

## 可满足性问题中信念传播算法的收敛性分析\*

王晓峰<sup>1,2</sup>, 许道云<sup>2</sup>, 杨德仁<sup>3</sup>, 姜久雷<sup>1</sup>, 李强<sup>1</sup>, 刘欣欣<sup>1</sup>



<sup>1</sup>(北方民族大学 计算机科学与工程学院, 宁夏 银川 750021)

<sup>2</sup>(贵州大学 计算机科学与技术学院, 贵州 贵阳 550025)

<sup>3</sup>(宁夏医科大学 理学院, 宁夏 银川 750004)

通讯作者: 杨德仁, E-mail: deren.yang@nxmu.edu.cn

**摘要:** 信念传播算法是基于因子图模型的消息传递算法, 通过图中的边, 将消息从一个结点传递给另一个结点, 以高概率地确定部分变量的取值, 这种方法被实验证明在求解可满足性问题时非常有效。然而, 目前还未对其有效性从理论角度给予解释。通过对信念传播算法的收敛性分析, 试图从理论上解释算法的有效性。在信息传播算法的信息迭代方程中, 参数的取值范围为(0,1), 将该取值范围扩展到整个实数空间, 即 $(-\infty, +\infty)$ 。利用压缩函数的数学原理, 得到了信息迭代方程收敛的判定条件。选取随机可满足性问题实例进行实验模拟, 验证了结论的正确性。

**关键词:** 信念传播算法; 收敛性; 可满足性问题; 因子图

**中图分类号:** TP301

中文引用格式: 王晓峰, 许道云, 杨德仁, 姜久雷, 李强, 刘欣欣. 可满足性问题中信念传播算法的收敛性分析. 软件学报, 2021, 32(5): 1360–1372. <http://www.jos.org.cn/1000-9825/5844.htm>

英文引用格式: Wang XF, Xu DY, Yang DR, Jiang JL, Li Q, Liu XX. Convergence analysis of belief propagation algorithm for satisfiability problem. Ruan Jian Xue Bao/Journal of Software, 2021, 32(5): 1360–1372 (in Chinese). <http://www.jos.org.cn/1000-9825/5844.htm>

### Convergence Analysis of Belief Propagation Algorithm for Satisfiability Problem

WANG Xiao-Feng<sup>1,2</sup>, XU Dao-Yun<sup>2</sup>, YANG De-Ren<sup>3</sup>, JIANG Jiu-Lei<sup>1</sup>, LI Qiang<sup>1</sup>, LIU Xin-Xin<sup>1</sup>

<sup>1</sup>(School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China)

<sup>2</sup>(School of Computer Science and Technology, Guizhou University, Guiyang 550025, China)

<sup>3</sup>(School of Science, Ningxia Medical University, Yinchuan 750004, China)

**Abstract:** Belief propagation (BP) algorithm is a message passing algorithm based on a factor graph, and it passes messages from one node to another through edges in the graph to determine the value of some variables with high probability. This method is experimentally proven to be very effective when solving the satisfiability (SAT) problem. However, there is no explanation for its effectiveness from a theoretical point of view at present. Through the analysis of convergence of the BP algorithm, the effectiveness of the algorithm is explained theoretically. In this study, the sufficient conditions are also derived for convergence of the BP for satisfiability problem, and extending message (0,1) to message  $(-\infty, +\infty)$ . Lastly, experimental results show that the conditions for convergence of BP are very effective in random SAT instances, and it proves the correctness of the conclusion.

**Key words:** belief propagation (BP) algorithm; convergence; satisfiability problem; factor graph

\* 基金项目: 国家自然科学基金(62062001, 61762019, 61762002, 61862051, 61962002); 宁夏自然科学基金(2020AAC03214, NZ17111, 2019AAC03120, 2019AAC03119, 2020AAC03219); 北方民族大学重大专项基金(ZDZX201901); 北方民族大学校级一般科学基金(2019XYZJK05)

Foundation item: National Natural Science Foundation of China (62062001, 61762019, 61762002, 61862051, 61962002); Natural Science Foundation of Ningxia (2020AAC03214, NZ17111, 2019AAC03120, 2019AAC03119, 2020AAC03219); Major Scientific Research Projects of North Minzu University (ZDZX201901); Scientific Research Projects of North Minzu University (2019XYZJK05)

收稿时间: 2018-06-22; 修改时间: 2018-11-22; 采用时间: 2019-03-25

一个子句  $C$  是若干个文字的析取(即  $C=L_1 \vee L_2 \vee \dots \vee L_n$ ),子句  $C$  可以视为文字的集合(即  $C=\{L_1, L_2, \dots, L_n\}$ ).一个合取范式(conjunctive normal form,简称 CNF)公式(以下简称公式) $F$  是若干个字句的合取(即  $F=C_1 \wedge C_2 \wedge \dots \wedge C_m$ ),公式  $F$  可以视为子句的集合(即  $F=\{C_1, C_2, \dots, C_m\}$ ).可满足性问题(satisfiability,简称 SAT)是指:给定一个公式  $F$ ,判定是否存在一组指派使得  $F$  为真<sup>[1]</sup>.SAT 问题是计算机科学中最为著名的 NP-完全问题,其理论及其应用研究是计算机与数理逻辑界众多学者共同关注的一个重大问题.人工智能中的大多数难解问题,通过编码技术都可转化为 SAT 问题.SAT 问题的 NP-完全性表明:在  $P \neq NP$  条件下,不太可能存在多项式时间算法求解该问题.近年来,随着人工智能的迅速发展,特别是大数据和物联网的发展,对 SAT 问题原有的求解算法需要重新审视和设计,这为研究 SAT 问题提供了新的机遇和突破口.

目前,对于 SAT 问题的研究主要集中在两个方面:实例产生模型和求解算法.在实例产生模型研究中,获取了一些影响实例可满足性的控制参数,如子句数目  $m$  和变元数目  $n$  之比值  $\alpha=m/n$ .在  $G(n,3,m)$  模型中,随机实例的可满足与不可满足之间似乎存在一个阈值  $\alpha_d$ .当  $\alpha$  的值小于  $\alpha_d$  时,实例高概率满足;反之,高概率不可满足.该阈值  $\alpha_d$  称为满足与不可满足的相变点,当  $\alpha$  的值接近  $\alpha_d$  时实例往往是难解的<sup>[2,3]</sup>.目前,尚无得到  $\alpha_d$  的精确值,但  $\alpha_d$  的下界为 3.52<sup>[4]</sup>,上界为 4.489 8<sup>[5]</sup>.许可等人的研究基于 B 模型和 D 模型,针对经典 CSP 实例产生模型的平凡无解性问题,提出了具有精确相变点的实例产生模型,分别为 RB 和 RD 模型<sup>[6-8]</sup>,该模型被广泛应用于国际算法竞赛测试.

