2I3-2

大学入試過去問による数学問題解答システムの評価と課題分析

Entrance Exam Math Problems, from the Viewpoint of Automatic Problem Solving

松崎 拓也[*1] 岩根 秀直[*1*2] 穴井 宏和[*1*2*3] 新井 紀子[*1]
Takuya Matsuzaki Hidenao Iwane Hirokazu Anai Noriko H. Arai

[*1]国立情報学研究所 [*2](株) 富士通研究所 [*3]九州大学
National Institute of Informatics Fujitsu Laboratories Ltd. Kyushu University

We analyzed real math problems taken from university entrance examinations from the viewpoint of automated reasoning. We further applied a prototype problem solving system to the real problems to see the practicality of our approach wherein a natural-to-formal language translation system is joined with automated reasoning technologies.

1. 言語の理解と数学問題を機械で解くこと

自然言語で書かれた数学問題を計算機で解く、というタスクは、言語理解技術のベンチマークとして特別な地位をもつ。これは以下の 2 つの事情による。まず、これまで提案されてきた「言葉の意味」の定義のうち最も成功したもののひとつは、「文の意味とは、その真理条件、すなわち、その文が真であるような世界の在り様を過不足なく述べたものである」、という定義であること（真理条件意味論）。真理条件を形式的に述べる手段としては論理が自然な選択である。次に、形式論理はそもそも数学の形式化のために開発された手段であり、かつ、論理による形式化が最も成功した知的活動はおそらく数学であること。このように、完全に形式的・明示的な分野オントロジーを持つ領域を対象に、統語・文脈・意味解析における曖昧性解消から語用論的効果の理解まで含めた高度な言語処理を実現することは、「一般的な言語理解技術」なるものが果たして成立しうるのかを占う上で当然取り組むべき最初の課題である。

我々は「ロボットは東大に入れるか」プロジェクト [新井 12] の一部として、大学入試数学問題を自動的に解くシステムを開発している。自動解答の最初のステップは、自然言語で書かれた問題を論理による表現へと翻訳することである。翻訳先の言語として、我々は Zermelo-Fraenkel (ZF) 集合論の（大幅な）保存拡大を選んだ。ZF 集合論は、その被覆範囲の広さ、また、数学問題中には集合・関数・タプル等々といった集合論上の概念が自由に表れることから、翻訳の目的言語として自然な選択と言える。また、システムの実現の側面から見ると、ZF 集合論を目的言語として選ぶことは、例えば、算術問題のための翻訳システム、幾何問題のための翻訳システム、といったように問題タイプ毎に異なる翻訳システムを用意するのではなく、完全に同一の翻訳システムを全ての問題に対して適用できることを意味する。

しかし、解答システムの中間言語としての ZF 集合論は、その柔軟性の一方、自動推論のための表現体系としては全く現実的でない、という明らかな欠点をもつ。このため、自然言語からの翻訳結果を実用的な推論技術へと接続するためには、さらにもう一段階の処理が必要となる。すなわち、実際に問題を解くためには、ZF 集合論の言葉で形式的に表示された問題から、自動推論が現実的に可能な別の表現を機械的に導く必要がある。この問題は AI の歴史的な難問である「問題表現の問題」[McCarthy 64] とも極めて近い関係にあり、たとえ入試数学問題に対象を限ったとしても、根本的な解決がすぐに見つかるとは考えにくい。

我々は上のことを踏まえた上で、なお、ZF 集合論を中間言語とするプロトタイプシステムを開発中である。このシステムでは ZF 集合論による問題の表現に対し、同値性を保ついくつかの書き換え規則を繰り返し適用することで自動推論が可能な問題表現を探す。書き換えが成功した時は、得られた問題表現に応じた推論器が呼び出される。「現実的な自動推論」の代表的な例は決定手続きである。このプロトタイプシステムの構成は実閉体（Real Closed Field, RCF）の理論は決定可能で、しかも限量子消去が可能である [Tarski 51] という古典的な結果と、実際の決定アルゴリズム（限量子消去; Quantifier Elimination, QE）の改良に関する近年の進展 [Caviness 98, 穴井 11] にヒントを得ている。

ごくおおまかに言えば、RCF に対する QE アルゴリズムの存在は「x の値を求めよ」というタイプの問題に対して、一階の実閉体の言語による問題の表現が得られれば、原理的には x の可能な値を全て見つけられることを意味している。これは自動問題解答の観点からは非常に好都合である。また、実閉体の理論によって表現できる問題は、例えば入試問題においては代数および幾何の問題のかなりの部分を含み、応用範囲が広い。

