



调查传播算法收敛的一个充分条件

王晓峰^{1,2}, 许道云², 姜久雷^{1*}, 唐延辉¹

1. 北方民族大学计算机科学系, 银川 750021

2. 贵州大学计算机科学系, 贵阳 550025

* 通信作者. E-mail: 1994029@nun.edu.cn

收稿日期: 2016–10–20; 接受日期: 2016–12–12; 网络出版日期: 2017–09–05

国家自然科学基金 (批准号: 61462001, 61650206, 61563001, 61402017) 和计算机应用技术自治区重点学科建设基金资助项目

摘要 信息传播算法求解可满足问题时有良好的有效性, 使得难解区域变窄. 然而, 信息传播算法不总有效, 常表现为不收敛. 对于这种现象, 至今缺少系统的理论解释. 调查传播 (survey propagation, SP) 算法是最为有效的信息传播算法, 对 SP 算法的收敛性研究是设计其他信息传播算法的重要基础, 并为信息传播算法的广泛应用提供理论依据. 为了分析 SP 算法的收敛性, 通过对消息更新方程取双曲正切, 将消息取值从 $[0, 1]$ 扩展为 $(-\infty, \infty)$, 利用压缩映射原理给出了 SP 算法收敛的一个充分条件. 基于随机 3-SAT 实例, 给出了 SP 算法收敛的一个概率条件. 最后, 选取了随机 3-SAT 实例进行数值实验模拟, 验证了该条件的有效性.

关键词 调查传播算法, 可满足性问题, 收敛性, 信息传播算法, 因子图

1 引言

约束可满足性问题 (constraint satisfiability problems, CSP) 是人工智能中的一个重要研究领域^[1~3]. 可满足性判定问题 (satisfiability, SAT) 是典型的 CSP 问题, 即给定一个合取范式公式 (conjunctive normal form, CNF) F , SAT 判定问题指是否存在一组指派使得 F 为真.

SAT 问题是第一个被证明了的 NP- 完全问题, 不存在多项式时间的确定性图灵算法判定 SAT 问题. 然而, 在实际应用中, 该问题又不可避免地存在. 如在工程技术、军事、人工智能、并发控制、交通运输、智能交通控制、社交网络数据处理和智能挖掘等领域中, 诸多带有协调性 (或一致性) 检验、组合优化等问题均可以编码到 SAT 问题. 目前, 对 SAT 问题的研究已进入深水区, 研究进展较为缓慢, 研究也越来越复杂. 梳理这几年的研究成果, SAT 问题研究主要集中在 SAT 实例产生模型研究和 SAT 判定算法研究两个方面. 一方面, 实例产生模型最为典型的是随机 3-SAT 实例产生模型, 该模型中子句与变元的个数之比值 α 是一个重要的参数, 它不仅影响实例的可满足性, 而且影响实例的

引用格式: 王晓峰, 许道云, 姜久雷, 等. 调查传播算法收敛的一个充分条件. 中国科学: 信息科学, 2017, 47: 1646–1661, doi: 10.1360/N112016-00248
Wang X F, Xu D Y, Jiang J L, et al. Sufficient conditions for convergence of the survey propagation algorithm (in Chinese). Sci Sin Inform, 2017, 47: 1646–1661, doi: 10.1360/N112016-00248

判定难度. 随机统计现象表明: 对于随机 3-SAT 实例, 存在可满足的相变点 α_d . 当 $\alpha < \alpha_d$ 时, 实例高概率可满足; 当 $\alpha > \alpha_d$ 时, 实例高概率不可满足; 把满足与不可满足之间出现的这种临界现象称为相变现象, α_d 称为相变点, 在相变点附近区域的实例求解难度较大^[4,5]. 尽管人们不知道 α_d 的确切值, 但研究表明: α_d 至少为 3.52^[6], 至多为 4.4898^[7]. 另一方面, 目前也出现了很多有效的判定算法, 如基于纯文字规则的算法能够有效求解 $\alpha < 1.6$ 区域的随机 3-SAT 实例 (如 DPLL 算法); 基于正文字或负文字出现次数最多而设计的算法可有效求解 $\alpha < 3.52$ 区域的随机 3-SAT 实例 (如 Zchaff 算法); WALKSAT 算法可有效求解 α 稍高于 3.95 区域的随机 3-SAT 实例^[8,9].

信息传播算法是求解 3-SAT 问题最为有效的算法, 文献 [10] 中设计了 3 种求解 3-SAT 问题的信息传播算法, 分别为警示传播 (warning propagation, WP) 算法、置信传播 (belief propagation, BP) 算法和调查传播 (survey propagation, SP) 算法. BP 算法能有效求解 $\alpha < 3.95$ 区域的随机 3-SAT 实例; SP 算法能有效求解 $\alpha < 4.26$ 区域的随机 3-SAT 实例. SP 算法是目前求解随机 3-SAT 问题最为有效的算法, 能够稍高于线性时间求解难解 SAT 区域的具备 10^7 变量规模的实例, 几乎能够有效求解接近相变点的可满足性实例^[11~13].

最近几年, 信息传播算法得到了广泛应用. Xu 等^[14] 在 NP-完全问题相变现象的研究中, 提出了具有精确相变和难解实例的 RB 模型. 随后, 文献 [15,16] 分别从理论上和实验上说明在相变的阈值附近, 求解 RB 模型随机实例的难度随着问题规模的增加以指数形式增长. 而文献 [17] 表明 BP 算法在求解 RB 模型实例时非常有效. 同样, 信息传播算法可用于求解量化布尔公式 (quantified Boolean formulae, 简称 QBF) 的可满足性. 殷明浩等^[18] 设计了一种基于 SP 的启发式算法 HSPQBF 用于求解 QBF 问题. Ravanbakhsh 和 Greiner^[19] 提出了求解 CSP 问题的 PBP (perturbed BP) 算法和 PSP (perturbed SP) 算法, 实验结果表明在随机 CSP 实例上, PBP 算法效果优于 BP 算法. 同样, 对于随机 SAT 实例, PSP 算法也优于 SP 算法. 特别是, 在组合优化问题方面, 信息传播算法的应用已取得丰硕的成果, 如 MAX-SAT 问题、Maximum Weight Matching 问题、Minimum Cost Network Flow 问题、图的着色、最大独立集等^[20~25].

事实上, 在信息传播算法设计中, 将变量的联合概率分布分解为变量子集上的局部函数的乘积形式. 称局部函数为因子 (factor), 每一个因子依赖于一个变量子集, 将变量联合分布的这一因子形式表示为图模型 — 因子图 (factor graph). 因子图是一个二部图, 图中有两类结点: 因子结点和变量结点. 信息传播算法是通过因子图上的边传递信息, 这种信息传递方式有两种: 变量结点传递给因子结点的信息和因子结点传递给变量结点的信息. 这种传递的信息可以理解为因子结点与变量结点之间的某种依赖程度, 从每个结点传出的信息由传入该结点的信息决定, 通过某种迭代策略对信息进行更新. 信息迭代过程可能收敛到某一固定点 μ^∞ , 或无限次进行迭代 (不收敛). 如果信息传播算法收敛, 则利用 μ^∞ 来近似计算变量的边缘概率分布, 这种方法已经被证明是非常有效的^[11,12,26]. 然而, 仍然有两个问题有待理论上深入考虑:

- (1) 信息传播算法是否能够收敛到固定点 μ^∞ ;
- (2) 信息传播算法收敛时得到的边缘概率是否有效.

