Probability Conditions for Convergence of the Warning Propagation Algorithm

Wang Xiao-feng, Ding Hong-sheng⁺, Yu Qian-cheng and Tian Jin-qin

Department of Computer Science, Beifang Minzu University, Yinchuan, China, 750021

Abstract. Message propagation algorithms are very effective in finding satisfying assignments for random SAT instances, and hard region become narrower. However, message propagation algorithms do not always converge for graphs with cycles. Unfortunately, rigorous theory proof of this phenomenon is still lacking. Warning Propagation algorithm is the most basic message propagation algorithm, we analyses convergence of the warning propagation algorithm, and gives the conditions for convergence.

Keywords: warning propagation algorithm, convergence, satisfiability problems.

1. Introduction

Satisfiability decision problem (SAT problem) refers to: given a Conjunction Normal Form (CNF) F, whether there exists a truth-values assignment that can make F true. Clause length is k for the SAT problem is called the k-SAT problem, when $k \ge 3$, the k-SAT problem is the famous NP-complete problem. Making NP completeness of SAT as seeds, and by the transforming of polynomial simplification, it has been found that many of the combinatorial optimization problems with certain characteristics ,the SAT problem with some constraint condition is still the NP-complete problem [1, 2], among them, the 3-SAT problem is a classic and basic NP complete problem.

The NP-completeness of SAT problem shows that the SAT problem cannot be solved intuitively in polynomial time. However, this problem cannot be avoided in practical applications. If we do not consider all the NP classes, and only consider the decision problem in certain special classes, it can be solved in polynomial time (even linear regression time). Such as the 2-CNF formulas, the Horn formulas, the CNF formulas with limiting some parameters etc. Intuitively, the solution time of the general SAT problem is exponential time. However, in practical applications, not every CNF problem needs exponential time. In the study of the phase transition phenomena of SAT problem, statistical phenomenon shows that there is a kind of critical phenomenon between satisfiability and unsatisfiability. For 3-CNF, the value of clauses in the formula and variable element number ratio is an important parameter, when the formula of the parameters in 4.25 nearby, formula satisfiability decision is indeed very difficult. However, when the formula of this parameter is far away from 4.25, the judgment of formula satisfiability may be in polynomial time completed [1]. Although the SAT problem is NP-complete, the statistics show that the hard NP class is not large proportion. It is proved that the approximation algorithm and the stochastic algorithm are reasonable in many practical applications. The phase transition of SAT problem has further promoted the research and application of probability method in SAT problem.

Since the birth of the DPLL algorithm for solving the SAT problem, people have designed a lot of excellent algorithms for SAT. Such as Zchaff algorithm [3], introduces strategies of conflict learning and two observation variables of Boolean Constraint Propagation (BCP) and other strategies, which have made a great breakthrough on DPLL algorithm. According to the phase change phenomena and statistical physics

⁺ Corresponding author. Tel.: +86-18109586231.

E-mail address: 78950228@qq.com.

theory and method, a lot of excellent probability algorithm and approximation algorithm are designed. Such as Huang Wenqi et al proposed the quasi-physical and quasi-human algorithm [4], some European scholars proposed the information propagation algorithm [5], and so on. For example, warning propagation(WP) algorithm [5] is a basic algorithm for information propagation, which can effectively solve the problem of the random 3-SAT instance when a < 3.50; Belief propagation (BP) algorithm is effective when a < 3.95; Survey propagation (SP) algorithm can extend the effective 3-SAT regional to a < 4.26 [5,6]; SP algorithm is the most effective information propagation algorithm.

Although the information propagation algorithm provides a very powerful tool to solve the SAT problem, there are still two problems need to be deeply studied in theory:

1) Whether the information propagation algorithm is convergent;

2) If the information propagation algorithm is convergent, whether the result is the effective approximation of the probability distribution of the variable edge. If the information propagation program does not converge, the algorithm is invalid.

