

米国における人工知能研究の一動向
—ニューメキシコ州立大学における
自然言語理解を中心として

岡田 直之

大分大学工学部

ニューメキシコ州立大学(NMSU)計算研究所(CRL)では人工知能研究の中でも、特にビジョン/ロボティクスと人間-機械インターフェイスの分野で卓越した研究を行い、地域経済の発展に貢献しようと努めている。本稿はCRLの最近の研究活動について自然言語理解プロジェクトを中心に述べたものである。最後にNMSUで開かれた人工知能の基礎に関する研究集会についてもふれる。

A TREND OF ARTIFICIAL INTELLIGENCE
RESEARCH IN THE UNITED STATES OF
AMERICA ————— CENTERING AROUND THE
NATURAL LANGUAGE UNDERSTANDING
IN NEW MEXICO STATE UNIVERSITY

Naoyuki OKADA

Information Science and Systems Engineering

Oita University

Computing Research Laboratory of New Mexico State University is developing excellence in two areas, vision/robotics and human-machine interface, of artificial intelligence, stimulating economic development in New Mexico. This paper describes the recent research activities of CRL, centering around the natural language understanding project. Lastly, the workshop on foundations of artificial intelligence held at NMSU is referred to cursorily.

1 まえがき

近時人工知能 (Artificial Intelligence) に関する研究は、一部のシステムの实用化の動きとも相まって、極めて活発なものとなっている。それと共に研究の対象や分野も多様化し、得られた成果に対する評価が人によって異なってくる。一例をあげると、IJCAI-83において一つの投稿論文に対する二人の査読者の意見が、一方は受理し他方は受理しないと二つに分れてしまったケースが50%以上に上っている⁽¹⁾。自然言語理解はAIにおける重要な柱の一つであり、上記傾向はそのまま自然言語理解の研究にも当てはまるものと言えよう。

このような混迷状態を打開する一つの方策は、個々の研究者/団体が、従来以上に研究の目的や態度を明確にすることである。筆者は1985年12月より約二月間、米国ニューメキシコ州立大学 (New Mexico State University) の計算研究所 (Computing Research Laboratory) において、言語と図形の理解に関するセミナを担当する機会を得た。CRLは地域社会の要請に基づいて設立される明確な方針の下にAIシステムの開発を行っている。即ち研究的なビジョン/ロボティクスと人間-機械インタフェイスに絞っている。自然言語理解は人間-機械インタフェイスにおいて重要な一部分を担っている。本稿はCRLにおけるAI研究、特に自然言語理解プロジェクトを紹介すると共に、CRLの主催で開かれたAIの基礎に関するワークショップについてもふれ、今後の自然言語理解研究の在り方に対する、一つの参考資料を提供しようとするものである。

2 CRLの概要

2.1 設立の背景

ニューメキシコ州は従来農業および工業依存型の経済から高度技術に基づく情報依存型経済へと脱皮を図っている。州政府は1981年に技術的優秀委員会 (Committee on Technical Excellence) を設置し、高度技術開発のための調査を依頼した。CTEはニューメキシコ州を南北に流れるリオグランデ川沿いに、ロスアラモスからラスクルーセスに至る約300マイルをリオグランデ研究回廊 (Rio Grande Research Corridor) と定めた。そしてCTEは10の主要研究所、三つの総合大学、七つの技術職業訓練所、80の科学技術商社、および44のエネルギー開発会社をRGRCの指定機関とした。それらにはロスアラモス国立研究所、アルバカーク地震研究所、ニューメキシコ太陽熱工業開発会社、ニューメキシコ鉱山鉱物資源局、USAF電子*光学観測所、NASAジョンソン宇宙センターホワイトサンズ試験場、NMSU, 等が含まれる。更にCTEは五つの技術的優秀センターの新設を州政府に進言し、1983年にこれに関する法律が州議会を通過した。

現在RGRCにおいては、五つの技術的優秀センターと二つの技術的革新センターが、科学技術諮問委員会 の指導、監督の下に互いに連絡を保ちながら地域経済の発展のために機能している。技術的優秀センターは次の通りである。

- *非侵入診断センター (アルバカーク、ニューメキシコ大学内)
- *高度技術材料センター (アルバカーク、ニューメキシコ大学内)
- *爆発技術開発センター (ソックロ、鉱山技術研究所内)
- *植物遺伝子工学センター (ラスクルーセス、NMSU内)
- *計算研究応用センター (ラスクルーセス、NMSU内)