信息传播算法的收敛性研究已取得了一些成果. 文献 [27] 中分析了因子图中含有一个环的实例上信息传播算法的收敛性, 表明信息传播算法能够有效收敛到某一固定点. 文献 [28] 表明, 对于任意结构的 Gauss 图模型实例, 信息传播算法能够正确收敛. Tatikonda 和 Jordan^[29] 分析了基于计算树的 Gibbs 测度序列的唯一性与 BP 算法收敛性之间的关系, 给出了 BP 算法收敛的一个充分条件. Heskies^[30] 利用 Bethe 自由能量最小值的唯一性, 给出了 BP 算法收敛到唯一固定点的充分条件. Ihler^[31] 直接从 BP 算法的信息更新函数出发, 获得了传递误差信息的界, 并给出了信息传播算法

收敛的一个充分条件. Mooij 和 Kappen [32] 利用向量空间中信息更新函数的压缩映射原理导出了 BP 算法收敛的充分条件. Shi 等 [33] 得到了传递误差信息的一个更紧致的界, 并利用这个界研究了信息传播算法的动态行为, 给出了算法收敛的一个充分条件. 文献 [34] 中给出了信息传播算法收敛的一个概率条件, 研究表明 BP 算法能够给出最大权重匹配问题和最小代价网络流问题的最优解, 且算法迭代次数高概率地有多项式界. 文献 [35] 中分析了 BP 算法的复杂性, 提出一个 SBP (stochastic BP) 算法, 该算法的消息更新复杂性较 BP 算法的复杂性低了一个指数级别. Feige 等 [36] 分析了 WP 算法的收敛性, 并给出了 WP 算法收敛性的一个概率条件. 但该条件主要局限于植入指派的随机可满足实例产生模型 $P_{n,p}^{\text{plant}}$, 在 $P_{n,p}^{\text{plant}}$ 模型中要求取概率 p 充分大, 其本质是高概率地最多只有一个赋值满足该实例. 文献 [37] 证明了 WP 算法在 $G(n, k, p)$ 模型产生的 3-CNF 公式上的收敛性, 并给出了算法收敛的一个概率条件. 文献 [38] 给出了 WP 算法收敛的一个充分条件. 文献 [39] 给出了 BP 算法在 RB 模型实例集上收敛的概率条件. 目前, 对 SP 算法的收敛性研究较少.

为此, 本文分析了 SP 算法的收敛性. 在 SP 算法中, 通过对消息更新方程取双曲正切, 将消息取值从 $[0, 1]$ 扩展为 $(-\infty, \infty)$, 利用压缩映射原理给出了 SP 算法收敛的一个充分条件. 更进一步, 基于随机 3-SAT 实例, 给出 SP 算法在该实例集上收敛的概率条件. 最后, 通过数值实验模拟, 验证了该条件的有效性.

2 调查传播算法

2.1 随机 3-SAT 实例产生模型

模型 $G(n, 3, m)$ 产生随机 3-SAT 实例 (或称作 3-CNF 公式), 其中 n 表示变元个数, m 表示子句个数, 子句大小为 3. 模型 $G(n, 3, m)$ 按如下方式产生一个随机 3-SAT 实例: 随机均匀地在可能的子句 (这样的子句共有 $2^3 \binom{n}{3}$ 个) 中选择 $m = \alpha n$ 个子句, 被选中的子句之集构成实例.

对于一个随机 3-SAT 实例, α 是一个重要的参数, 它影响实例的可满足性和判定难度. 存在可满足性的相变点 α_d , 当 $\alpha < \alpha_d$ 时, 实例高概率可满足; 当 $\alpha > \alpha_d$ 时, 实例高概率不可满足; 同时存在第二相变点 α_c , $\alpha < \alpha_c$ 为易解 SAT 区域, $\alpha > \alpha_c$ 为难解 SAT 区域. 大量的研究表明, α_d 至少为 3.52, 至多为 4.506, $\alpha_c \approx 3.921$.

一个随机 3-CNF 公式有 αn 个子句, 每个子句有 3 个变元, 共有 $3\alpha n$ 次变元出现. 因此, 每个变元平均出现的次数为 3α . 在 3-CNF 公式中, 每个变元对应两种文字, 即正文字和负文字. 因此, 每个文字平均出现的次数为 $\frac{3}{2}\alpha$.

2.2 因子图

设 $F = \{C_1, C_2, \dots, C_m\}$ 为一个 CNF 公式, 含有 n 个变元 x_1, x_2, \dots, x_n , 用 i 代表 x_i . 公式 F 可以用一个二分图 $G = (C \cup X, E)$ 表示, 称 G 为因子图. 其中, 变元结点集为 $X = \{1, 2, \dots, n\}$, 子句结点集为 $C = \{C_1, C_2, \dots, C_m\}$. 图 G 中的边分为两类:

实边: $(C_i, j) \in E \Leftrightarrow$ 子句 C_i 含正文字 x_j ;

虚边: $(C_i, j) \in E \Leftrightarrow$ 子句 C_i 含负文字 $\neg x_j$.

为了书写方便, 不妨用字母 a, b, \dots 分别代表子句 C_1, C_2, \dots

例1 公式 $F = \{(x_1 \vee \neg x_2 \vee \neg x_3), (\neg x_1 \vee x_2 \vee x_4), (\neg x_2 \vee x_3 \vee x_5), (\neg x_2 \vee x_4 \vee x_5)\} = \{a, b, c, d\}$ 的因子图 $G = (C \cup X, E)$ 如图 1 所示.

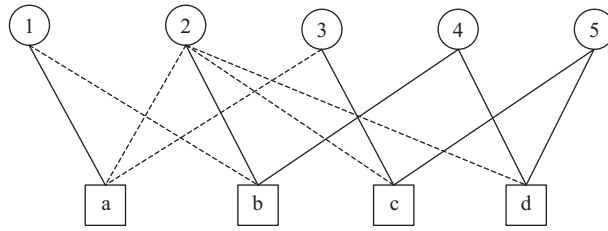


图 1 因子图

Figure 1 Factor graph

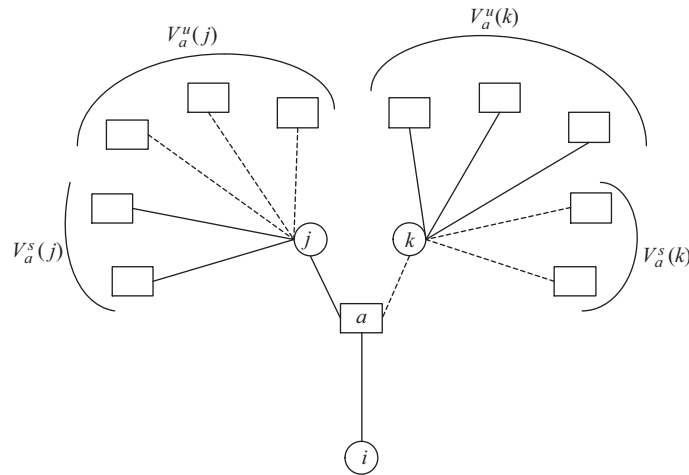


图 2 局部因子图

Figure 2 Part of factor graph

$V(a)$ 表示出现在子句 a 中的变元集合, $V(a) = V_+(a) \cup V_-(a)$. 其中, $V_+(a)$ 表示子句 a 中正文字对应的变元集合; $V_-(a)$ 表示子句 a 中负文字对应的变元集合; $V(a) \setminus i = V(a) - \{i\}$.

如图 1 中, $V(a) = \{x_1, x_2, x_3\}$, $V_+(a) = \{x_1\}$, $V_-(a) = \{x_2, x_3\}$.

$V(j)$ 表示包含变元 x_j 的子句的集合, $V(j) = V_+(j) \cup V_-(j)$. 其中, $V_+(j)$ 表示变元 x_j 正出现的子句集合; $V_-(j)$ 表示变元 x_j 负出现的子句集合; $V(j) \setminus a = V(j) - \{a\}$.

