

SCSat3 によるラムゼーグラフ探索について

Search for Ramsey Graphs Using SCSat3

藤田 博*¹

Hiroshi Fujita

*¹九州大学大学院システム情報科学研究院

Graduate School of Information Science and Electrical Engineering, Kyushu University

We try to find Ramsey graphs using a SAT solver SCSat3. The sought-after graphs should have some inherent (though not necessarily well-defined) regularity properties. SCSat3 is able to do such a search efficiently with the help of soft constraints. Some effective soft constraints are given as geometric patterns in adjacency matrices of Ramsey graphs. Recent results include a *RamseyGraph*(4, 11, 100) that proves $R_{maseyNumber}(4, 11) \geq 101$.

1. はじめに

2012年, 著者は2色ラムゼー数 $R(4, 8)$ の最良下界を56から58に改善した [Fujita 12, Fujita 13a, Fujita 13b]. この記録更新に用いたのはSATソルバーSCSat(以降, SCSat1)である. その後, SCSat3に改訂し, 研究を継続している.

n 頂点の完全グラフ K_n を2色で辺彩色したとき, n が十分大きいと第1色の部分完全グラフ K_s , あるいは, 第2色の部分完全グラフ K_t が必ず含まれる (Ramseyの定理) が, そのような n の最小値のことをラムゼー数とよび, $RamseyNumber(s, t)$ (以降, $R(s, t)$) と表す. ラムゼー数未満の n については第1色の K_s も第2色の K_t も含まないような K_n が作れるはずで, これをラムゼーグラフとよび, $RamseyGraph(s, t, n)$ (以降, $R(s, t, n)$) と表す. $R(s, t, n)$ を具体的に示せば, $R(s, t)$ は $n+1$ 以上, すなわち下界が与えられる. ラムゼー数の下界改善に関しては, [Exoo, McKay] が詳しい. いずれも主として計算機利用の探索手法に基づく. [Radziszowski 14] では網羅的で最新のサーベイが随時更新されている. なお, 上界の改善に本稿のようなSATアプローチは不適であろう.

SCSat1もSCSat3もともに完全なSATソルバーである. すなわち, 充足解が存在する場合には必ず有限時間内にその1個を発見できる. SCSat1はソフト制約をハード制約と同格に扱って解探索を開始し, 求解失敗 (UNSAT) の際はソフト制約を緩和 (最終的には全削除) して再試行する. 一方, SCSat3はソフト制約を探索分岐 (decision, case splitting) の選好にのみ利用し, 確定的な推論 (implication, unit propagation) はハード節のみによって遂行し, SCSat1のようなソフト制約の緩和に伴う推論再試行は行わない.

2. ラムゼーグラフの規則性

ラムゼーグラフ $R(s, t, n)$ の存在は, 以下のCNF (Ramsey節集合とよび, $C(s, t, n)$ と表す) が充足可能であることを示せばよい.

$$\left(\bigwedge_{K_s \subset K_n} \bigvee_{e_{ij} \in K_s} \neg e_{ij} \right) \wedge \left(\bigwedge_{K_t \subset K_n} \bigvee_{e_{ij} \in K_t} e_{ij} \right)$$

連絡先: 藤田博, 九州大学大学院システム情報科学研究院,
〒819-0395 福岡市西区元岡744,
Tel:092-802-3609, Fax:092-802-3600,
fujita@inf.kyushu-u.ac.jp

ここで, e_{ij} は2頂点 i と j の間の辺の色を表す命題変数で, 真のとき第1色, 偽のとき第2色を表すものとする. Ramsey節集合 $C(s, t, n)$ の充足性判定には, MiniSat[Eén] のような高性能SATソルバーが適用できる.

しかし, Ramsey節集合をそのまま解くのは極めて難しい. 変数の個数は $\binom{n}{2}$, 節の本数は $\binom{n}{s} + \binom{n}{t}$ で, 例えば, $C(5, 5, 43)$ においては903変数, 1925196節となる. この程度ならば十分SATソルバーの適用範囲内ではあるが, 莫大な探索空間中に充足解の数は (存在したとしても) 極めて少ない.