基于实例的不可满足性是否能够给出完备性证明,SAT 算法分为完备算法和不完备算法.不完备算法相对于完备算法而言,能够在短时间内求出问题的解,而完备算法的时间成本较大.因此,不完备算法得到了更为广泛的应用.DPLL 算法是一种典型的完备求解算法,主要是利用单字句规则和分裂规则对公式进行化简,其执行过程是一棵二叉搜索树<sup>[9]</sup>,增加了算法的执行时间.为此,人们提出了各种各样的策略,进而出现了许多其他有效的算法,如,基于 VSIDS 策略的 ZChaff 算法<sup>[10]</sup>、BerkMin 算法<sup>[11]</sup>、MiniSat 算法<sup>[12]</sup>、SATO 算法<sup>[13]</sup>、GRASP 算法<sup>[14]</sup>、GSAT 算法<sup>[15]</sup>和 WALKSAT 算法<sup>[16]</sup>等等.

信息传播算法起源于统计物理领域<sup>[17]</sup>.该算法是一种高效的边缘概率计算方法,在许多领域得到了广泛应用<sup>[18-29]</sup>.文献[30]中根据不同的随机 SAT 问题分别提出了 3 种基于因子图的信息传播算法,即警示传播(warning propagation,简称 WP)算法、信念传播(belief propagation,简称 BP)算法和调查传播(survey propagation,简称 SP)算法.在 3-SAT 随机实例中,该算法被验证非常有效,特别是 SP 算法与局部搜索算法相结合,几乎能够有效求解难解可满足性实例<sup>[30-33]</sup>.信息传播算法这种优秀的性能,引起了众多学者的关注,主要集中在两个方面:(1) 信息传播算法的收敛性;(2) 信息传播算法的有效性.

所谓收敛性是指:信息传播算法两次信息迭代得到的值不变;有效性是指:信息传播算法能够有效求解问题.目前,对信息传播算法的性能分析已取得一些重要的研究成果.文献[34-40]分析了信息传播算法的收敛性和有效性,并给出了算法收敛的充分条件.文献[41,42]基于具体的实例产生模型分析了信息传播算法的收敛性,并给出了算法收敛的概率条件.文献[43]利用压缩函数的相关性质给出了算法收敛的充分条件,该判定条件由于受消息取值的限制(消息只能取 0 或 1),导致其应用范围较窄.在此基础上,文献[44]分析了调查传播算法的消息取值在  $[0,1]$  区间上的收敛性,并给出算法收敛的一个充分条件.在高斯图模型中,信息传播算法常被用于计算均值和方差.文献[45]基于高斯线性模型给出了信息传播算法收敛的充分必要条件,但局限于相邻变量不能共享同一个因子.文献[46]基于信息矩阵分析了信息传播算法的收敛性,给出了一个充分必要条件,但局限于消息的信息矩阵为正半定矩阵,并且主要用于判定高斯 BP 算法的收敛性.文献[47]得到了信息传播算法的一个收敛条件,但该条件的证明需要基于双因素线性高斯模型,而该条件在一般问题上并不适用.文献[48]主要分析了求解高斯模型中方差的信息传播算法的收敛性,但要求用一个半定规划来检查收敛性,检查效率较低且非常困难.文献[49]分析了求解均值的信息传播算法的收敛性,要求估计有限尺度矩阵的谱半径,在实际应用中可行性不高.文献[50]分析了求解线性规划问题的信息传播算法的收敛性,给出算法收敛到线性规划问题最优解的一个标准框架.文献[51]分析了求解块内插问题(block interpolation problem)的信息传播算法的有效性,当因子图是一个  $N \times N$  的网格图时,算法能够正确收敛.在附加噪音的高斯图模型网络线性系统中,利用 WLS(weighted least

squares)标准进行状态估计,文献[52]基于协方差分析了信息传播算法收敛的动态行为.特别是在文献[53]中,利用赋范空间中函数的压缩映射原理,给出了信念传播算法收敛的一个统一框架.然而,对于具体的 SAT 问题,该条件验证复杂,可判定的范围有限,在大多数 SAT 实例上判定条件失效.不难看出:上述关于信息传播算法的收敛性分析都是基于某个具体模型,或者使用范围窄,或者判定条件验证复杂等.因此,对信息传播算法的收敛性分析仍然不完善.

综上,本文针对 BP 算法在求解 SAT 问题中的信息传播策略,对该算法的性能表现进行了系统的理论分析,将信息迭代方程的取值由(0,1)扩大到整个实数空间,得出了验证 BP 算法收敛的一个有效条件,检查该条件相对简单.最终,通过实验分析了该条件的有效性.至此,我们系统分析了 3 种求解 SAT 问题的信息传播算法的收敛性,这对于基于信息传播算法的求解器设计以及信息传播算法的广泛应用具有重要意义.

## 1 信念传播算法

### 1.1 因子图

一个 CNF 公式  $F=\{C_1, C_2, \dots, C_m\}$ , 公式由变元  $x_1, x_2, \dots, x_n$  构成, 为了方便表示, 将  $x_i$  简记  $i$ . 则 CNF 公式能够表示为一个二部网络  $G=(C \cup X, E)$ , 该二部网络称为因子图. 图中包含两类结点, 分别为由所有变元构成的变元结点集, 记为  $X=\{1, 2, \dots, n\}$ , 以及由变元的析取构成的子句结点集, 记为  $C=\{C_1, C_2, \dots, C_m\}$ . 图中的虚边和实边包含原 CNF 公式中两种不同的信息:

- 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  $\neg x_j$ .

为了方便在后文进行表示, 本文用  $a, b, \dots$  表示子句  $C_1, C_2, \dots$ .

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

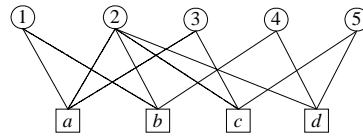


Fig.1 Factor graph

图 1 因子图

- $V(a)$ : 子句  $a$  所包含的所有变元的集合,  $V(a)=V_+(a) \cup V_-(a)$ . 其中,
  - $V_+(a)$ : 在子句  $a$  中正出现的变元集合;
  - $V_-(a)$ : 在子句  $a$  中负出现的变元集合;
  - $V(a) \setminus i$  表示集合  $V(a)$  中剔除结点  $i$ .

