих $\mathbf{M}$ (**конституентные множества**). Нумерацию естественно производить, зафиксировав порядок индексов модельных множеств. При этом каждой конституенте (7) сопоставляется набор из нулей и единиц $\langle\sigma_1, \sigma_2, ..., \sigma_n\rangle$ и соответствующее ему десятичное число, которое будем называть номером конституенты (см. рисунок 2). Рассмотрим модель для представления n-арной модельной схемы. Обозначим через $B(\widetilde{X}_n)$ семейство из всех конституентных подмножеств универсума, которые могут быть построены из конечной системы $\widetilde{X}_n = \langle X_1, X_2, \ldots, X_n\rangle$ конституентных множеств модельной схемы (6), посредством операций объединения, пересечения и дополнения до универсума. Семейство $B(\widetilde{X}_n)$ включает также конституентное множество универсум (U) и пустое множество. Рассмотрим алгебраическую систему (АС) (10), где $\mathrm{W}_F{=}\{+,\cdot,'\}$, $\quad W_R{=}\{=,\subset\}$[1],

$$\langle B(\widetilde{X}_n), \mathrm{W}_F, \mathrm{W}_R\rangle. \tag{10}$$

Модельной схеме (6) однозначно сопоставляется кортеж

$$I_n = \langle U, X_1, X_2, \ldots, X_n\rangle, \tag{11}$$

выражающий ее через конституентные множества универсума и модельных множеств. Например, для рисунка 2

$$I_3 = \langle U = \{0,1,2,3,4,6\},\ X_1 = \{4,6\},\ X_2 = \{2,3,6\},\ X_3 = \{1,3\}\rangle\,.$$

Определение 1. *Единицей M алгебраической системы* (10), *называется множество номеров ее непустых конституент $M = U$. Нулем называется множество $N = U^0 \setminus M$, $U^0 = \{0, 1, \ldots, 2^n - 1\}$. Конституентные множества $X_1, X_2, \ldots, X_n$ также будем называть модельными. Кортеж* (11) *будем называть алгебраической онтологией* (*А-онтологией*).

---

[1] $\mathrm{W}_F{=}\{+,\cdot,'\}$ – операции объединения, пересечения, дополнения до универсума алгебры множеств

А-онтология $I_n^0 = \langle U^0, X_1^0, X_2^0, ..., X_n^0 \rangle$; $M(I_n^0) = U^0$ *называется канонической, если* $U^0 = \{0, 1, ..., 2^n - 1\}$. *Все конституенты соответствующей ей модельной схемы непусты* (*см. рисунок 4*). *Будем называть модельную схему для канонической А-онтологии* I_n^0 *отношением независимости в совокупности ее модельных множеств.*

Замечание 1. *АС* (10) *с выбранной нумерацией модельных множеств задается А-онтологией* (11) *которая является формой, выражающей* n*-арную модельную схему* (6)[2].

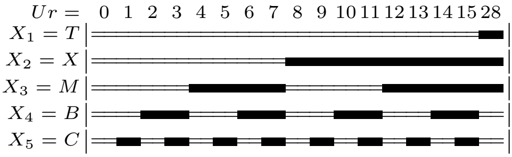

Рисунок 3. Контекст понятия «тигры» $T \subset X \cap M \cap B' \cap C'$

Например, понятие «все тигры —хищные млекопитающие, не живущие в воде и не приспособленные к жизни в условиях Крайнего Севера», выражаемое атомным суждением $A(T, X \cdot M \cdot B' \cdot C')$, представляется А-онтологией (рисунок 3). Она иллюстрирует его неаристотелевское строение [4]. Достоинством такой модели понятия является то, что оно содержит контекст (хотя и ограниченный); ему можно сопоставить алгоритм вычисления объема по логическому содержанию. Алгоритмическая составляющая таких понятий проиллюстрирована в [2]. Контекст понятия можно изменять, дополняя его другими модельными множествами.

