Research on the Solution Space of 2-SAT and Max-2-SAT

Bai-Feng Li[1,a], Wei -Wei[2] and Chao-Qun Liu[3]

[1] *Beihang University, Beijing, China*

Abstract. We study the properties of 2-SAT and Max-2-SAT problems by analyzing the node adding process on the factor graph. Two important structures, backbones and mutual-determinations are investigated, and the reduced solution graph for the expression of solution space of 2-SAT and Max-2-SAT is defined. For 2-SAT problem, a complete evolution process for the reduced graph is discussed and corresponding algorithm is obtained. For the Max-2-SAT problem, the analysis shows it's backbone number can evolve in a much harder way by which it can increase or decrease. The research in this paper provide a new view point for understanding the solution space of 2-SAT and Max-2-SAT, which will be benefit for recognizing the complexity nature of the NP-hard problems.

1 Introduction

As the first NPC problem (Cook, 1971), K-SAT is concerned by mathematicians, physicists and computer scientists [1]. It has a wide variety of applications such as consistency in expert systems knowledge bases ([2], [3]), integrity constraints in databases ([4], [5]).

A K-SAT formula contains N Boolean variables and M clauses, each of which is disjunction of K literals (variables or its negation). An assignment (to the N variables) forms by fixation of each Boolean variable. If there is at least one assignment that makes all clauses be true, the K-SAT formula is satisfiable (SAT); otherwise it's unsatisfiable (UNSAT) [6].

For the UNSAT instances, a natural question, of great practical importance, arises, i.e., how close can one get to satisfiability? Making that question precise leads to the Maximum Satisfiability problem ([7], [15]). This problem can also be expressed as an optimization problem Max-K-SAT and then takes the form: given a formula of M clauses involving N logical variables x1, x2, … , xn, determine an assignment for the formula that satisfies the maximum number of clauses.

In the literature of computational complexity [13], Max-2-SAT is NP-hard [8] and even hard to approximate [9], which makes a strong contrast with 2-SAT which can be solvable in linear time [10]. At present, there are some methods to detect the structure of solution space. Like spin glass approach, cavity method, phase transition ([11], [12], [14]). In [17], a new method is proposed to detect the whole solution space of the minimum vertex cover problem, and we want to generalize the corresponding analysis to 2-SAT and Max-2-SAT problem.

Following [17], in this paper, we focus on the solution space of Max-2-SAT. Backbones and mutual determination are discussed to understand the structure of solution space of 2-SAT and Max-2-sat problems. By the reduced solution graph, we provide a complete analysis of the solution space evolution for the 2-SAT problem and a corresponding algorithm is given, by which we can detect the correlation of the nodes in the solution space and a clear prospect of the solution organization. Furthermore, some analysis is provided for the Max-2-SAT instances, which shows the complicated evolution of it's solution space and results from the NP-hard intricacy. The analysis in this paper illustrates that the solution space for 2-SAT can be expressed by a simple graph (the reduced solution graph) and it may not be the case for the Max-2-SAT problem.

2 Model Description

In this paper, we mainly use the technique of the nodes-states evolution on the factor graph of the satisfiability problem [12].

2.1 Factor graph

We investigate the solution space of Max-2-SAT problem by constructing a factor graph: each variable node xi represents a Boolean variable, and each pair of linked nodes represents a clause (constraint) [12]. On the graph, there are three types of connection — dashed to dashed, dashed to dotted, dotted to dotted: dashed lines mean that variables appear as their negations in the corresponding clauses, and dotted lines mean themselves in clauses, like that in Figure 1.

[a] libaifeng@buaa.edu.cn

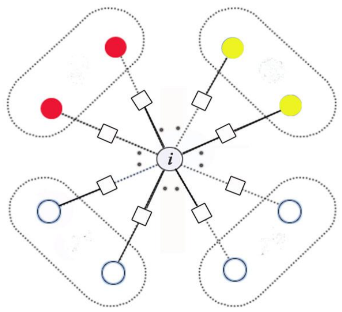

Figure 1. Circle Node Means Variable and Square node Means clause. Dash Line Means that Variables Appears in Its Connected Clause as A Negation and Dotted Line Means Itself.