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

- $V(j)$ : 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;
  - $V(j) \setminus a = V(j) - \{a\}$ .

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

定义两个一致性集合  $V_a^u(j)$  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.

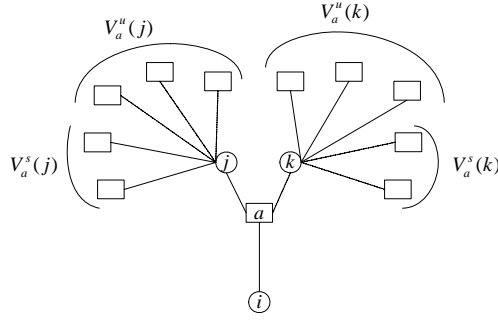


Fig.2 Part of factor graph

图 2 局部因子图

1.2 信念传播算法

基于消息传递而设计的信念传播算法,求解可满足性问题时有良好的有效性.文献[30]给出了求解可满足性问题的 BP 算法,其特征是基于因子图的消息传递过程.具体地讲,在因子图的每条边 $(a,i)$ 上定义消息 $\delta_{a \rightarrow i} \in [0,1]$ ,消息更新方程为

$$\delta_{a \rightarrow i} = \prod_{j \in V(a) \setminus i} \left( \frac{P_{j \rightarrow a}^u}{P_{j \rightarrow a}^u + P_{j \rightarrow a}^s} \right) \tag{1}$$

其中,

$$P_{j \rightarrow a}^u = \prod_{b \in V_a^s(j)} (1 - \delta_{b \rightarrow j}) \tag{2}$$

$$P_{j \rightarrow a}^s = \prod_{b \in V_a^u(j)} (1 - \delta_{b \rightarrow j}) \tag{3}$$

求解 CNF 公式的 BP 算法如算法 1 所示.

算法 1. BP algorithm.

Input: factor graph  $G$ , max step  $t_{\max}$ , precision  $\varepsilon$ .

Output: fix message  $\delta_{a \rightarrow i}^*$ .

1.  $t=0$ , initialize  $\delta_{a \rightarrow i}(t=0) \in [0,1]$  for random order;
2. **For**  $t=1$  to  $t=t_{\max}$ :
  - 2.1. call BP-UPDATE to generating new message  $\delta_{a \rightarrow i}(t)$ ;
  - 2.2. **If**  $|\delta_{a \rightarrow i}(t) - \delta_{a \rightarrow i}(t-1)| < \varepsilon$ , set  $\delta_{a \rightarrow i}^* = \delta_{a \rightarrow i}(t)$ ,  
go to Step 3;
3. **If**  $t=t_{\max}$   
return UN-CONVERGENCE.  
**Else if**  $t < t_{\max}$   
return  $\delta_{a \rightarrow i}^* = \delta_{a \rightarrow i}(t)$ ;

算法 2. BP-UPDATE.

Input: variable node  $j \in V(a) \setminus i$ .

Output:  $\delta_{a \rightarrow i}$ .

1. compute  $P_{j \rightarrow a}^u, P_{j \rightarrow a}^s$  ;  
**If**  $V_a^s(j) = \emptyset$ , set  $P_{j \rightarrow a}^u = 1$ ;
2. return  $\delta_{a \rightarrow i}$  by formula (1);

在信息传播算法中,因子图上的信息 $\delta_{a \rightarrow i}$  定义为:子句结点向变元结点发送的一个概率值,即变元的取值使

子句不可满足的概率信息.若BP算法进行若干次迭代后能够收敛,则可以利用全局信息引导可满足指派的搜索过程,这种方法被证明是非常有效的<sup>[30]</sup>.然而,对于因子图含有回路的SAT实例,BP算法有时候表现为不收敛,因此无法返回SAT实例的可满足性结果.是什么原因导致BP算法不收敛,目前还未给出理论解释.因此,仍然有必要对BP算法的收敛性从理论上做更为深入的研究.

## 2 主要结论及证明

设 $(V, \|\cdot\|)$ 为一个有限维实向量空间, $\mathbb{R}^N$ 上的 $l_1$ -范数定义为<sup>[44]</sup>

$$\|x\|_1 = \sum_{i=1}^N |x_i| \tag{4}$$

$\mathbb{R}^N$ 上的 $l_\infty$ -范数定义为

$$\|x\|_\infty = \max_{i \in \{1, 2, \dots, N\}} |x_i| \tag{5}$$

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.

设函数 $f: V \rightarrow V$ ,当 $0 \leq K < 1$ ,任意 $v, w \in V$ ,都有 $d(f(v) - f(w)) \leq Kd(v, w)$ ,那么 $f$ 为压缩函数.函数的压缩原理是指:如果函数 $f: V \rightarrow V$ 是完备度量空间 $(V, d)$ 中的压缩函数,对于 $\forall x \in V, f$ 收敛到唯一点 $x_\infty$ .设 $(V, \|\cdot\|)$ 是一个赋范空间,映射 $A: V \rightarrow V$ 的矩阵范数:

$$\|A\| = \sup_{v \in V} \|Av\| \tag{6}$$

$\mathbb{R}^N$ 上的 $A_1$ -范数定义为

$$\|A\|_1 = \max_{j \in \{1, 2, \dots, N\}} \sum_{i=1}^N |A_{ij}| \tag{7}$$

$\mathbb{R}^N$ 上的 $A_\infty$ -范数定义为

$$\|A\|_\infty = \max_{i \in \{1, 2, \dots, N\}} \sum_{j=1}^N |A_{ij}| \tag{8}$$

**引理 1**<sup>[54]</sup>. 设 $(V, \|\cdot\|)$ 是赋范空间,函数 $f: V \rightarrow V$ 可微,那么对于任意的 $x, y \in V$ ,则可以得到

$$\|f(x) - f(y)\| \leq \|x - y\| \cdot \sup_{z \in V} \|f'(z)\| \tag{9}$$

由上述引理以及函数压缩映射可以知道,若 $\sup_{z \in V} \|f'(z)\| < 1$ ,则函数 $f$ 能够收敛到 $x_\infty$ .如果函数 $f: V \rightarrow V$ 是可微的,那么通过函数 $f$ 的雅可比矩阵 $f'(z)$ 来构造矩阵范数 $A$ .

