УДК 51.681.3

УСОВЕРШЕНСТВОВАННЫЙ МЕТОД ПРОВЕРКИ ВЫПОЛНИМОСТИ МНОЖЕСТВА ДИЗЪЮНКТОВ В ЯЗЫКЕ L

Кривой С.Л., Чеботарев А.Н.

Институт кибернетики им. В.М.Глушкова НАН Украины,
пр-т Глушкова,40, г.Киев, Украина, 03187
e-mail: *krivoi@i.com.ua*

Abstract

The language L is used for specifying finite automata, and is a fragment of a first order language with monadic predicates. Checking specification for satisfiability plays an important role in the development of reactive algorithms. Restricted syntax of this language and interpreting it over the integers make it possible to substantially improve resolution-based methods for satisfiability checking. In this paper, we present an improvement to the method based on the restriction of the type of atoms upon which the resolution is allowed.

Введение

Язык L используется в качестве языка спецификаций в системе доказательного проектирования реактивных алгоритмов исходя из их логической спецификаций [1]. Этот язык является подмножеством языка первого порядка с одноместными предикатами интерпретируемыми на м ножестве целых чисел. Проверка выполнимости спецификаций играет важную роль в процессе доказательного проектирования. Соответствующие процедуры используются не только для определения внутренней непротиворечивости спецификации, но также и для верификации решений, принимаемых проектировщиком (изменений в спецификации) при интерактивном проектировании алгоритма [2, 3]. Следовательно, такой процесс проектирования требует разработки высокоэффективных алгоритмов проверки выполнимости формул. Алгоритм, предлагаемый в настоящей работе, основан на резолюционном методе поиска вывода. Основная причина неэффективности проверки выполнимости формулы резолюционным методом заключается в порождении большого числа избыточных дизъюнктов, т.е. дизъюнктов, не влияющих на получаемый результат.

Анализ последних достижений и публикаций,посвященных этой проблеме, приводит к следующим выводам. Уменьшение числа порождаемых дизъюнктов может быть получено благодаря ограничениям, на применение правила резолюции, сохраняющим полноту метода. Эффективный метод проверки выполнимости формул языка L был предложен в [4], где упрощение соответствующей процедуры достигалось путем наложения ограничения на вид резольвируемых атомов. Правило резолюции в этом методе было названо R-резолюцией. Дополнительные возможности улучшения этого метода были предложены в работе [5], где множество дизъюнктов разбивалось на несколько классов и резольви-рование осуществлялось только между дизъюнктами, принадлежащими одному и тому же классу. Эти ограничения позволяют существенно уменьшить число получаемых резольвент по сравнению с методом R-резолюции.

*Целью настоящей работы является*усовершенствование метода R-резолюции, связанное с дополнительным ограничением, налагаемым на вид атомов, по которым допускается резольвирование.

ОСНОВНЫЕ ПОНЯТИЯ И ОБОЗНАЧЕНИЯ

Приведем сначала краткое описание понятий, относящихся к языку L и R - резолюции. Более подробное описание этого языка и метода R-резолюции можно найти в [4]. Пусть T- класс формул, построенных с помощью логических связок из атомарных формул вида $p(t+k)$, где p - одноместный предикатный символ, t - переменная, принимающая значения из множества целых чисел $\mathbf{Z}$, и k-целочисленная константа, называемая рангом атома. Язык L состоит из формул вида $\forall tF(t)$, где $F(t) \in T$, интерпретируемых на множестве $\mathbf{Z}$. Примером такой формулы может служить формула $\forall t(y(t-1)x(t) \to y(t))$, где y и x-предикатные символы, а $(t-1)$ - сокращение для $(t+(-1))$. Формула $\forall tF(t)$ называется выполнимой, если для нее существует непустая модель, т. е. интерпретация, в которой она принимает истинное значение. Поскольку $F(t)$ интерпретируется на множестве целых чисел, то имеет место эквивалентность $\forall tF(t) \Leftrightarrow \forall tF(t+k)$, где $F(t+k)$ означает формулу, полученную из $F(t)$ путем добавления целого числа k к рангу всех атомов. Таким образом, можно ограничиться рассмотрением только таких формул, максимальный ранг атомов в которых равен 0. Такие формулы будем называть нормализованными вправо. Предполагается, что формула $F(t)$ в спецификации записана в конъюнктивной нормальной форме, представленной в виде множества дизъюнктов, т. е. дизъюнкций литер, где литерой является атом или его отрицание. Дизъюнкт, не содержащий литер, называется пустым (обозначается $\square$). Множество дизъюнктов называется нормализованным вправо, если каждый дизъюнкт этого множества нормализован вправо.