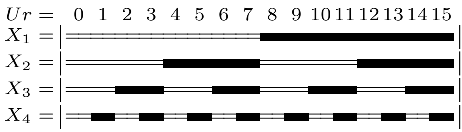

Рисунок 4. Каноническая А-онтология для $n = 4$

Множество всех бинарных отношений между модельными множествами для заданной А-онтологии I_n будем называть *полным бинарным инвариантом* $BIN(I_n)$ [1]. Конъюнкцию составленную из бинарных отношений входящих в $BIN(I_N)$, обозначим $FBIN(I_n)$. Для сокращения $FBIN(I_n)$ из $n \cdot (n-1)/2$ бинарных отношений

[2]Наглядно А-онтологию будем изображать в виде линейной диаграммы (смотри рисунки 2, 4, 5).

составляющих $BIN(I_n)$ будем опускать $IO(X_i, X_j)$ —отношения независимости пары модельных множеств. Каждому BIN соответствует не менее одной А-онтологии. **Максимальной** среди них называется та, которая имеет единицу с наибольшим числом номеров конституент. Добавление к ее единице любого не входящего в нее номера выводит эту А-онтологию из данного BIN (см. рисунки 3, 5). Добавление конституенты с номером 29 к А-онтологии Тигры искажает понятие. Тигры при этом могут жить на Крайнем Севере.

Замечание 2. *Логическое содержание любой А-онтологии I_n выражается атомарным суждением $NSKF = U$. В левой части стоит нормальная совершенная форма Кантора* (12)*, составленная из непустых конституент. При этом любое утверждение $F(\tilde{X^n}) = U$ в котором левая часть эквивалентная формула для $NSKF$ также выражает логическое содержание I_n. Логическое содержание немаксимальной А-онтологии I_n^{nomax} выражается формулой в виде конъюнкции логического содержания максимальной А-онтологии I_n с конъюнкцией дизституент выражающей пустоту не входящих в I_n конституент.*

Бинарный инвариант $BIN = \{A(X_1', X_2);\ IO(X_1, X_2);\ A(X_1', X_4);\ A(X_2', X_3');\ IO(X_2, X_3);\ A(X_3, X_4)\}$ имеют только две А-онтологии, одна из которых максимальная (см. рисунок 5)*. Это осуществляется непосредственной проверкой. Логическое содержание для не максимальной А-онтологии с этого рисунка выражаются КФ составленной из КФ* (12) *и дополнения конституенты $\big((X_1 \cdot X_2 \cdot X_3' \cdot X_4)' = U\big)$ выражающего дизституенту $\big(X_1' + X_2' + X_3 + X_4' = U\big)$.*

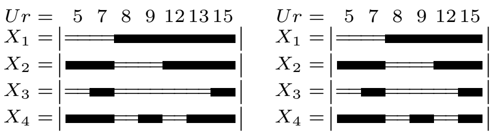

Рисунок 5. Две А-онтологии и их $FBIN = A(X_1', X_2) \cdot A(X_1', X_4) \cdot A(X_2', X_3) \cdot A(X_3, X_4)$

Таким образом, А-онтология I_n позволяет выразить модельную схему (6) в виде атомарного суждения $F(\widetilde{X}_n) = M(I_n)$, здесь $F(\widetilde{X}_n)$ — ППФ алгебры множеств, эквивалентная совершенной нормальной форме Кантора $SNFK$, сопоставляемой $M(I_n)$. Например, максимальная А-онтология рисунка 5 представляется суждением $NOBS$ с левой частью в форме $SNFK(FBIN(\widetilde{X}_4))$ (12), либо с левой частью в виде равносильной ей ППФ $-X_1 \cdot {X_3}' + X_2 \cdot X_4 = U$.

$$
\begin{aligned}
(12) \qquad & X_1' \cdot X_2 \cdot X_3' \cdot X_4 + X_1' \cdot X_2 \cdot X_3 \cdot X_4 + X_1 \cdot X_2' \cdot X_3' \cdot X_4' \\
& + X_1 \cdot X_2' \cdot X_3' \cdot X_4 + X_1 \cdot X_2 \cdot X_3' \cdot X_4' + X_1 \cdot X_2 \cdot X_3' \cdot X_4 \\
& + X_1 \cdot X_2 \cdot X_3 \cdot X_4 = U = \{5, 7, 8, 9, 12, 13, 15\}.
\end{aligned}
$$

Изменять А-онтологии позволяет исчисление конституентных множеств (теоремы 1, 2).