如图 1 中, $V(2) = \{a, b, c, d\}$, $V_+(2) = \{b\}$, $V_-(2) = \{a, c, d\}$.

定义集合 $V_a^u(j)$ 和 $V_a^s(j)$, 如果 $x_j \in V_+(a)$, 那么 $V_a^u(j) = V_-(j)$, $V_a^s(j) = V_+(j) \setminus a$; 如果 $x_j \in V_-(a)$, 那么 $V_a^u(j) = V_+(j)$, $V_a^s(j) = V_-(j) \setminus a$. 集合 $V_a^u(j)$ 和 $V_a^s(j)$ 如图 2 所示.

2.3 调查传播算法

基于消息传递而设计的信息传播算法, 求解可满足性问题时有良好的有效性. 特别地, SP 算法能够有效求解难解区域的随机 3-SAT 实例. 下面给出 SP 算法的基本过程.

在因子图的每条边 (a, i) 上, 定义 SP 算法中的消息传递 $\eta_{a \rightarrow i} \in [0, 1]$. SP 算法的消息更新迭代方程如下:

$$\eta_{a \rightarrow i} = \prod_{j \in V(a) \setminus i} \left(\frac{\Pi_{j \rightarrow a}^u}{\Pi_{j \rightarrow a}^u + \Pi_{j \rightarrow a}^s + \Pi_{j \rightarrow a}^0} \right), \quad (1)$$

其中,

$$\Pi_{j \rightarrow a}^u = \left(1 - \prod_{b \in V_a^u(j)} (1 - \eta_{b \rightarrow j}) \right) \prod_{b \in V_a^s(j)} (1 - \eta_{b \rightarrow j}), \quad (2)$$

$$\Pi_{j \rightarrow a}^s = \left(1 - \prod_{b \in V_a^s(j)} (1 - \eta_{b \rightarrow j}) \right) \prod_{b \in V_a^u(j)} (1 - \eta_{b \rightarrow j}), \quad (3)$$

$$\Pi_{j \rightarrow a}^0 = \prod_{b \in V(j) \setminus a} (1 - \eta_{b \rightarrow j}) = \left(\prod_{b \in V_a^u(j)} (1 - \eta_{b \rightarrow j}) \right) \left(\prod_{b \in V_a^s(j)} (1 - \eta_{b \rightarrow j}) \right). \quad (4)$$

求解 CNF 公式的 SP 算法如算法 1 所示.

算法 1 SP algorithm

Input: the factor graph of formula F ; a maximal number of iterations t_{\max} ; a requested precision ε .

Output: unconvergence if SP has not converged after t_{\max} sweeps. If it has converged: the messages $\eta_{a \rightarrow i}^*$.

- 1: At time $t = 0$: For every edge $a \rightarrow i$ of the factor graph, randomly initialize the messages $\eta_{a \rightarrow i}(t = 0) \in [0, 1]$;
 - 2: For $t = 0$ to $t = t_{\max}$:
 - 2.1. Sweep the set of edges in a random order, and update sequentially the warning on all the edges of the graph, generating the values $\eta_{a \rightarrow i}(t)$, using subroutine SP-UPDATE;
 - 2.2. If $|\eta_{a \rightarrow i}(t) - \eta_{a \rightarrow i}(t - 1)| < \varepsilon$ on all the edges, the iteration has converged and generated $\eta_{a \rightarrow i}^* = \eta_{a \rightarrow i}(t)$: go to 3;
 - 3: If $t == t_{\max}$ return UN-CONVERGENCE. If $t < t_{\max}$ return the set of fixed point survey $\eta_{a \rightarrow i}^* = \eta_{a \rightarrow i}(t)$.
-

算法 2 Subroutine SP-UPDATE ($\eta_{a \rightarrow i}$)

Input: set of all messages arriving onto each variable node $j \in V(a) \setminus i$.

Output: new value for the message $\eta_{a \rightarrow i}$.

- 1: For every $j \in V(a) \setminus i$, compute the three numbers $\Pi_{j \rightarrow a}^u, \Pi_{j \rightarrow a}^s, \Pi_{j \rightarrow a}^0$; If a set like $V_a^s(j)$ is empty, the corresponding product takes value 1 by definition;
 - 2: Using these numbers, compute and return the new survey $\eta_{a \rightarrow i}$.
-

在统计物理中, 调查信息 $\eta_{a \rightarrow i}$ 表示子句 a 传递给变元 i 的警示概率, 也即变元 x_i 在某个取值下使得子句 a 不可满足的概率. 因此, 当 SP 算法收敛到某一固定点 η^* 时, 利用某种规则寻找可满足的指派.

假设

$$W_i^+ = \left(\frac{\Pi_i^+}{\Pi_i^+ + \Pi_i^- + \Pi_i^0} \right), \quad (5)$$

$$W_i^- = \left(\frac{\Pi_i^-}{\Pi_i^+ + \Pi_i^- + \Pi_i^0} \right), \quad (6)$$

$$W_i^0 = 1 - W_i^+ - W_i^-, \quad (7)$$

其中,

$$\Pi_i^+ = \left(1 - \prod_{a \in V_+(i)} (1 - \eta_{a \rightarrow i}^*) \right) \prod_{a \in V_-(i)} (1 - \eta_{a \rightarrow i}^*), \quad (8)$$

$$\Pi_i^- = \left(1 - \prod_{a \in V_-(i)} (1 - \eta_{a \rightarrow i}^*) \right) \prod_{a \in V_+(i)} (1 - \eta_{a \rightarrow i}^*), \quad (9)$$

$$\Pi_i^0 = \prod_{a \in V(i)} (1 - \eta_{a \rightarrow i}^*). \quad (10)$$

文献 [10] 给出一种用于寻找可满足性指派的算法 (survey inspired decimation, SID) 如算法 3 所示.

算法 3 SID algorithm

Input: the factor graph of formula F ; a maximal number of iterations t_{\max} ; a requested precision ε .

Output: one assignment which satisfies all clauses, or SP UN-CONVERGENCE, or probably UNSAT.

- 1: Random initial condition for the surveys;
 - 2: Run SP. If SP does not converge, return SP UN-CONVERGENCE and stop. If SP converges, use the fixed-point surveys $\eta_{a \rightarrow i}^*$ in order to:
 - 3: Decimate:
 - 3.1. If non-trivial surveys ($\eta \neq 0$) are found, then
 - (a) Evaluate, for each variable node i , the three 'biases' $\{W_i^+, W_i^-, W_i^0\}$;
 - (b) Fix the variable with the largest $|W_i^+ - W_i^-|$ to the value $x_i = 1$ if $|W_i^+| > |W_i^-|$, to the value $x_i = 0$ if $|W_i^+| < |W_i^-|$.
Clean the graph;
 - 3.2. If all surveys are trivial ($\eta = 0$), then output the simplified sub-formula and run on it a local search process (e.g., WALKSAT);
 - 4: If the problem is solved completely by unit clause propagation, then output SAT and stop. If no contradiction is found then continue the decimation process on the smaller problem, go to 1, else stop.
-

SP 算法是目前求解随机 3-SAT 问题最为有效的算法,能够在稍高于线性时间内求解难解 SAT 区域的具有 10^7 变量规模的实例,几乎能够有效求解接近相变点的可满足性实例.注意到,如果算法不收敛,则算法失效.因此,SP 算法的关键在于算法是否收敛.然而,对于难解实例,SP 算法可能不收敛.目前,对 SP 算法是否收敛的理论分析较少.因此,有必要对 SP 算法的收敛性给予理论解释.

