

一阶逻辑中约束求解的局部搜索法*

张健

(中国科学院软件研究所计算机科学实验室 北京 100080)

摘要 以一阶谓词逻辑为基础,讨论约束满足问题,着重研究一阶逻辑公式可满足性的局部搜索法,并与命题逻辑中的可满足性过程加以比较,以皇后问题和哈密顿回路问题为例,说明基于一阶逻辑的方法能处理较大的问题实例。

关键词 约束满足问题,一阶谓词逻辑,局部搜索。

中图法分类号 TP18

命题逻辑中的可满足性问题 SAT (the satisfiability problem) 是计算机科学中的著名问题。20 多年前,它被证明为 NP-难问题。近几年,很多人仍然对它感兴趣。我国于 1996 年举行了国际 SAT 算法比赛。与之相对照,关于一阶逻辑中的可满足性问题却很少有人研究。其原因之一是,在最一般情况下,这个问题是不可判定的。

尽管如此,我们仍可在一定的限制条件下,研究一阶谓词逻辑中的可满足性问题。比如,只考虑特定形式的逻辑公式,或者只考虑公式在有限集合上的可满足性。事实上,在很多情况下,可满足公式具有有限模型。在这种限制下,人们通常将上述问题转换成 SAT 来求解。我们认为,有必要直接研究一阶逻辑公式的可满足性。本文提出以一阶谓词逻辑作为约束满足问题的基础。我们将讨论解决一阶可满足性问题的局部搜索法,并与相应的 SAT 过程进行比较。

1 约束满足问题与一阶谓词逻辑

在不少实际应用中,要求对象之间满足一定的关系或条件。人工智能中将这类问题归为约束满足问题 CSP (constraint satisfaction problems)。一个 CSP 包含一组变量 $\{X_i | 1 \leq i \leq n\}$, 其中每个变量 X_i 有一定的取值范围 D_i , 而且变量之间要满足一些约束条件 $\{C_j | 1 \leq j \leq m\}$ 。求解 CSP 就是要找出这些变量的值,使得所有约束条件都成立。在本文中,我们只考虑有限的 CSP, 也就是说,假定所有的 D_i 都是有限集合。

至于约束条件的表示方式,迄今为止并没有一个公认的好办法。在 CSP 的经典定义中,只是直接地列出满足条件的所有变量取值的组合。但在很多情况下,用一阶逻辑公式描述约束条件比较自然、简练。Mackworth^[1] 建议将有限 CSP 看成是一阶逻辑中的定理证明问题,或者命题逻辑中的可满足性问题。我们认为,将它看成是一阶逻辑公式在有限集合上的可满足性问题更合适。

具体说来,每个 CSP 对应于多种类 (Many-sorted) 逻辑公式 φ 在域 $D_1 \times D_2 \times \dots \times D_n$ 上的可满足性,其中 φ 的形式为 $\exists X_1 \dots \exists X_n (C_1 \wedge \dots \wedge C_m)$ 。

这里, X_1, \dots, X_n 是带类型 (Typed or Sorted) 的变量,而 C_1, \dots, C_m 是子句形式的一阶公式。不失一般性,可将 D_i 表示成整数集合,如 $\{1, \dots, m\}$ 。注意, φ 是一种特殊形式的一阶逻辑公式。此外,我们还限制它不含有未解释的 (Uninterpreted) 函数符号和谓词符号 (如未定义的 f, g, P, Q)。换句话说,公式中所有函数符和谓词符都是有特定含义的 (如 $\leq, \neq, +, -$ 等)。

例 1: 图的顶点着色问题。设图中有 n 个顶点,那么就有 n 个变量 X_1, \dots, X_n 。其类型是所有颜色的集合。如果有 m 条边,就有 m 个子句表示约束条件。对于连接顶点 i 和 j 的边,有子句: $X_i \neq X_j$ 。

例 2: 皇后问题。对于 n -皇后问题,有 n 个变量 X_1, \dots, X_n 。其中每个变量的取值范围是整数 $1 \sim n$ 。约束条件包括: 对任何 i, j ($1 \leq i < j \leq n$), $X_i \neq X_j$, 并且 $X_i \sim X_j \neq i - j$ 。这里, \sim 是定义在整数集合上的二元函数,表示两个数的差 ($x \sim y$ 等于 $x - y$ 的绝对值)。

* 本文研究得到国家 863 高科技项目基金和中国科学院择优支持回国工作基金资助。作者张健, 1969 年生, 博士, 副研究员, 主要研究领域为自动推理与搜索技术。

本文通讯联系人: 张健, 北京 100080, 中国科学院软件研究所计算机科学实验室

本文 1997-03-31 收到原稿, 1997-07-30 收到修改稿

例 3:哈密顿回路问题.若图中有 n 个顶点,可设 n 个变量 X_1, \dots, X_n ,其中 X_i 表示顶点 i 在回路中的序号.设序号集为 $\{1, \dots, n\}$.约束条件有:对任何 $i, j (1 \leq i < j \leq n)$, $X_i \neq X_j$;而且,如果没有边从顶点 i 到 j ,则 $X_i \neq X_j \bmod n + 1$.