Теорема 1. *Для А-онтологии I_n с n модельными множествами X_i и универсумом $U(I_n)$ выполняются соотношения (13)*

$$
(13) \qquad
\begin{aligned}
& M(I_n) = \bigcap_{i \in \ N} D(i'); \quad i' = 2^n - i - 1; \\
& X_i = U(I_n) \cdot X_i^0 = M(I_n) \cdot {X_i}^0, \quad i = 1, n;
\end{aligned}
$$

X_i^0-модельные множества канонической А-онтологии. N – множество номеров констуент

Доказательство теоремы 1. Для доказательства достаточно заметить, что множество $D(i') = U^0 \setminus \{i\}$. Здесь $U^0 = M(I_n^0)$— единица канонической А-онтологии. $D(i') = K(i)'$, $i' = (2^n - 1 - i)$ —дизституента, двойственная i-ой конституенте. Принцип нумерации дизституент такой же как у конституент. $K(2)' = (X_1' \cdot X_2 \cdot X_3')' = D(5) = X_1 + X_2' + X3$. $\square$

Из соотношения (13) теоремы 1 следует, что объемы модельных множеств А-онтологии I полностью определяются ее единицей $M(I)$. Теорема 2 позволяет последовательно вводить в исходную А-онтологию бинарные отношения (суждения NOB_S), если они в ней отсутствуют.

Теорема 2. *1. Чтобы ввести в n-арную А-онтологию отношение $X \subset U$, достаточно добавить к ее единице хотя бы одну конституенту с номером из конституентного множества $X\ ' = U^0 \setminus X$, где U^0-универсум n-арной канонической А-онтологии.*
2. Чтобы ввести в А-онтологию отношение $X = U$, необходимо и достаточно убрать из нее все конституенты с номерами из конституентного множества $X\ ' = U \setminus X$.
3. Чтобы ввести в А-онтологию отношение $X \subset Y$, необходимо и достаточно убрать из нее все конституенты с номерами, образующими конституентное множество $X \cdot Y\ '$.
4. Чтобы ввести в А-онтологию отношение $X = Y$, необходимо и достаточно убрать из нее все конституенты с номерами, образующими контституентное множество $X\ ' \cdot Y + X \cdot Y\ '$.
5. Чтобы ввести в А-онтологию отношение $IO(X, Y)$, достаточно добавить в нее хотя бы по одному номеру из констиуентных множеств $X\ ' \cdot Y\ '$, $X\ ' \cdot Y$, $X \cdot Y\ '$, $X \cdot Y$, которые являются в ней пустыми.

Справедливость теоремы доказывается рассмотрением ее утверждений для всех пятнадцати бинарных модельных схем.

Теорема 3. *Имеет место соотношение.*

$$F(\widetilde{X}_n^0) = U = F(\widetilde{X}_n), \quad X_i = {X^0}_i \cdot U, \quad i = \overline{1, n}; \tag{14}$$

Доказательство теоремы 3. Для доказательства достаточно указать на то, что равенство $F(\widetilde{X}_n^0) = U$ является атомом силогистики L_{S_2} и к нему как к конъюнктивной формуле применима теорема 2. □

На основе теорем 1, 2 и 3 разработан M-алгоритм, использующий утверждения 2-4 теоремы 2 для вычисления А-онтологии $I(Q)$, сопоставленной конъюнкции Q суждений (1) [1, 2] и ее единицы $M(I(Q))$. Данный алгоритм реализован программно на языке Free Pascal. Сложность вычислений экспоненциально зависит от числа $n-8$, где $n-$ число модельных множеств. Объем используемой памяти также экспоненциально зависит от $n-8$, например, для 22 переменных канонический универсум занимает $2^{22-8} \cdot 2^5 = 0,5$ мегабайт дисковой памяти и представлен массивом из базовых множеств $[0, ..., 255]$, содержащим 16384 элемента. Если число переменных будет 40, то объем будет $2^{40-8} \cdot 2^5 = 2^37 = 2^{10} \cdot 2^{10} \cdot 2^{10} \cdot 128 = 128$ гигабайт. Время вычисления единицы линейно зависит от количества K_{op} операций и равно

$$T = K_{op} \cdot t_0 \cdot 2^{22-8},$$

кроме того, к этому числу нужно добавить время для потоковой записи и считывания множеств из массивов. Здесь $t^0 = \max\{t(+), t(\cdot), t^0\}$ это максимальное время выполнения операций объединения, пересечения и дополнения до универсума для базового множества. Указан способ эффективного распараллеливания алгоритма вычислений [1].

