

ランダムに生成された和積形論理式が充足不能となるしきい値について

坊野 博典 岩間 一雄

Hironori BOUNO Kazuo IWAMA

京都大学工学研究科情報工学専攻

Department of Information Science, Graduate School of Engineering, Kyoto University

1 はじめに

ϕ をランダムに生成された 3-CNF 式とする. 変数の数を n , 節の数を m とするとき m/n を節変数比と呼び, 本稿では r で表す. n を十分大きくしたとき $r = 4.2 \sim 4.3$ の付近で ϕ の充足可能確率がほぼ 100% 充足可能からほぼ 100% 充足不能へと急激に変化することが主に AI 研究者の実験によって広く確かめられている. このように充足可能性が節変数比の値を (存在すれば) しきい値と呼び κ で表す.

急激に変化する κ の値が 4.2 付近に存在することは多くの実験結果が一致しており疑いの余地はない. しかし, それを理論的に検証することは多くの研究者が試みているにもかかわらずあまり成功していない [1, 2, 3, 4, 5]. 最近 Kirousis らは極大選択法と呼ばれる方法により κ の上限が 4.601 であることを示した [6]. 本稿では BDT 法と呼ばれる方法を提案し, 別の方向からの改良を試みる.

2 基本的な考え方

κ の値を理論的に求める基本的な考え方は以下の様なものである. n 変数の CNF 式 ϕ の各項は 2^n 個の真理値割当てを一部を“覆う”ことによってその割当てに対する ϕ の値を 0 にする. このとき 1 つの割当てが 1 つの項によって覆われる確率は $(1/2)^3 = 1/8$ であり, 逆に覆われない (“生き残る”と呼ぶ) 確率は $7/8$ である. 結局 1 節増やすことによって各割当てが生き残る期待値は $7/8$ 倍に減少する. よって m 項をランダムに生成したときの生き残っている割当ての期待値は $2^n (7/8)^m$ である. この値が 0 に収束するときの $m/n = r$ が κ の上限になることは Malkov の不等式から説明できる.

しかし, この方法 (真理値法と呼ばれる) によって求めた κ の上限は約 5.191 であり実験値からはかけ離れている. その理由は以下のように説明できる [6]. ランダム式 ϕ が充足可能となる確率は

$$\sum_{\phi} \Pr[\phi] \cdot I_{\phi}$$

で表わされる. ここで $\Pr[\phi]$ は ϕ が生成される確率 I_{ϕ} は $I_{\phi} = 1$ (ϕ が充足可能), $I_{\phi} = 0$ (ϕ が充足不能) で定義される. ところで, ϕ について生き残る割当ての個数を S_{ϕ} としたとき, 明らかに $I_{\phi} \leq S_{\phi}$ である. よって ϕ が充足可能となる確率は

$$\sum_{\phi} \Pr[\phi] \cdot S_{\phi}$$

より小さいことがわかる. この式は ϕ の期待値そのものであり, 先の $\kappa \leq 5.191$ の議論はこの S_{ϕ} の期待値が 0 に収束するときの m/n の値を求めたことになる.

$I_\phi \leq S_\phi$ であるが ϕ によっては $I_\phi \ll S_\phi$ となることは明らかであり、それが上限値を高くしている大きな原因である。そこで S_ϕ に代わる何か ($S_\phi^\#$ とおく) を発見することが先の上限の改良につながると考えられる。ただし $S_\phi^\#$ は $I_\phi \leq S_\phi^\# \leq S_\phi$ を満たす必要がある。Kirousis らは [6] でこの $S_\phi^\#$ として生き残っている変数割当ての中で“極大”となっているものの数を採用し、その値を評価している。この方法を極大選択法と呼ぶ。割当てが極大であるとは ϕ を充足する割当てのうち 0 であるような任意の変数値を 1 に変化させた割当ては ϕ を充足させなくなるような割当てのことである。

本論文では割当ての数ではなく二分決定木の 1 の葉に至る経路の数を評価する。この方法を BDT 法と呼ぶ。二分決定木とは最初は 1 個の葉節点から始まり、節が加わることによって以下のように変化する。今 i 項のときの二分決定木を T_i とすると節が 1 つ加わることにより T_i の各経路は

1. 消滅する (葉の値が 0 に変化)
2. 変化しない
3. 高さが 1 増加する
4. 高さが +1, +2 の 2 本の経路に分裂
5. 高さが +1, +2, +3 の 3 本の経路に分裂

の様な変化を起こす。したがって高さが n に達していない 1 本の経路は数多くの割当てに相当している。このような経路の数は先の $S_\phi^\#$ の条件を満たしている。

