

4 SATとAI

井上克巳 (国立情報学研究所/総合研究大学院大学/東京工業大学)

AAAI-16

アリゾナ州フェニックスは砂漠の真ん中に作られた都会である。2016年2月12日～17日、日中の気温が摂氏30度を超えるこの乾燥した都市で第30回AAAI^{☆1}会議(AAAI-16)が熱く開催された。AAAIは人工知能(AI)のトップカンファレンスであり、AIがかつてないほどの注目を浴びる中、AAAI-16では論文投稿数が過去最高を記録した。会議全体では、ディープラーニング(深層学習)を中心とした機械学習・応用に関する講演や論文発表がかなり目立っていた。特にGoogle DeepMind社のDemis Hassabis氏が「汎用人工知能に向けて」と題した招待講演で、碁プログラムAlphaGoについて語り、本会議の目玉となった。彼は3月に対戦予定だった世界トップ級囲碁棋士との対戦に向けた抱負も自信満々に語っていたが、実際に約1カ月後AlphaGoは4勝1敗で見事勝利を収めた。

SATはというとわずか1セッションしかなく、それも解集合プログラミング(ASP)と併せて1つであった。SATだけでなく全体的に知識表現・推論分野における理論論文が減少し、研究の流行が前年のAAAI-15と比べて変化したため、AI研究が新たな段階に入ったと感じさせる会議となった。このディープラーニングを中心とした学習は一時的な流行なのだろうか？このままSATを始めとする記号的AIは廃れていくのだろうか？そもそもSATはAIと呼んでよかったのだろうか？

本稿ではこうした問いに答えつつ、最新のAIから見たSATの現状と将来について論じてみたい。まず、

^{☆1} AAAIはAssociation for the Advancement of Artificial Intelligenceの略。かつてはAmerican Association for Artificial Intelligenceの意味であった。

SATとAIの深い関係は歴史的事実から始まる。あまり知られていないことであるが、今をときめくディープラーニングの立役者であるトロント大学・GoogleのGeoffrey Hinton(1947～)は、実はGeorge Boole(1815～1864)の玄孫なのである。

George Boole 200

Booleはデジタル計算機の基礎をなすブール代数で知られる数学者である。SATとはブール式の充足可能性判定問題であるから、BooleはSATの父であるとも言える。とりわけ、Boole展開式^{☆2}

$$f(x, y) = f(1, y) \cdot x + f(0, y) \cdot (1-x) \quad (1)$$

は、SATソルバーの基礎をなすDPLL手続きにおける分割規則^{☆3}そのものを表現した式となっている。ここで式(1)の $f(x, y)$ はブール変数 x およびその他のブール変数列 y を含むブール式であり、 $f(1, y), f(0, y)$ は x にそれぞれ1, 0を代入して得られる式を表す。文献1)はタイトルが示すように、思考法則(laws of thought)について論じているが、論理的思考だけでなく現在の確率推論に通ずる貢献も見られる。さらにBooleは、計算機の父と呼ばれているCharles Babbage(1791～1871)との会話において、「思考する機械」(thinking machine)のアイデアを交換していたことも分かっており、彼をAIの父と呼んでも過言ではないだろう。

アイルランドにあるコーク大学(現University

^{☆2} 式(1)は文献1)に明示されているが、日本ではBoole展開を回路形式で解釈したClaude Shannon(1916～2001)にちなんでシャノン展開(定理)と呼ばれている。

^{☆3} DPLLはアルゴリズムの考案に寄与したDavis, Putnam, Logemann, Lovelandの4名の頭文字を取ったものである。DPLLの詳細は文献2)を参照されたい。

Dana Scott (計算機科学), Michael Wooldridge (AI: マルチエージェントシステム), Eugene Freuder (AI: 制約プログラミング), Geoffrey Hinton (AI: 機械学習), Oren Etzioni (AI: 情報検索), Robert Kowalski (AI: 論理), Francesca Rossi (AI: 制約プログラミング), John Hooker (オペレーションズリサーチ), Barry O'Sullivan (AI: 制約充足)