Замечание 3. *Доказано, что А-онтология, получаемая в по M-алгоритму для случаев 2–4 теоремы 2, является* ***максимальной*** *по числу номеров непустых конституент.*

Теорема 4. *Имеет место функциональная полнота атомарных суждений (1), то есть любая А-онтология может быть выражена КФ в L_{S_2}.*

Доказательство теоремы 4. Для того, чтобы представить единицу любой А-онтологии I, как КФ логики L_{S_2}, достаточно записать ее $FBIN$ и найти $M(FBIN(I))$ используя теорему 2. Получится максимальная А-онтология с данным BIN. Если исходная А-онтология не совпадает с максимальной, необходимо достроить $FBIN$ до КФ

$FBIN \cdot D(i_1') \cdot D(i_2') \ldots \cdot D(i_k')$, где $\{i_1, i_2, \ldots, i_k\} = M(FBIN(I)) \setminus M(I)$. Тогда согласно теореме 1 $M\big(I\big) = M\big(FBIN \cdot D(i_1') \cdot D(i_2') \ldots \cdot D(i_k')\big)$ $\square$

Например, для того, чтобы выразить логическое содержание правой А-онтологии на рисунке 5 в $FBIN$, нужно добавить множитель $D(2) = U$, который утверждает пустоту конституенты $K(13)$, $13' = 2^4 - 13 - 1 = 2$, $D(2) = (X_1' + X_2' + X_3 + X_4') = U$.

Таким образом, каждой конъюнктивной ППФ логики L_{S_2} ставится в соответствие множество номеров непустых конституент, которое называется семантическим значением этой ППФ. Каждому множеству конституентных номеров соответствует одна модельная схема. Каждому утверждению $F(\widetilde{X}^n) = U$ сопоставляется логическое уравнение $F(\tilde{x}^n) = 1$, множество решений которого определяются единицей утверждения $F(\widetilde{X}^n) = U$. Например, равенству $X_1 \cdot {X_3}' + X_2 \cdot X_4 = U$ сопоставляется уравнение $x_1 \cdot {x_3}' + x_2 \cdot x_4 = 1$ и множество его выполняющих подстановок заданное номерами конституент из $M(X_1 \cdot {X_3}' + X_2 \cdot X_4 = U) = \{5, 7, 8, 9, 12, 13, 15\}$ смотри (12).

3. Фронтальный алгоритм решения SAT задачи

Проблема SAT часто описывается в формате DIMACS-CNF. Для фронтального алгоритма мы примем новую нотацию, которую легко получить из этого формата. Рассмотрим уравнение

$$f(\tilde{x}_4) = (x_2' + x_3 + x_4') \cdot (x_1' + x_3') \cdot (x_3 + x_4) \cdot (x_1 + x_2') \cdot (x_3) \cdot (x_1' + x_3' + x_4) = 1 \tag{15}$$

Сопоставим ему атомарное суждение из $NOBS$,

$$(X_2' + X_3 + X_4') \cdot (X_1' + X_3') \cdot (X_3 + X_4) \cdot (X_1 + X_2') \cdot (X_3) \cdot (X_1' + X_3' + X_4) = U, \tag{16}$$

которое можно рассматривать как конъюнкцию атомарных формул $NOBS$ вида

$$(X_2' + X_3 + X_4') = M1; \tag{17}$$

$$(X_1' + X_3') = M2; \tag{18}$$

$$(X_3 + X_4) = M3; \tag{19}$$

$$(X_1 + X_2') = M4; \tag{20}$$

$$(X_3) = M5; \tag{21}$$

$$(X_1' + X_3' + X_4) = M6. \tag{22}$$

Универсум семантического значения конъюнкции суждений формул (17)–(22) вычисляется по формуле

$$U = (X_2^{0\prime} + X_3^0 + X_4^{0\prime}) \cdot (X_1^{0\prime} + X_3^{0\prime}) \cdot (X_3^0 + X_4^0) \cdot (X_1^0 + X_2^{0\prime}) \\ \cdot (X_3^0) \cdot (X_1^{0\prime} + X_3^{0\prime} + X_4^0),$$