2 可满足性求解的局部搜索法

如前所述,CSP 的搜索空间为 $D_1 \times D_2 \times \dots \times D_n$.当域 D_i 比较小时,我们可以用穷举的办法(如回溯过程)来寻找解.但回溯法的时间复杂度比较高,对规模较大的问题往往无能为力.近年来,国内外很多研究人员用不完备的局部搜索技术(如爬山法)解决大规模的 SAT 问题实例,取得一定的成功.^[2~5]本文着重讨论这类方法.

搜索空间中的每个点 P 可用 m 维向量 $\langle v_1, \dots, v_n \rangle$ 表示.我们在点集上定义二元的“相邻”关系如下: P_1 和 P_2 相邻,如果它们有且仅有一个分量不一样.例如, $n=4$ 时,点 $\langle 1, 2, 3, 4 \rangle$ 的邻点包括 $\langle 2, 2, 3, 4 \rangle, \langle 1, 2, 1, 4 \rangle, \langle 1, 2, 3, 2 \rangle$ 等.我们还在点集上定义一个“价值”函数: $value(\langle v_1, \dots, v_n \rangle)$ 表示当 $X_1=v_1, \dots, X_n=v_n$ 时被满足的约束条件的个数.

对于上节中定义的可满足性问题,我们可用局部搜索过程找到它的解.有很多种做法.下面描述的是一种爬山法.

```

for i=1 to MaxTries do {
  为每个变量  $x_i$  随机地选一个值
  (即,随机地在搜索空间中选一点作为当前点  $P$ );
  for j=1 to MaxChanges do {
    如果当前点有一邻点,其价值比当前点大,
    则修改某个变量的值,转到那个邻点;
    否则,当前点的价值比其所有邻点都大(即,当前点为局部极大点),
    于是跳出该循环,去选另外的初始点;
  }
}

```

这里的 MaxTries 和 MaxChanges 是事先给定的两个正整数.

除了爬山法外,我们也可以采用贪心策略(即每次在相邻点中选价值最大的那个)、随机游走(Random Walk)等技术.^[6]在此,我们不一一叙述.

基于上述思想,我们用 C 语言实现了一个约束求解工具.它所接受的问题描述分为两部分:变量说明和约束条件.前者包括变量名和取值范围,例如, $x_i: 1..10$.后者由一组子句形式的公式表示.公式由项构成,变量和正整数是项,而几个简单的项通过运算(包括比较运算和算术运算)可构成新的更复杂的项.一个子句由若干个项“或”起来而得到.例如, $x + y \leq 3 \vee x = y$.

我们的工具是基于一阶逻辑的,因此比较通用,可以在高层次上直接描述.表 1 给出它在 SPARCstation 10 上解央皇后问题和逆皇后问题的运行时间.

表 1

n	皇后问题(s)	逆皇后问题(s)
100	38	16
150	491	113
200	1 926	211
300	2 131	1 623

这里所谓的逆皇后问题(Inverse Queens Problem)^[7]是由皇后问题演变而来的.它也要求每行只有 1 个皇后,但是任何两个皇后要么在同一列,要么在同一对角线上.与皇后问题不同的是,逆皇后问题的解比较少,只有 $n+2$ 个(当 $n > 3$ 时).由此可见,解的个数少并不意味着问题要难一些.

3 相关工作

我们将本文中有限的 CSP 看成是判定一阶逻辑公式在有限论域中的可满足性.从理论上讲,它也可转换成命题逻辑公式的可满足性判定(即 SAT)问题.其标准做法是,为每一种可能的赋值 $x=v$ 引入一个命题变元,然后将约束条件翻译成命题公式.反过来,SAT 也是本文所讨论的可满足性问题的特例.其特殊之处在于,每个变元只能取布尔值,而且每个项必须是文字(Literal),即变元或其否定.