Some research results have been obtained in the theoretical analysis of the convergence of the information propagation algorithm at present. For example, [7] analyzes the convergent of instance of factor diagram of information propagation algorithm, which the graph structure consists of a ring. It indicates that the information propagation algorithm has convergence, and the obtained result is an effective approximation for the distribution of the variables. In the literature [8] shows that the information propagation algorithm is able to converge correctly to the Gauss graph model with arbitrary structure. Tatikonda and Jordan analyzes the relationship between the uniqueness of the Gibbs measure sequence based on the computation tree and the convergence of the BP algorithm, and derive a sufficient condition of the BP algorithm for convergence [9]. Using the uniqueness of the Bethe free energy minimum value, Heskes gives the BP algorithm a effective condition with the only fixed point, and analyze the relationship of the free energy function between the information propagation algorithm and statistical physics. He points out the sufficient conditions for the BP algorithm convergence to the unique fixed point, however the theoretical analysis of the convergence of the information propagation algorithm hasn't done[10]. According to the characteristic of the information renewal function, some researchers have given some sufficient conditions for convergence of the information propagation algorithm by using the function of contraction mapping principle. For example, using the BP algorithm update function, Ihler obtains the boundary of the transfer error information, and gives a sufficient condition for convergence of information propagation algorithm taking advantage of the error information boundary [11]. Using vector space information update function of the compression mapping principle, Mooij and Kappen derive the sufficient conditions for the BP algorithm convergence, and point out that if the coefficient matrix spectral radius of the update function strictly less than 1, the BP algorithm is convergent [12]. Shi and Schonfeld et al. get a more compact boundary on the error information transmission, study the dynamic behavior of the information propagation algorithm by using this bound, and give a sufficient condition for the convergence of the algorithm. [13]. BP algorithm convergence for the minimum cost network flow (MCF) is analyzed in the paper [14], the results show that in the sub polynomial time, the BP algorithm receives the optimal solution of the MCF problem, and the optimal solution is unique. At the same time, this paper also provides a simplified BP algorithm, which can give a fully polynomial-time randomized approximation scheme (FPRAS) for the MCF problem. Reference [15] gives a conditional probability of the BP algorithm convergence. The results show that BP algorithm can give the optimal solution of the maximum weight matching problem and the minimum cost network flow problem, and the times of algorithm iterations is polynomial bound on high probability. [16] Analyzes the complexity of BP algorithm and BP Stochastic (SBP) algorithm, message update complexity of SBP is lower an exponential level than BP algorithm.

Propagation WP algorithm is the most basic information propagation algorithm, and the convergence analysis of WP algorithm is helpful to analyze the convergence of other information propagation algorithm. The convergence of the WP algorithm is analyzed in [17], which the main limitation is random satisfiability instances of the plant assignment model $P_{n,p}^{plant}$, that is, given an assignment and randomly selected clause sets with probability P of satisfied assignment to constitute instances. Specifically, in the model $P_{n,p}^{plant}$, p is sufficiently large, whose essence is, at most, only one

assignment meet examples with high probability, WP algorithm convergence is proved by using this property, and the corresponding convergence of a conditional probability is given. In the literature [18], we have analyzed the convergence of the warning propagation algorithm in the 3-SAT instance, that is, when $p < 1/8n^2$, the WP algorithm in the 3-CNF formula generated by the model G(n, k, p) is convergent with a high probability.

It can be seen that the information propagation algorithm is convergent on the cases with some special properties or structure. Therefore, the theoretical analysis of the convergence of the information propagation algorithm is still not perfect.

We draws lessons on the genetic algorithm convergence proof technique [19], which the message update process of the warning propagation algorithm is mapped to the state migration process of the Markov chain. Based the theory of Markov chain and the relevant conclusions in [19], we gives a probability condition of the warning propagation algorithm convergence. The experimental results show that the decision condition is valid.

2. Warning Propagation Algorithm

Set $F = \{C_1, C_2, \dots, C_m\}$ as a CNF formula, which contains n variables x_1, x_2, \dots, x_n . The formula *F* can be represented by a bipartite graph $G = (C \cup X, E)$, called a factor graph. The variable nodes set is $X = \{1, 2, \dots, n\}$, the clause nodes set is $C = \{C_1, C_2, \dots, C_m\}$. The edges in *G* are divided into two types: real edge and virtual edge.