3 主要结论及证明

设 $(V, \|\cdot\|)$ 是一个有限维实向量空间. \mathbb{R}^N 上的 ℓ_1 -范数定义为

$$\|x\|_1 = \sum_{i=1}^N |x_i|. \quad (11)$$

\mathbb{R}^N 上的 ℓ_∞ -范数定义为

$$\|x\|_\infty = \max_{i \in \{1, 2, \dots, N\}} |x_i|. \quad (12)$$

向量空间 V 上的范数可诱导一个 V 上的度量 $d(v, w) = \|v - w\|$, $v, w \in V$. 显然,度量空间 (V, d) 是完备的 [40].

定义 1 (压缩函数) 设函数 $f : V \rightarrow V$, 如果存在 $0 \leq K < 1$, 对于任意的 $v, w \in V$, 使得 $d(f(v) - f(w)) \leq Kd(v, w)$, 则称函数 f 是压缩函数.

定义 2 (压缩映射原理) 如果 $f : V \rightarrow V$ 是完备度量空间 (V, d) 中的压缩函数, 那么对于 $\forall x \in V$, f 收敛到唯一固定点 x_∞ . 也即, $x, f(x), f^2(x), \dots, x_\infty$.

设 $(V, \|\cdot\|)$ 是一个赋范空间, 诱导关于映射 $A : V \rightarrow V$ 的矩阵范数:

$$\|A\| = \sup_{v \in V} \|Av\|. \quad (13)$$

\mathbb{R}^N 上的 A_1 -范数定义为

$$\|A\|_1 = \max_{j \in \{1, 2, \dots, N\}} \sum_{i=1}^N |A_{ij}|. \quad (14)$$

\mathbb{R}^N 上的 A_∞ -范数定义为

$$\|A\|_\infty = \max_{i \in \{1, 2, \dots, N\}} \sum_{j=1}^N |A_{ij}|. \quad (15)$$

引理1 (中值定理) 设 $(V, \|\cdot\|)$ 是赋范空间, 函数 $f: V \rightarrow V$ 可微, 那么对于任意的 $x, y \in V$, 有式 (16) 成立:

$$\|f(x) - f(y)\| \leq \|x - y\| \cdot \sup_{z \in V} |f'(z)|. \quad (16)$$

根据引理 1 和压缩映射原理, 如果 $\sup_{z \in V} |f'(z)| < 1$, 那么函数 f 收敛到固定点 x_∞ . 若函数 $f: V \rightarrow V$ 可微, 可利用 f 的 Jacobi 矩阵 $f'(z)$ 来构造矩阵范数 A .

由于 $\eta_{a \rightarrow i} \in [0, 1]$, 一个很自然的参数化过程是取双曲正切 (常被用于统计物理), 不妨假设 $\frac{1}{2}(1 + \tanh \lambda_{a \rightarrow i}) = \eta_{a \rightarrow i}$, 参数 $\lambda_{a \rightarrow i} \in \mathbb{R}$. SP 算法的迭代式 (1) 被写成如下形式:

$$\tanh \lambda_{a \rightarrow i} = 2 \prod_{j \in V(a) \setminus i} \left(\frac{\Pi_{j \rightarrow a}^u}{\Pi_{j \rightarrow a}^u + \Pi_{j \rightarrow a}^s + \Pi_{j \rightarrow a}^0} \right) - 1, \quad (17)$$

$$\Pi_{j \rightarrow a}^u = \left(1 - \prod_{b \in V_a^u(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) \right) \prod_{b \in V_a^s(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}), \quad (18)$$

$$\Pi_{j \rightarrow a}^s = \left(1 - \prod_{b \in V_a^s(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) \right) \prod_{b \in V_a^u(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}), \quad (19)$$

$$\Pi_{j \rightarrow a}^0 = \left(\prod_{b \in V_a^u(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) \right) \left(\prod_{b \in V_a^s(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) \right). \quad (20)$$

为了书写方便, 假设

$$x = \prod_{b \in V_a^u(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}), \quad y = \prod_{b \in V_a^s(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}),$$

因此,

$$\tanh \lambda_{a \rightarrow i} = 2 \prod_{j \in V(a) \setminus i} \left(\frac{(1-x)y}{x+y-xy} \right) - 1. \quad (21)$$

设实例的因子图中边的序列集合为 $D = \{a \rightarrow i : (a, i) \in E\}$, 假设算法中的消息更新并行执行, 定义函数 $f: \mathbb{R}^{|D|} \rightarrow \mathbb{R}^{|D|}$, 函数 f 的分量 $f(\lambda)_{a \rightarrow i}$ 定义为 $f(\lambda)_{a \rightarrow i} = \lambda_{a \rightarrow i}$. 自然, SP 算法中消息更新方程 (17) 的并行执行过程可以用函数 f 的迭代过程来刻画. 因此, 为了给出 SP 算法收敛的条件, 根据引理 1, 我们只要求 $\sup_{\lambda \in \mathbb{R}^{|D|}} |f'(z)| < 1$.

由第 2 节可知, $V_a^u \cap V_a^s = \emptyset$. 下面分两种情形进行讨论.

情形 1 设 $b \in V_a^u(j)$, 根据式 (21) 有

$$f'(\lambda_{a \rightarrow i, b \rightarrow j}) = \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} = \frac{-x'}{(1-x)(x+y-xy)(1-\tanh \lambda_{a \rightarrow i})}. \quad (22)$$

由于

$$x' = -\frac{1}{2}(1 - \tanh^2 \lambda_{b \rightarrow j}) \prod_{c \in V_a^u(j) \setminus b} \frac{1}{2}(1 - \tanh \lambda_{c \rightarrow j}), \quad (23)$$

所以

$$\frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} = \frac{\frac{1}{2}(1 - \tanh^2 \lambda_{b \rightarrow j}) \prod_{c \in V_a^u(j) \setminus b} \frac{1}{2}(1 - \tanh \lambda_{c \rightarrow j})}{(1-x)(x+y-xy)(1 - \tanh \lambda_{a \rightarrow i})} = \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{(1-x)(x+y-xy)(1 - \tanh \lambda_{a \rightarrow i})}. \quad (24)$$

又因为

$$x = \prod_{b \in V_a^u(j)} \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) = \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}) \prod_{c \in V_a^u(j) \setminus b} \frac{1}{2}(1 - \tanh \lambda_{c \rightarrow j}), \quad (25)$$

而式 (25) 的后一项

$$0 \leq \prod_{c \in V_a^u(j) \setminus b} \frac{1}{2}(1 - \tanh \lambda_{c \rightarrow j}) \leq 1, \quad (26)$$

所以

$$x \leq \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}). \quad (27)$$

由式 (21) 可知

$$\begin{aligned} \tanh \lambda_{a \rightarrow i} &= 2 \prod_{j \in V(a) \setminus i} \left(\frac{(1-x)y}{x+y-xy} \right) - 1 \\ &= 2 \left(\frac{(1-x)y}{x+y-xy} \right) \prod_{k \in V(a) \setminus i, j} \left(\frac{\Pi_{k \rightarrow a}^u}{\Pi_{k \rightarrow a}^u + \Pi_{k \rightarrow a}^s + \Pi_{k \rightarrow a}^0} \right) - 1, \end{aligned} \quad (28)$$

其中,

$$0 \leq \prod_{k \in V(a) \setminus i, j} \left(\frac{\Pi_{k \rightarrow a}^u}{\Pi_{k \rightarrow a}^u + \Pi_{k \rightarrow a}^s + \Pi_{k \rightarrow a}^0} \right) \leq 1, \quad (29)$$

所以

$$\tanh \lambda_{a \rightarrow i} \leq 2 \left(\frac{(1-x)y}{x+y-xy} \right) - 1. \quad (30)$$