2.1 Z制約

Ramsey節集合の規模は莫大でも形は極めて単純なので, 充足解にもその単純さが反映するであろうことが期待される. 実際, 以下のような制約 (Z制約とよぶ) を追加すると, 単純な解が容易に得られる場合がある.

$$e_{ij} \equiv z_k \quad (0 \leq j < i < n, i - j = k)$$

ここで, z_k ($1 \leq k \leq n-1$) は新たな命題変数 (Z変数とよぶ) である. 上式を以下のようにCNFで表したものをZ節集合とよび, Z^n と表す.

$$\bigwedge_{0 \leq j < i < n} ((\neg e_{ij} \vee z_{i-j}) \wedge (e_{ij} \vee \neg z_{i-j}))$$

実際には, MaxSAT問題で用いられる形式と同様のWCNF (重みつき主乗法標準形) で与える. 節集合 $C(s, t, n) \wedge Z^n$ が充足可能ならば, $C(s, t, n)$ も充足可能で, このときの充足解をZラムゼーグラフとよび, $R^Z(s, t, n)$ と表す. 独立命題変数は実質的に $n-1$ 個となり, 問題ははるかに易くなる.

ラムゼー数より1小さい頂点数のラムゼーグラフを臨界ラムゼーグラフという. 容易に求めることができるZラムゼーグラフ $R^Z(3, 3, 5)$, $R^Z(3, 4, 8)$, $R^Z(3, 5, 14)$, $R^Z(3, 9, 35)$, $R^Z(4, 4, 17)$, $R^Z(4, 5, 24)$ はすべて臨界ラムゼーグラフである. しかし, 臨界ラムゼーグラフ $R(3, 6, 17)$, $R(3, 7, 22)$, $R(3, 8, 28)$ はいずれもZラムゼーグラフとしては存在しないことが分かっている. さらに, 上記より大きなラムゼー数に関する臨界ラムゼーグラフで, Zラムゼーグラフのものは存在しないと予想される.

しかし, 完璧なZラムゼーグラフとしては得られない $R(5, 5, 42)$, $R(3, 12, 51)$, $R(4, 11, 100)$ などでも, Z制約の大部分を満たすものとしてなら得ることができる.

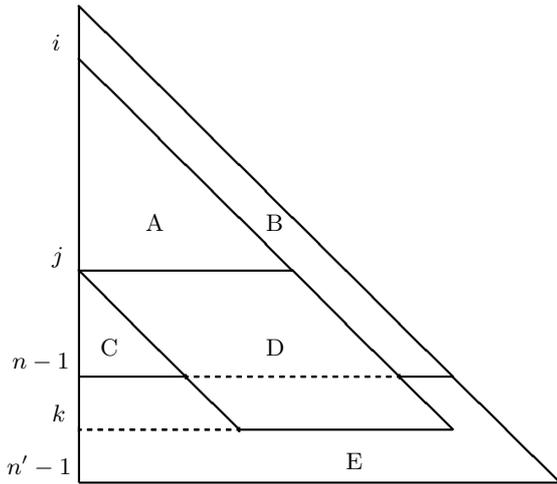


図 1: 隣接行列の拡張操作 $X_n^{n'}(i, j, k)$

1									
2	$-b_{11}$	$-b_{12}$							
3	$-b_9$	$-b_{10}$							
4	b_{11}	b_{12}	$-b_{11}$	$-b_{12}$					
5	b_9	b_{10}	$-b_9$	$-b_{10}$					
6	$-b_3$	$-b_4$	b_3	b_4	$-b_7$	$-b_8$			
7	$-b_1$	$-b_2$	b_1	b_2	$-b_5$	$-b_6$			
8	b_3	b_4	$-b_3$	$-b_4$	b_7	b_8	$-b_7$	$-b_8$	
9	b_1	b_2	$-b_1$	$-b_2$	b_5	b_6	$-b_5$	$-b_6$	
	0	1	2	3	4	5	6	7	8

図 2: $B^{10}(2, 2)$ 制約

2.2 ラムゼーグラフの拡張

ラムゼーグラフ $R(s, t, n)$ の真部分完全グラフ $K_{n'}$, ($n > n' \geq t$)^{*1} はすべてラムゼーグラフ $R(s, t, n')$ である。逆に、ラムゼーグラフ $R(s, t, n)$ をそのまま含むように拡張した完全グラフ $K_{n'}$, ($n' > n$) がラムゼーグラフ $R(s, t, n')$ になるとは限らない。