RCF-QE による解法の深刻な限界は、その計算コストにある。現在、実用上もっとも高速な RCF-QE アルゴリズムの時間計算量は入力中の変数の数の 2 重指数オーダーである [Collins 75]。ゆえに、RCF-QE による解法が現実的に実行可能かどうかは、問題の表現に大きく依存する。このため、上記アプローチの現実性は、実際に言語処理の結果として得られる意味表示に対して RCF-QE を適用することで初めて明らかになる。

以下では、開発中の数学解答システムについて簡単に紹介した後、まず、国立大学 7 大学、1 年分の 2 次試験数学問題を自動推論の観点から分析した結果を示す。この分析では、解答の主要な部分で必要となる演繹がいずれの数学理論（theory）のもとで行われるべきか、また、問題からその理論を自動的に決定することが現実的なのか、という 2 つの観点から問題を分析する。さらに、実閉体の理論のもとで表現可能とみられる問題にたいして、実際にプロトタイプシステムを適用した結果を示す。最後に、現在のシステムの性能上限を示すひとつの指標として、予備校による東大入試模試の過去問に対する結果を示し、実際の模試受験者の平均点と比較する。

現在のシステムでは、言語解析の一部を入力問題文に対する言語アノテーションで代用している。この意味で本稿の実験

連絡先: 松崎拓也、takuya-matsuzaki@nii.ac.jp

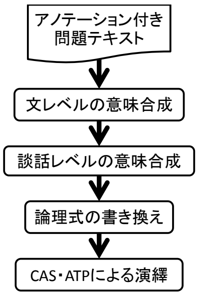

図 1: 処理の流れ

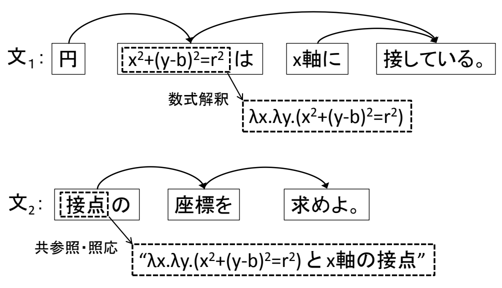

図 2: 問題に対する言語アノテーション

表 1: Theory による問題の分類

Theory		問題数
実閉体の理論 (RCF)		47
ペアノ算術 (PA)		10
ZF 集合論	超越関数	23
	RCF+PA	15
	その他	4
total		99

結果はやや楽観的な性能上限見積もりというべきものである。しかし、実験の結果は我々のアプローチの有効性を示唆するものであり、今後の課題のいくつかが明らかになった。

2. アノテーション付き入力から解答まで

図 1 に解答に至る処理の流れを示す。現在のプロトタイプシステムは、言語アノテーション付きの問題文を入力とし、まず、構文解析を経由して単語レベルの意味表示から文レベルの意味表示を合成する。次に、アノテーションで指定された文間の論理関係に従って談話レベルの意味合成を行い、問題全体の意味表示を得る。さらに、問題全体の意味表示に対して同値性を保つ書き換え処理を行い、現実的に自動推論が可能な論理式を得る。最後に、書き換え結果に従って、計算代数システム（CAS）ないし自動証明器（ATP）による演繹を行い、解を得る。以降、本節では、言語アノテーションと、処理の各ステップについて概要を述べる。文レベル・談話レベルの意味合成の詳細については別稿 [Matsuzaki 13] を参照されたい。

2.1 言語アノテーション

現在のプロトタイプシステムは、(1) 数式の解釈、(2) 係り受け構造、(3) 共参照・照応関係、(4) 文間の論理関係、の 4 種類のアノテーションが施された問題テキストを入力とする。図 2 に、文 $_1$ と文 $_2$ の 2 文からなる問題に対するアノテーションの例を示す。この例では、数式「$x^2+(y-b)^2=r^2$」の部分に、この等式は xy 平面上で定義された特性関数（ないし点集合）として解釈すべきであることを指定するアノテーション「$\lambda x.\lambda y.(x^2+(y-b)^2=r^2)$」が施されている。また、文 $_2$ における「接点」にはゼロ代名詞の先行詞を補って言い換えた形「$\lambda x.\lambda y.(x^2+(y-b)^2=r^2)$ と x 軸の接点」がアノテートされている。また、文には文節区切りおよび文節間の係り受け構造がアノテートされている。最後に、図には示さないが、2 つの文の間の論理関係が「文 $_1$ & 文 $_2$」である、すなわち 2 つの文は連言の関係にあることもアノテートされている。

