第25卷 第2期
2002年2月

计 算 机 学 报
CHINESE J. COMPUTERS

Vol. 25 No. 2
Feb. 2002

求解SAT问题的拟人退火算法

张德富 黄文奇 汪厚祥
(华中科技大学计算机学院 武汉 430074)

摘 要 该文利用一个简单的变换,将可满足性(SAT)问题转换为一个求相应目标函数最小值的优化问题,提出了一种用于跳出局部陷阱的拟人策略.基于模拟退火算法和拟人策略,为SAT问题的高效近似求解得出了拟人退火算法(PA),该方法不仅具有模拟退火算法的全局收敛性质,而且具有一定的并行性、继承性.数值实验表明,对于本文随机产生的测试问题例,采用拟人策略的模拟退火算法的结果优于局部搜索算法、模拟退火算法以及近来国际上流行的WALKSAT算法,因此拟人退火算法是可行的和有效的.

关键词 SAT问题,模拟退火算法,拟人
中图法分类号: TP18

Personification Annealing Algorithm for Solving SAT Problem

ZHANG De-Fu HUANG Wen-Qi WANG Hou-Xiang
(*School of Computer Science, Huazhong University of Science and Technology, Wuhan* 430074)

Abstract The satisfiability (SAT) problem is core topic of the fields of artificial intelligence and computer science. Therefore, algorithms to solve the SAT problem play an important role in the development of computing theory and systems. Traditional algorithms treat the SAT problem as a constrained decision problem. In this paper, we transform the SAT problem into a global optimization problem to the objective function by a simple transformation, thus many algorithms can be used to solve it. The SA algorithm is a general stochastic search algorithm for combinatorial optimization problems, however, this algorithm need often cost too much time for finding a solution, which prevents it from being applied to many practical problems. How to improve this algorithm for solving the SAT problem is what this paper concerns. In this paper, the personification strategies obtained by observing and learning from the social and nature phenomena are presented. These strategies are generally straightforward and intuitive, and are helpful for the search process jumping out of local minimum, thus allow simulated annealing process to converge fast. These strategies explained how to select variables to flip in each iterative step and how to raise the system temperature when the search process got stuck the local minimum. Combining the simulated annealing algorithm and proposed personification strategies, we present a personification annealing (PA) algorithm for solving the SAT problem. The PA algorithm inherits the global convergence property from simulated annealing algorithm and has the property of parallelism and inheritance. In order to compare the PA algorithm with local search algorithm, simulated annealing algorithm and WALKSAT algorithm, a C implementation of these algorithms was tested on random generated 3-SAT problem instances. The actual

收稿日期: 2000-04-18; 修改稿收到日期: 2001-06-06. 本课题得到国家"九七三"重点基础研究发展规划项目(G1998030600)资助.
张德富,男,1972年生,博士研究生,研究方向为人工智能、组合优化、NP难问题求解. E-mail: zdfld@263.net. 黄文奇,男,1938年生,教授,博士生导师,研究方向为NP难问题现实求解、算法优化. 汪厚祥,男,1960年生,博士研究生,研究方向为网络、算法设计.

computational results show that the PA algorithm outperforms completely local search algorithm, simulated annealing algorithm and WALKSAT algorithm which is very popular recently, therefore the PA algorithm is feasible and efficient.