図-1 George Boole 200 記念講演会での主な講演者

College Cork) は Boole が 1849 年以降教鞭をとっていた伝統ある大学である。2015 年にコーク大学全学を挙げて Boole 生誕 200 年祭 (George Boole 200) が行われ、8 月末には記念講演会が開催された。このときの主な講演者のリストを図-1 に示すが、その中に Boole の直系子孫である Geoffrey Hinton がいた。Hinton はこのとき、「真実の思考法則」(The Real Laws of Thought) と題した講演を行った。これは文献 1) のタイトルの最重要部分を借用したもので、わざわざ「真実の」という修飾を付けている。Hinton によれば、記号に基づく形式的思考は人間の脳による思考の中では最も遅く身につく能力^{☆4}であり、直感的思考や通常の脳の判断プロセスである分散した神経回路パターン^{☆4}の発現による認知活動とは異なっている。Hinton はある意味、Boole が提唱した思考法則に対する現代的解釈を与えている。すなわち、SAT のような記号的推論とディープラーニングのようなパターン学習は対立するものではなく、相補的であり深く関連しているものである。

AI における SAT の歴史

George Boole 200 で AI 研究者がこれだけ集まったのは偶然ではない。AI の歴史の中で SAT が重要な役割を果たしてきたからこそ、人々はその起源を回顧しようとしたのである。しかし AI において SAT が注目を浴び始めるのは 1990 年代以降であり、それまでの道のりは決して平坦ではなかった。

★証明の機械化

SAT は反駁法、すなわち証明すべき定理の否定を

加えて矛盾を導く方法、のための推論方法として AI の初期から研究されていた。詳細は文献 2) に譲るが、1962 年に開発された DPLL 手続きは現在の SAT ソルバーの基礎をなしており、AI のテキストに掲載されることが多い。当時、推論の研究者は数学の定理の自動証明を目指しており、DPLL はその中で述語論理式を直接扱うのではなく、変数を具体化して命題論理式として扱う方式として考案された。ほどなく 1965 年に一階述語論理式を命題論理式に落とさずに扱うための変数ユニフィケーションを用いた導出原理が John Alan Robinson により発明され、DPLL は非効率であるとされ不要となってしまった。その後第 1 次 AI ブーム後期の 1969 年に、もう 1 つの AI トップ会議である IJCAI^{☆5} の第 1 回会議が開催されたが、SAT に関する論文はすでに消えていた。その頃に推論研究は定理証明だけでなくプログラム検証や質問応答に応用され始め、論理プログラミング研究にも展開された。

★冬眠

DPLL が定理証明で使われなくなって以降、SAT は AI 研究では見られなくなってしまった。一方で、計算機科学においては SAT の NP 完全性が Stephen Cook により 1971 年に証明され、SAT にはもっぱら NP 完全問題の代表としての存在価値が与えられてしまう。その後、知識ベースシステムの研究が盛んになり第 2 次 AI ブームが訪れ、AAAI の第 1 回会議が 1980 年に開催されるが、この頃になっても SAT に関する研究は AI では見られなかった。

しかしながら、後の SAT 技術につながるような推論に関する研究が 1970 年代以降次々と生まれた。特にアルゴリズム研究とも関連が深い制約充足問題 (Constraint Satisfaction Problem : CSP) が盛んに取り上げられるようになる。ブール制約式とは CSP の特殊形でもあるため、CSP 研究で培われた推論・探索技術を SAT に応用するようになるのは自然な流れであった。こうして 1988 年、SAT ソルバーの効率改善に関する研究が Ramin Zabih と David McAllester によ