2.2 Two basic structures

For each node's states of Max-2-SAT on its solution space, there are three types of possible values: x=1 backbones, x=0 backbones or free nodes, which always take value of x=1, x=0 or can take any one in the solution space. Similar as the analysis in the minimum vertex-Cover problem, we mainly concern with two structures on the solution space: backbones and mutual-determinations [17].

For $(x \vee y) \wedge (\neg x \vee y)=1$, to get the solution of this formula, we need to solve $(x \vee y)=1$ and $(\neg x \vee y)=1$ respectively. $(x \vee y)=1$can be expressed by the form of $(\neg x \rightarrow y)$ that means $\neg x=1$ required $y=1$ and $(\neg y \rightarrow x)$. $(\neg x \vee y)=1$ can be expressed by the form of $(x \rightarrow y)$ that means $x=1, y=1$ and $(y \rightarrow x)$. Then, if $(x \rightarrow y)$ and $(y \rightarrow z)$, we can have $(x \rightarrow z)$ and this process is named as logical deduction.

Backbones: In 2-SAT problem, backbones are a set of variables which occupy fixation in every satisfying truth assignment, i.e., if one variable x and its negation ¬x are both connected to another variable y by the logical deduction, we have $(\neg x \rightarrow y)$ and $(x \rightarrow y) \Rightarrow$ y is backbone; especially, $(\neg x \rightarrow x)$ and $(x \rightarrow x) \Rightarrow$ x is backbone [16].

Proof: In $(\neg x \rightarrow y)$, if $x=0 \Rightarrow y=1$; in $(x \rightarrow y)$, if $x=1 \Rightarrow y=1$, so no matter what x values, y is fixed. In a similar way, $(\neg x \rightarrow x)$ and $(x \rightarrow x) \Rightarrow$ x is fixed. These fixed variables are backbones.

For the instance in Figure 2, x1 =x2=x3=x4=1, and x1 =x2=x3=0, x4=1 are two solutions. According to the definition of backbones, x4 is a backbone.

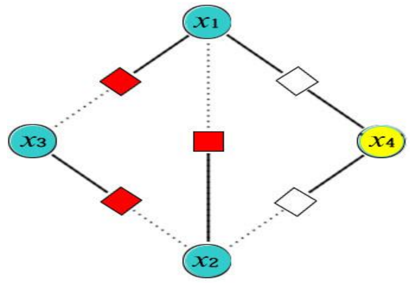

Figure 2. The Instance of Backbones

Mutual-determination: In 2-SAT problem, we named mutual-determination as an interactive relation of unfrozen variables in strong component in the viewpoint of logical deduction, i.e., a group of nodes are in a mutual-determination if and only if one node in them taking any value can force the others to be frozen in the solution space [7].

Proof: If the nodes form a mutual-determination, the fixation of anyone will result in the fixation of the others in strong component. Indeed, it is a special relation implied by the constraints that unfrozen variables can be mutually determined by the others. If some unfrozen variables $x \rightarrow y \rightarrow z \rightarrow x$ form a strong component, $x \rightarrow y \rightarrow z \rightarrow x \Rightarrow (\neg x \vee y), (\neg y \vee z)$ and $(\neg z \vee x)$, if $x=1 \Rightarrow y=1, z=1$.

Conversely, for $\neg x \leftarrow \neg y \leftarrow \neg z \leftarrow \neg x$, if $x=0 \Rightarrow z=0$, y=0. So we call x, y, z mutually determined.

For the instance in Figure 2, x1, x2, x3 are mutually determined.

2.3 Reduced solution graph for Max-2-SAT

Similar as the organization of solution space of Vertex-cover [17], we define the reduced solution graph of Max-2-SAT, which is mainly based on the two important structures backbones and mutual-determination and the evolution of them. The reduced graph is based on the factor graph with the x=1 backbones yellow circles, x=0 backbones red circles, free nodes blue circles and the mutual-determination red squares.