см. теорему 3, где ${X^0}_i$ – модельные множества канонической А-онтологии I_n. Он также может быть вычислен как пересечение единиц суждений (17)–(22). Каноническую А-онтологию с модельными множествами X_i^0, $i = 1, 4$ будем задавать в алгебре кортежей [5] в виде C-системы:

$$\begin{matrix} X_1^0 = \\ X_2^0 = \\ X_3^0 = \\ X_4^0 = \end{matrix} \begin{bmatrix} 1 & * & * & * \\ * & 1 & * & * \\ * & * & 1 & * \\ * & * & * & 1 \end{bmatrix}$$

здесь $* = \{0, 1\}$, единицей обозначено множество $\{1\}$ каждая строка задает конституентные множества получаемые из декартова произведения множеств одного кортежа.

$$X_1^0 = \{8, 9, 10, 11, 12, 13, 14, 15\} \qquad X_2^0 = \{4, 5, 6, 7, 12, 13, 14, 15\}$$
$$X_3^0 = \{2, 3, 6, 7, 10, 11, 14, 15\} \qquad X_4^0 = \{1, 3, 5, 7, 9, 11, 13, 15\}$$

Например, последнее модельное множество X_4^0 получается из наборов нулей и единиц выраженных выше десятичными числами.

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Они получаются из декартова произведения множеств, стоящих в строке обозначенной как X_4^0. Смотри также рисунок 4. Зададим нуль конъюнкции суждения (17)–(22) как объединение нулей каждого из них.

Эти нули могут получиться как C-система состоящая из C-кортежей (нулей) сопоставленных каждому клозу.

$$N(f(\tilde{x}_4)) = N(1) + N(2) + N(3) + N(4) + N_5 + N_6$$

$$= \begin{bmatrix} * & 1 & 0 & 1 \\ 1 & * & 1 & * \\ * & * & 0 & 0 \\ 0 & 1 & * & * \\ * & * & 0 & * \\ 1 & * & 1 & 0 \end{bmatrix} = \begin{bmatrix} N(1) \\ N(2) \\ N(3) \\ N(4) \\ N(5) \\ N(6) \end{bmatrix}$$

Принцип кодирования C-кортежом нуля —невыполняющих подстановок каждого клоза весьма прост. Клозу $(x'_2 + x_3 + x'_4)$ сопоставляется C-кортеж $N(1) = [* \quad 1 \quad 0 \quad 1]$ определяющий конституентное множество $\{5, 13\}$ соответствующее двум подстановкам 0101 и 1101, на которых клоз невыполним. Объединение нулей представим как объединение непересекающихся (ортогональных) множеств по методу, предложенному Порецким П.С. Соответственно искомая единица получается как $U_4^0 \setminus N(f(\widetilde{X}_4))$.

$$N(f(\tilde{X}_4)) = \begin{bmatrix} N(1) \\ N(2) \cdot N(1)' \\ N(3) \cdot N(1)' \cdot N(2)' \\ N(4) \cdot N(1)' \cdot N(2)' \cdot N(3)' \\ N(5) \cdot N(1)' \cdot N(2)' \cdot N(3)' \cdot N(4)' \\ N(6) \cdot N(1)' \cdot N(2)' \cdot N(3)' \cdot N(4)' \cdot N(5)' \end{bmatrix}'$$

Более предпочтительным является способ вычисления единицы $N(f(\tilde{X}_4)$ как пересечения всех единиц суждений (17)–(22). Каждая единица $M(i) = N(i)'$ вычисляется как дополнение нуля. Это дополнение можно получить как объединение ортогональных множеств выражаемых C-кортежами объединенными в C-систему. Алгоритм получения единицы клоза является эффективным и имеет линейную сложность относительно количества нулей и единиц в C-кортеже дополнение которого находится. Тогда имеем:

$$N_1{}' = M_1 = \begin{bmatrix} * & 0 & * & * \\ * & 1 & 1 & * \\ * & 1 & 0 & 0 \end{bmatrix}, \qquad N'_2 = M_2 = \begin{bmatrix} 0 & * & * & * \\ 1 & * & 0 & * \end{bmatrix},$$