SAT 是计算机科学中著名的 NP-难问题,在这方面有大量的理论结果和实际算法.由于命题逻辑比较简单,可以利用它的某些特性来研究高效率的实现技术.这些特性包括,命题变元只有真假值;而且可以假定每个变元在同一个子句中只出现 1 次.在域比较小时,转换成命题逻辑的办法使我们能利用现有的高效 SAT 工具进行约束求解.对一些

规模较小的问题,我们的工具从运行时间上看要比目前流行的 SAT 工具慢,但当域比较大或者约束条件较复杂时,转换成 SAT 会带来很多的命题变元和子句,往往会出现存储空间不够的问题。例如,100-皇后问题需要 10 000 个变量,160 多万条子句,这对目前的很多计算机来说,在空间上已达到极限。而采用一阶逻辑则避免了这个困难。对 n -皇后问题,只需 n 个变量, $n(n-1)$ 个子句。类似地,对哈密顿回路问题,大约需要 $O(n^2)$ 个子句。而如果将该问题表示成命题逻辑公式,则需要 $O(n^3)$ 个子句。我们的工具可处理随机生成的几百个顶点的图,大大超过了一般 SAT 程序的能力。

近几年有不少人研究任意一阶理论有限模型的自动构造,我们也先后研制了 FALCON 和 SEM 工具^[8],并解决了组合学和泛世代数中的一些公开问题。这类问题中有不少是不可满足的,因此,通常采用完备的搜索方法。在文献[9]中,我们也讨论了局部搜索法,但主要是用在数学问题的求解方面。现有的有限模型生成器往往要求完全用逻辑公理来描述问题,处理算术运算不太方便。本文针对实际中常见的一些 CSP,假定公式中不含未解释的函数符,这种限制从理论上讲并不是本质的,因为他问题可以转换过来。

与约束逻辑程序设计 CLP (constraint logic programming) 系统不同的是,我们的工具允许非 Horn 子句。这使得某些问题的描述更简练。另外,CLP 系统中一般不考虑局部搜索方法。

4 结束语

本文提出在一阶谓词逻辑的基础上研究有限 CSP,以一阶逻辑公式描述约束条件。从理论上讲,一阶逻辑中的有限可满足性问题也可转换成 SAT。但实际上,对于较大规模的问题实例来说,这种转换未必合适。本文采用局部搜索方法进行约束求解。在不少实际应用中,只要求满足大部分(而不是全部)约束条件。局部搜索技术比较适合解决这类问题。本文没有讨论如何改进局部搜索过程(如局部极大点的处理)以及如何提高实现效率。在这些方面,我们目前仍采用通常的做法。本文也没有着重考虑难解的随机问题实例。我们对常见的实际问题更感兴趣。一阶逻辑公式的可满足性判定是个很重要,而且具有挑战性的问题。期望今后在这一方向上会有更大的进展。

参考文献

- 1 Mackworth A. The logic of constraint satisfaction. *Artificial Intelligence*, 1992, 58(1): 3~20
- 2 Gu J. Efficient local search for very large-scale satisfiability problems. *SIGART Bulletin*, 1992, 3(1): 8~12
- 3 Selman B *et al.* A new method for solving hard satisfiability problems. In: *Proceedings of the 10th National Conference on Artificial Intelligence (AAAI-92)*. Menlo Park, CA: AAAI Press, 1992. 440~446
- 4 李未, 黄文奇. 一种求解合取范式可满足性问题的数学物理方法. *中国科学(A 辑)*, 1995, 38(1): 116~128
(Li Wei, Huang Wen-qi. A mathematic-physical approach to the satisfiability problem. *Science in China*, 1995, 38(1): 116~128)
- 5 刘涛, 李国杰. 求解 SAT 问题的分级重排搜索算法. *软件学报*, 1996, 7(4): 201~210
(Liu Tao, Li Guo-jie. Multi-stage search rearrangement algorithm for solving SAT problem. *Journal of Software*, 1996, 7(4): 201~210)
- 6 Selman B *et al.* Noise strategies for improving local search. In: *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)*. 1994. 337~343
- 7 Nadel B. Tree search and arc consistency in constraint satisfaction algorithms. In: Kanal L, Kumar V eds. *Search in Artificial Intelligence*. New York: Springer-Verlag, 1988. 287~342
- 8 Zhang J, Zhang H. SEM: a system for enumerating models. In: Mellish C S ed. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann Publishers, 1995. 298~303
- 9 Zhang J, Zhang H. Combining local search and backtracking techniques for constraint satisfaction. In: *Proceedings of the 13th National Conference on Artificial Intelligence and the 8th Innovative Applications of Artificial Intelligence Conference*. Cambridge, MA: MIT Press, 1996. 369~374

Local Search Methods for Constraint Solving in First-Order Logic

ZHANG Jian

(Laboratory of Computer Science Institute of Software The Chinese Academy of Sciences Beijing 100080)

Abstract In this paper, the author discusses constraint satisfaction problems in the framework of first-order logic. Local search methods for satisfying first-order formulas are studied, and compared with satisfiability procedures in the propositional logic. Experimental results on the Queens problem and the Hamiltonian circuit problem show that the framework is suitable for dealing with quite large problem instances.

Key words Constraint satisfaction problems, first-order predicate logic, local search.