Keywords SAT problem, simulated annealing algorithm, personification

1 引 言

命题逻辑中合取范式(CNF)的可满足性问题(SAT)是当代理论计算机科学的核心问题, 是一典型的NP完全问题. 由于现代科技、军事以及经济管理的大量重要应用都可归结为求解NP完全问题, 因此, 它的快速求解不仅具有重要的理论意义, 而且在软件自动开发技术、逻辑推理机、VLSI设计以及知识库维护等许多领域都有重要的实际应用价值. 正是由于SAT问题的重要性, 各国学者对它进行了广泛而深入的研究, 提出了完全和不完全两类算法. 李未等在文献[1]中对前者做了总结性研究. 虽然完全算法能够保证正确判定公式的可满足性, 但是它的计算效率太低, 基本上不能实用. 因此, 人们更多地是去寻找求解SAT问题的不完全但又快速实用的算法. 许多不完全算法是基于局部搜索算法[2,3]的. 李未等提出数学物理方法[4]、Selman等提出的WALKSAT算法[5]、黄文奇等提出拟物拟人法[6]极大地丰富了不完全算法. 然而这些基于局部搜索或梯度下降的算法都存在一定的缺陷, 它们往往陷入局部极小而达不到全局最优, 即使有些算法(例如WALKSAT)加入了随机移动策略, 也是如此, 因此必须借助其它能逃离局部极小的策略. 模拟退火(SA)算法[7]是一个求解组合优化问题的通用随机搜索算法, 然而这个算法也常常需要花很长时间才能找到问题的解, 这就限制了该算法在实际问题中的应用, 因此还必须结合其它的策略. 通过观察和学习人类社会和自然现象, 我们提出了能逃离局部极小的拟人策略, 它能使模拟退火过程快速收敛. 基于此, 本文独辟蹊径, 结合SA算法和提出的拟人策略, 给出了求解SAT问题的拟人退火算法(PA). 计算结果表明, 对于本文随机产生的测试问题例, 采用拟人策略的模拟退火算法的结果优于局部搜索算法(SAT1.3)、改进前的SA算法以及近来国际上流行的WALKSAT, 因此PA用于求解SAT问题是可行的和有效的.

2 问题的表示及转换

考虑CNF

$$A = C_1 \wedge \cdots \wedge C_i \wedge \cdots \wedge C_n \qquad (1)$$

子句C_i具有如下形式

$$P_{i,1} \vee P_{i,2} \vee \cdots \vee P_{i,k_i} \vee \overline{P}_{ri,1} \vee \overline{P}_{ri,2} \vee \cdots \vee \overline{P}_{ri,k_{ri}},$$

其中$P_{i,1}$, $P_{i,2}$, …, P_{i,k_i}, $\overline{P}_{ri,1}$, $\overline{P}_{ri,2}$, …, $\overline{P}_{ri,k_{ri}}$是两两不同的文字, $P_{i,j}$为命题变元集$\{P_1, P_2, \cdots, P_m\}$中的一个变元, 文字$\overline{P_i}$表示变元$P_i$的非, m表示命题变元的个数, n表示子句的个数.

一个SAT问题是指: 对于给定的CNF是否存在一组关于命题变元的真值指派使得A为真. 显然, 如果A为真, 则CNF的每个子句中必有一个命题变元为1(真), 将每个子句中的每个命题变元取反, 则CNF的每个子句中必有一个命题变元为0(假), 然后将∧看成加, 将∨看成乘, 将变元P_i看成实参数x_i, 则SAT问题就可以转换为一个求相应实函数最小值的优化问题. 令T表示这种转换, 它可递归地定义为

$$T: A \to R^m \to R,$$

$$T(C_1 \wedge \cdots \wedge C_i \wedge \cdots \wedge C_n) = T(C_1) + \cdots + T(C_i) + \cdots + T(C_n),$$

$$T(C_i) = T(P_{i,1} \vee P_{i,2} \vee \cdots \vee P_{i,k_i} \vee \overline{P}_{ri,1} \vee \overline{P}_{ri,2} \vee \cdots \vee \overline{P}_{ri,k_{ri}})$$

$$= T(P_{i,1}) T(P_{i,2}) \cdots T(P_{i,k_i}) T(\overline{P}_{ri,1}) T(\overline{P}_{ri,2}) \cdots T(\overline{P}_{ri,k_{ri}}),$$

$$i = 1, \cdots, n.$$

$$T(P_i) = 1 - x_i, \ T(\overline{P_i}) = x_i, \ x_i \in [0, 1], \ i = 1, \cdots, m,$$

$$T(T) = 1, \ T(F) = 0.$$

例如,

$$T((P_1 \vee \overline{P_2}) \wedge (\overline{P_1} \vee \overline{P_2})) = T(P_1 \vee \overline{P_2}) + T(\overline{P_1} \vee \overline{P_2})$$

$$= T(P_1) T(\overline{P_2}) + T(\overline{P_1}) T(\overline{P_2})$$

$$= (1 - x_1) x_2 + x_1 x_2,$$

用$E(x_1, x_2, \cdots, x_m)$表示$T(A)$在点$(v(P_1), \cdots, v(P_m))$的值, 则有下面定理.