$$N'_3 = M_3 = \begin{bmatrix} * & * & 1 & * \\ * & * & 0 & 1 \end{bmatrix}, \qquad N'_4 = M_4 = \begin{bmatrix} 1 & * & * & * \\ 0 & 0 & * & * \end{bmatrix},$$

$$N'_5 = M_5 = \begin{bmatrix} * & * & 1 & * \end{bmatrix}, \qquad N'_6 = M_6 = \begin{bmatrix} 0 & * & * & * \\ 1 & * & 0 & * \\ 1 & * & 1 & 1 \end{bmatrix},$$

$M(f(\tilde{x}_4)) = M_1 \cdot M_2 \cdot M_3 \cdot M_4 \cdot M_5 \cdot M_6;$

Автором разработан эффективный алгоритм вычисления произведений для данного частного случая C-систем. Количество операций умножения C-кортежей зависит от порядка выполнения умножений единиц M_i. Для порядка 123456 число произведений равно 30. Для порядка 256413 число произведений равно 11, для порядка 253416 оно равно 12. Общая единица равна $\{2, 3\}$. То есть количество выполняющих подстановок для уравнения

$$(x_2' + x_3 + x_4') \cdot (x_1' + x_3') \cdot (x_3 + x_4) \cdot (x_1 + x_2') \cdot (x_3) \cdot (x_1' + x_3' + x_4) = 1$$

равно двум $\{\langle 0010\rangle, \langle 0011\rangle\}$. Минимизация числа операций умножения достигается за счет приоритетного выполнения перемножений единиц, которые представлены меньшим числом ортогональных множеств.

При решении задачи нахождения всех выполняющих подстановок сведенной к поиску единицы сопряженного с логическим уравнением $F(x_1, x_2, ..., x_n) = 1$, суждения $F(X_1, X_2, ..., X_n) = U$ возникает проблема быстрого возрастания объема промежуточных результатов вычислений для числа переменных свыше 40. Данная проблема может быть решена при последовательных вычислениях выгрузкой этих промежуточных результатов из оперативной памяти в долговременную.

Ограничив объем используемых в оперативной памяти промежуточных результатов допустимой порцией, мы приходим к классическому алгоритму с возвратами, дерево вычислений которого развивается в глубину. При этом на каждом шаге на выполнимость для очередного клоза тестируется сразу множество подстановок.

Для решения SAT задачи достаточно, используя стратегию поиска в глубину, обнаружить хотя бы одну выполняющую подстановку для всех клозов. Проверка на выполнимость очередного клоза осуществляется путем, примерки к нему на предмет выполнимости сразу множества подстановок. «Ширину фронта» проверки (объем порции проверяемых подстановок) можно увеличивать и уменьшать.

Алгоритм реализован на персональном компьютере с процессором $Intel(R)Core\ i5 - 2430\ CPU$ @ 2.40 GHz. Оперативная память 4.00 ГБ. Операционная система Windows 7, 64 разрядная. Логическое уравнение в форме КНФ=1 генерируется случайны образом.

При вычислениях замечена закономерность, заключающаяся в том, что уменьшение ширины фронта алгоритма приводит к увеличению времени поиска выполняющей подстановки для заданного уравнения. Время вычислений выполняющих подстановок в общем случае экспоненциально зависит от числа булевых переменных в логическом уравнении. В случае отсутствия решений придется перебрать все возможные подстановки, используя стратегию поиск в глубину по дереву.

Предлагаемый алгоритм можно подвергнуть крупноблочному распараллеливанию и получать решение SAT с помощью грид вычислений. Об этом будет изложено в следующей публикации.