In order to detect the structure of solution space, directly applying the analysis of [12] and [17], we add one new node (called i) to previous N-1 system (N-1 nodes), new node connects to other k nodes x1, x2, … , xk. We suppose k1 nodes free, k2 nodes frozen to 1, k3 nodes frozen to 0 in previous system, by which k1 + k2 +k3=k.

When xi =1, all the clauses connected to xi with dotted lines are already satisfied and the variables connected to it are free. All the clauses connected to it with dashed lines must be satisfied by the variable connected to it to keep the minimum energy.

When xi =0, it is similar to the above case: all the clauses connected to xi with dashed lines are already satisfied and the variables connected to it are free. All the clauses connected to it with dotted lines must be satisfied by the variable connected to it to keep the minimum energy.

3 Solution Space Expression On 2-Sat

In this section, we will provide a complete process to obtain the reduced solution graph of 2-SAT problem, i.e., the whole solution space can be expressed by a graph. Considering an N-1 system and adding a new node i, we detect the solution space of new system. According to the types of nodes that connected to i, we have three cases below (here we only consider the instances in the satisfiable phase):

Case 1: If the nodes connected to i are only free in the N-1 system, no matter how i values, the whole system is satisfied. Now we want to detect the relationship between node i and the free nodes.

Subcase 1: If the nodes connected to i are free and no free neighbor can influence other free neighbors, no matter what value the node i takes, the new added clauses can be satisfied by its free neighbors. Then, in this case node i is also free in new system. Here, influence means that one free neighbor taking some value forces some other free neighbors to be fixed, and the mutual-determination relation is not included.

Subcase 2: If the nodes connected to i are free and some free neighbors can influence other free neighbors (no mutual-determination), the relationship of new node and the former free nodes may be backbone, mutual-determination or free.

Eg.1

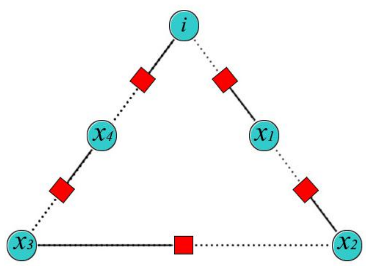

Figure 3. Eg.1

Analysis: In Eg.1, the former problem without node i have four free nodes, (x1 free, x2=x3=x4=1) and (x1 =x2=x3=0, x4 free) are all solutions of the problem. Some of them are influence others, but they are not mutual-determination. When add node i, we find the new node i and former free nodes form mutual-determination. Node i is free node in new problem.x1=x2=x3=x4=i=1 and x1 =x2=x3=x4=i=0 are all solutions.

Subcase 3: If the nodes connected to i are free and there are mutual-determinations in the neighborhood, as they can influence each other, the new node i may be backbone, mutual-determination or free in new system.

Eg.2

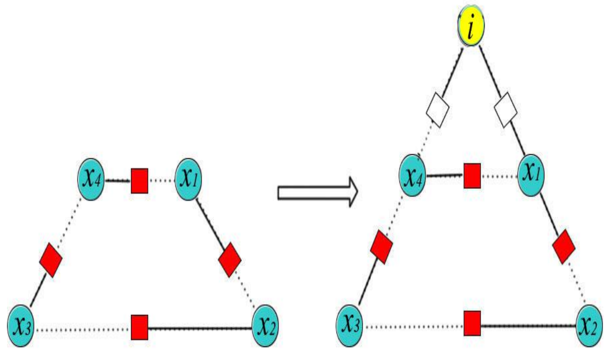

Figure 4. Eg.2

Analysis: In Eg.2, the former problem have four free nodes, x1 =x2=x3=x4=1 and x1 =x2=x3=x4=0 are all solutions of the problem. They are already mutual-determination. When add node i, we find that free nodes x1 and x4 can't satisfy the clauses they connected. So node i must be backbone valued 1 to satisfy the clauses.

Case 2: If the nodes connected to i are only backbones in N-1 system and the backbones have already satisfy the new clauses, the new added node i should be free node.