☆4 たとえば、Jean Piaget による認知発達では最終段階。

☆5 International Joint Conference on Artificial Intelligence の略。

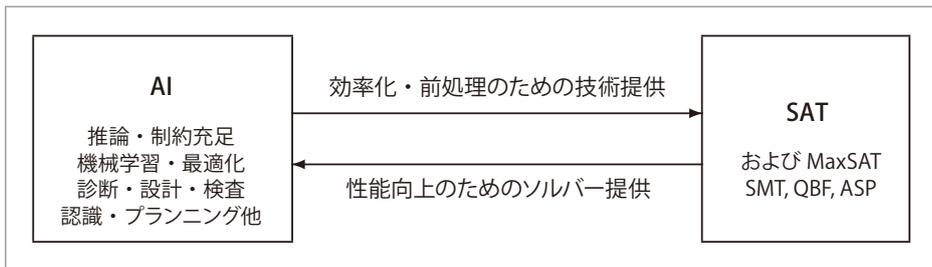


図-2 AIとSATの相互発展

りAAAI会議で初めて発表され、SATが長かった冬眠から目覚めた。1990年代に入ると、Henry KautzとBart Selmanが確率的SATソルバーの研究を進展させ、SATはAI研究の中で一躍脚光を浴び始めた。

★発展・AI 応用

1990年代以降SATソルバーの改良が続いた結果、21世紀に入ってCDCL (Conflict-Driven Clause Learning) ソルバーとして飛躍的進歩を遂げることになる。CDCLで使われた基本技術は、CSPで用いられたnogood学習と呼ばれる、矛盾する組合せ(nogood)の記録により探索領域を削減する手法である。このnogoodという概念は、非年代順バックトラッキング(non-chronological backtracking)と併せて、1970年代の真偽維持システム(TMS)^{☆6}や論理プログラミングなどの推論研究で開発された技術³⁾が受け継がれて開花したものである。

SATソルバーである程度の変数サイズが扱えるようになってくると応用も増える。AIでは最初に、KautzとSelmanがSATをプランニング分野に応用した。プランニングとは、与えられた2つの状態に対して片方(初期状態)から他方(目標状態)に到達することができるアクション列を求めることである。プランニング問題は、あるステップ数で目標が達成できるかどうかという判定問題であればSATで定式化して解くことができる。また最短プランを求める最適化問題でも、ステップ数の上限をインクリメンタルに増やしていくことで、最初に充足不能から充足可能になった上限値を持つプランを最短プランとして求めることができる。一般には

このような解き方は何度もSATソルバーを呼び出す必要があるため効率が悪く思えるが、呼び出し回数は多項式または対数オーダーで済むため計算自体は可能である。こうして、ハードウェアやソフトウェアのモデル検査や、スケジューリング等の組合せ問題でもSATは有効な手段となり、時には専用のソルバーを凌駕する性能を発揮し、オープン問題まで解いてしまう。中には1秒間に数千回ものSATソルバー呼び出しを行うような応用も存在しており、そこではSATがまさにNPオラクルのごとき役割を果たしているのである。

こうなると面白いことに、一度は使われなくなっていたSATによる解法が定理証明研究で蘇り、探索の効率化や学習節の利用で影響を与えたCSP研究にも高速SATソルバー技術が逆輸入されるようになった。また機械学習を使って問題の特徴から最適なソルバーやパラメータを学習しておきSAT問題を解く前処理で利用する研究がある一方、最近ではSATソルバーの機械学習への応用も注目されている。すなわち図-2のように、AIとSATは互いに影響を及ぼし合いながら進歩し続けている。

SATソルバーの中の宇宙

SATソルバーを用いた問題解決については、従来から各分野で専用アルゴリズムを開発してきた研究者にとっては当初受け入れがたいものがあった。たとえば、SATを用いたプランナーは2005年頃には国際プランニング競技会で上位を占めるようになっていたが、以降はコンテストの対象問題がヒューリスティクスを使用しないと解けない、すなわちSATベースでは不利なものに変更された経緯がある。また、CSPソルバー競技会も日本製のSAT型ソルバーであるSugarが優

^{☆6} TMSでは信念や言明間の依存関係を記録・監視しておき、矛盾が生じた際にはその原因となった信念の真偽値を変更する。TMSの意味論は後にASPにも継承されている。

勝してからは2010年以降開催されなくなりました。

SAT型問題解決では、問題をSATで記述した後はSATソルバーが高速に解いてくれるが、SATソルバーはブラックボックスとして扱われる。このように、外から見ると中で何が起きているか分からないが、実際には賢く問題を解いている、ということはAIシステムにはよく見られる。ディープブルーやAlphaGoが人間のチャンピオンに勝った際にも、なぜあの場面であの手を打ったのか理解できないというコメントが多かった。SATの推論過程でも億単位の学習節が生成されることがあり、もはや人間が簡単に読み取れるようなものではない。だからこそSATはAIと言えるのである。