为了分析方便,假设 $\delta_{a \rightarrow i} \in (0, 1)$ .更进一步,为了利用实向量空间中的压缩映射原理,我们将 $(0, 1)$ 扩展为 $(-\infty, +\infty)$ ,利用双曲正切函数进行参数化, $\frac{1}{2}(1 + \tanh \lambda_{a \rightarrow i}) = \delta_{a \rightarrow i}$ ,参数 $\lambda_{a \rightarrow i} \in \mathbb{R}$ .BP算法的迭代公式(1)被写成如下形式:

$$\tanh \lambda_{a \rightarrow i} = 2 \prod_{j \in V(a), i} \left( \frac{P_{j \rightarrow a}^u}{P_{j \rightarrow a}^u + P_{j \rightarrow a}^s} \right) - 1 \tag{10}$$

其中,

$$P_{j \rightarrow a}^u = \prod_{b \in V_a^s(j)} \frac{1}{2} (1 - \tanh \lambda_{b \rightarrow j}) \tag{11}$$

$$P_{j \rightarrow a}^s = \prod_{b \in V_a^u(j)} \frac{1}{2} (1 - \tanh \lambda_{b \rightarrow j}) \tag{12}$$

为了书写方便,假设

$$x_j = \prod_{b \in V_a^s(j)} \frac{1}{2} (1 - \tanh \lambda_{b \rightarrow j}), y_j = \prod_{b \in V_a^u(j)} \frac{1}{2} (1 - \tanh \lambda_{b \rightarrow j}).$$

因此,

$$\tanh \lambda_{a \rightarrow i} = 2 \prod_{j \in V(a), i} \left( \frac{x_j}{x_j + y_j} \right) - 1 \tag{13}$$

将CNF公式对应的因子图的消息边进行随机排序,构造序列集合 $D = \{a \rightarrow i: (a, i) \in E\}$ ,定义函数 $f: \mathbb{R}^{|D|} \rightarrow \mathbb{R}^{|D|}$ ,则其分量为 $f(\lambda)_{a \rightarrow i} = \lambda_{a \rightarrow i}$ .那么函数 $f$ 的迭代过程能够表示为算法中的消息并行更新过程.因此,由引理1可知,计

算  $\sup_{\lambda \in \mathbb{R}^D} |f'(\lambda)| < 1$ , 则可以得到判定算法收敛的条件. 由前文第 1 节可知,  $V_a^u(j) \cap V_a^s(j) = \emptyset$ . 下面分两种情形进行讨论.

- 情形 1: 设  $b \in V_a^u(j)$ , 根据公式(13), 有:

$$\begin{aligned} f'(\lambda_{a \rightarrow i, b \rightarrow j}) &= \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \\ &= 2 \left( \frac{x_j}{x_j + x_j} \right)' \frac{1}{1 - \tanh^2 \lambda_{a \rightarrow i}} \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right) \\ &= 2 \frac{-x_j y_j'}{(x_j + y_j)^2 (1 - \tanh^2 \lambda_{a \rightarrow i})} \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right). \end{aligned}$$

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

所以,

$$\begin{aligned} \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} &= 2 \frac{x_j y_j (1 + \tanh \lambda_{b \rightarrow j})}{(x_j + y_j)^2 (1 - \tanh^2 \lambda_{a \rightarrow i})} \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right) \\ &= \frac{y_j (1 + \tanh \lambda_{b \rightarrow j}) (1 + \tanh \lambda_{a \rightarrow i})}{(x_j + y_j) (1 - \tanh^2 \lambda_{a \rightarrow i})} \\ &= \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \tanh \lambda_{a \rightarrow i}}. \end{aligned}$$

由公式(13)可知:

$$\tanh \lambda_{a \rightarrow i} = 2 \prod_{j \in V(a) \setminus i} \left( \frac{x_j}{x_j + y_j} \right) - 1 = 2 \left( \frac{x_j}{x_j + y_j} \right) \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right) - 1,$$

其中,

$$0 \leq \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right) \leq 1,$$

所以,

$$\tanh \lambda_{a \rightarrow i} \leq 2 \left( \frac{x_j}{x_j + y_j} \right) - 1.$$

进一步有:

$$\frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} = \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \tanh \lambda_{a \rightarrow i}} \leq \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \left( 2 \frac{x_j}{x_j + y_j} - 1 \right)} = \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{2 \frac{y_j}{x_j + y_j}} = \frac{1}{2} (1 + \tanh \lambda_{b \rightarrow j}) < 1,$$

因此,

$$\left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| < 1.$$

- 情形 2: 设  $b \in V_a^s(j)$ , 根据公式(13)有:

$$\begin{aligned} f'(\lambda_{a \rightarrow i, b \rightarrow j}) &= \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \\ &= 2 \left( \frac{x_j}{x_j + x_j} \right)' \frac{1}{1 - \tanh^2 \lambda_{a \rightarrow i}} \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right) \\ &= 2 \frac{x_j' y_j}{(x_j + y_j)^2 (1 - \tanh^2 \lambda_{a \rightarrow i})} \prod_{k \in V(a) \setminus i, j} \left( \frac{x_k}{x_k + y_k} \right). \end{aligned}$$

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

所以,

$$\begin{aligned} \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} &= -2 \frac{x_j y_j (1 + \tanh \lambda_{b \rightarrow j})}{(x_j + y_j)^2 (1 - \tanh^2 \lambda_{a \rightarrow i})} \prod_{k \in V(a) \setminus \{i, j\}} \left( \frac{x_k}{x_k + y_k} \right) \\ &= -\frac{y_j (1 + \tanh \lambda_{b \rightarrow j}) (1 + \tanh \lambda_{a \rightarrow i})}{(x_j + y_j) (1 - \tanh^2 \lambda_{a \rightarrow i})} \\ &= -\frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \tanh \lambda_{a \rightarrow i}}, \end{aligned}$$

因此,

$$\begin{aligned} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| &= \left| -\frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \tanh \lambda_{a \rightarrow i}} \right| \\ &= \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \tanh \lambda_{a \rightarrow i}} \\ &\leq \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{1 - \left( 2 \frac{x_j}{x_j + y_j} - 1 \right)} \quad (0 < x_j, y_j < 1) \\ &= \frac{y_j}{x_j + y_j} \frac{1 + \tanh \lambda_{b \rightarrow j}}{2 \frac{y_j}{x_j + y_j}} \\ &= \frac{1}{2} (1 + \tanh \lambda_{b \rightarrow j}) < 1. \end{aligned}$$

关于情形 1 和情形 2,公式(21)和公式(25)表明:对于  $\forall \lambda_{b \rightarrow j} \in (-\infty, +\infty)$ , 都有:

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

所以,

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

注意: $\lambda_{a \rightarrow i}$  和  $\lambda_{b \rightarrow j}$  是  $\lambda$  的分量.