Real edge: $(C_i, j) \in E \Leftrightarrow$ clause C_i containing positive literal x_j ; Virtual edge: $(C_i, j) \in E \Leftrightarrow$ clause C_i containing negative literal $\neg x_j$; V(a): to indicate a variable set that appears in clause a; $V^+(a)$: to indicate a variable set for positive literal that appears in clause a; $V^-(a)$: to indicate a variable set for negative literal that appears in clause a; $V(a) \setminus i := V(a) - \{i\}$; V(j): To indicate a set of clauses containing variable x_j , $V(j) := V^+(j) \cup V^-(j)$; $V^+(j)$: To indicate a set of clauses that the variable x_j is positive; $V^-(j)$: To indicate a set of clauses that the variable x_j is positive; $V^-(j)$: To indicate a set of clauses that the variable x_j is positive; $V(j) \setminus a := V(j) \setminus a = V(j) \setminus a$

 J_j^a is a identification parameter, if $x_j \in a$, then $J_j^a = -1$; if $\neg x_j \in a$, then $J_j^a = 1$. On each side (a,i) of the factor graph, we define the message transfer $u_{a\rightarrow i}$ (also known as warning information) of the WP algorithm. $u_{a\rightarrow i}$ represents the satisfiability of the clause *a* for variable x_i value tendency. The WP algorithm iterative equation is as follows:

$$u_{a \to i}(t) = \prod_{j \in V(a) \setminus i} \theta(-J_j^a(\sum_{b \in V(j) \setminus a} J_j^b u_{b \to j}(t-1)))$$
(1)

t is the number of iterations, $\theta(x)$ is the truncated function. If $x \le 0$, then $\theta(x) = 0$, else $\theta(x) = 1$. If *a* contains only variables x_i , then set $u_{a \to i} = 1$. When the WP algorithm convergence, according to the fixed variable of warning information x_i to assign.

$$H_i = -\sum_{b \in V(i)} J_i^b u_{b \to i}^*$$
⁽²⁾

If $H_i > 0$, then $x_i = 1$; if $H_i < 0$, then $x_i = 0$; else x_i is temporarily not assigned. Generally, (1) can be written as follows:

$$u_{a\to i}(t) = \prod_{j\in V(a)\setminus i} \theta(J_j^a(\sum_{b\in V^+(j)\setminus a} u_{b\to j}(t-1) - \sum_{b\in V^-(j)\setminus a} u_{b\to j}(t-1)))$$
(3)

$$h_{j \to a} = \sum_{b \in V^+(j) \setminus a} u_{b \to j}(t-1) - \sum_{b \in V^-(j) \setminus a} u_{b \to j}(t-1)$$
(4)

Named $h_{j\to a}$ as the cavity domain. If the variable x_j is only in a, then set $h_{j\to a} = 0$. The WP algorithm for 3-CNF formula F is as follow:

Warning Propagation (3-CNF formula F)

1. Construct the corresponding factor graph G(F);

2. Random assignment 1 or 0 for all message edges $u_{a\to i}(t=0)$ of the factor graph;

3. Repeat the following procedure, until the algorithm converges (We also set the maximum iterative step t_{max} to force algorithm end):

3.1 Random permutation edges of G(F);

3.2 According to random edges sequence, using (1) to update message $u_{a \to i}$;

4 according to H_i to calculate the partial assignment ψ , simplify the formula F;

5 return ψ ;

WP algorithm convergence refers to the n-th algorithm iterative information $u_{a\to i}(t)$ and the (n-1)-th information $u_{a\to i}(t-1)$ is consistent. The fixed point of Warning information is obtained when the algorithm is convergent, and determine the partial variable assignment furthermore, return the partial assignment ψ ; otherwise, the algorithm is not convergent, and return fail. It is obvious that the algorithm convergence plays the key role on the performance. A conclusion is as follow:

Theorem 1[5]: If the factor graph corresponding to the formula is a tree structure, the WP algorithm is convergent.

However, WP algorithm does not always be convergent if the factor graph instance is a ring graph structure. At present, the convergence analysis of the algorithm is still not perfect. Therefore, it is necessary to analyze the WP algorithm convergence.

3. A Probability Condition for the Convergence of the Warning Propagation Algorithm

In the WP algorithm, the message values is relaxed from $\{0,1\}$ to [0,1], and the convergence of the algorithm is analyzed by using the property of the function. There are the following conclusions:

Theorem 2: Let f has a finite fixed point, φ is a continuous function, which is satisfied $x \neq f(x) \Rightarrow \varphi(x) > \varphi(f(x))$, then f is convergent.