Определение 1. Пусть $c_1 = c \vee p(t)$, $c_2 = c^{'} \vee \neg p(t)$ нормализованые вправо дизъюнкты, где $p(t)$ атом ранга 0. Дизъюнкт $c \vee c^{'}$ называется R-резольвентой дизъюнктов c_1 и c_2 по атому $p(t)$.

R-резолюция (ограниченная резолюция) является правилом вывода, которое допускает резольвирование только по атомам нулевого ранга.

Определение 2. R-выводом дизъюнкта c из множества дизъюнктов C называется конечная последовательность дизъюнктов $c_1, ..., c_m$ такая, что $c_m = c$ и каждый дизъюнкт c_i $(i = 1, ..., m)$ либо принадлежит C, либо является R-резольвентой c_j и c_k для $j, k < i$, либо есть результат нормализации вправо дизъюнкта c_{i-1}.

Определение 3. Дизъюнкт $c_1(t)$ поглощает дизъюнкт $c_2(t)$, если существует такое $k \in \mathbf{Z}$, что все литеры дизъюнкта $c_1(t+k)$ содержатся среди литер дизъюнкта $c_2(t)$.

Множество дизъюнктов C называется противоречивым, если оно задает невыполнимую формулу $\forall t F(t)$. В [4] доказано следующее предложение.

Предложение 1. Множество C нормализованных вправо дизъюнктов невыполнимо тогда и только тогда, когда существует R-вывод пустого дизъюнкта из C.

Соответствующая процедура проверки выполнимости называется процедурой R-пополнения. В этой процедуре после добавления каждого нового дизъюнкта в текущее множество дизъюнктов, выполняется процедура поглощения, т.е. дизъюнкты, поглощаемые другими дизъюнктами, удаляются.

ОПИСАНИЕ РЕЗОЛЮЦИОННОГО МЕТОДА

Пусть $p_1 < p_2 < \ldots < p_n$ упорядоченное множество всех предикатных символов, входящих в множество C нормализованных вправо дизъюнктов. С этим упорядочением предикатных символов ассоциируется разбиение множества C на подмножества $C_i (i = 1, \ldots, n)$, где C_n состоит из всех дизъюнктов, содержащих атом $p_n(t)$ (литеру $p_n(t)$ или $\neg p_n(t)$), а $C_i (i = 1, \ldots, n-1)$ состоит из всех дизъюнктов, которые не принадлежат $C_j (j > i)$ и содержат атом $p_i(t)$.

Определение 4. S-выводом дизъюнкта c из множества дизъюнктов C называется такой R-вывод c из C, при котором правило R-резолюции применяется только к дизъюнктам из одного и того же подмножества C_i, причем единственным атомом, по которому осуществляется резольвирование дизъюнктов из этого подмножества, является $p_i(t)$.

Метод основывается на следующей теореме.

Теорема 1. *Если множество нормализованных вправо дизъюнктов C невыполнимо, то для произвольного упорядочения предикатных символов в множестве C существует S-вывод пустого дизъюнкта из C.*

Сначала докажем следующее предложение.

Предложение 2. Для произвольного множества дизъюнктов C с упорядоченными предикатными символами и дизъюнкта c, не содержащего атомов ранга 0, из существования R-вывода c из C следует существование S-вывода этого дизъюнкта из C.

Справедливость теоремы непосредственно вытекает из этого предложения, поскольку пустой дизъюнкт не содержит атомов нулевого ранга.

Для доказательства предложения 2, достаточно рассмотреть R-вывод, который не содержит дизъюнктов, полученных путем применения операции нормализации вправо. Такой R-вывод будем называть простым R-выводом. Действительно, если определить подходящим образом глубину вывода (например, как число последовательных применений операции нормализации вправо в дереве вывода), то с помощью индукции по глубине вывода можно легко показать, что если предложение 2 верно для простого R-вывода, то оно также верно и для любого другого R-вывода.