利用函数  $f$  的雅可比矩阵构造矩阵范数  $A$ , 也即  $A$  中的元素为  $\tau_{a \rightarrow i, b \rightarrow j} = \sup_{\lambda \in V} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right|$ . 我们给出如下结论.

**定理 1.** 对于  $A_1$ -范数, 如果  $\max_{b \rightarrow j} \sum_{a \rightarrow i} \tau_{a \rightarrow i, b \rightarrow j} < 1$  成立, 那么 BP 算法在  $A_1$ -范数下能够有效收敛到固定点, 且与初始信息无关.

证明: 根据公式(6)、公式(7)和公式(9), 有:

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

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 BP 算法在  $A_1$ -范数下收敛, 且与初始信息无关. □

同理可得以下定理.

**定理 2.** 对于  $A_\infty$ -范数, 如果  $\max_{a \rightarrow i} \sum_{b \rightarrow j} \tau_{a \rightarrow i, b \rightarrow j} < 1$  成立, 则 BP 算法在  $A_\infty$ -范数下收敛到固定点, 且与初始信息无关.

证明: 根据公式(6)、公式(8)和公式(9), 有:

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

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 BP 算法在  $A_\infty$ -范数下收敛,且与初始信息无关. □

为了书写方便,不妨设  $f^2(x)=f(f(x)), f^3(x)=f(f(f(x))), \dots$  有如下引理成立.

**引理 2**<sup>[44]</sup>. 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  $x_\infty$ ,也即对于任意的  $x \in V$ ,函数  $f$  的迭代序列  $x, f(x), f^2(x), \dots$  收敛到  $x_\infty$ .

设  $\gamma_1, \gamma_2, \dots, \gamma_n$  是方阵  $M$  的特征值,构成  $M$  的谱  $\gamma(M) = \{\gamma_1, \gamma_2, \dots, \gamma_n\}$ ,  $M$  的谱半径为  $\rho(M) = \sup\{|\gamma_i|: \gamma_i \in \gamma(M)\}$ . 假设  $\tau = \max_{a \rightarrow i, b \rightarrow j} \{\tau_{a \rightarrow i, b \rightarrow j}\}$ , 一个  $|D| \times |D|$  的方阵  $M$  中的元素定义为  $M_{a \rightarrow i, b \rightarrow j} = 1_{V(j)a}(b) 1_{V(a)i}(j) \tau$ . 若  $x \in X$ , 则  $1_X(x) = 1$ ; 若  $x \notin X$ , 则  $1_X(x) = 0$ .

**定理 3.** 如果方阵  $M$  的谱半径  $\rho(M) < 1$ , 那么 BP 算法一定能达到收敛点,并且收敛过程不依赖于初始条件.

证明:文献[44]中的定理 4 表明:对于任意的  $x \in \mathbb{R}^{|D|}$ ,  $M$  是一个不依赖于  $x$  的非负方阵,一个可微函数  $f: \mathbb{R}^{|D|} \rightarrow \mathbb{R}^{|D|}$ , 假设  $|f'(x)|$  小于等于矩阵  $M$  中的对于元素.如果  $M$  的谱半径  $< 1$ , 那么对于任意的  $x \in \mathbb{R}^{|D|}$  收敛到唯一

固定点  $x_\infty$ , 也即  $x, f(x), f^2(x), \dots, x_\infty$ . 由前文中关于函数  $f$  的构造及情形 1 和情形 2 的推导可知,  $|f'(x)| \leq M$ . 因此, BP 算法能够有效收敛,收敛过程不依赖于初始条件. □

例 2: 设命题公式为  $F_1 = \{(x_1 \vee x_2), (x_2 \vee x_3), (x_3 \vee x_4), \dots, (x_{n-1} \vee x_n), (x_n \vee x_1)\}$ , 该公式的对应因子图中包含一个环路结构. 所以, 由上文中所提构造方阵  $M$  的过程可知, 方阵中任意行中除了一个值可能不是 0 的元素之外, 所有元素的值都是 0, 将该可能值为非 0 的元素设为  $\tau$ . 通过圆盘定理可以知道:  $|\lambda| \leq \tau$ , 其中  $\lambda$  为  $M$  的特征值. 则利用上文中的证明可知:  $|\lambda| \leq \tau < 1$ . 因此, 有  $\rho(M) < 1$ . 又由定理 3 可得, 在公式  $F_1$  中, BP 算法能够迭代至收敛状态. 同理, 在  $F_2 = \{(\neg x_1 \vee \neg x_2), (\neg x_2 \vee \neg x_3), (\neg x_3 \vee \neg x_4), \dots, (\neg x_{n-1} \vee \neg x_n), (\neg x_n \vee \neg x_1)\}$  中, BP 算法同样能够到达收敛的状态.

注意到: 定理 1 和定理 2 分别是针对范数  $\|A\|_1$  和  $\|A\|_\infty$  得出的结论, 不同的范数对应的收敛条件不同. 然而, 目前还不清楚是否存在一个最优范数, 使得 BP 算法收敛的判定条件更加有效.

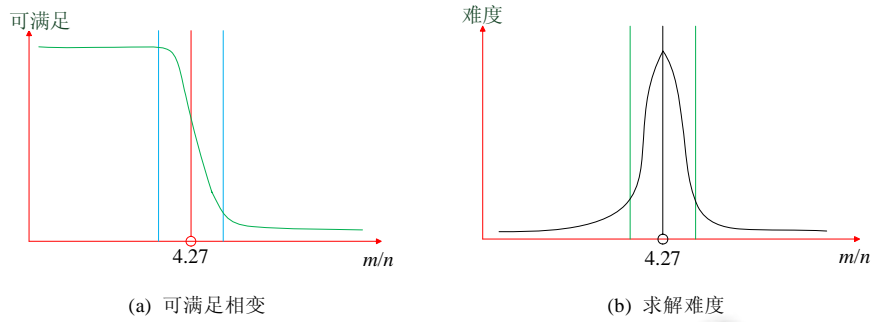
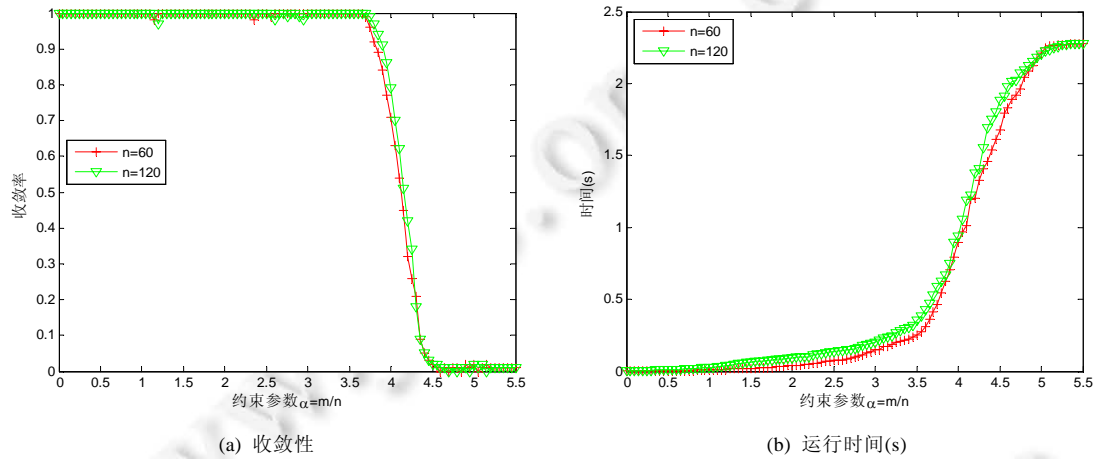
### 3 数值实验及分析