Список литературы

[1] Ю. М. Сметанин *Непарадоксальное логическое следование и проблема решения МЛ-уравнений* // Программные системы: теория и приложения.– 2016.– Т. **7**.– № 1(28).– с. 99–115. URL ↑164, 165, 169, 172

[2] Ю. М. Сметанин *Верификация логического следования с использованием исчисления конституентных множеств и соответствий Галуа* // Программные системы: теория и приложения.– 2017.– Т. **8**.– № 2(33).– с. 69–93. doi URL ↑164, 165, 169, 172

[3] Ю. М. Сметанин *Верификация логического следования в неклассической многозначной логике* // Изв. ИМИ УдГУ.– 2017.– Т. **50**.– с. 62–82. doi ↑164, 165

[4] В. К. Финн *О неаристотелевском строении понятий* // Логические исследования.– 2015.– № 21(1).– с. 9–48. doi ↑169

[5] Б. А. Кулик, А. А. Зуенко, А. Я. Фридман *Алгебраический подход к интеллектуальной обработке данных и знаний*.– СПб.: Изд-во Политехн. ун-та.– 2010.– ISBN 9785742228363.– 235 с. ↑174

Поступила в редакцию	14.10.2022;
одобрена после рецензирования	09.11.2022;
принята к публикации	10.11.2022.

Рекомендовал к публикации *д.ф.-м.н. Н. Н. Непейвода*

Информация об авторе:

Юрий Михайлович Сметанин

Кандидат физико-математических наук, доцент. Область научных интересов: прикладная логика, силлогистика

iD 0000-0002-2508-4939

e-mail: ***gms1234gms@rambler.ru***

Автор заявляет об отсутствии конфликта интересов.

ISSN 2079-3316 **PROGRAM SYSTEMS: THEORY AND APPLICATIONS** **vol. 13**, No 4(55), pp. 163–179
Research Article THEORETICAL COMPUTER SCIENCE

UDC 519.68:510.64
10.25209/2079-3316-2022-13-4-163-179

Front-end algorithm for solving the SAT problem

Yurii Mikhailovich **Smetanin**
Udmurt State University, Izhevsk, Russia
✉*gms1234gms@rambler.ru* *(learn more about the author in Russian on p. 177)*

Abstract. The algorithm for calculating the semantic value of conjunctive formulas of the form $U = F(X_1, X_2, ..., X_n)$ in non-classical propositional logic L_{S_2} also calculates the set of all solutions of the logical equation

$$F(x_1, x_2, ..., x_n) = 1.$$

Where $F(X_1, X_2, ..., X_n)$ — a formula of Boolean algebra of the sets making a discrete Venn diagram. The elements of these sets are non-negative integers. Based on this algorithm, a new algorithm is built to solve the SAT problem. A significant difference between it and a family of algorithms based on $DPLL$ and $CDCL$ is the replacement of Boolean variables with sets. This allows you to effectively check the feasibility of not one, but many sets of values of the logical variables $x_1, x_2, ..., x_n$. (*In Russian*).

Key words and phrases: non-classical propositional logic based on a model with non-degenerate Boolean algebra, calculus of discrete Venn diagrams, problem SAT

2020 *Mathematics Subject Classification:* 08A70; 03G05,93B60

For citation: Yurii M. Smetanin. *Front-end algorithm for solving the SAT problem* // Program Systems: Theory and Applications, 2022, **13**:4(55), pp. 163–179. (*In Russian*). http://psta.psiras.ru/read/psta2022_4_163-179.pdf

References

[1] Yu. M. Smetanin. "Non-paradoxical logical consequence and the problem of solving ML-equations", *Program Systems: Theory and Applications*, **7**:1(28) (2016), pp. 99–115 (in Russian). URL ↑164, 165, 169, 172

[2] Yu. M. Smetanin. "Verification of logical consequence, using the calculus of constituent sets and correspondences of Galois", *Program Systems: Theory and Applications*, **8**:2(33) (2017), pp. 69–93 (in Russian). doi URL ↑164, 165, 169, 172

[3] Yu. M. Smetanin. "Verification of the logical sequence in nonclassical multivalued logic", *Izv. IMI UdGU*, **50** (2017), pp. 62–82 (in Russian). doi ↑164, 165

[4] V. K. Finn. "On the non-Aristotelian concept structure", *Logicheskiye issledovaniya*, 2015, no. 21(1), pp. 9–48 (in Russian). doi ↑169

[5] B. A. Kulik, A. A. Zuyenko, A. Ya. Fridman. *Algebraic approach to intelligent data and knowledge processing*, Izd-vo Politekhn. un-ta, SPb., 2010, ISBN 9785742228363 (in Russian), 235 pp. ↑174