しかし、ラムゼーグラフ $R(s, t, n)$ を拡張して異なる対 (s, t) に対するラムゼーグラフ $R(s', t', n')$, ($s \leq s', t \leq t', n < n'$) を効果的に得られる場合がある。典型的には $s = s', t + 1 = t'$ の場合である。拡張の方法として図 1 に示す操作 $X_n^{n'}(i, j, k)$ を考える。図は隣接行列の下三角の領域のみを表している^{*2}。領域 A は $R(s, t, n)$ の彩色の通り、領域 B, C, E は未彩色、領域 D は領域 A の“自然な拡張”とする。

2.3 様々な規則性制約

Z 制約以外に、ラムゼーグラフに内在すると期待される何らかの対称性、周期性を表すような制約を考えてみる。ただし、いずれも数学的根拠の薄い“任意な制約”である。

- D 制約 (ダイヤ格子): 主対角線に並行なスリット状の同色ライン、および、副対角線に並行なスリット状の同色ラインを形成。

$$e_{ij} \equiv z_k, \quad (0 \leq j < i < n, i - j = k = 1, 3, \dots)$$

$$e_{ij} \equiv s_l, \quad (0 \leq j < i < n, i + j = l = 2, 4, \dots)$$

- I 制約 (市松): 4 辺ずつの組 (代表辺 e_{ij}) ごとに新変数 a_{ij} を導入し、以下を満たすこと。

$$e_{ij} \equiv e_{(i-1)(j+1)} \equiv a_{ij}, \quad e_{i(j+1)} \equiv e_{(i-1)j} \equiv \neg a_{ij},$$

$$(n > i > 2, 0 \leq j < n)$$

- $B(r, c)$ 制約 (ブロック市松): 隣接行列を r 行 c 列の小行列 O_{ij} に分割し、市松状の以下の制約を満たすこと。

$$O_{ij} \equiv O_{(i-1)(j+1)}, \quad O_{i(j+1)} \equiv O_{(i-1)j} \equiv \neg O_{ij},$$

$$(\lfloor n/r \rfloor > i > 2, 0 \leq j < \lfloor n/c \rfloor)$$

*1 一般に $R(s, t) = R(t, s)$ 故、 $s \leq t$ とする。

*2 扱うのは無向グラフであり、その隣接行列は対称である。また、主対角線上の要素も考えない。

例として、 $B^{10}(2, 2)$ 制約を図 2 に示す。

- S 制約 (副対角線対称): 副対角線を軸として対称であること。

$$e_{ij'} \equiv e_{j'i}, \quad (j' = n - j - 1,$$

$$0 \leq j' < i < n, 0 < i + j' < n - 1)$$

- C 制約 (サイクル): n 頂点 $\{v_0, v_1, \dots, v_{n-1}\}$ の完全グラフにおいて、頂点 v_i から始め、定数 d 飛びに頂点を巡ると、必ず v_i に回帰して閉路を成す。 $1 \leq d \leq \lfloor n/2 \rfloor$ で得られる異なる閉路 $\langle v_i, v_{(i+d) \bmod n}, v_{(i+2d) \bmod n}, \dots, v_i \rangle$ のそれぞれに以下のような制約のいずれかを課す。

- 同一閉路に属する辺たちを同色とする。
- 閉路長 m が合成数 $m = p \times q$ のとき、 p 個の辺を 1 グループとして q 個のグループに分割し、同一グループに属する辺を同色とする。

上記いずれかの単独ではなく、複数を合わせて適用することも考える。たとえば、I 制約と S 制約の組合せ $\langle I, S \rangle$ や Z 制約と S 制約の組合せ $\langle Z, S \rangle$ などが実際に効果的であることが分かっている。制約対において第 1 制約は第 2 制約より優先される。その優先度は、制約を WCNF 形式で与える際に、節への適切な重み付けによって表現される。

2.4 規則性制約の静的緩和

SCSat1 では、規則性制約が強すぎて本来の充足解を失うことがあり、緩和と推論やり直しが必須であった。SCSat3 においても、強すぎる規則性制約による推論誘導が不適切で逆効果となるような場合には緩和が必要である。

規則性制約を作成する時点で、たとえば図 3 に示すように、適用範囲を隣接行列全域から部分的に“間引く”だけでも効果的なことが分かっている。間引きも何らかの規則性に従うことにより、ラムゼーグラフに内在する真の規則性(あるいは非規則性?)の“透かし出し”に効果があるものと考えられる。

たとえば、ラムゼーグラフ $R(6, 7, 112)$ を求める際に、D 制約に対する間引き $M_{r(2,1),c(2,1)}^{112}$ を用いた。ここで、下付きの $r(2, 1), c(2, 1)$ は「2 行適用 1 行間引き, 2 列適用 1 列間引き」を表している。