これらのアノテーションに相当する情報をテキストから自動的に認識する技術・ツールの開発は活発に行われている。現在のアノテーションはこれらの言語処理ツールの代用であり、既存の技術の数学テキストへの適応は今後の研究課題である。

2.2 意味合成

文レベルの意味合成は、問題中の各文の構文解析を行い、文の統語構造に従って、個々の単語の意味表示から文全体の意味表示を得る処理である。実際の解析処理では、問題文にアノテートされた文節間の係り受け関係を、組み合わせ範疇文法（Combinatory Categorial Grammar, CCG）[Steedman 01] による文の導出木の構造に関する制約として用い、制約下で可能な CCG 導出木の中から、簡単なスコア関数によって選ばれたものを文の意味構造として出力する。

談話レベルの意味合成は、個々の文の意味表示を問題文にアノテートされた文間の論理関係に従って組み合わせ、問題全体の意味表示を得る処理である。これまでの観察によると、国立大学 2 次試験の記述解答式の問題に現れる文間の論理関係は、条件を表す含意関係（文 $_1 \rightarrow$ 文 $_2$）と、累加を表す連言（文 $_1$ & 文 $_2$）の 2 種類でほぼ尽くされる。問題全体の論理構造は、これら 2 種の関係（$\rightarrow$ および &）を中間ノード、文を葉とする 2 分木の形でアノテートされている。

2.3 論理式の書き換え

問題の意味表示は、最終的には高階の（ZF 集合論の）述語論理式となる。先述のように、ここから現実的に自動演繹が可能な問題表現を機械的に得ることは一般的には非常に困難である。現在のシステムでは、さしあたっての第一歩として、ペアノ算術ないし RCF の式として解釈できるものが見つかるまで以下のような同値書き換えの規則を貪欲に（バックトラック無しで）適用し続ける、という単純な手続きを用いている：

- 論理式の単純な同値変形、例えば $\exists x(x=\alpha \wedge \phi(x)) \Rightarrow \phi(\alpha)$ や $\forall x(x=\alpha \rightarrow \phi(x)) \Rightarrow \phi(\alpha)$（ただし、いずれにおいても x は α に自由に現れない）。
- 知識ベース（公理集）を用いた、述語および関数の定義による書き換え。例：
  $\mathrm{is\_divisible\_by}(n,m) \Rightarrow \exists k \in \mathbb{Z}(n=km)$
  $\mathrm{distance}((x_1,y_1),(x_2,y_2)) \Rightarrow \sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$
- CAS を用いた、積分や微分を含む式の簡単化
- 述語および関数の手続き的な評価

最後のタイプの規則は、定義による書き換えで実現するのが現実的でないような複雑な概念、例えば「多項式で定義された複数の曲線に囲まれた図形の面積」[岩根 13] などの関数・述語の評価を CAS 上に実装したプログラムで実現するものである。

2.4 RCF-QE の高速化

現在の解答システムは 2 つの推論器と接続されている。ひとつは Mathematica の Reduce コマンドに実装されている整数ドメインの論理式に対する証明器、もうひとつは RCF-QE の実装 SyNRAC[Iwane 13] である。以下 RCF-QE とその高速化について述べる。

RCF の言語の原子論理式は有理数係数多項式間の等式および不等式であり、それらの等式・不等式をブール演算（$\wedge$、$\vee$、$\rightarrow$、$\neg$ など）および限量子（$\forall$、$\exists$）によって組み合わせたものが一階の RCF の式となる。RCF に対しては限量子消去（QE）が可能である [Tarski 51]： 任意の RCF の式に対し、それと等価で限量子を含まない式を見つけることができる。これは、問題を RCF の式で表すことができさえすれば、QE を行い、

表 2: RCF で表現可能な問題に対する解答結果

	翻訳	
	Manual	Semi-Auto
正答	28	23
時間切れ	12	14
CAS との接続	4	2
RCF 式の導出失敗	2	2
未知の指示	1	1
翻訳失敗	None	3
文法設計	None	2
total	47	47