本論文では以下で述べるように $S_\phi^\#$ を上記のような経路の数としたとき、その $S_\phi^\#$ の期待値が 0 に収束する r の値として $r < 5.040$ を得た。さらに BDT 法と極大選択法を併用することにより $r < 4.472$ を得た。

3 定義

二分決定木とは二分木による論理関数の表現法である。二分決定木の根節点を除く各節点の入次数は 1 である。二分決定木は各節点の出次数は 0 または 2 である。出次数 2 の節点は変数 x_i をラベルに持ち、変数節点と呼ばれる。出次数 0 の節点は定数 0, 1 をラベルに持ち、定数節点と呼ばれる。二分決定木の変数節点から出ている 2 本の枝にはラベル 0, 1 が付けられており、それぞれ 0-枝、1-枝と呼ばれる。二分決定木 T と入力 x が与えられたとき、その出力は T の根節点から各変数節点の入力と同じラベルを持つ枝を進んでいったときの定数節点の値となる。二分決定木の任意の経路に対して各変数をラベルに持つ節点は同じ順序で必ず一回ずつ出現する。

ある二分決定木 T の根節点から 1 のラベルを持つ節点までの経路を 1-path と定義する。任意の 1-path p に対して 1 のラベルを持つ節点までの変数節点の個数を p の長さとして定義し $len(p)$ と書く。以下では 1-path p の表現として $\widehat{x_{i_1}} \cdot \widehat{x_{i_2}} \cdots \widehat{x_{i_l}}$ を採用する。ここで $\widehat{x_{i_k}}$ は p が x_{i_k} をラベルとして持つ節点において 0-枝 (1-枝) を選択したとき $\overline{x_{i_k}}$ (x_{i_k}) とする。ある論理式 ϕ に対してそれを表現する二分決定木を $BDT(\phi)$ と表す。2つの二分決定木 T_1, T_2 に対してその積を $T_1 \cdot T_2$ と表す。

ある n 変数 3-CNF 式 $\phi = C_1 \cdots C_m$ に対して $r = m/n$ を節変数比率と定義する。 $n = 50$ のとき節変数比率とランダムに生成された 3-CNF 式が充足可能となる確率の関係は実験的に図 1 のようになる。よって n 変数 3-CNF 式 $\phi = C_1 \cdots C_m$ に対して実験的に以下のような Conjecture が成立すると考えられている。

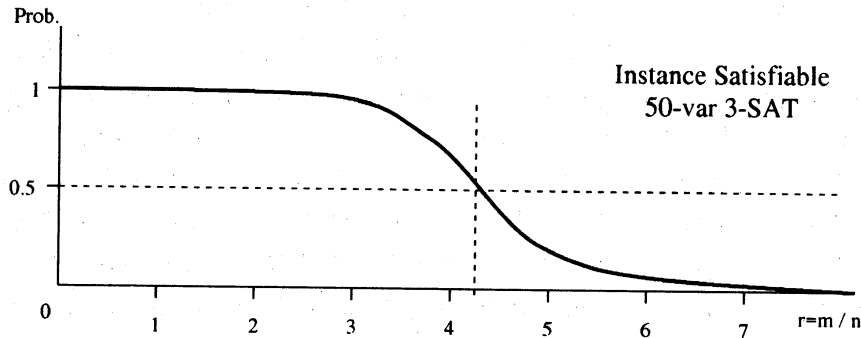


図 1: 節変数比率とランダムに生成された 3-CNF 式が充足可能となる確率の関係

Conjecture 3.1. $n \rightarrow \infty$ となるとき以下の条件を満たすしきい値 κ が存在する.

$$\lim_{n \rightarrow \infty} \Pr[\phi \text{ が充足可能}] = \begin{cases} 1 & m/n < \kappa \text{ のとき} \\ 0 & m/n > \kappa \text{ のとき} \end{cases}$$

このような κ を Unsatisfiability Threshold と呼ぶ. 実験的に κ は $\kappa \approx 4.24$ であることが知られている.