定理1 赋值v为使A可满足的充要条件是$E(x_1, x_2, \cdots, x_m)$达到最小值0.

证明 因为式(1)依变换T给出的实目标函

数为

$$E(X) = E(x_1, x_2, \cdots, x_m) = \sum_{i=1}^{n} e_i(x_1, x_2, \cdots, x_m) \quad (2)$$

其中, $X = (x_1, x_2, \cdots, x_m)$, $e_i(x_1, x_2, \cdots, x_m) = (1- x_{i,1})\cdots(1- x_{i,k_i}) x_{ri,1}\cdots x_{ri,k_{ri}}$. 显然, 若 A 可满足, 则必有一真值指派使 CNF 的值为 1, 由 $E(X)$ 及 T 的定义, 知 $E(X)$ 在该点的值必为 0; 反之, 若目标函数 $E(X)$ 的最小值为 0, 又 $E(X)$ 为非负函数 (由定义知), 必有一最小值点 $X \in \{0, 1\}^m$, 由 T 的定义, 知存在一组真值指派使 A 可满足. 证毕.

3 模拟退火算法

在组合优化问题中, 常用某种目标函数的全局最优作为算法搜索的目标. 然而基于局部搜索或梯度下降的算法往往容易陷入局部极小而达不到全局最优, 即使有些算法加入了随机移动策略, 也是如此, 因此必须求助其它随机搜索算法. SA 算法用于求解组合优化问题是基于固体物质的退火过程与组合优化问题求解过程的相似性. 组合优化问题的解和目标函数分别与固体的一个微观状态及其能量相对应. 与其它算法不同的是, SA 算法利用一个概率机制来控制跳坑的过程. 在搜索过程中, SA 算法不仅接受优化解, 而且以一定的概率接受恶化解, 在高温时, 接受优化解的概率比较大, 随着温度的降低, 接受恶化解的概率也随之下降, 当温度趋于零值时, 就不再接受任何恶化解, 这就使得 SA 算法有更多的机会逃离局部最优的"陷阱", 避免了其它局部搜索算法或梯度算法存在的缺陷. 一般地, SA 算法可以描述如下.

(1) 任给初始状态 X_0, 给定一个初始温度 T_0.

(2) 随机扰动产生状态 X_1, 计算 $\Delta E = E(X_1) - E(X_0)$.

(3) 如果 $\Delta E < 0$, 转(4); 否则, 若 $\exp(- \Delta E / T_0) \le \mathrm{random}(0, 1)$, 转(2).

(4) $X_0 = X_1$, $E(X_0) = E(X_1)$.

(5) 在温度 T_0 下检验能量是否达到平衡, 若不平衡转(2).

(6) $T_0 = \alpha T_0$, 退火过程是否结束, 是就停止, 否则转(2).

为了提高 SA 算法的性能, 必须合理的选择所需要的参数[8], 例如 T_0, α.

4 拟人策略

在组合优化领域里, 当算法搜索到局部极小点时, 一般来说, 就不再往下搜索了. 但是对于 SAT 问题, 就这样停止搜索是不可取的, 其原因在于我们只对 SAT 问题的全局最优点感兴趣. 因为只有全局最优点才能回答该问题是否可满足. 因此我们必须采取有效的策略尽可能的避免陷入局部极小点, 或者到了局部极小点, 能够有效地跳出. 人类在长期的实践中, 积累了丰富的经验, 受他们启发, 我们可以找到一些解决问题的方法. 我们把从人类社会或自然现象中学习到的经验和方法称之为拟人策略[6, 10].