数值实验中用到的数据由模型  $G(n, k, m)$  产生,  $n$  表示变元个数,  $m$  表示子句个数,  $k$  表示每个子句的长度. 实验中, BP 算法的最大迭代次数  $t_{\max} = 10^3, \epsilon = 0.001$ .

当  $k=3$  时,  $G(n, k, m)$  模型被用于产生随机 3-SAT 实例. 研究表明: 对于随机 3-SAT 实例, 存在可满足性相变现象. 如图 3(a) 所示, 当子句个数与变元个数的比值  $m/n$  小于 3.5 时, 随机 3-SAT 实例高概率可满足<sup>[4]</sup>; 当子句个数与变元个数的比值  $m/n$  大于 4.5 时, 随机 3-SAT 实例高概率不可满足<sup>[5]</sup>; 对于子句个数与变元个数的比值  $m/n$  介于 3.5 与 4.5 之间的随机 3-SAT 实例, 其可满足性不能从概率上给出较为确切的回答. 同时, 子句个数与变元个数的比值  $m/n$  介于 3.5 与 4.5 之间的随机 3-SAT 实例, 其求解难度也较大, 大多数求解算法在该区域的实例集上效果不明显, 称该区域为随机 3-SAT 实例的难解区域, 如图 3(b) 所示. 在  $\alpha = 4.27$  附近, 求解难度最大. 因此, 常被国际竞赛用于构造难解实例来测试算法的性能.

在实验中, 我们选取了两组规模不同的随机 3-SAT 实例, 分别为  $n=60$  和  $n=120$ , 每次利用随机实例产生模型生成 100 个随机实例作为一组数据, 并统计每组数据的实验结果, 作为图中的每个数据点. 相关研究表明: 随机实例的难解程度与  $\alpha$  的值密切相关. 信息传播算法是基于消息迭代的近似算法, 收敛性直接影响算法的性能. 图 4(a) 是 BP 算法在随机 3-SAT 实例集上的收敛情况, 纵坐标是统计了所有实例中, 收敛实例所占的比值, 即实例能够达到收敛状态的所占的概率; 横坐标为  $\alpha$  的变化情况. 从图中可以看出: 大约  $\alpha < 3.8$  时, BP 算法高概率收敛; 大约  $\alpha > 3.8$  时, BP 算法的收敛性开始下降; 大约  $\alpha > 4.5$  时, BP 算法高概率不收敛. 这也表明 BP 算法主要用于求解可满足性实例, 在求解不可满足的实例时, 算法常常会失效. 图 4(b) 是 BP 算法的运行时间随约束参数  $\alpha$  的变化情况, 纵坐标表示运行时间(单位: s). 当  $\alpha > 3.52$  时, 算法的运行时间急剧增大; 随着  $\alpha$  的增大, 实例的因子图的结构变得复杂, 导致 BP 算法的计算时间增加.



Fig.3 Properties change of random 3-SAT instance with parameter  $m/n$ 图 3 随机 3-SAT 实例的特征随参数  $m/n$  的变化情况Fig.4 Properties change of convergence and time of BP with parameter  $m/n$ 图 4 BP 算法的收敛性和运行时间随参数  $m/n$  的变化情况

定理 3 中,要求计算矩阵的谱半径.图 5 为随机 3-SAT 实例的约束参数  $\alpha$  与对应矩阵的谱半径之间的关系.参数  $\rho$  随着约束参数  $\alpha$  的变化情况,纵坐标表示谱半径  $\rho$ ,横坐标表示约束参数  $\alpha$ .当  $\alpha < 2.0$  时,实例的谱半径小于 0.7,由定理 3 可知,BP 算法能够达到收敛状态.当约束参数  $\alpha > 2.0$  时,谱半径  $\rho$  几乎不小于 1,定理 3 失效,无法判定 BP 算法的收敛性.直观上, $\alpha$  越小,因子图的结构就越简单,对应的方阵  $M$  越稀疏,谱半径  $\rho$  越接近 0;相反, $\alpha$  越大,因子图的结构就越复杂,谱半径  $\rho$  也越大,可能导致定理 3 中的判定条件失效.

本文与文献[53]类似,都可用于约束可满足性问题的收敛性判定.其他文献主要用于其他问题,与本文没有可比性.所以,我们把本文结论与文献[53]中的结论做了比较,见表 1.由于可满足性问题实例的求解难度与问题的规模无关,实验中取  $n=120$ .结合图 3(a)和图 4,选取约束参数  $\alpha$  的值分别为 1.0,2.0,3.0,4.0,5.0 进行比较,每个数据点随机生成 100 个实例.表 1 中的数据表示能够判定 BP 算法收敛的实例数占总实例数的比例,可以看出:当  $\alpha < 2.0$  时,本文结论优于文献[53]中的结论;当  $\alpha=3.0$  和  $\alpha=4.0$  时,本文结论比文献[53]结论略差.然而,通过实验发现,验证文献[53]中判定条件远比本文条件复杂,在实际应用中可操作性不高.