Proof: set $\omega_i (j \le N_0)$ is the finite fixed point in the V about f.

$$M_{i} = \{x \in V : ||x - \omega_{i}|| < \varepsilon\} \quad (j \le N_{0})$$

$$\tag{5}$$

Let ε sufficiently small, when $x \in M_j$, $f(x) \notin M_k$ $(k \neq j, k \leq N_0)$. Using absurdity to prove. Assume that f does not converge, then $f^n(x)$ will have unlimited times to fall into the complement \overline{M} of $M = \bigcup_{i=1}^{N_0} M_j$. So there is a limit point of $z \notin M$ of $f^n(x)$. And because $\varphi(z) > \varphi(f(z))$, by φ continuity, $\exists^1 \varepsilon > 0$, $M' = \{y \in V : || y - z || \leq \varepsilon\}$.

$$\varphi(z) > \varphi(f(M')) = \max_{y \in M'} \varphi(f(y)) \tag{6}$$

Thus, $\varphi(z) > \varphi(f(M')) \ge \varphi(f^2(M'))\cdots$, however z is the limit point of $f^n(x)$, it must be fall in $\bigcup_k f^k(M')$ (contradiction). So, f is convergent.

Let *F* is a CNF formula, whose set of sequences of factor graph edge is $D = \{a \rightarrow i : (a,i) \in E\}$, warning messages $u_{a\rightarrow i}$ on each side $a \rightarrow i$ is to be updated by (1). Noticed that warning information is $u_{a\rightarrow i} \in \{0,1\}$, we suppose message update of WP algorithm is performed in parallel, iteration message vector is $u = (u_{a\rightarrow i}, u_{b\rightarrow j}, \cdots) \in V = \{0,1\}^{|D|}$. *u* is considered as a state, u_0 is the initial state, and u_t is the *t* state by *t* steps iteration.

Theorem 3: The message iterative process of the WP algorithm is a finite state Markov chain.

Proof: Let $\{X_n : n \ge 0\}$ as a column of discrete random variables, the entire discrete value marked with S = V, *S* is the state space, which is limited. $X_0 = u_0$ indicates the initial iterative state of the warning propagation algorithm, $X_n = u_n$ indicates the state by *n* steps iteration, set $f : V \rightarrow V$, for arbitrary $u \in V$, the message update equation (1) can be represented components form $f(u)^{a \rightarrow i} = u_{a \rightarrow i}$. Parallel update of WP algorithm is described as follow :

$$X(n+1) = f(X(n)) \tag{7}$$

Therefore, X_{n+1} is only relevant with X_n . Apparently, the sequence $\{X_n : n \ge 0\}$ constitutes a finite state Markov chain on *S*. According to Theorem 3, we can use correlation theories of Markov chain to study the WP algorithm convergence.

 $s_i = (s_{i,1}, s_{i,2}, \dots, s_{i,|D|}) \in S$ and $s_j = (s_{j,1}, s_{j,2}, \dots, s_{j,|D|}) \in S$ are the states set. The distance between the two iterative information is defined as $d(s_i, s_j) = \sum_{i=1}^{|D|} |s_{i,k} - s_{j,k}|$. For any integer $t \ge 0$,

If the distance between *t* iteration information s_t and t+1 iteration information s_{t+1} is $d(s_t, s_{t+1}) = 0$, then the WP algorithm convergence. Defining a collection $B \subset S$: for any $X(n) \in B$, there is d(X(n), X(n+1)) = 0, we call *B* as the fixed point set of WP algorithm.

$$\alpha_n^B = \Pr(X(n+1) \cap B = \emptyset / X(n) \cap B \neq \emptyset)$$
(8)

$$\beta_n^B = \Pr(X(n+1) \cap B = \emptyset / X(n) \cap B = \emptyset)$$
(9)

Theorems 4:If α_n^B , β_n^B meet:

(1)
$$\sum_{n=1}^{\infty} (1 - \beta_n^B) = \infty$$
; (2) $\lim_{n \to \infty} \frac{\alpha_n^B}{1 - \beta_n^B} = 0$.