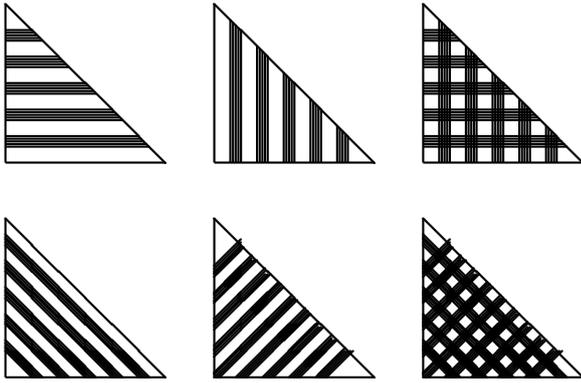


図 3: 静的緩和 M のパターン 6 例 (黒帯は制約非適用領域)

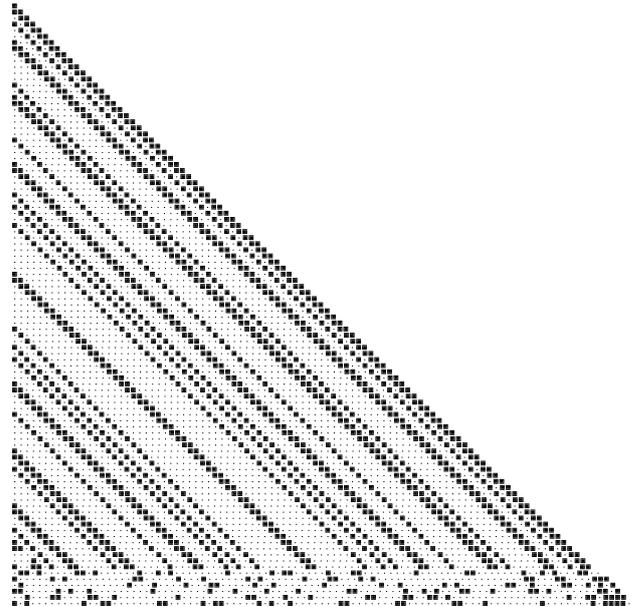


図 4: ラムゼーグラフ $R(4, 11, 100)$ の隣接行列

3. SCSat3

Z 節 (集合) のように, 所与の問題の探索空間を削減することが主目的で, それ自体の充足が必ずしも要請されないものをソフト節 (集合) とよぶ. ソフト節に対する反則点賦課と緩和等の機能を備えた SAT ソルバーを MiniSat[Eén] を基に開発し, SCSat とよんでいる [Fujita]. ソフト節集合の最大充足を目的とする MaxSAT ソルバーとは似て非なるものである. SCSat3 では, ソフト節の推論誘導としての用途がより鮮明となっているので, “誘導節 (集合)” (guiding clause(s)) の呼称が相応しいかもしれない.

ソフト節集合は, MaxSAT 問題と同形式の重み付き CNF (問題ファイルの拡張子は wcnf) として表される. 重みは正整数^{*3} である. 重みは当該節の嗜好の度合い (値が大きいほど優先される) を表す.

3.1 ソフト節による推論誘導

ソフト節が一般節 (MaxSAT におけるハード節) と如何に異なる取り扱いを受けるかについて述べる.

- あるソフト節の非充足が判明したとき, これをコンフリクトとは看做せず, 推論を継続する.
- あるソフト節において未割りテラルが 1 個のみ (単位節) であることが判明したとき, これをただちに真とは看做せず, バッファに登録する.
- 単位節伝播 (unit propagation) 処理が行き詰まり, 場合分け (decision) が必要となったとき, バッファ中に未割りテラルがあれば, 一つを選択して割当を行う. 無ければ, 通常通りの decision 処理を行う.

バッファは, MiniSat における trail と同様のデータ構造を有する. また, これに登録されるリテラルは, 決定水準 (decision level), および, 起源となったソフト節に付与された重みによって, 順位付け / 選択が可能になっている. リテラル選択に際し, 下記に示す観点のいずれか, あるいはそれらの組合せの中からいくつかのオプションが用意されており, SCSat3 の起動時に指定できる.

- 決定水準優先か重み優先か

- 決定水準の浅いもの優先か深いもの優先か
- 起源となったソフト節の反則点の小さいもの優先か

実験によれば, 「決定水準の浅いものを第 1 優先とし, 同一水準の場合はソフト節重みに拠る」のが一般的に良好である.