我们知道, SA 算法能否找到 $E(X)$ 的全局最优解, 取决于 T_0 是否足够高和 T_0 下降得是否足够慢, 而这些正好与计算时间相反. 虽然合理选择参数可以使得算法的实现只需问题规模的多项式时间, 但是随着问题规模的增大, 所需时间也随之增长, 合理选择参数并不能从根本上提高算法的效率. 怎样合理地产生状态 X_1, 才是提高效率的关键. 为了提高退火过程的效率, 根据 SAT 问题的特点, 我们提出如下拟人策略, 即在每一个迭代步对每个非零子句都随机选择一个变元, 将其取反, 来合理地产生状态 X_1, 当然, 这里我们必须注意不能选择重复的变元, 即在一个子句里随机选择一个变元后, 即使该变元再在其它非零子句里出现, 也不能被选择. 这就像一所学校最初制定一项分配政策, 总有满意和不满意的单位, 然后根据具体情况, 修改令某个单位不满意的一条规则, 使其满意. 当然这条规则在一个时期应统一, 不能这个单位一套, 对别的单位又是另一套. 这样才能成为规则, 政策也才得以实施. 该策略可以避免盲目搜索以及减少陷入局部极小的机会, 而且具有一定的并行性. 此外, 给定一个具体的退火机制后, SA 算法并不保证能够找到全局最优点, 例如在 SAT 问题里, 能找到一组赋值, 使它满足所有的子句. 前面说过, 我们只对 SAT 问题的全局最优点感兴趣, 当搜索过程陷入局部陷阱时, 就需要某种策略使搜索过程能够跳出, 从而使搜索过程尽可能的找到全局最优. 在这种情况下, 有些算法采用重新随机初始化变元的值, 例如下面我们要比较的 SA, WALKSAT 和局部搜索算法. 而 PA 却不是这样, 当温度趋于零时, 如果还未找到最优解, 可采用以下拟人策略, 即人为地对退火过程进行升温. 具体来

说, 只需在 SA 的(6)步后加上(7)如果 $E(X)\neq 0$, 则 $T_0=n$, 转(2). 这相当于给一个跳坑没有劲的人以能量, 从而使他跳出来. 本策略能有效地继承一些变元在局部极小点的值, 同时又能将搜索引向未曾搜索过的区域.

为了显示采用拟人策略的效果, 我们将提出的 PA 算法与改进前的 SA 算法[5,9]作了比较. 后者在每一个迭代步只随机选择一个变元, 退火结束后, 如没有达到全局最优, 则重新随机初始化变元的初值, 再开始模拟退火过程.

5　计算结果及结论

为了测试 PA 的计算效果, 我们用随机产生的 3-SAT 模型(每个子句的长度 $l=3$, 且子句里的变元两两不同) 做实例, 并且将 PA 与改进前的 SA[5]、局部搜索算法的代表——SAT1.3[2]以及近来国际上比较流行的WALKSAT[5,11]相比较, 从计算的平均执行时间以及可满足样例数两个方面进行分析, 这应该是比较全面和具有说服力的.

表 1　PA 和 SA,WALKSAT, SAT1.3 的比较结果

合取范式			平均执行时间(s)				可满足样例数			
m	n	l	PA	SA	WALKSAT	SAT1.3	PA	SA	WALKSAT	SAT1.3
100	200	3	0.00	2.97	0.17	3.83	20	20	20	20
100	250	3	0.05	5.50	9.09	24.83	20	20	20	20
100	300	3	0.19	8.23	7.23	62.47	20	20	20	16
100	320	3	0.38	8.46	26.50	106.01	20	20	18	11
100	340	3	1.95	18.75	57.24	158.37	20	20	17	7
100	360	3	6.52	38.65	94.31	151.91	20	20	17	3
100	380	3	24.50	49.21	137.79	203.64	20	18	13	1
100	400	3	55.86	90.45	115.33	> 300	18	18	9	0
100	410	3	51.33	76.56	173.94	> 300	14	13	7	0
100	425	3	63.46	65.97	120.07	> 300	12	9	5	0
100	430	3	97.54	146.88	198.42	> 300	9	7	5	0

表 1 给出了四种算法的计算结果, 从中可以看出, 当子句数在 370 以下时, PA 的计算速度(从 20 个样例中可满足样例的平均执行时间来看) 比 SAT1.3 快 20 倍以上, 比WALKSAT 快 10 倍以上, 比 SA 快 5 倍以上; 当子句数在 370 以上时, PA 的平均执行时间与其它算法相差不多, 但可满足样例的次数却超过其它三种算法, 即当子句个数增加时, SAT1.3,WALKSAT 和 SA 计算成功的可能性大大降低, PA 的优越性却愈加明显, 这是因为 PA 可以从局部最优的“陷阱”中跳出, 更有可能求得 SAT 问题的全局最优解, 可见 PA 是可行的和有效的. 另外在实现WALKSAT 的过程中, 我们还发现一个奇特的现象, 在可满足的情形,WALKSAT 对大部分问题例算得很快, 但是有一小部分问题例算得很慢, 经过分析, 我们发现在跳坑的过程中, 该算法容易走回头路. 以上算法都用 C 语言编程, 并在 PC486 微机上运行通过, 运行时间设定为300(s), 每个测试例的样本数为 20. 为获得好的实验性能, 每个算法所用到的参数都进行了优化选择, SAT1.3 所用到的参数设置见文献[2], SA 和WALKSAT 所用到的参数见文献[5]. PA 选择 $T_0=n\times l$, $\alpha\in[0.9, 0.95]$.