Then the WP algorithm is high probability convergent.

$$\lim_{n \to \infty} \Pr(X(n) \cap B \neq \emptyset) = 1 \tag{10}$$

Proof: Set $P_0(n) = \Pr(X(n) \cap B = \emptyset)$, according to the Bayesian formula:

$$P_{0}(n+1) = \Pr(X(n+1) \cap B = \emptyset)$$

$$= \Pr(X(n+1) \cap B = \emptyset / X(n) \cap B \neq \emptyset)$$

$$\cdot \Pr(X(n) \cap B \neq \emptyset)$$

$$+ \Pr(X(n+1) \cap B = \emptyset / X(n) \cap B = \emptyset)$$

$$\cdot \Pr(X(n) \cap B = \emptyset)$$

$$\leq \alpha_{n}^{B} + \beta_{n}^{B} P_{0}(n)$$
(11)

By condition (2), for any $\varepsilon > 0$, there exists N_1 , when $n \ge N_1$, there is

$$\frac{\alpha_n^{\scriptscriptstyle D}}{1 - \beta_n^{\scriptscriptstyle B}} \le \frac{\varepsilon}{2} \tag{12}$$

$$\begin{pmatrix}
P_0(n+1) - \frac{\varepsilon}{2} \\
-\beta_n^B \left(P_0(n) - \frac{\varepsilon}{2} \right) \\
\leq P_0(n+1) - \beta_n^B P_0(n) - \alpha_n^B \\
\leq 0$$
(13)

thereby

$$P_0(n+1) - \frac{\varepsilon}{2} \le \beta_n^B \left(P_0(n) - \frac{\varepsilon}{2} \right)$$
(14)

By using the recursive method

$$P_0(n+1) \le \frac{\varepsilon}{2} + \prod_{k=1}^n \beta_n^B \left(P_0(n) - \frac{\varepsilon}{2} \right)$$
(15)

The condition (1) is equivalent to $\prod_{k=1}^{n} \beta_{n}^{B} = 0$, then exists N_{2} , when $n \ge N_{2}$, $\prod_{k=1}^{n} \beta_{k}^{B} \le \frac{\varepsilon}{2}$, then $n \ge \max(N_{1}, N_{2})$, when $n \ge \max(N_{1}, N_{2})$, $P_{0}(n+1) \le \varepsilon$. According to arbitrary nature of ε , we can get following results:

 $\lim_{n \to \infty} \Pr(X(n) \cap B = \emptyset) = \lim_{n \to \infty} P_0(n) = 0$

Thereby

 $\lim \Pr(X(n) \cap B \neq \emptyset) = 1.$

4. Numerical Experiments and Analysis

WP algorithm to solve the satisfiability problem is more effective. But WP algorithm may not converge to the instance of the factor graph with ring structure. The experimental results show that the WP algorithm

cannot converge in random 3-SAT instances G(n,k,m) on condition a > 3.91, and can't give the effective solution of the problem. In the numerical experiments, we use the model G(n,k,m) to generate three different scale random 3-SAT instances set. In the WP algorithm, the maximum number of iterations is $t_{max} = 10^3$.

When the problem scale is n = 20, the convergence of the WP algorithm in the random 3-SAT instance set G(n,k,m) generated is shown in Fig. 1. It shows that the probability of the WP algorithm convergence changes with the change of the constraint parameter a = m/n. Each data point in the graph is composed of 1000 random instances generated by the model G(n,k,m). When the parameter *a* increases to 3.91, WP algorithm is changed from the convergence to not convergent. When the parameter a < 3.91, the WP algorithm is convergent with high probability. We know that the structure of factor graphs directly affects the convergence of WP algorithm.



Fig. 1: Problem scale is n = 20, the convergence of the WP algorithm on random instances of the model G(n, k, m).



Fig. 2: Problem scale is n = 40, the convergence of the WP algorithm on random instances of the model G(n, k, m)



Fig. 3: Problem scale is n = 60, the convergence of the WP algorithm on random instances of the model G(n, k, m)

In fact, when the parameters a > 3.91, the random instances by G(n,k,m) generating are basically in the hard solution region and more difficult to solve, and the instance structure of the factor diagram is more complex. As a result, the WP algorithm does not converge. Fig. 2 and Fig. 3 show respectively the convergence of the WP algorithm on scale n = 40 and n = 60 of G(n,k,m) instance set. As the scale *n* of the problem increases, the probability of WP algorithm convergence on the instance G(n,k,m) tend to a stable value. As shown in fig 2 and fig3. The convergence of WP algorithm has a mutation at the constraint parameter a = 3.91, and the probability curve of convergence almost unchanged with the *n* increase. The

experimental results show that the convergence of the algorithm is independent of the scale n of the problem. In this convergent region, the validity of the conclusion in Theorem 4 is explained from the probability.