4 BDT 法

4.1 概要

ある n 変数 CNF 式 ϕ に対して ϕ を表現する二分決定木の 1-path の集合を S_ϕ^p とおく. また ϕ が充足可能ならば 1, そうでなければ 0 を返す関数を I_ϕ とおく. このとき明らかに以下の不等式が成立する.

$$\begin{aligned} I_\phi &\leq |S_\phi^p| && \phi \text{ が充足可能} \\ I_\phi &= |S_\phi^p| = 0 && \phi \text{ が充足不能} \end{aligned}$$

よって $|S_\phi^p|$ の期待値を $\mathbf{E}[|S_\phi^p|]$ とおくと, 以下のように Markov の不等式が成立する.

$$\begin{aligned} \Pr[\text{ランダムに選ばれた論理式が充足可能}] &= \sum_{\phi} \Pr[\phi] \cdot I_\phi \\ &\leq \sum_{\phi} \Pr[\phi] \cdot |S_\phi^p| \\ &\leq \mathbf{E}[|S_\phi^p|] \end{aligned}$$

ここで $\mathbf{E}[|S_\phi^p|] = 0$ に収束するときの節変数比率を r' とおくと $r > r'$ を満たす任意の節変数比率 r に対して

$$\Pr[\text{ランダムに選ばれた論理式が充足可能}] = 0$$

が成立する. よってこのような r' は Unsatisfiability Threshold κ の上限を与えていると考えられる.

4.2 解析

ある n 変数 m 節の CNF 式 $\phi_m = C_1 \cdots C_m$ に対してそれを表現する二分決定木 $BDT(\phi)$ を以下のように再帰的に構成することを考える. ただし $\phi_i = C_1 \cdots C_i$ とする.

1. $i = 0$ のとき, $BDT(\phi_0) = BDT(1)$.
2. $i > 0$ のとき, $BDT(\phi_i) = BDT(\phi_{i-1}) \cdot BDT(C_{i-1})$.

S_ϕ^p の中に含まれる経路のうち長さが l のものの集合を $N_\phi(l, i)$ とおき, 以下のように $\mathbf{N}_\phi(i)$, $\|\mathbf{N}_\phi(i)\|$ をおく.

$$\mathbf{N}_\phi(i) = (|N_\phi(0, i)|, \dots, |N_\phi(n, i)|)^T$$

$$\|\mathbf{N}_\phi(i)\| = \sum_{l=0}^n N_\phi(l, i).$$

このとき, 各 $\mathbf{N}_\phi(i)$ は以下のように表現できる.

場合 (1) $i = 0$ の場合. $BDT(\phi_0) = BDT(1)$ であることから

$$\mathbf{N}_\phi(0) = (1, 0, 0, \dots, 0).$$

場合 (2) $i > 0$ の場合. $\mathbf{N}_\phi(i)$ を $\mathbf{N}_\phi(i-1)$ を用いて表現する. $C_i = (u + v + w)$, $p \in S_{\phi_{i-1}}^p$, $l = \text{len}(p)$ とおくと以下の 5 つの場合が考えられる.

場合 (a) p は u, v, w を全て含んでいる. p は除去される (p の葉節点が 0 になる).
この場合の起こる確率は

$$r_3(l) = \frac{\binom{l}{3}}{8\binom{n}{3}} = \frac{1}{8} \left(\frac{l}{n}\right)^3.$$

場合 (b) p は u, v, w のうち 2 変数を含んでいる. p は v, w を含んでいると仮定する. p は $p_1 = p \cdot \bar{u}$, $p_2 = p \cdot u$ を表現する 2 つの 1-path に分解され, そのうち p_2 が除去される. この場合の起こる確率は

$$r_2(l) = \frac{2\binom{l}{2}\binom{n-l}{1}}{8\binom{n}{3}} = \frac{3}{4} \left(\frac{l}{n}\right)^2 \left(1 - \frac{l}{n}\right).$$

場合 (c) p は u, v, w のうち 1 変数を含んでいる. p は w を含んでいると仮定する. p は $p_1 = p \cdot \bar{u}$, $p_2 = p \cdot u \cdot \bar{v}$, $p_3 = p \cdot u \cdot v$ を表現する 3 つの 1-path に分解され, そのうち p_3 が除去される. この場合の起こる確率は

$$r_1(l) = \frac{4\binom{l}{2}\binom{n-l}{1}}{8\binom{n}{3}} = \frac{3}{2} \left(\frac{l}{n}\right) \left(1 - \frac{l}{n}\right)^2.$$

場合 (d) p は u, v, w を全て含んでいない. p は $p_1 = p \cdot \bar{u}$, $p_2 = p \cdot u \cdot \bar{v}$, $p_3 = p \cdot u \cdot v \cdot \bar{w}$, $p_4 = p \cdot u \cdot v \cdot w$ を表現する 4 つの 1-path に分解され, そのうち p_4 が除去される. この場合の起こる確率は

$$r_1(l) = \frac{8\binom{l}{2}\binom{n-l}{1}}{8\binom{n}{3}} = \left(1 - \frac{l}{n}\right)^3.$$