結果得られた連立方程式・不等式を解くことで全ての解が得られることを意味する。しかし、逐語訳的に問題文を翻訳して得られた意味表示は往々にして非常に冗長で多くの変数を含む。現在もっとも一般に用いられている RCF-QE アルゴリズムは式中の変数の数の 2 重指数オーダーの計算量を持つため、入力の冗長性は RCF-QE の適用の上で大きな問題となる。

この問題に RCF-QE アルゴリズムの側から対応するために、我々は様々な高速化技法を SyNRAC に加えた。具体的には、(1) 例えば、限量子が付いた変数について線形な式のみを含む、など特別な形を持つ入力式に対する効率的なアルゴリズムの利用、(2) 一部の限量子を消去した結果である中間式に対する種々の簡単化、(3) 論理式の単純な同値変形（§ 2.3）による限量子消去が効果的に行われるように問題を分割、そして (4) 変数の順序付けによって大きく計算量が変わるため、複数の変数順序について並列に計算を行う、という改良を行った。

3. 自動演繹の観点からの入試数学問題の分析とプロトタイプシステムによる半自動解答

本節では、まず自動推論の観点から入試数学問題を分類した結果について述べ、このうち RCF で表現可能な問題に対して解答システムを適用した結果について報告する。実験に用いた問題は、1999 年度の国立大学 7 大学（北大、東北大、東大、名大、阪大、京大、九大）の 2 次試験問題 124 題である（異なり小問数 ; 大問数としては 65 題）。このうち、現在の我々の意味表示の体系では直接翻訳できない確率および組み合わせ論の問題を除いた 99 題を分析の対象とした。この 99 題には、代数（自然数、実数、複素数）、平面および空間幾何、線形代数、初等関数および微積分など様々な分野の問題が含まれる。

3.1 Theory による問題の分類

表 1 に、各問題で必要とされる演繹処理を表現しうる数学理論（theory）は何か、という観点から問題を分類した結果を示す。最初の行の 47 題は主要な演繹処理が RCF-QE と方程式・不等式の求解として表現できること、2 行目の 10 題はペアノ算術の定理証明問題として表現できることが問題文からほぼ明らかだった問題である。これらは全体の約 60%を占める。ここで、自然言語による問題記述、ないしはその ZF 集合論への直訳から、その問題で必要な演繹を行うための数学理論を機械的に決定することは一般には容易ではないということを注意しておきたい。たとえば、円や垂直二等分線など初等幾何の対象は RCF で記述可能であるが、だからといって、初等幾何の対象の性質のすべてが RCF で記述可能なわけではない。それは次のような例を考えれば明らかである :

Q1: 周の長さが 4 である正方形の一辺の長さを求めよ。

Q2: 半径が 1 である円周の長さを求めよ。

Q3: 半径が 1 である円周上の 2 点間の最大距離を求めよ。

Q1-Q3 は語彙および概念のレベルでは大きな重なりをもつが、Q1 および Q3 は RCF で表現可能、Q2 はそうではない、という違いがある。これは Q2 には超越数 π が本質的に現れるためであり、RCF の中で問題 Q2 を表しえないことは Q2 をどう翻訳するか（たとえば線積分の定義まで遡って円周長を表現するなど）にはよらない。

表で「ZF 集合論」に分類した、残りの 40%の問題は、それらを直接表現でき、しかも現実的に自動演繹が可能な理論が見当たらないという意味でより問題となる問題である。これらの問題はおおむね 3 つのグループに分類できそうである。一つ目は、表で「超越関数」に分類した 23 題で、三角関数とその逆関数、指数・対数関数に関する何らかの推論を必要とする問題である。これらのうち半数は、超越関数を変数でおきかえる（$x := \sin\theta, y := \cos\theta$ として制約 $x^2 + y^2 = 1$ を加える、など)、また、CAS の機能を用いて微積分・極限・最大・最小化の計算を行う、などの操作によって RCF で表現可能な問題に帰着できた。しかし、RCF へと帰着させるための操作は必然的に発見的なものであり、どの程度まで系統化・自動化できるか見極めるのは今後の課題である。

「RCF+PA」に分類された 15 題は、自然数（ないし整数・有理数）と実数をともに含むような問題である。これらの内ほとんどは、「…である実数 x は有理数（あるいは無理数）であることを示せ」というタイプの証明問題か、自然数と実数をともに含む命題を帰納法で示すタイプの問題であった。