CRLは上記計算研究応用センターとして1983年7月に発足した。その目的は、計算機応用、とりわけAIにおける幾つか特定の分野で、高度な専門技術を開発することにある。

2.2 組織とプロジェクト

CRLの所員は専任、併任およびパートタイムを含めて、およそ60名であ

る。専任は所長の Y. ウイルクス博士、8名の研究系所員、3名のソフト技術系所員、1名のハード技術系所員、および3名の事務系所員、計15名である。併任は NMSU の教授、助教授ら14名である。その他はパートタイムの研究助手(主に大学院、学部学生)である。所員の専門は大部分が計算機科学であるが、生物学、数学、心理学等、NMSU の三つのカレッジと14の学科に関する係している。その他、客員研究所員、奨学研究生(NMSU の学生でありながら、CRL で高度技術の訓練を受けると)等のポストがある。

研究は主として専任の研究系所員、および併任の教授、助教授等によって推進されるが、半数近くを占めるパートタイム所員が新鮮で柔軟なプロジェクトの構成を可能にしている。

CRL のプロジェクトには大きくビジョン/ロボティクスと人間-機械インタフェイスとがある。以下に、主なサブプロジェクトと研究者を示す。

ビジョン/ロボティクス

*ビジョン関係

形状解析とパターン認識(フーリエ級数を用いる)、縁解析理論(MIT と意見交換)、面解析理論、農業地帯の監視(航空写真を解析し農作業上の意志決定を行う。民間企業と契約)、昆虫認識(農作物の被害検出)など。---

K. フィリップ博士

*ロボティクス関係

認知ロボット(学習/慣れおよび推論機能を持つ) --- D. パトリック博士および V. ジョンソン博士

工業ロボット(マルチセンサを持つ) --- S. カマ博士

人間-機械インタフェイス

*認知モデルと知識表現関係

リンク重みづけネットワーク(意味記述のグラフ表現) --- D. デアホルト博士

*利用者インタフェイス関係

利用者の知識に基づくインタフェイス設計(NASA と契約) --- J. マクドナルド博士他

インタフェイスの迅速なプロトタイプ化(サンジャ国立研究所と協力) --- J. マクドナルド博士

音声*点字出力ターミナル(視覚障害者用。低価格) --- A. シェマ博士
科学ワークベンチ(問題解決用グラフィック) --- W. コンレ博士

*自然言語理解関係

プリファレンスセマンティクス(幾つかの可能性の中から最良のものを選び、他を排除しない)、コラティブセマンティクス(語の意味の照合と選り分け)、機械翻訳(英中)、並列構文解析(Intel iPSC/D6による)など --- Y. ウイルクス博士

以上の組織とプロジェクトは、MIT の J. プラディ博士、コロラド大の W. キンシ博士、デック社のストックエブランド氏等5名からなる諮問委員会の指導、監督を受けている。

最後に CRL の主要設備を示す。

主機

Symbolics 3670 LISP Computer (400Mb ディスク)

Sun Microsystems 2/120 desktop unit X 2

TI Explorer LISP machine X 2

Intel iPSC/D6 (ハイバキューブ並列プロセッサ Intel 80286(512k バイト) X 64)

付属機器

Merlin robot, ASEA*IRB-6 robot, IRI*P256 vision system(SUN 170, SUN

120, Intel 310, DEC VAX 11/750, SONY DXC-M3 TVカメラ他) など

3 自然言語理解プロジェクト

C R Lにおける自然言語理解プロジェクトの内、特に機械翻訳と並列構文解析について述べる。

3.1 機械翻訳

X T R A (English Chinese Sentence TRAnslator) は一般分野での英語文を文法的に正しい中国語文に翻訳することを意図した、全自動翻訳システムである(2)。X T R Aは辞書、解析器および生成器から成り、以下では前2者について述べる。

1) 辞書

X T R AはC P R O L O Gで書かれている。辞書は意味素辞書、文法辞書、固定表現(熟語)辞書および意味辞書から成る。

意味素辞書

Y. ウィルクスによって提案された意味素(3)を基にして、14の意味クラスと各クラス毎に2、3の意味素が与えられている。

文法辞書

形態ならびに品詞情報を与えるもので、以下に幾つかの例を示す。

```
determiner(the)
noun(coach,[coach1,coach2])
noun(star,[star1,star2])
adjective(tough,[tough1,tough2])
verb(marry,[marry1,marry2])
```