Case 3: If the nodes connected to i are only backbones in N-1 system and there exists backbones which don't satisfy the new clauses, the new added node i should be backbone to satisfy the new added clauses.

Eg.3

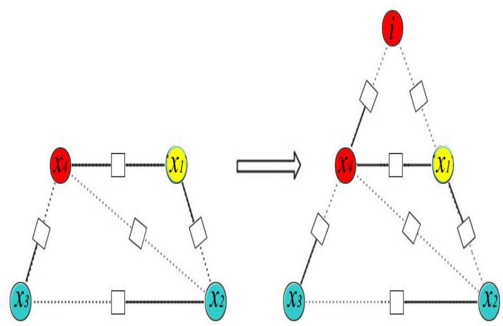

Figure 5. Eg.3

Analysis: In Eg.3, the former problem is satisfied with x1 =1 is backbone, x4=0 is backbone, x2 and x3 are all free. When add node i, we find that backbones nodes x1 and x4 are can't satisfy the clauses they connected. So node i must be backbone valued 0 to satisfy the clauses.

In a node adding process, all the above environment can be met with and the comprehensive analysis can be done using the above results. The new added node can be one backbone only when it faces Subcase 1-2, 1-3 and Case 3; the new added node can belong to some mutual-determination only when it faces Subcase 1-2, 1-3; and in the rest cases, the new added node should be a free node. In the following, the whole process for determining the reduced solution graph of 2-SAT is stated as an algorithm:

The Solution Space Expression Algorithm for 2-SAT

INPUT: Satisfiable 2-SAT Factor Graph *G*

OUTPUT: The Reduced Solution Graph $S(G)$ of 2-SAT

```
Freezing (S(G), j)
begin
  for free neighbors l of j
    if state of j doesn't satisfy the clause (l, j)
      S(G)=Freezing (S(G), l)
    end
  end
 return(S(G))
end
main ( )
begin
initial S(G)=Ø
for i=1: N
  add node i with the edges to S(G)
  use analysis in Case 1-3 to determine state of i
  Nb ={neighbors of i change from free to backbones}
  for j ∈Nb
    S(G)=Freezing (S(G), j)
  end
end
end
```

4 Solution Space Expression On MAX-2-Sat

In this section, we will provide some simple discussion on the Max-2-SAT problem (NP hard field), we can get the solution space change with the increase of variables and the corresponding algorithm, but the details will be quite complicated and will be discussed in our future work.

By the analysis in the above section, the solution space can be exactly expressed when there is no energy increase, so we only concern on the energy increase cases as the increase of variables.

In the 2-SAT problem, the number of backbones is monotonic increasing during the node-adding process and the number of free variables is monotone decreasing, but free variables are not always mutual-determination, which is quite different with that of minimum vertex cover problem. In the Max-2-SAT problem, it is not the case and the backbones' number can increase or decrease during the node-adding process, evenly the solution space can't be expressed by one reduced solution graph and evolution of the solution space will be rather unpredictable and intricate.

Eg.4

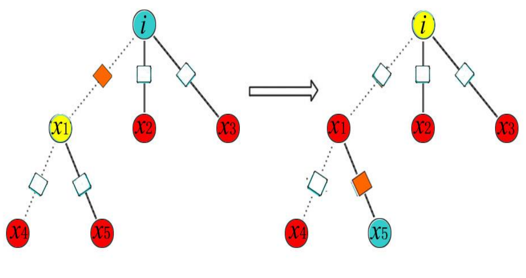

Figure 6. Unsatisfiable Situation

Analysis: In Eg.4, x1=1 and x2=x3=x4=x5=0 are backbones in the original problem, when we add a new node i, an unsatisfiable clause occurs. This causes energy increase, so node i should be xi=1 backbone and the clause of (i, 1) is unsatisfied. In order to satisfy the clauses added by node i, we can alternatively change the state of x1 to be x1=0, and the clause (i, 1) is satisfied but the clause (1, 5) is not.