場合 (e) p と C_i は交わらない. p は変化しない. この場合の起こる確率は

$$\begin{aligned} r_i(l) &= 1 - (r_0(l) + r_1(l) + r_2(l) + r_3(l)) \\ &= \frac{7}{8} \left(\frac{l}{n} \right)^3 + \frac{9}{4} \left(\frac{l}{n} \right)^2 \left(1 - \frac{l}{n} \right) + \frac{3}{2} \left(\frac{l}{n} \right) \left(1 - \frac{l}{n} \right)^2. \end{aligned}$$

以上のことを表にまとめると表 1 のようになる. 表 1 から $\mathbf{N}_\phi(i)$ は $\mathbf{N}_\phi(i-1)$ を用いて以

出現変数	$ N_\phi(l, i) $ の変化量				確率
	l	$l+1$	$l+2$	$l+3$	
3	-1	0	0	0	$r_3(l)$
2	-1	+1	0	0	$r_2(l)$
1	-1	+1	+1	0	$r_1(l)$
0	-1	+1	+1	+1	$r_0(l)$
交わらない	0	0	0	0	$r_i(l)$

表 1: 出現変数と各 $N_\phi(l, i)$ の変化量の関係

下のように表現できる.

$$\mathbf{N}_\phi(i) = A_n \cdot \mathbf{N}_\phi(i-1),$$

$$A_n = (a_{ij}),$$

$$a_{ij} = \begin{cases} p_0(j) = r_i(j) & i = j \text{ のとき} \\ p_1(j) = r_0(j) + r_1(j) + r_2(j) & i = j + 1 \text{ のとき} \\ p_2(j) = r_0(j) + r_1(j) & i = j + 2 \text{ のとき} \\ p_3(j) = r_0(j) & i = j + 3 \text{ のとき} \\ 0 & \text{それ以外} \end{cases}$$

よって $\mathbf{E}[|S_\phi^p|]$ は以下のような式で表現できる.

$$\begin{aligned} \mathbf{E}[|S_\phi^p|] &= \sum_{\phi} \Pr[\phi] \cdot \|\mathbf{N}_\phi(m)\| \\ &= \|\mathbf{N}_\phi(m)\| \\ &= \|A^m \cdot \mathbf{N}_\phi(0)\| \end{aligned}$$

$n = 1000$ において $\mathbf{E}[|S_\phi^p|] = 0$ を満たす $r = m/n$ の数値解を求めると $r < 5.040$. よって κ の上限は $\kappa < 5.040$. また $n = 1000$ のときの $\ln \mathbf{E}[|S_\phi|] = (7/8)^r$ と $\ln \mathbf{E}[|S_\phi^p|]$ のグラフは図 2 のようになる.

5 BDT 法と極大選択法の併用

4 章で述べた BDT 法に極大選択法を適用する. ある 1-path $p = \widehat{x}_{i_1} \cdots \widehat{x}_{i_l}$ に対して $\widehat{x}_{i_k} = \overline{\widehat{x}_{i_k}}$ を満たすような変数 x_{i_k} のうちの 1 個を x_{i_k} に変更したパスを 1-flip されたパスと呼ぶ. このとき,