Пусть c - дизъюнкт, не содержащий атомов нулевого ранга. Рассмотрим простой R-вывод c из множества дизъюнктов $C = \{c_i, ..., c_n\}$. В таком выводе все дизъюнкты, кроме последнего, содержат атомы нулевого ранга. Каждому дизъюнкту c_i из этого вывода поставим в соответствие дизъюнкт c_i' состоящий из всех литер нулевого ранга, содержащихся в c_i. Последовательность дизъюнктов c_i' соответствующая R-выводу дизъюнкта c является R-выводом $\Box$ из множества дизъюнктов $C' = \{c_1', ..., c_n'\}$. Очевидно, что если существует S-вывод $\Box$ из C', то также существует S-вывод c из C. Таким образом, задача сводится к пропозициональному случаю.

Рассмотрим теперь невыполнимое множество дизъюнктов C' атомы которого являются пропозициональными переменными и упорядочены следующим образом: $p_1 < p_2 < ... < p_q$. Существование S-вывода пустого дизъюнкта $\Box$ из C' вытекает из правила расщепления метода Дэвиса-Патнема [6], которое для рассматриваемого случая можно переформулировать в терминах резолюций в виде следующего предложения.

Предложение 3. Пусть C' - множество пропозициональных дизъюнктов и q-произвольная пропозициональная переменная, встречающаяся в C'. Если все резольвенты по переменной p, не являющиеся тавтологиями, добавить к C', а затем все дизъюнкты содержащие p или $\neg p$ удалить, то полученное множество дизъюнктов невыполнимо тогда и только тогда, когда невыполнимо исходное множество C'.

Соответствующее правило преобразования множества дизъюнктов C' будем называть исключением переменной p. Пусть $W_1, W_2, ..., W_q$- разбиение множества C', соответствующее приведенному выше упорядочению переменных.

Переменная p_q содержится только в дизъюнктах из W_q, причем каждый дизъюнкт содержит эту переменную. Применяя к C' правило исключения переменной p_q, получим новое невыполнимое множество дизъюнктов, заданное разбиением $W_1, W_2, ...W_{q-1}$. Затем, исключим переменную p_{q-1} и т.д. до тех пор, пока в результате применения этого правила не будет получен пустой дизъюнкт. Здесь уместно прокомментировать особенности применения правила исключения переменной.

При исключении переменных в описанном порядке (обратном их упорядочению) исключение переменной p_i приводит к удалению множества дизъюнктов W_i. При этом оставшиеся подмножества разбиения множества дизъюнктов могут измениться в результате добавления к ним новых дизъюнктов. Для произвольного множества дизъюнктов C' процесс последовательного применения рассматриваемого правила завершается в двух случаях: а) получен пустой дизъюнкт, что свидетельствует о противоречивости множества C'; б) получено пустое множество дизъюнктов, что свидетельствует о выполнимости множества C', так как пустое множество дизъюнктов является тождественно истинной формулой.

Поскольку в нашем случае рассматривалось невыполнимое множество дизъюнктов C', то процесс завершается получением пустого дизъюнкта. По результатам этого процесса несложно построить вывод $\Box$ из C'. Поскольку все резоль-вирования выполнялись между дизъюнктами, принадлежащими одному и тому же подмножеству разбиения, полученный вывод является S-выводом. Следовательно, если существует простой R-вывод дизъюнкта c, не содержащего атомов нулевого ранга, из множества дизъюнктов C, то существует S- вывод $\Box$ из C' и, следовательно, существует S-вывод c из C. Это завершает доказательство предложения 3, а вместе с ним и теоремы 1.

Теперь можно резюмировать основные особенности предлагаемого резолюционного метода.

1) R-резольвирование допускается между дизъюнктами только внутри од ного и того же подмножества разбиения, соответствующего выбранному упорядочению предикатных символов. 2. В каждом подмножестве разбиения резольвирование выполняется толь ко по атому, который соответствует этому подмножеству.
2) В каждом подмножестве разбиения резольвирование выполняется толь ко по атому, который соответствует этому подмножеству.
3) Порядок, в котором обрабатываются подмножества дизъюнктов, несу ществен, поскольку эти подмножества не удаляются после генерации всех резольвент по соответствующему атому.
4) S-резольвента c, не являющаяся тавтологией и не поглощаемая ни каким из существующих дизъюнктов, добавляется в соответствующее подмножество (согласно правилу разбиения) в нормализованном вправо виде, а все дизъюнкты, поглощаемые с, удаляются из текущего множества дизъюнктов.

ПРИМЕР

Пусть множество дизъюнктов представлено в виде разбиения, соответствующего следующему упорядочению предикатных символов: $x < u < y < z$.