5 Conclusion

Given a 2-sat formula, the objective is to judge whether it is satisfiable or not, and if it is, to find a satisfying solution. However, not all suitable results can be found for 2-sat formulas. For this case, we find the truth assignment that satisfies the maximum possible number of its clauses as solution [5]. We call it Max-2-sat problems.

Based on the long range frustration of nodes in Max-2-sat problem, backbones and mutual determination are proposed to study the solution space. Backbones are the fixed nodes in graph. It is the basic framework of the solution. Mutual determination shows how the nodes influence each other.

To study the correlation among the nodes, we add a new node to the former graph. By changing the state of some nodes to see the variation of solution, we find that when we change the state of some nodes, if the energy increase of the system less than the number of changed nodes, we take the new assignment for solution.

When add new node to the previous system, we know that may lead to unsatisfied of all system. At this moment, the energy of all system is increase.

References

[1] S. A. Cook, "The complexity of theorem-proving procedures," Proc. Third Annual ACM Sympos. Theory of Computing (1971), pp. 151–158.
[2] Kleer J D. Building expert systems: F. Hayes-Roth, D. A. Waterman and D. B. Lenat (Addison-Wesley Reading, MA, 1983); 444 pages, $32.50[J]. Artificial Intelligence, 1985, 25(1):105-107.
[3] Nguyen T A, Perkins W A, Laffey T J, et al. Checking an Expert Systems Knowledge Base for Consistency and Completeness.[C]// International Joint Conference on Artificial Intelligence. 1985:375-378.
[4] Asirelli P, Santis M D, Martelli M. Integrity constraints in logic databases[J]. Journal of Logic Programming, 1985, 2(3):221-232.
[5] Gallaire H, Minker J, Nicolas J M. Logic and Databases: A Deductive Approach[J]. Readings in Artificial Intelligence & Databases, 1989, 16(2):231-247.
[6] Scott Kirkpatrick, Bart Selman, Critical Behavior in the Satisfiability of Random Boolean Expressions, Science, Vol. 264, p. 1297-1301 (1994)
[7] Karp, R. M., Reducibility among combinatorial problems, in Complexity of Computer Computations,

R. E. Miller and J. W. Thatcher (eds.), Plenum Press, New York, 1972, 85-104.
[8] M.R. Garey, D.S. Johnson, L. Stockmeyer, Some simplified NP complete graph problems, Theoret. Comput. Sci. 1 (3) (1976) 237–267.
[9] J. Håstad, Some optimal inapproximability results, J. ACM 48 (4) (2001) 798–859.
[10] B. Aspvall, M.F. Plass, R.E. Tarjan, A linear-time algorithm for testing the truth of certain quantified Boolean formulas, Inform. Process. Lett. 8 (3) (1979) 121–123.
[11] Zecchina R. Analytic and Algorithmic Solution of Random Satisfiability Problem [J]. 2002.
[12] Zhou H. Long Range Frustration in Finite--Connectivity Spin Glasses: Application to the random K-satisfiability problem [J]. 2004.
[13] Rémi Monasson, Riccardo Zecchina, Scott Kirkpatrick, Bart Selman& Lidror Troyansky, Determining computational complexity from characteristic 'phase transitions',Nature,Vol.400,p.133-137(1999)
[14] William Wells Adams, Philippe Loustaunau, An Introduction to GrÄobner Bases. [M], Amer.Math.Society (1994).
[15] Coppersmith D, Gamarnik D, Hajiaghayi M and Sorkin G B 2004 Random MAX SAT, random MAX CUT, and their phase transitions Random. Struct. Algor. 24 4 502-545.
[16] Kilby P, Slaney J, Thiébaux S, et al. Backbones and Backdoors in Satisfiability.[C]// Conference on. 2005:1368-1373.
[17] Wei W, Zhang R, Guo B, et al. Detecting the solution space of vertex cover by mutual determinations and backbones. [J]. Physical Review E Statistical Physics Plasmas Fluids & Related Interdisciplinary Topics, 2012, 86(2):1-22.