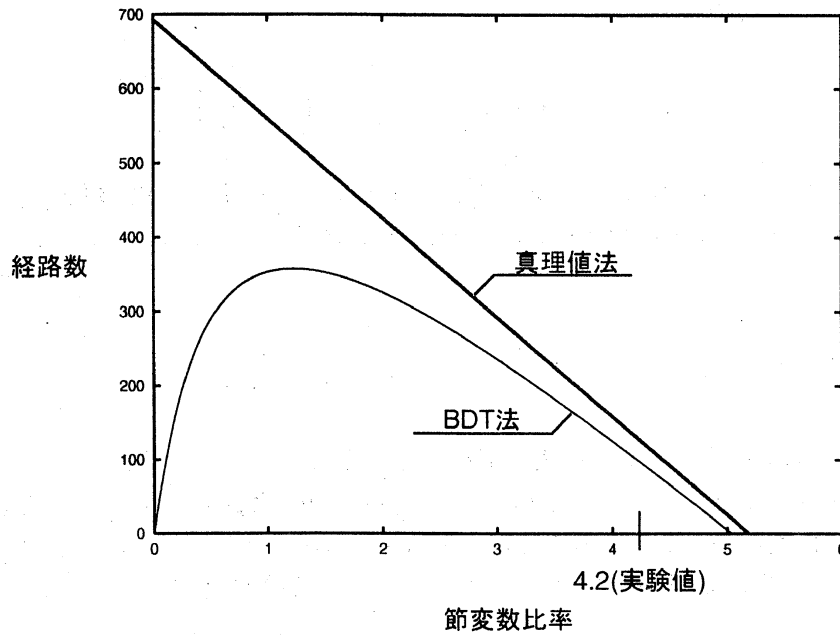


図 2: $\ln \mathbf{E}||S_\phi||$ と $\ln \mathbf{E}||S_\phi^p||$ のグラフ

図 3 のように太い実線で表された 1-path p に対して点線で表わされた 1-flip されたパス p' を考える. このとき p' は変数節点 x_4 で終了している (0-節点で終了していない) ので p は BDT 法と極大選択法を併用した方法では除去できる. しかし, 極大選択法のみでは p' が 1-節点で終了していないので除去できない. よってこのような場合においては極大選択法のみでは除去できない節点が除去可能となる.

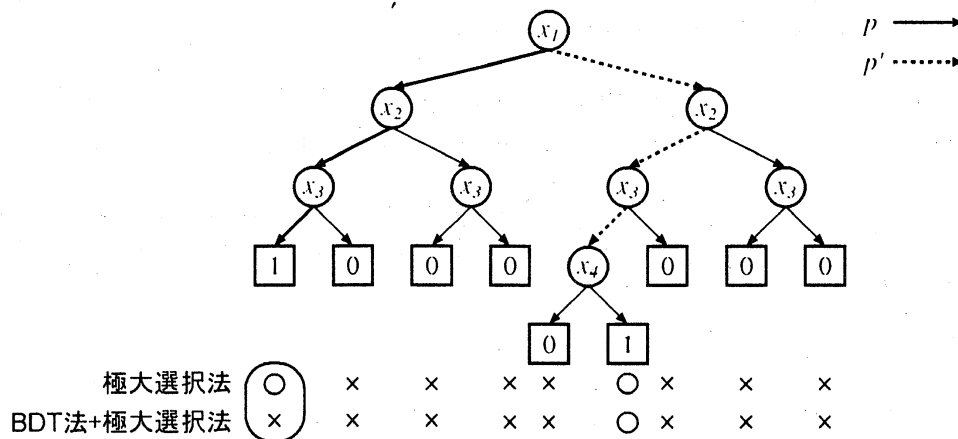


図 3: 極大選択法のみでは除去できないが BDT 法と併用することにより除去できる経路の例

S_ϕ^p に含まれる経路のうち極大であるものの集合を $S_\phi^{p\#}$. $N_\phi(l, i)$ に含まれる経路のうち極大であるものの集合を $N_\phi^\#(l, i)$ とおく. $l \geq 3$ のとき $|N_\phi^\#(l, i)|$ の期待値は [6] の場合と同様に考えること

により以下のように表現できる.

$$|N_{\phi}^{\sharp}(l, i)| = |N_{\phi}(l, i)| \cdot 2^{-l} \left(2 - \left(1 - \frac{3}{7l} \right)^m \right)^l.$$

4章の場合と同様に $n = 1000$ とおき, $\mathbf{E}[|S_{\phi}^{\sharp}|] = 0$ となる $r = m/n$ の数値解を求めると $r < 4.472$ となる. よって Unsatisfiability Threshold κ の上限は $\kappa < 4.472$. なお, これらの結果の詳細については別稿にて述べる.

6 まとめ

本稿では二分決定木を用いた BDT 法を解析することにより κ の上限として $\kappa < 5.040$ を示した. これは最も基本的な方法によって求められる上限 5.191 と比較して 0.151 程改善されている. また BDT 法と極大選択法を併用することにより κ の上限として $\kappa < 4.472$ を示した. これは極大選択法のみによって求められる上限 4.601 と比較して 0.129 程改善されている.

参考文献

- [1] N. Anon, J. H. Spencer, and P. Erdős, *The Probabilistic Method*, J. Wiley, New York, 1992
- [2] B. Bollobás, *Random Graphs*, Academic Press, London, 1985
- [3] V. Chvátal and B. Reed, "Mick gets some(the odds are on his side)," *Proc. 33rd IEEE Symposium on Foundation of Computer Science*, pp 620–627, 1992
- [4] S. Kirkpatrick and B. Selman, "Critical behaviour in the satisfiability of random Boolean expressions," *Science* 264, pp 1297–1031, 1994
- [5] J. Franco and M. Paull, "Probabilistic analysis of the Davis-Putnam procedure for solving the satisfiability problem," *Discrete Applied Mathematics* 5, pp 77–87, 1983
- [6] Leftris M. Kirousis, Evangelos Kranakis, Danny Krizanc, and Yannis C. Stamation, "Approximating the Unsatisfiability Threshold of Random Formulas," *Manuscript*, 1997