Подмножество C_4 (соответствует предикатному символу z):
$(y(t-2) \vee \neg y(t-1) \vee z(t) \vee \neg u(t))$ 1, $(\neg z(t-2) \vee \neg y(t-1) \vee \neg z(t) \vee y(t))$ 2,

Подмножество C_3 (соответствует предикатному символу y):
$(z(t-1) \vee y(t) \vee \neg u(t))$ 3, $(y(t-2) \vee \neg y(t-1) \vee \neg y(t) \vee u(t))$ 4,
$(z(t-1) \vee \neg y(t-1) \vee y(t) \vee \neg x(t))$ 5, $(z(t-1) \vee y(t) \vee x(t))$ 6.
Подмножество C_2 (соответствует предикатному символу u):
$(z(t-1) \vee \neg u(t-1) \vee \neg u(t) \vee x(t))$ 7,
$(z(t-1) \vee \neg y(t-1) \vee u(t-1) \vee u(t) \vee \neg x(t))$ 8.
Подмножество C_1 (соответствует предикатному символу x):
$(\neg z(t-2) \vee \neg \mathrm{y}(t-2) \vee u(t-1) \vee x(t))$ 9.

Номер дизъюнкта записывается справа от него, а пары чисел, записанные слева от резольвент, означают номера дизъюнктов, по которым осуществлялось резольвирование.

Процесс S-пополнения протекает следующим образом.
(1, 2) $(\neg z(t-2) \vee y(t-2) \vee \neg y(t-1) \vee \neg u(t) \vee \mathrm{y}(t))$ 10, добавляется в C_3.
(4, 5) $(y(t-2) \vee z(t-1) \vee \neg y(t-1) \vee \neg x(t) \vee u(t))$ 11, добавляется в C_2.
(4, 6) $(y(t-2) \vee z(t-1) \vee \neg y(t-1) \vee x(t) \vee u(t))$ 12, добавляется в C_2.
(7, 12) $(y(t-2) \vee z(t-1) \vee \neg u(t-1) \vee \neg y(t-1) \vee x(t))$ 13, добавляется в C_1.

Процесс заканчивается после генерации 4-х резольвент, в то время как процесс R-вывода генерирует 35 резольвент.

ЗАКЛЮЧЕНИЕ

Основным результатом данной статьи является эффективный резолюционный метод проверки выполнимости спецификаций, представленных формулами языка L. Этот метод приводит к существенному уменьшению количества дизъюнктов, порождаемых в процессе проверки выполнимости формулы, по сравнению с методом R-резолюции. Еще одним фактором, обеспечивающим эффективность метода, является уменьшение количества пар дизъюнктов, проверяемых на возможность их резольвирования. Основной результат этой работы, можно рассматривать как доказательство полноты стратегии, сочетающей упорядочение предикатных символов с R-резолюцией. Заметим, что различные упорядочения предикатных символов дают различные разбиения множества дизъюнктов, что может сказаться на количестве дизъюнктов, генерируемых в процессе проверки выполнимости формул. Кроме того, различные порядки обработки подмножеств разбиения дизъюнктов, могут давать различные времена выполнения процедуры. В качестве одной из возможных эвристик мы рекомендуем обрабатывать подмножества дизъюнктов в порядке уменьшения их индексов.

Представляется перспективным дальнейшее изучение возможных подходов к выбору упорядочения множества предикатных символов с целью повышения эффективности процедуры проверки выполнимости формулы.

Список литературы

1. ChebotarevA. Provably-correct development of reactive algorithms. Proc. Int. Workshop "Rewriting Techniques and Efficient Theorem Proving"(RTETP-2000). - 2000. - P. 117 - 133.
2. Чеботарев А. Н. Детерминизация логических спецификаций автоматов. Кибернетика и системный анализ. - 1995. - N1, C. 3-12.
3. Мороховец М. К., Чеботарев А. Н. Революционный подход к проверке согласованности вза имодействующих автоматов. Кибернетика и системный анализ. - 1994. - N. 6, C. 36 - 50.
4. Чеботарев А. Н. Проверка непротиворечивости простых спецификаций автоматных си стем. Кибернетика и системный анализ. - 1993. - N. 3, C. 3 - 11.
5. Чеботарев А. Н. Метод раздельного резольвирования для проверки выполнимости формул языка L. Кибернетика и системный анализ. - 1998. N.6, C. 13 - 20.
6. ЧеньЧ., Ли Р. Математическая логика и автоматическое доказательство теорем. М.: На ука. 1973, 360 с.