致　谢　对评审人提出的有助于改进和完善本文的中肯意见, 对参与过讨论的何大华博士的帮助, 致以深切的谢意.

参　考　文　献

1　Li Wei, Huang Xiong. The analysis of algorithms for the proposition logic satisfiability. Computer Science, 1999, 26(3): 1-9(in Chinese)
(李　未, 黄　雄. 命题逻辑可满足性问题的算法分析. 计算机科学, 1999, 26(3): 1-9)

2　Gu J. Local search for satisfiability (SAT) problem. IEEE Trans Systems, Man and Cybernetics, 1993, 23(4): 1108-1128

3　Liu Tao, Li Guo-Jie. Local search for solving SAT problems and its average time complexity. Chinese Journal of Computers, 1997, 20(1): 18-26(in Chinese)
(刘　涛, 李国杰. 求解SAT 问题的局部搜索算法及其平均时间复杂性分析. 计算机学报, 1997 20(1): 18-26)

4　Li Wei, Huang Wen-Qi. A physic-mathematical method for solving conjunctive normal form satisfiability problem. Science

in China, Series A, 1994, 11: 1208- 1217(in Chinese)
(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, 1994, 25(11): 1208- 1217)

5 Selman B, Kautz H, Cohen B. Noise strategies for improving local search. In: Proc 12th National Conference on AI, American Association for Artificial Intelligence, 1994. 337-343

6 Huang Wen-Qi, Jin Ren-Chao. The quasi-physical personification algorithm for solving SAT problem —— Solar. Science in China, Series E, 1997, 2: 179- 186(in Chinese)
(黄文奇, 金人超. 求解 SAT 问题的拟物拟人算法——Solar. 中国科学(E 辑), 1997, 2: 179- 186)

7 Kirkpatrick S *et al*. Optimization by simulated annealing. Science, 1983, 220: 671- 680

8 Kang Li-Shan, Xie Yun, You Shi-Yong, Luo Zu-Hua. Non-Numerical Parallel Algorithms (1st Volume): Simulated Annealing. Beijing: Science Press, 1998 (in Chinese)
(康立山, 谢 云, 尤矢勇, 罗祖华. 非数值并行算法(第一册)模拟退火算法. 北京: 科学出版社, 1998)

9 Johnson D S, Aragon C R, McGeoch L A, *et al*. Optimization by simulated annealing: An experimental evaluation; part 11, graph coloring and number partitioning. Operations Research, 1991, 39(3): 378- 406

10 Zhang De-Fu, Ying Ai-Hua, Wang Hou-Xiang. Personificational neural network algorithm for SAT problem. Journal of Nanjin University, 2000, 36(10): 46- 50(in Chinese)
(张德富, 尹爱华, 汪厚祥. 求解 SAT 问题的拟人神经网络算法. 南京大学学报, 2000, 36(10): 46- 50)

11 Holger H Hoos, Thomas Stützle. Towards a characterization of the behavior of stochastic local search algorithms for SAT. Artificial Intelligence, 1999, 112(1-2): 213- 232

ZHANG De-Fu, male, Born in 1972, Ph.D. candidate in school of computer science at Huazhong University of Science and Technology. His research interests include artificial intelligence, combinatorial optimization, NP hard problem solving.

HUANG Wen-Qi, male, born in 1938, professor and Ph.D. supervisor. His research interests include NP hard problem solving, algorithm optimization.

WANG Hou-Xiang, male, born in 1960, Ph.D. candidate, vice-professor. His research interests include network, multi-media, algorithm design.