进一步有

$$\begin{aligned} \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} &= \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{(1-x)(x+y-xy)(1 - \tanh \lambda_{a \rightarrow i})} \\ &\leq \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{(1 - \frac{1}{2}(1 - \tanh \lambda_{b \rightarrow j}))(x+y-xy) \left(1 - \left(2 \left(\frac{(1-x)y}{x+y-xy} \right) - 1 \right) \right)} \\ &= \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{\frac{1}{2}(1 + \tanh \lambda_{b \rightarrow j})(x+y-xy) \left(2 \left(1 - \frac{(1-x)y}{x+y-xy} \right) \right)} \\ &= \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{\frac{1}{2}(1 + \tanh \lambda_{b \rightarrow j})(x+y-xy) \left(2 \frac{x}{x+y-xy} \right)} = 1, \end{aligned} \quad (31)$$

因此

$$\left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| = \left| \frac{x(1 + \tanh \lambda_{b \rightarrow j})}{(1-x)(x+y-xy)(1 - \tanh \lambda_{a \rightarrow i})} \right|$$

$$\begin{aligned}
 &= \frac{x(1 + \tanh\lambda_{b \rightarrow j})}{(1-x)(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})} \quad \because 0 < x, y < 1 \\
 &\leq 1.
 \end{aligned} \tag{32}$$

情形 2 设 $b \in V_a^s(j)$, 根据式 (21) 有

$$f'(\lambda_{a \rightarrow i, b \rightarrow j}) = \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} = \frac{xy'}{y(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})}. \tag{33}$$

由于

$$y' = -\frac{1}{2}(1 - \tanh^2\lambda_{b \rightarrow j}) \prod_{c \in V_a^s(j) \setminus b} \frac{1}{2}(1 - \tanh\lambda_{c \rightarrow j}), \tag{34}$$

所以

$$\frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} = -\frac{\frac{1}{2}x(1 - \tanh^2\lambda_{b \rightarrow j}) \prod_{c \in V_a^s(j) \setminus b} \frac{1}{2}(1 - \tanh\lambda_{c \rightarrow j})}{y(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})} = -\frac{x(1 + \tanh\lambda_{b \rightarrow j})}{(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})}. \tag{35}$$

因此

$$\begin{aligned}
 \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| &= \left| -\frac{x(1 + \tanh\lambda_{b \rightarrow j})}{(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})} \right| \\
 &= \frac{x(1 + \tanh\lambda_{b \rightarrow j})}{(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})} \quad \because 0 < x, y < 1 \\
 &\leq \frac{x(1 + \tanh\lambda_{b \rightarrow j})}{(1-x)(x+y-xy)(1-\tanh\lambda_{a \rightarrow i})} \\
 &\leq 1.
 \end{aligned} \tag{36}$$

关于情形 1 和 2, 式 (32) 和 (36) 表明: 对于 $\forall \lambda_{b \rightarrow j} \in (-\infty, \infty)$, 都有

$$\left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| \leq 1, \tag{37}$$

所以

$$\sup_{\lambda \in V} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| \leq 1. \tag{38}$$

利用函数 f 的 Jacobi 矩阵构造矩阵范数 A . 由压缩映射原理及引理 1 可知, 如果 A 的范数严格小于 1, 则 f 为压缩映射, 收敛到唯一固定点, 且与初始信息无关. 不妨假设 $\tau_{a \rightarrow i, b \rightarrow j} = \sup_{\lambda \in V} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right|$, 有如下结论.

定理 1 对于 $\|A\|_1$ - 范数, 如果

$$\max_{b \rightarrow j} \sum_{a \rightarrow i} \tau_{a \rightarrow i, b \rightarrow j} < 1,$$

则 SP 算法在 $\|A\|_1$ - 范数下收敛, 且与初始信息无关.

证明 根据式 (13), (14) 和 (16), 有

$$\|A\|_1 = \|f'(\lambda)\|_1 = \max_{b \rightarrow j} \sum_{a \rightarrow i} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| = \max_{b \rightarrow j} \sum_{a \rightarrow i} \tau_{a \rightarrow i, b \rightarrow j}.$$

由压缩映射原理和引理 1 可知, 如果

$$\max_{b \rightarrow j} \sum_{a \rightarrow i} \tau_{a \rightarrow i, b \rightarrow j} < 1$$

成立, 则函数 f 收敛到某一固定点, 且与初始信息无关. 所以, SP 算法在 $\|A\|_1$ -范数下收敛, 且与初始信息无关.

同理可证如下结论.

定理2 对于 $\|A\|_\infty$ -范数, 如果

$$\max_{a \rightarrow i} \sum_{b \rightarrow j} \tau_{a \rightarrow i, b \rightarrow j} < 1,$$

则 SP 算法在 $\|A\|_\infty$ -范数下收敛, 且与初始信息无关.

证明 根据式 (13), (15) 和 (16), 有

$$\|A\|_\infty = \|f'(\lambda)\|_\infty = \max_{a \rightarrow i} \sum_{b \rightarrow j} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| = \max_{a \rightarrow i} \sum_{b \rightarrow j} \tau_{a \rightarrow i, b \rightarrow j}.$$

由压缩映射原理和引理 1 可知, 如果

$$\max_{a \rightarrow i} \sum_{b \rightarrow j} \tau_{a \rightarrow i, b \rightarrow j} < 1$$

成立, 则函数 f 收敛到某一固定点, 且与初始信息无关. 所以, SP 算法在 $\|A\|_\infty$ -范数下收敛, 且与初始信息无关.

定义 $1_X(x)$ 为标识函数, 若 $x \in X$, 则 $1_X(x) = 1$. 若 $x \notin X$, 则 $1_X(x) = 0$. 对于一个随机 3-SAT 实例, 定理 1 可描述为如下形式.

推论1 对于 $\|A\|_1$ -范数, 如果

$$\max_{b \rightarrow j} (\max_{a \rightarrow i} (1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) \tau_{a \rightarrow i, b \rightarrow j})) < \frac{1}{2},$$

则 SP 算法在 $\|A\|_1$ -范数下收敛, 且与初始信息无关.

在随机 3-SAT 实例中, 每个变元高概率地平均出现 3α 次, α 为子句数与变元数之比值. 因此, 定理 2 可以描述为如下推理 2 的概率条件.

推论2 对于 $\|A\|_\infty$ -范数, 如果

$$\max_{a \rightarrow i} (\max_{b \rightarrow j} (1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) \tau_{a \rightarrow i, b \rightarrow j})) < \frac{1}{2(3\alpha - 1)},$$

则 SP 算法在 $\|A\|_\infty$ -范数下收敛, 且与初始信息无关.

设 M 是一个方阵, $\gamma_1, \gamma_2, \dots, \gamma_n$ 是 M 的特征值, 称 $\gamma(M) = \{\gamma_1, \gamma_2, \dots, \gamma_n\}$ 为 M 的谱, $\rho(M) = \sup\{|\gamma_i| : \gamma_i \in \gamma(M)\}$ 为 M 的谱半径.

引理2 设函数 $f : V \rightarrow V$, d 是 V 上的度量. 对于某个 $N \in \mathbb{N}$, 假设 f^N 是关于 d 的压缩函数, 那么 f 收敛到唯一固定点 x_∞ , 并且对于任意 $x \in V$, 函数 f 的迭代序列 $x, f(x), f^2(x), \dots$ 收敛到 x_∞ .

证明 对于任意的 $x \in V$, 考虑如下 $N - 1$ 个序列:

$$x, f^N(x), f^{2N}(x), \dots$$

$$\begin{aligned}
 & f(x), f^{N+1}(x), f^{2N+1}(x), \dots \\
 & \quad \vdots \\
 & f^{N-1}(x), f^{2N-1}(x), f^{3N-1}(x), \dots
 \end{aligned}$$

因为 f^N 是关于 d 的压缩函数, 所以上述每一个序列都收敛到 x_∞ . 因此, 序列 $x, f(x), f^2(x), \dots$ 一定收敛到 x_∞ . 所以, 函数 f 收敛到唯一固定点 x_∞ , 且与初始信息无关.