最後に、「その他」に分類された 4 題は自動演繹で処理するための現実的な方策が差し当たり見いだせなかった問題である。例えば、このうちの 1 題は、直円錐を考え、その側面上で底面の直径の 2 端点を結ぶ最短経路の長さを求める問題であった。「その他」に分類されたものを含め、表 1 に示した分類は、問題文および人間の解法の非形式的な観察によるものであることを注意しておく。このため、「RCF」や「PA」に分類された問題であっても、解答に至る完全に形式的な演繹過程のどこかで、より強い公理系が必要になり、現在の我々のアプローチでは扱いきれない可能性は残る。このことは、ジョルダン曲線定理（「平面上の自己交差を持たない閉曲線は平面を『内側』と『外側』に分ける」）のようなごく「当たり前」の事実の証明がようやく 20 世紀になって行われたことを考えれば特に驚くべきことではない。「RCF」に分類された問題については、次節の実験がこの点の検証を行うものとなる。

3.2 RCF で表現できる問題に対する半自動解答

上述のように、表 1 で「RCF」「PA」以外に分類された 40%の問題は現在の我々のシステムでは解けない。また、「PA」に分類された 10 題はいずれも Mathematica 9.0 に含まれる証明器では証明できなかった。ここでは、残りの、「RCF」に分類された問題に対して解答システムを適用した結果を示す。

我々はシステムへの入力として 2 種類を用意した。ひとつは ZF 集合論に基づく意味表示を手で書き下したもの（Manual）で、もうひとつはアノテーション付きの問題文から意味合成を経て半自動的に得た（同じく ZF 集合論に基づく）意味表示（Semi-Auto）である。現在、意味解析で用いている辞書は未だ被覆率が十分でないため、1 問につき平均して数語、未知の語ないし既知の語の未知の用法が含まれていた。このため、必要な辞書項目 67 項目を辞書に後から追加した。このうちほとんどは、既に知識ベースで定義済みの概念に対する未知の表現であった。Manual、Semi-Auto の 2 種類の意味表示に対する結果の比較により、逐語訳的な意味合成に起因する意味表示の冗長性が計算コストに与える影響を見積もることができる。

表 2 に 2 種の意味表示に対する結果の分類を示す。手で書

表 3: 東大模試 数学問題に対する結果

理系 (120 点満点)			
Test set	翻訳		受験者
	Manual	Semi-Auto	平均点
2013 年 7 月	40	40	21.8
2012 年 11 月	40	40	32.5
2012 年 7 月	25	18	29.8

文系 (80 点満点)			
Test set	翻訳		受験者
	Manual	Semi-Auto	平均点
2013 年 7 月	40	40	24.9
2012 年 11 月	20	8	25.7
2012 年 7 月	40	40	30.9

いた意味表示では約 60%、半自動的に導出した意味表示では約 50%の問題が正しく解けた。もっとも多い失敗の原因は時間切れ（最大 15 分）であり、このうち大多数が RCF-QE の計算コストによるものであった。Manual と Semi-Auto の 2 種の意味表示間で、時間切れになった問題の数の差はそれほど大きくはない。(ただし、Manual では時間切れだが Semi-Auto では他の原因で失敗した問題が 3 題あった)。このことから、Semi-Auto の意味表示の冗長性が計算コストに与える影響は、§2.4 で述べた改良を含めた RCF-QE アルゴリズムの洗練によって低く抑えられていることが伺える。このことは、後段の計算コストを気にせずに、文法および意味合成処理を単純に保つための前提条件であり、完全に End-to-End の解答システムへ向けた好結果と言える。

3 行目の「CAS との接続」に分類された問題は、解答システムからの CAS の利用における実装上の問題によって失敗したものである。ここまで先頭 3 行のいずれかに分類された問題が、ZF 集合論に基づく意味表示から、§ 2.3 で示した書き換えアルゴリズムによって RCF での意味表示が得られた問題となる。書き換えアルゴリズムの限界が原因で RCF での意味表示が得られなかった問題は Manual/Semi-Auto 共通で 2 題のみであった（「RCF 式の導出失敗」)。そのうちの 1 題では無限に長い三角柱の形状（不等式による定義）をその断面の形状だけから求めることが必要であった。このための演繹はそれ自体は難しくはないが、現在の書き換えアルゴリズムでは ad-hoc な書き換えルールを用意しない限りは実現しづらい操作である。