3.2 ソフト節の不活化 (規則性制約の動的緩和)

Minisat においては, 学習節が推論 (analyze 中) に関与するたびに活性度 (activity) を加点する. この活性度に基づき, 学習節集合が許容量を超えた時点などを契機として, その一定割合が削除される.

これに倣い, SCSat3 においては推論の進行に伴い不適合と看做されたソフト節を不活化することができる. そのため, ソフト節が非充足となるたびに適合度を減点する. 実際には反則点 (penalty) を加点する^{*4}. SCSat3 で用意されている penalty 算式のオプションのいくつかを以下に示す.

$$(\text{decisionLevel}() + 1)^{-1} \quad (1)$$

$$(\text{trail.size}() + 1)^{-1} \quad (2)$$

$$(\text{nVars}() - \text{trail.size}()) / \text{nVars}() \quad (3)$$

$$\exp(-\text{decisionLevel}()) \quad (4)$$

$$\exp(-\text{trail.size}()) \quad (5)$$

特に指定しない場合には, 実績の多い (3) 式を既定としている.

ソフト節自体がコンフリクト節となる時以外に, ソフト節から含意されたリテラルによる decision が失敗 (コンフリクト) したときにも反則点を課すことが考えられるが, その有無に拠る効果に特段の差異は認められない.

また, 制約緩和の際, 1 回あたりの不活化ソフト節集合の割合について, 反則点の分布等の統計を基に適切な値を求めるのが望ましいが, たとえば, 毎回 50% と固定しても概ね効果的である.

*3 SCSat3 の現実装においては, 1,2,3 に限られる.

*4 学習節は加点主義だが, ソフト節は減点主義がよい. ソフト節の加点主義的実装は, 実験では良好な結果が得られていない.

表 1: SCSat3 で得られたラムゼーグラフ

solution	hard clauses				soft clauses			runtime
	base sol.	extention op.	vars	cls	type	vars	cls	
$R(3, 7, 22)$	$R^Z(3, 6, 16)$	$X_{16}^{22}(1, 17, 22)$	112	13,168	-	0	0	0.016 s
$R(3, 8, 27)$	$R^Z(3, 7, 21)$	$X_{21}^{27}(0, 21, 24)$	81	4,887	-	0	0	0.0065 s
$R(3, 12, 51)$	$R^Z(3, 11, 45)$	$X_{45}^{51}(0, 41, 46)$	255	1,800 K	Z^{51}	50	510	31 s
$R(4, 6, 34)$	$C(4, 6, 34)$		561	1,391 K	$\langle I^{34}, S^{34} \rangle$	442	2,244	132.5 h
$R(4, 7, 48)$	$C(4, 7, 48)$		1128	73,823 K	$\langle Z^{48}, S^{48} \rangle$	599	4,464	36.1 h
$R(4, 8, 57)$	$R(4, 7, 48)$	$X_{48}^{57}(0, 48, 57)$	468	3,489 K	$B^{57}(4, 4)$	16	936	75 h
$R(4, 11, 100)$	$R^Z(4, 11, 91)$	$X_{91}^{100}(0, 89, 94)$	594	78,446 K	Z^{100}	99	1,188	102.8 h
$R(5, 5, 42)$	$C(5, 5, 42)$		861	1,701 K	Z^{42}	41	1,722	2 s
$R(6, 7, 112)$	$R^Z(6, 6, 101)$	$X_{101}^{112}(0, 100, 103)$	969	14,351 K	$D^{112} \cdot M_{r(2,1),c(2,1)}^{112}$	164	848	1,071 s

4. ラムゼーグラフ各論

これまでに SCSat3 によって得られたラムゼーグラフのうち、重要なもののいくつかを表 1 に示す。

$R(3, 7, 22)$, $R(3, 8, 27)$ は小ラムゼーグラフの拡張によれば、規則性制約に頼らずとも易しく得られる。新たな規則性制約を考案したとき、その効果をまずは小ラムゼーグラフ拡張なしでテストしてみる問題としても手頃である。

$R(5, 5, 42)$ は小ラムゼーグラフの拡張に頼らずとも、 Z 制約を適切に緩和すれば比較的易しく発見できる。続く $R(5, 5, 43)$ は様々な試みにも拘らず、発見の兆しはない (UNSAT にも至ってない)。