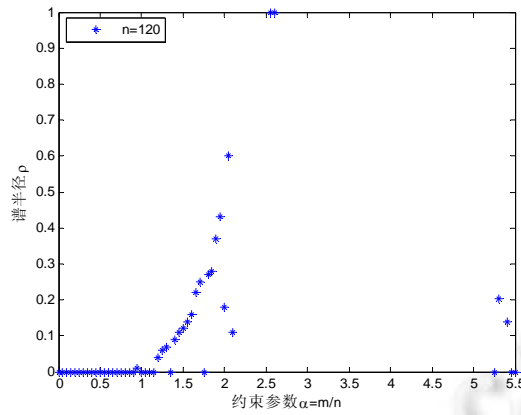


Fig.5 Properties change of spectral radius  $\rho$  with parameter  $m/n$

图 5 谱半径  $\rho$  随参数  $m/n$  的变化情况

Table 1 Comparison with other methods

表 1 与其他方法比较

判定方法	$\alpha=1.0$	$\alpha=2.0$	$\alpha=3.0$	$\alpha=4.0$	$\alpha=5.0$
文献[53]中的定理 4	0.91	0.84	0.27	0.07	0.13
本文定理 3	0.95	0.88	0.21	0.02	0.22

#### 4 结束语

本文针对用于求解 SAT 问题的 BP 算法的收敛性进行了分析,在 BP 算法中,通过将算法的消息迭代方程取双曲正切,并且将信息迭代方程的标准化函数的取值由(0,1)扩展为 $(-\infty, +\infty)$ ,得到了 BP 收敛的一个充分条件,该判定条件相对简单,并通过实验验证了该条件的有效性.至此,本文基于一般条件,完整分析了求解 SAT 问题的信息传播算法的收敛性,该结果对于理论研究或者实际应用提供了很大的参考价值.BP 算法求解 SAT 问题的不收敛现象大致有两种:算法通过迭代无法收敛于一个固定点或者算法达到了预定的迭代次数上限.由于本文的理论基础均建立在算法能够有效收敛的前提下.因此,下一步工作需要考虑是否存在收敛的唯一固定点、固定点的分布等问题、从而更为深入地对信息传播算法性能开展相关研究.

#### References:

- [1] Biere A, Heule M, Maaren H. van Walsh T. Handbook of Satisfiability. IOS Press, 2009.
- [2] Kirkpatrick S, Selman B. Critical behavior in the satisfiability of random Boolean formulae. Science, 1994,264(5163):1297-1301.
- [3] Mertens S, Mezard M, Zecchina R. Threshold values of random  $k$ -SAT from the cavity method. Random Structure Algorithms, 2006,28(3):340-373.
- [4] Kaporis A, Kirousis L, Lalas E. The probabilistic analysis of a greedy satisfiability algorithm. Random Structures&Algorithms, 2006,28(4):444-480.
- [5] Josep D, Kirousis L, Dieter M, Xavier PG. On the satisfiability threshold of formulas with three literals per clause. Theoretical Computer Science, 2009,410(30):2920-2934.
- [6] Xu K, Li W. Exact phase transitions in random constraint satisfaction problems. Journal of Artificial Intelligence Research, 2000, 12(1):93-103.
- [7] Xu K, Li W. Many hard examples in exact phase transitions. Theoretical Computer Science, 2006,355(3):291-302.
- [8] Xu K, Boussemart F, Hemery F. Random constraint satisfaction: Easy generation of hard satisfiable instances. Artificial Intelligence, 2007,171(8):514-534.
- [9] Davis M, Logemann G, Lovelan D. A machine program f or theorem proving. Communications of the ACM, 1962,5(7):394-397.

- [10] Moskewicz MW, Madigan CF, Zhao Y. Chaff: Engineering an efficient SAT solver. In: Proc. of the 38th Design Automation. 2001. 530–535.
- [11] Goldberg E, Novikov Y. BerkMin: A fast and robust SAT-solver. In: Proc. of the Design Automation and Test in Europe. 2002. 142–149.
- [12] Eén N, Sörensson N. An Extensible SAT-solver: Theory and Applications of Satisfiability. Berlin: Springer-Verlag, 2004. 502–518.
- [13] Zhang H. SATO: An efficient propositional prover. In: Proc. of the Int'l Conf. on Automated Deduction. 1997. 272–275.
- [14] Silva JP, Sakallah KA. GRASP: A search algorithm for propositional satisfiability. IEEE Trans. on Computers, 1999,48(5): 506–521.
- [15] Selman B, Kautz H, Cohen B. Noise strategies for improving local search. In: Proc. of the AAAI. 1994. 337–343.
- [16] Selman B, Levesque HJ, Mitchell D. A new method for solving hard satisfiability problem. In: Proc. of the AAAI. 1992. 440–446.
- [17] Pearl J. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufman Publishers, 1988.
- [18] Zhao CY, Zheng ZM. A belief-propagation algorithm based on variable entropy for constraint satisfaction problems. SCIENTIA SINICA Informationis, 2012,42(9):1170–1180 (in Chinese with English abstract).
- [19] Yin MH, Zhou JP, Sun JG, Gu WX. Heuristic survey propagation algorithm for solving QBF problem. Ruan Jian Xue Bao/Journal of Software, 2011,22(7):1538–1550 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3859.htm> [doi: 10.3724/SP.J.1001.2011.03859]
- [20] Ravanbakhsh S, Greiner R. Perturbed message passing for constraint satisfaction problems. Artificial Intelligence, arXiv Preprint arXiv:1401.6686, 2014.
- [21] Chieu HL, Lee WS. Relaxed survey propagation for the weighted maximum satisfiability problem. Journal of Artificial Intelligence Research, 2009,36(1):229–266.
- [22] Bayati M, Shah D, Sharma M. Max product for maximum weight matching: Convergence, correctness, and LP duality. IEEE Trans. on Information Theory, 2007,54(3):1241–1251.
- [23] Coja-Oghlan A, Mossel E, Vilenchik D. A spectral approach to analyzing belief propagation for 3-colouring. Combinatorics, Probability and Computing, 2009,18(6):881–912.
- [24] Gamarnik D, Shah D, Wei Y. Belief propagation for min-cost network flow: Convergence and correctness. Operations Research, 2012,60(2):410–428.
- [25] Sanghavi S, Malioutov DM, Willsky AS. Belief propagation and LP relaxation for weighted matching in general graphs. IEEE Trans. on Information Theory, 2011,57(4):2203–2212.
- [26] Sanghavi S, Shah D, Willsky AS. Message passing for maximum weight independent set. IEEE Trans. on Information Theory, 2009, 55(11):4822–4834.
- [27] Hetterich S. Analysing survey propagation guided decimation on random formulas. arXiv Preprint arXiv:1602.08519, 2016.
- [28] Amin CO. On belief propagation guided decimation for random  $k$ -SAT. In: Proc. of the 22nd Annual ACMSIAM Symp. on Discrete Algorithms. San Francisco, 2016. 957–966.
- [29] Marino R, *et al.* The backtracking survey propagation algorithm for solving random  $K$ -SAT problems. Nature Communications, 2016,7:Article No.12996. [doi: 10.1038/ncomms12996(2016)]
- [30] Braunstein A, Mezard M, Zecchina R. Survey propagation: An algorithm for satisfiability. Random Structures and Algorithms, 2005,27(2):201–226.
- [31] Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations. Journal of the ACM, 2007,54(4): 1089–1098.
- [32] Yedidia JS, Freeman WT, Weiss Y. Understanding belief propagation and its generalizations. Artificial Intelligence, 2003,8(1): 239–269.
- [33] Braunstein A, Zecchina R. Survey and belief propagation on random  $k$ -SAT. LNCS, 2004,2919(1):519–528.
- [34] Weiss Y. Correctness of local probability propagation in graphical models with loops. Neural Computation, 2000,12(1):1–41.
- [35] Weiss Y, Freeman WT. Correctness of belief propagation in gaussian graphical models of arbitrary topology. Neural Computation, 2001,13(10):2173–2200.