设 $\tau = \max_{a \rightarrow i, b \rightarrow j} \{\tau_{a \rightarrow i, b \rightarrow j}\}$, 定义 $|D| \times |D|$ 的方阵 M . M 中的元素 $M_{a \rightarrow i, b \rightarrow j} = 1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) \tau$. 有如下结论.

定理3 如果方阵 M 的谱半径 $\rho(M)$ 严格小于 1, 则 SP 算法收敛, 且与初始信息无关.

为了证明定理 3, 我们需要证明如下定理 4.

定理4 设函数 $f : \mathbb{R}^{|D|} \rightarrow \mathbb{R}^{|D|}$ 可微, 对于任意 $x \in \mathbb{R}^{|D|}$, 假设 $|f'(x)| \leq M$. M 是一个不依赖于 x 的非负方阵. 如果 $\rho(M) < 1$, 那么对于任意 $x \in \mathbb{R}^{|D|}$, f 收敛到唯一固定点 x_∞ . 也即, $x, f(x), f^2(x), \dots, x_\infty$.

注意: $|f'(x)| \leq M$ 表示由 $f'(x)$ 构成的 Jacobi 矩阵中的每个元素的绝对值不大于 M 中对应元素的值.

证明 对于任意的 $x \in \mathbb{R}^{|D|}$, 假设 $n = 1, 2, \dots$, 有

$$\begin{aligned}
 \|(f^n)'(x)\| &= \|\prod_{i=1}^n f'(f^{i-1}(x))\| = \|\|\prod_{i=1}^n f'(f^{i-1}(x))\|\|_1 \\
 &\leq \|\|\prod_{i=1}^n |f'(f^{i-1}(x))|\|\|_1 \\
 &\leq \|\|\prod_{i=1}^n M\|\|_1 \\
 &= \|M^n\|_1.
 \end{aligned}$$

由 Gelfand 定理: $\lim_{n \rightarrow \infty} \|M^n\|_1^{\frac{1}{n}} = \rho(M)$. 选取 $\varepsilon > 0$, 使得 $\rho(M) + \varepsilon < 1$, 对于某个 $N \in \mathbb{N}$, $\|M^N\|_1 = (\rho(M) + \varepsilon)^N < 1$. 因此, $\|(f^N)'(x)\| < 1$. 所以, $\|(f^N)(x)\|$ 是压缩函数. 由引理 2 可知, 函数 $f(x)$ 收敛到唯一固定点 x_∞ , 且与初始信息无关.

根据式 (38) 和定理 4, 函数 f 收敛到唯一固定点, 且与初始信息无关. 所以, SP 算法收敛, 且与初始信息无关. 至此, 定理 3 已证明.

假设 M 的特征值为 γ , 根据代数知识, 有 $|\gamma| \cdot \|x\| = \|\gamma x\| = \|Mx\| \leq \|M\| \cdot \|x\|$. 因此, 定理 3 的条件弱于定理 1 和 2. 因此, 使用定理 3 判定 SP 算法的收敛性.

例2 设 CNF 公式 $F_1 = \{(x_1 \vee x_2), (x_2 \vee x_3), (x_3 \vee x_4), \dots, (x_{n-1} \vee x_n), (x_n \vee x_1)\}$, 其因子图含有一个环. 因此, 根据前文, 所构成的矩阵 M 的每一行只有一个元素可能不为 0, 其他元素均为 0, 不妨设这个元素为 τ . 根据圆盘定理, 特征值 λ 满足 $|\lambda| \leq \tau$. 从第 3 节的推导可知: λ 满足 $|\lambda| \leq \tau < 1$. 所以, 谱半径 $\rho(M) < 1$. 根据定理 3, 则 SP 算法在公式 F_1 上收敛. 同样, SP 算法在公式 $F_2 = \{(\neg x_1 \vee \neg x_2), (\neg x_2 \vee \neg x_3), (\neg x_3 \vee \neg x_4), \dots, (\neg x_{n-1} \vee \neg x_n), (\neg x_n \vee \neg x_1)\}$ 上也收敛.

4 数值实验及分析

在数值实验中, 使用模型 $G(n, 3, m)$ 产生随机 3-SAT 实例, 其中 n 表示变元个数, m 表示子句个数, 子句大小为 3. 在算法中, 取最大迭代次数 $t_{\max} = 10^3$, $\varepsilon = 0.001$.

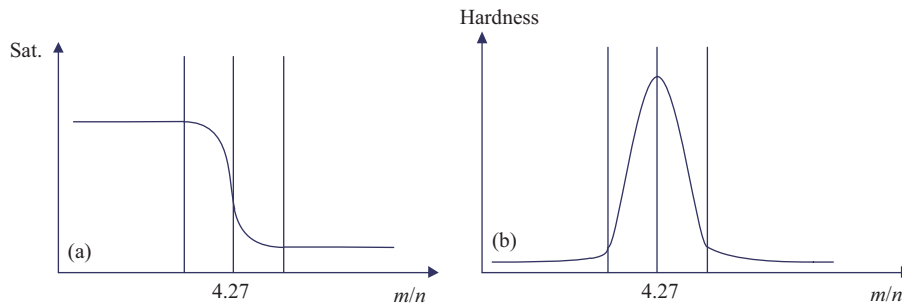


图 3 随机 3-SAT 实例的性质随约束参数 $\alpha = \frac{m}{n}$ 的变化情况

Figure 3 Properties change of random 3-SAT instance with parameter $\alpha = \frac{m}{n}$. (a) The phase transition of satisfaction; (b) the hardness of solving random 3-SAT instance

对于一个随机 3-SAT 实例, $\alpha = \frac{m}{n}$ 是一个重要的参数, 它影响实例的可满足性和判定难度. 存在可满足性的相变点 α_d . 当 $\alpha < \alpha_d$ 时, 实例高概率可满足; 当 $\alpha > \alpha_d$ 时, 实例高概率不可满足. 同时存在第二相变点 α_c , $\alpha < \alpha_c$ 为易解 SAT 区域; $\alpha_c < \alpha < \alpha_d$ 为难解 SAT 区域, 如图 3 所示. 大量的研究表明, α_d 至少为 3.52, 至多为 4.4898, $\alpha_c \approx 3.921$. 图 3(a) 为随机 3-SAT 实例的可满足相变现象, 在 $\alpha < 3.52$ 时, 实例高概率可满足; 在 $\alpha > 4.4898$ 时, 实例高概率不可满足; 而 $3.52 < \alpha < 4.506$ 区域的实例的可满足性不能从概率上给出确切的回答. 图 3(b) 为实例的求解难度, 在 $\alpha > 3.52$ 时, 实例的求解难度开始增大; 在 $\alpha = 4.27$ 附近, 求解难度最大. 事实上, 当 $3.921 < \alpha < 4.27$ 时, 实例已经位于难解区域, 其因子图的结构较为复杂, 大部分求解算法将失效. 因此, 常被国际竞赛构造实例用于测试算法的性能.