最後の 3 行（「未知の指示」「翻訳失敗」「文法設計」）に分類された問題は言語理解の問題によって、意味表示を手で書くあるいは半自動で導出する時点で失敗したものである。「未知の指示」は想定していなかった問題指示（図形の形状について言葉で説明する）を含む問題、「翻訳失敗」は文法が許す複数の意味表示のうち、スコア関数によって適切なものを選ぶことができなかったことによる失敗、「文法設計」は現在の文法の基本的な設計では扱えない構文を含む問題である。

4. 東大模試問題による評価

代々木ゼミナール主催の東大入試模擬試験の過去問を用いたシステムの評価を行った。前節の実験と同じく、手で書き下した意味表示とアノテーション付きの問題文から半自動的に導出した意味表示の 2 種類を用いた。また、これも前節と同様、辞書に欠けていた問題中の語については後から辞書項目を追加することを許した。理系・文系それぞれ 3 回分の過去問について評価を行った結果を表 3 に示す。各回とも文系の問題は大問 4 問、理系は大問 6 問からなり、大問 1 問の配点は全て 20 点である。部分点の配点は過去問付属の解説に従って行った。この文系・理系 3 回分の過去問で解けた問題は Mathematica の整数ドメインの論理式に対する証明器で解けた 1 問を除き全て RCF-QE を用いて解けたものである。

人手によって翻訳した意味表示（Manual）と半自動的に得たもの（Semi-Auto）で結果に差が出た問題は 2012 年 11 月の文系、2012 年 7 月の理系にそれぞれ小問 1 問ずつあった。このうち 1 問は文法がカバーしていない構文のため意味表示が得られなかった。もう 1 問では意味表示は得られたもののスコア関数で正しい解釈が選べず、正解が得られなかった。

5. まとめ

本稿では、深い言語処理技術と計算代数アルゴリズムを中心とする自動演繹の接合による数学問題の自動解答について、その意義と見通しを述べ、現時点での性能評価の結果を示した。現在のプロトタイプシステムで解ける問題は、限量子消去が可能な RCF で表現できるタイプの問題にほぼ限られる。また、実験では問題文に対する言語アノテーションを意味合成に利用していること、また、不足だった辞書項目の追加を許したことから、実験結果は最終的な End-to-End の自動解答システムの性能に対する現時点での上限見積もりと考えるべきものである。しかし、東大模試問題に対する結果では、6 回分の模試のうち 4 回でシステムの得点が受験者全体の平均点をすでに上回っている。意味解析における曖昧性解消の高度化、RCF-QE 手続きのさらなる効率化、決定手続きが存在しないタイプの問題に対する発見的解法の蓄積や、初等関数を含む問題のうち特定のタイプのものに対する系統的解法の実現により、この差はさらに開く可能性がある。

参考文献

[Caviness 98] Caviness, B. and Johnson, J. eds.: *Quantifier Elimination and Cylindrical Algebraic Decomposition*, Springer-Verlag (1998)

[Collins 75] Collins, G. E.: Quantifier elimination for real closed fields by cylindrical algebraic decomposition, in *Automata Theory and Formal Languages 2nd GI Conference*, Vol. 33 of *LNCS*, Springer-Verlag (1975)

[Iwane 13] Iwane, H., Yanami, H., Anai, H., and Yokoyama, K.: An effective implementation of symbolic-numeric cylindrical algebraic decomposition for quantifier elimination, *Theoretical Computer Science* (2013)

[Matsuzaki 13] Matsuzaki, T., Iwane, H., Anai, H., and Arai, N.: The Complexity of Math Problems – Linguistic, or Computational?, in *Proc. IJCNLP-2013* (2013)

[McCarthy 64] McCarthy, J.: A tough nut for proof procedures, Ai memo, MIT (1964)

[Steedman 01] Steedman, M.: *The Syntactic Process*, Bradford Books, Mit Press (2001)

[Tarski 51] Tarski, A.: *A Decision Method for Elementary Algebra and Geometry*, University of California Press, Berkeley (1951)

[岩根 13] 岩根 秀直, 松崎 拓也, 穴井 宏和, 新井 紀子 : 数式処理による入試数学問題の解法と言語処理との接合における課題, 人工知能学会全国大会論文集 (2013)

[穴井 11] 穴井 宏和, 横山 和弘 : QE の計算アルゴリズムとその応用 ー 数式処理による最適化, 東京大学出版会 (2011)

[新井 12] 新井 紀子, 松崎 拓也 : ロボットは東大に入れるか?—国立情報学研究所「人工頭脳」プロジェクト—, 人工知能学会誌, Vol. 27, No. 5 (2012)