- [36] Tatikonda S, Jordan MI. Loopy belief propagation and Gibbs measure. In: Proc. of the 18th Annual Conf. on Uncertainty in Artificial Intelligence. 2002. 493–500.
- [37] Heskes T. On the uniqueness of loopy belief propagation fixed points. *Neural Computation*, 2004,16(11):2379–2413.
- [38] Ihler AT, Lii JWF, Willsky AS. Loopy belief propagation: Convergence and effects of message errors. *Machine Learning Research*, 2005,6(1):905–936.
- [39] Shi XQ, Schonfeld D, Tuninetti D. Message error analysis of loopy belief propagation for the sum-product algorithm. *Computer and Information Science*, 2010,1009:1–30.
- [40] Feige U, Mossel E, Vilenchik D. Complete convergence of message passing algorithms for some satisfiability problems. *Theory of Computing*, 2013,9(19):617–651.
- [41] Wang XF, Xu DY, Wei L. Convergence of warning propagation algorithms for random satisfiable instances. *Ruan Jian Xue Bao/Journal of Software*, 2013,24(1):1–11 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4213.htm> [doi: 10.3724/SP.J.1001.2013.04213]
- [42] Wang XF, Xu DY. Convergence of the belief propagation algorithm for RB model instances. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(11):2712–2724 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4877.htm> [doi: 10.13328/j.cnki.jos.004877]
- [43] Wang XF, Xu DY. Sufficient conditions for convergence of the warning propagation algorithm. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(12):3003–3013 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4940.htm> [doi: 10.13328/j.cnki.jos.004940]
- [44] Wang XF, Xu DY, Jiang JL, Tang YH. A sufficient condition for survey propagation convergence. *SCIENTIA SINICA Informationis*, 2017,47(12):1646–1661 (in Chinese with English abstract).
- [45] Du J, Ma S, Wu YC, Kar S, Moura JMF. Convergence analysis of distributed inference with vector valued Gaussian belief propagation. *Journal of Machine Learning Research*, 2018,18(172):1–38.
- [46] Du J, Ma S, Wu YC, Kar S, Moura JMF. Convergence analysis of the information matrix in Gaussian belief propagation. In: Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. 2017. 4074–4078.
- [47] Du J, Ma S, Wu YC, Kar S, Moura JMF. Convergence analysis of belief propagation for pairwise linear Gaussian models. In: Proc. of the IEEE Global Conf. on Signal and Information Processing. 2017. 548–552.
- [48] Su Q, Wu YC. Convergence analysis of the variance in Gaussian belief propagation. *IEEE Trans. on Signal Processing*, 2014, 62(19):5119–5131.
- [49] Su Q, Wu YC. On convergence conditions of Gaussian belief propagation. *IEEE Trans. on Signal Processing*, 2015,63(5): 1144–1155.
- [50] Park S, Shin J. Convergence and correctness of max-product belief propagation for linear programming. *Siam Journal on Discrete Mathematics*, 2017,31(3):2228–2246.
- [51] Wang Y, Reyes MG, Neuhoff DL. Correct convergence of min-sum loopy belief propagation in a block interpolation problem. *arXiv Preprint arXiv:1702.06391*, 2017.
- [52] Sui T, Marelli D, Fu M. Convergence analysis for gaussian belief propagation: Dynamic behaviour of marginal covariances. In: Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. 2016. 2599–2602.
- [53] Mooij JM, Kappen HJ. Sufficient conditions for convergence of the sum-product algorithm. *IEEE Trans. on Information Theory*, 2007,53(12):4422–4437.
- [54] Dieudonne J. *Foundations of Modern Analysis*. New York: Academic Press, 1960.

#### 附中文参考文献:

- [18] 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(9):1170–1180.
- [19] 殷明浩,周俊萍,孙吉贵,谷文祥.求解 QBF 问题的启发式调查传播算法. *软件学报*,2011,22(7):1538–1550. <http://www.jos.org.cn/1000-9825/3859.htm> [doi: 10.3724/SP.J.1001.2011.03859]
- [41] 王晓峰,许道云,韦立.随机实例集上警示传播算法的收敛性. *软件学报*,2013,24(1):1–11. <http://www.jos.org.cn/1000-9825/4213.htm> [doi: 10.3724/SP.J.1001.2013.04213]

- [42] 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(11):2712–2724. <http://www.jos.org.cn/1000-9825/4877.htm> [doi: 10.13328/j.cnki.jos.004877]
- [43] 王晓峰,许道云.警示传播算法收敛的充分条件.软件学报,2016,27(12):3003–3013. <http://www.jos.org.cn/1000-9825/4940.htm> [doi: 10.13328/j.cnki.jos.004940]
- [44] 王晓峰,许道云,姜久雷,唐延辉.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(12):1646–1661.



王晓峰(1980—),男,博士,副教授,CCF 专业会员,主要研究领域为机器学习,人工智能,算法分析与设计.



姜久雷(1972—),男,博士,教授,CCF 专业会员,主要研究领域为业务流程建模,服务计算,大数据分析.



许道云(1959—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算复杂性,可计算性分析.



李强(1987—),男,博士,副教授,CCF 专业会员,主要研究领域为协同工程,机器学习.



杨德仁(1964—),男,博士,教授,主要研究领域为软件理论,软件工程.



刘欣欣(1994—),女,硕士生,主要研究领域为算法分析与设计.