5. Ending

The convergence of WP algorithm is analyzed in this paper. By mapping the iterative process of the WP algorithm to the state transfer process of Markov chain, and using the properties of stable distribution of Markov chain, we give the conditions for convergence. Further work is that the study of belief propagation algorithm and survey propagation algorithm convergence based on the WP algorithm convergence analysis method, gives a sufficient condition for the algorithm convergence and strict theoretical proof.

6. Acknowledgements

This work is supported by National Nature Science Foundation of China(NO.61462001), Ningxia Colleges Projects Research Foundation (NO.NGY2013093), North University for Nationalities Projects Research Foundation (No.2014XYZ03), North University for Nationalities Projects Research Foundation :Agricultural IOT Innovation Team Project Foundation of North University for Nationalities.

7. References

- [1] Kleine B. H and Lettman T. Propositional Logic: Deduction and Algorithms, Cambridge University Press, 1999
- [2] Papadimitriou C. H. Computational Complexity, Addison-Wesley Publishing Company, 1994
- [3] Moskewicz M.W, Madigan C.F, Zhao Y.Chaff:Engineering an efficient SAT solver[C].//Proceedings of the 38th annual Design Automation Conference, ACM, New York,2001:530-535
- [4] Huang Wenqi, Jin Chao, Quasiphysical and quasisociological algorithm Solar for solving SAT problem [J]. Chinese Science (series E), 1997 (2): 179-186
- [5] Braunstein A, Mezard M, Zecchina R. Survey propagation: an algorithm for satisfiability[J]. Random Structures and Algorithms, 2005, 27(2): 201-226
- [6] Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations[J]. Journal of the ACM, 2007,54(4):1089-1098
- [7] Weiss Y. Correctness of local probability propagation in graphical models with loops[J]. Neural Computation, 2000,12(1):1-41
- [8] Weiss Y, Freeman WT. Correctness of belief propagation in Gaussian graphical models of arbitrary topology[J]. Neural Computation, 2001, 13(10):2173-2200
- [9] Tatikonda S, Jordan M I. Loopy belief propagation and Gibbs measure[C]. In Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence,2002,493-500.
- [10] Heskes T. On the uniqueness of loopy belief propagation fixed points[J]. Neural Computation, 2004, 16:2379-2413
- [11] Ihler A T. Loopy belief propagation: Convergence and effects of message errors[J]. Machine Learning Research, 2005,6:905-936.
- [12] Mooij J M, Kappen H J. Sufficient conditions for convergence of the sum-product algorithm[J]. IEEE Transactions on Information Theory, 2007,53:4422- 4437
- [13] Shi X Q, Schonfeld D, Tuninetti D. Message error analysis of loopy belief propagation for the sum-product algorithm[J]. Computer and Information Science, 2010, 1009:1-30
- [14] Gamarnik D, Shah D, Wei Y. Belief propagation for min-cost network flow: Convergence and correctness[J]. Operations Research, 2012,60(2): 410-428.
- [15] Brunsch T, Cornelissen K, Manthey B, Roglin H. Smoothed analysis of BP for minimum cost flow and matching[J]. Journal of Graph Algorithms and Applications, 2013,17(6):647-670.
- [16] Norshams N, Wainwright M.J. Stochastic belief propagation: A low complexity alternative to the sum product algorithm[J]. IEEE Transactions on Information Theory, 2013,59(4):1981-2000

- [17] Feige U, Mossel E, Vilenchik D. Complete convergence of message passing algorithms for some satisfiability problems[J]. Theory of computing, 2013,9(19):617-651.
- [18] XF Wang,XU Dao Yun,L Wei. Convergence of warning propagation algorithms for random satisfiable instances[J]. Journal of Software, 2013, 24(24):1-11.
- [19] Zhang Wenxiu, Liang Yi. Mathematical foundation of genetic algorithms [M]. Xi'an, Xi'an Jiao Tong University press, 2005.5