$R(4, 6, 34)$ の探索は長らく難航したが、2014 年 8 月漸く発見に成功した。狙い通り、 I 制約 (市松) が功を奏した。続く $R(4, 6, 35)$ の探索は依然として難航している。これには C 制約 (サイクル) が効果的のようで、発見は近いかもしれない。

$R(4, 8, 57)$ については 2012 年 10 月 SCSat1 の原型となったソルバーで既に得られており、ラムゼー数 $R(4, 8)$ の最良下界更新記録に寄与したが、SCSat3 での追試で新たなインスタンスが得られた。 $R(4, 7, 48)$ は $R(4, 8, 57)$ を求める上でベースとされる重要なグラフで、それ自体発見が困難なものであるが、これも SCSat3 による別インスタンスを得ている。

特筆すべきは 2014 年 5 月に発見した $R(4, 11, 100)$ で、ラムゼー数 $R(4, 11)$ の既知最良下界 98 を 101 まで一気に +3 改善したことになる。現時点で世界記録と思われる。その隣接行列を図 4 に示す。

4.1 未解決問題

ラムゼー数の既知最良下界を更新するために発見が期待されるラムゼーグラフは、

$$R(3, 10, 40), R(3, 11, 47), R(3, 12, 52), R(3, 13, 59), \\ R(4, 6, 36), R(4, 7, 49), R(4, 8, 58), R(4, 11, 101), \\ R(5, 5, 43), R(5, 6, 58), R(6, 6, 102), R(6, 7, 113)$$

である。

$R^Z(6, 6, 101)$ は唯一個 (2 色を入換えると 2 個) しかなく、 $R(6, 6, n)$, ($n \geq 102$) としての Z ラムゼーグラフは存在しない。 $R^Z(6, 6, 101)$ は $R(6, 7, n)$, ($n \geq 113$) の探索のベースとして貴重なグラフでもある。

なお、 $R(3, 11, 46)$, $R(3, 13, 58)$, $R(4, 6, 35)$, $R(4, 9, 72)$, $R(5, 6, 57)$ は文献で存在が知られているが、SCSat3 では未確認である。

5. おわりに

Erdős によれば $R(5, 5)$ は確定の可能性があり、Exoo らは $R(5, 5) = 43$ と予想している。しかし、 $R(6, 6)$ となると到底人知の及ばないものようである。ラムゼー数 $R(s, t)$ を確定すること、あるいはその良い下界を与えるラムゼーグラフ $R(s, t, n)$ を具体的に見つけること、それ自体の意義について著者には不明である。しかし、近年発展著しい SAT 技術の実力テストとして貴重な挑戦課題の一つと言えよう。

SCSat のさらなる改善のためには、ソフト節に関する様々な実行時パラメタの自動チューニング、より大規模な節集合 (億以上のオーダー) への対応などが課題である。また、一般の SAT ソルバーには、ソフト制約に頼らずとも元問題に含まれる潜在的規則性や特殊構造を自動的に認識し、充足解に敏感に察知できるような探索能力が期待される。

本稿の内容に関する詳細および関連情報については、著者の web ページ [Fujita] を参照されたい。

参考文献

- [Eén] Eén, N. and Sörrenson, N.: The MiniSat Page, <http://minisat.se>
- [Exoo] Exoo, G.: Ramsey numbers, <http://ginger.indstate.edu/ge/RAMSEY>
- [Fujita 12] Fujita, H.: A New Lower Bound for the Ramsey Number $R(4, 8)$, arXiv:1212.1328 [cs.DM] (2012).
- [Fujita 13a] 藤田 博: SCSat を用いたラムゼー数の下界更新について, 第 27 回人工知能学会全国大会, (2013).
- [Fujita 13b] Fujita, H., Koshimura, M. and Hasegawa, R.: SCSat: A Soft Constraint Guided SAT Solver, Proc. of 16th Int'l Conf. on Theory and Applications of Satisfiability Testing, (2013).
- [Fujita] Fujita, H.: Hirohi Fujita's WEB site, <http://opal.inf.kyushu-u.ac.jp/~fujita>
- [McKay] McKay, B. D.: Ramsey Graphs, <http://cs.anu.edu.au/~bdm/data/ramsey.html>
- [Radziszowski 14] Radziszowski, S. P.: Small Ramsey numbers, Electron. J. Combin., DS1, <http://www.combinatorics.org/issue/view/Surveys> (2014).