高速SATソルバーの内部では、膨大な組合せの探索空間を効率よく解くために、リスタート戦略やある一定数以上溜まった学習節の消去など、さまざまな工夫が施されている。これらの探索技術は、短期記憶の制御や試行錯誤などの脳の活動と対比することもできる。現在は種々のパラメータを設定することで実現されているが、探索のメタな制御が自動化されればSATソルバーはさらに知的になるだろう。またCDCLソルバーについて、効率がよい理由を科学的に解明しようとした研究もある。たとえば、産業応用分野から作られるSAT問題のインスタンスでは、変数集合の中にコミュニティが存在し、その構造とCDCLの性能の間に相関があることが分かっている⁴⁾。

SATの近未来

順調に発展してきたSAT研究だが、昨今の第3次AIブームではディープラーニングを中心とした認識・学習研究の陰に隠れてしまっている。CDCLソルバーのこれ以上の大きな発展も現時点では難しいと考えられている。本稿の最初の問いに戻ると、SAT研究は今後減速に向かうのだろうか？ 筆者はそのようには考えておらず^{☆7}、さらなる発展を期待している。人間の脳による処理の中で、パターン認識はディープラーニング研究の成果が上がり始

☆7 2016年7月開催のIJCAI-16ではSATおよび知識表現・推論に関する論文は一定数をキープしている。

めてはいるが、知識の抽出・表現や推論といった高次処理とはまだ結び付いていない。「George Boole 200」の章で述べたように、記号的推論は認識・学習能力とは相補的であり、より強いAIの実現のためには必要である。ただし、今後の記号的推論が現在のSATソルバーの形をとるとは限らない。

AAAI-16では、SAT問題の入力形式を記号のパターンとみなし、ポートフォリオ型問題解決において、パターンごとに得意なSATソルバーをディープラーニングで学習する試みが発表された⁵⁾。この研究はSATとディープラーニングの1つの結合形となっているが、実際に解くソルバーは既存のCDCLソルバーである。前述した、問題に内在するコミュニティ構造についても、同様の分類はできるだろう。これとは根本的に別の解法も考えられる。SATで扱われてきた問題には、グラフ探索やパズルや配置などのパターンに関連したCSP問題も多く、一部はニューラルネットで解くことが可能になるかもしれない。ここで問題の階層構造まで捉えて解くことができれば「ディープSAT」と呼べるものにもなる。さらにアナログ計算や代数的計算または量子計算がブール代数や論理計算をいかにして扱うことができるか今後注目したい。

参考文献

- 1) Boole, G. : *An Investigation of the Laws of Thought, on Which are Founded the Mathematical Theories of Logic and Probabilities*, Macmillan (1854).
- 2) 井上克巳, 田村直之 : SATソルバーの基礎. 人工知能学会誌, 25 (1), 57-67 (2010).
- 3) Stallman, R. M. and Sussman, G. J. : Forward Reasoning and Dependency-Directed Backtracking in a System for Computer-Aided Circuit Analysis, *Artificial Intelligence*, 9, pp.135-196 (1977).
- 4) Newsham, Z., Ganesh, V., Fischmeister, S., Audemard, G. and Simon, L. : Impact of Community Structure on SAT Solver Performance, *Proc. SAT 2014*, LNCS 8561, pp.252-268, Springer (2014).
- 5) Loreggia, A., Malitsky, Y., Samulowitz, H. and Saraswat, V. : Deep Learning for Algorithm Portfolios, *Proc. AAAI-16*, pp.1280-1286 (2016).

(2016年4月29日受付)

井上克巳 (正会員) inoue@nii.ac.jp

国立情報学研究所情報学プリンシプル研究系教授・総合研究大学院大学複合科学研究科情報学専攻教授 (併任)・東京工業大学情報理工学研究院情報工学系特任教授。AIによる科学的思考の実現に興味を持つ。