文献 [10] 中给出的 SP 算法能够有效求解 $3.921 < \alpha < 4.27$ 区域的实例, 几乎能够求解 $\alpha = 4.27$ 附近的实例. SP 算法是基于因子图结构的一种信息迭代算法, 之所以能够有效求解难解实例, 原于算法能够在实例上正确收敛, 如图 4 所示. 图 4 选取了两组规模不同的随机 3-SAT 实例, 图中每个数据点均由 $G(n, 3, m)$ 模型产生的 100 个随机实例构成. 我们知道, 实例的求解难度与约束参数 $\alpha = \frac{m}{n}$ 有关, 而与变量的规模无关. 当 $\alpha < 4.15$ 时, SP 算法高概率收敛; 在 $\alpha = 4.27$ 附近, SP 算法的收敛性有所下降, 而且下降幅度较大, 甚至不收敛, 如图 4(a) 所示, 纵坐标表示成功收敛的概率, 也即能够成功收敛的实例数占总实例数的比值, 横坐标表示约束参数 α . 同时, 我们也注意到, 实例的规模对算法收敛性的影响不大. 当约束参数 $3.921 < \alpha < 4.27$ 时, 实例的因子图的结构比较复杂, 从图 4(a) 中可以看出, 在某些数据点上, 存在实例使得 SP 算法不收敛. 图 4(b) 是 SP 算法收敛的运行时间随约束参数 α 的变化情况, 当 $\alpha > 3.921$ 时, 算法的运行时间急剧增大, 这主要与约束参数 α 有关. 当 α 增大时, 实例的因子图的结构变得复杂, SP 算法可能不收敛或计算次数增加, 使之计算时间增加.

图 5 表示定理 3 中的谱半径 ρ 随着约束参数 α 的变化情况. 当约束参数 $\alpha < 2.2$ 时, 谱半径 $\rho < 1$, 根据定理 3, SP 算法收敛; 当约束参数 $\alpha > 4.9$ 时, 存在实例使得谱半径 $\rho < 1$; 而当 $2.2 < \alpha < 4.9$ 时, 谱半径 ρ 几乎不小于 1, 因此, 在该区域定理 3 失效.

因子图的结构决定了算法的性能, 对于有多个回路的因子图实例, SP 算法可能不收敛. 分析 SP 算法可知, 算法不收敛有两种情况: 算法本身不收敛和算法在 t_{\max} 步之后被强制结束. 同样, SP 算法收敛也有两种情况: 算法的确收敛到唯一固定点和算法收敛时两次信息迭代误差为 ϵ , 本文中的相关收敛的结论都是基于算法收敛到唯一固定点. 因此, 为了更深入地分析算法的收敛性, 需要考虑是否存在固定点、固定点有几个等问题.

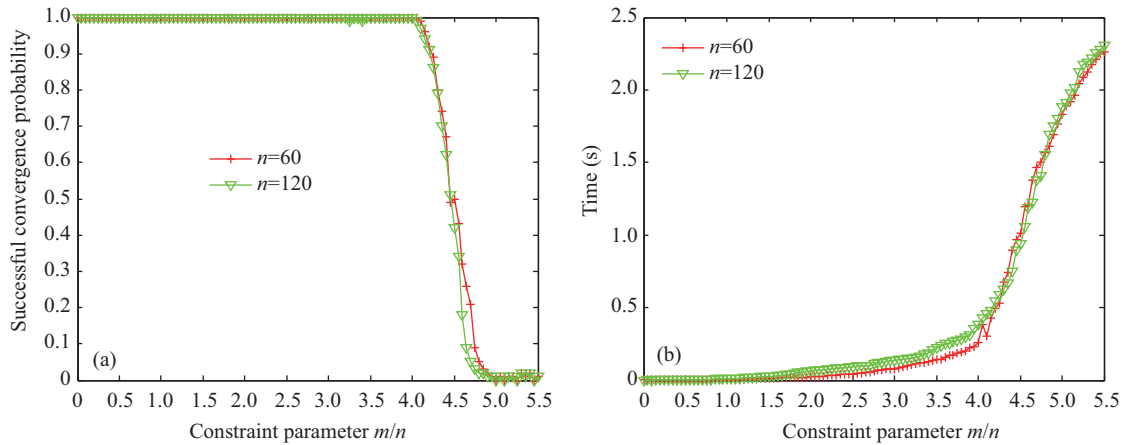


图 4 (网络版彩图) SP 算法的收敛性和计算时间随约束参数 $\alpha = \frac{m}{n}$ 的变化情况

Figure 4 (Color online) Properties change of convergence (a) and time (b) of SP with parameter $\alpha = \frac{m}{n}$.

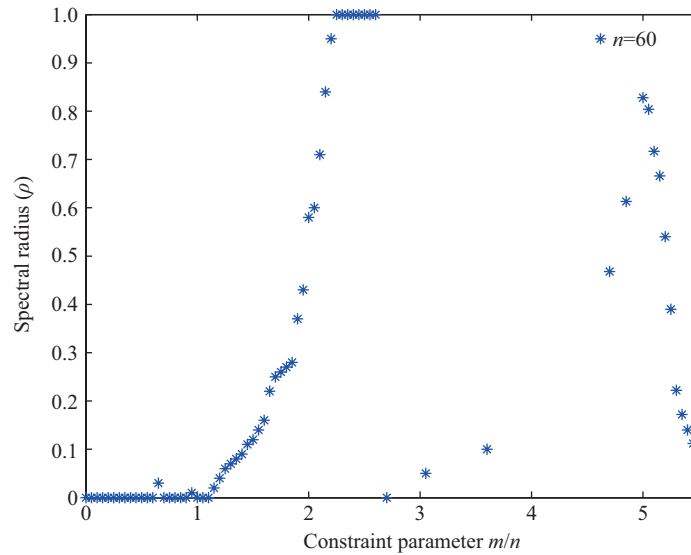


图 5 (网络版彩图) 谱半径 ρ 随约束参数 $\alpha = \frac{m}{n}$ 的变化情况

Figure 5 (Color online) The change of spectral radius ρ with parameter $\alpha = \frac{m}{n}$

5 结论

本文分析了 SP 算法的收敛性, 为了利用向量空间中压缩函数的性质, 对信息更新方程取双曲正切, 将信息取值从 $[0, 1]$ 松弛为 $(-\infty, \infty)$, 给出了 SP 算法收敛的一个充分条件, 为设计其他信息传播算法提供理论依据. 信息传播算法是一种概率算法, 利用信息传播算法可近似计算变量的边缘概率. 因此, 进一步的工作是: (1) 设计用于个性化推荐的信息传播算法; (2) 当信息传播算法收敛时, 分析算法的有效性; (3) 算法存在固定点的个数等.

参考文献

- 1 Martin D, Alan F, Michael M. A probabilistic analysis of randomly generated binary constraint satisfaction. *Theory Comput Sci*, 2003, 290: 1815–1828
- 2 Creignou N, Eaude H. The SAT-UNSAT transition for random constraint satisfaction problems. *Discrete Math*, 2009, 309: 2085–2099
- 3 Martin O C, Monasson R, Zecchina R. Statistical mechanics methods and phase transitions in optimization. *Theory Comput Sci*, 2001, 265: 3–67
- 4 Kirkpatrick S, Selman B. Critical behavior in the satisfiability of random Boolean formulae. *Science*, 1994, 264: 1297–1301
- 5 Mertens S, Mezard M, Zecchina R. Threshold values of random k-SAT from the cavity method. *Random Structure Algorithms*, 2006, 28: 340–373
- 6 Kaporis A, Kirousis L, Lalas E. The probabilistic analysis of a greedy satisfiability algorithm. *Random Structures Algorithms*, 2006, 28: 444–480
- 7 Díaz J, Kirousis L, Mitsche D, et al. On the satisfiability threshold of formulas with three literals per clause. *Theory Comput Sci*, 2009, 410: 2920–2934
- 8 Moskewicz M W, Madigan C F, Zhao Y. Chaff: engineering an efficient SAT solver. In: *Proceedings of the 38th Annual Design Automation Conference*. New York: ACM, 2001. 530–535
- 9 Aurell E, Gordon U, Kirkpatrick S. Comparing beliefs, surveys, and random walks. In: *Proceedings of the 17th International Conference on Neural Information Processing Systems*, Vancouver, 2004. 49–56
- 10 Braunstein A, Mezard M, Zecchina R. Survey propagation: an algorithm for satisfiability. *Random Structures Algorithms*, 2005, 27: 201–226
- 11 Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations. *J ACM*, 2007, 54: 1089–1098
- 12 Yedidia J S, Freeman W T, Weiss Y. Understanding belief propagation and its generalizations. *Artif Intell*, 2003, 8: 239–269
- 13 Braunstein A, Zecchina R. Survey and belief propagation on random k-SAT. In: *Theory and Applications of Satisfiability Testing*. Berlin: Springer, 2004. 519–528
- 14 Xu K, Li W. Exact phase transitions in random constraint satisfaction problems. *J Artif Intell Res*, 2000, 12: 93–103
- 15 Xu K, Li W. Many hard examples in exact phase transitions. *Theory Comput Sci*, 2006, 355: 291–302
- 16 Xu K, Boussemart F, Hemery F, et al. Random constraint satisfaction: easy generation of hard instances. *Artif Intell*, 2007, 171: 514–534
- 17 Zhao C Y, Zheng Z M. A belief-propagation algorithm based on variable entropy for constraint satisfaction problems. *Sci Sin Inform*, 2012, 42: 1170–1180 [赵春艳, 郑志明. 一种基于变量熵求解约束满足问题的置信传播算法. *中国科学: 信息科学*, 2012, 42: 1170–1180]
- 18 Yin M H, Zhou J P, Sun J G, et al. Heuristic survey propagation algorithm for solving QBF problem. *J Softw*, 2011, 22: 1538–1550 [殷明浩, 周俊萍, 孙吉贵, 等. 求解 QBF 问题的启发式调查传播算法. *软件学报*, 2011, 22: 1538–1550]
- 19 Ravanbakhsh S, Greiner R. Perturbed message passing for constraint satisfaction problems. arXiv:1401.6686
- 20 Chieu H L, Lee W S. Relaxed survey propagation for the weighted maximum satisfiability problem. *J Artif Intell Res*, 2009, 36: 229–266
- 21 Bayati M, Shah D, Sharma M. Max product for maximum weight matching: convergence, correctness, and LP duality. *IEEE Trans Inf Theory*, 2007, 54: 1241–1251
- 22 Coja-Oghlan A, Mossel E, Vilenchik D. A spectral approach to analyzing belief propagation for 3-colouring. *Comb Probab Comput*, 2009, 18: 881–912
- 23 Gamarnik D, Shah D, Wei Y. Belief propagation for min-cost network flow: convergence and correctness. *Oper Res*, 2012, 60: 410–428
- 24 Sanghavi S, Malioutov D M, Willsky A S. Belief propagation and LP relaxation for weighted matching in general graphs. *IEEE Trans Inf Theory*, 2011, 57: 2203–2212
- 25 Sanghavi S, Shah D, Willsky A S. Message passing for maximum weight independent set. *IEEE Trans Inf Theory*, 2009, 55: 4822–4834
- 26 Kschischang F, Frey B, Loeliger H. Factor graph and the sum-product algorithm. *Inform Theory*, 2001, 47: 498–519

- 27 Weiss Y. Correctness of local probability propagation in graphical models with loops. *Neural Comput*, 2000, 12: 1–41
- 28 Weiss Y, Freeman W T. Correctness of belief propagation in Gaussian graphical models of arbitrary topology. *Neural Comput*, 2001, 13: 2173–2200
- 29 Tatikonda S C, Jordan M I. Loopy belief propagation and Gibbs measure. In: *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence, Alberta, 2002*. 493–500
- 30 Heskes T. On the uniqueness of loopy belief propagation fixed points. *Neural Comput*, 2004, 16: 2379–2413
- 31 Ihler A T. Loopy belief propagation: convergence and effects of message errors. *Mach Learn Res*, 2005, 6: 905–936
- 32 Mooij J M, Kappen H J. Sufficient conditions for convergence of the sum-product algorithm. *IEEE Trans Inf Theory*, 2007, 53: 4422–4437
- 33 Shi X Q, Schonfeld D, Tuninetti D. Message error analysis of loopy belief propagation for the sum-product algorithm. *Comput Inf Sci*, 2010, 1009: 1–30
- 34 Brunsch T, Cornelissen K, Manthey B, et al. Smoothed analysis of BP for minimum cost flow and matching. *J Graph Algorithms Appl*, 2013, 17: 647–670
- 35 Norshams N, Wainwright M J. Stochastic belief propagation: a low complexity alternative to the sum product algorithm. *IEEE Trans Inf Theory*, 2013, 59: 1981–2000
- 36 Feige U, Mossel E, Vilenchik D. Complete convergence of message passing algorithms for some satisfiability problems. *Theory Comput*, 2013, 9: 617–651
- 37 Wang X F, Xu D Y, Wei L. Convergence of warning propagation algorithms for random satisfiable instances. *J Softw*, 2013, 24: 1–11 [王晓峰, 许道云, 韦立. 随机可满足实例集上警示传播算法的收敛性. *软件学报*, 2013, 24: 1–11]
- 38 Wang X F, Xu D Y. Sufficient conditions for convergence of the warning propagation algorithm. *J Softw*, 2016, 27: 3003–3013 [王晓峰, 许道云. 警示传播算法收敛的一个充分条件. *软件学报*, 2016, 27: 3003–3013]
- 39 Wang X F, Xu D Y. Convergence of the belief propagation algorithm for RB model instances. *J Softw*, 2016, 27: 2712–2724 [王晓峰, 许道云. RB 模型实例集上置信传播算法的收敛性. *软件学报*, 2016, 27: 2712–2724]
- 40 Dieudonné J. *Foundations of Modern Analysis*. New York: Academic Press, 1960

Sufficient conditions for convergence of the survey propagation algorithm

Xiaofeng WANG^{1,2}, Daoyun XU², Jiulei JIANG^{1*} & Yanhui TANG¹

1. *Department of Computer Science, Beifang University of Nationalities, Yinchuan 750021, China;*

2. *Department of Computer Science, Guizhou University, Guiyang 550025, China*

* Corresponding author. E-mail: 1994029@nun.edu.cn

Abstract Message propagation algorithms are very effective at finding satisfying assignments for SAT instances, and hard regions of a random SAT have become narrower. However, message propagation algorithms do not always converge for some random SAT instances. Unfortunately, a rigorous theoretical proof of this phenomenon is still lacking. The survey propagation (SP) algorithm is very effective at solving SAT instances, and a theoretical analysis of SP is very important in designing other message passing algorithms. Through this study, we derived the sufficient conditions for convergence of SP with extending message $[0, 1]$ to message $(-\infty, \infty)$. Finally, the experiment results show that the conditions for the convergence of SP are very effective in random 3-SAT instances.

Keywords survey propagation algorithms, satisfiability problems, convergence, message passing algorithms, factor graph



Xiaofeng WANG was born in 1980. He received his M.S. degree from Guizhou University, Guiyang in 2010, and his Ph.D. degree from Guizhou University, Guiyang in 2013. He is currently an associate professor at the College of Computer Science and Engineering in Beifang University of Nationalities. His research interests include algorithm design and analysis.



Daoyun XU was born in 1959. He received his Ph.D. degree from Nanjing University, Nanjing in 2002. He is currently a professor at the College of Computer Science and Technology in Guizhou University. His research interests include algorithm design and analysis, and complexity theory.



Jiulei JIANG was born in 1972. He received his MS degree from Donghua University, Shanghai in 2006, and his Ph.D. degree from Northwest University, Xi'an in 2014. He is currently a professor at the School of Computer Science and Engineering in Beifang University of Nationalities. His research interests include information systems analysis and business process modeling.