注) []内は多義を示す。

固定表現辞書

熟語用の辞書で、例えばfor the time beingの場合、以下のように表される。
fixed_expression([for, the, time, being], [for_the_time_being])

意味辞書

幾つかの例を示す。

```
sem(coach1,[head(thing)]). (eg. 'a passenger coach')
sem(coach2,[head(man)]). (a trainer)
sem(star1,[head(thing)]). (a celestial object)
sem(star2,[head(man)]). (eg. 'a singing star')
sem(tough1,[poss(thing)]). (modifies 'thing' as in 'a tough material')
sem(tough2,[poss(man)]). (modifies 'man' as in 'a tough mountainer')
sem(marry1,[subj(man),obj(man),head(do)]). (eg. 'John married Mary')
```

2) 解析器

解析器はD C G (Definite Clause Grammer)を拡張したS D C G (Semantic Definite Clause Grammer)に基づいている。1、2の規則を示す。

```
sentence(Subj_Np, vp(v(Verb_sense), Obj_Np))→
noun_phrase(Subj_Np), is_verb(Verb), subject_verb_match(Subj_Np,
Verb, Verb_sense), noun_phrase(Obj_Np), verb_object_match(Verb_sense, Obj_Np). [1]
```

この規則は、もし入力系列の最初の句Subj_Npが名詞句で、次の語verbが動詞で、動詞の意味verb_senseがSubj_Npの意味と矛盾せず、次の句Obj_Npが名詞句で、かつVerb_senseがObj_Npの意味と矛盾しないなら、その入力系列はs(Subj_Np, vp(v(Verb_Sense), Obj_Np))

なる構造をもつ文であることを示している。

もう一つ名詞句に関する規則を示す。

```
noun_phrase(np(det((Det), adj(Adj_sense), n(Noun_sense)))→
  determiner(Det), adjective(Adjective), noun(Noun), adj_noun_match
  (Adjective, Noun, Adj_sense, Noun_sense). [2]
```

解析の手順を前述の辞書を用いながら、以下の例文に沿って述べる。

S1 The tough coach married a star.

解析器は規則[1]に従って、まず noun_phrase の解析を試みる。そのため規則[2]に注目する。the, tough, および coach の品詞がそれぞれ Det, Adjective および noun なので、規則[2]の適用のため adj_noun_match を試みる。tough と coach の意味としては、意味辞書によりまず tough1+coach1 の組み合わせが得られる。ここで coach1 の意味クラス thing は tough1 のスロット poss() を満足するので、tough1+coach1 は可能な解釈の一つであることが分る。

次に規則[1]に戻り is_verb の解析に移る。動詞 marry を見出すと、Subject_verb_match のチェックを試みる。しかし Subj_Np の意味 tough1+coach1 は marry1 においてスロット Subj() を満たさない。そこで解析器はバックトラックして再び規則[2]における adj_noun_match を行い、次に可能な解釈として tough2+coach2 を得る。この解釈は marry1 を満たす。解釈結果を以下に示す。

```
s
  type(decl)
  tense(past)
  aspect([])
  neg([])
  v
    marry1
      agent(np(det([the]), pre_mods([tough2]), n([coach2], post_mods([]),
        rel([])))
        object(np(det([a]), pre_mods([]), n([star1], post_mods([]))))
        pre_verb_mods([])
        verb_mods([]),
        post_verb_mods([]),
        adv_mods([])
```

以上のように解析器はトップダウン、デブスファーストでバックトラックしながら解析を進める。XTRA はサンワークステーション上に組み込まれている。システムの容量は語彙を除いて約 114 K バイトで、語彙数は約 1,000 である。1 文当たりの処理時間は秒または分のオーダーである。英中システムの他、英西システムも開発中で、更に英日システムも予定されている。

3.2 並列構文解析

CRL では、前節で述べた逐次構文解析を並列構文解析へと改良しつつある(4)。並列構文解析器は以下の六つのマスタから成る。

1) 文マスタ (S マスタ)

トップダウンで文の構文的骨組みを解析するマスタで、文頭モニタ (SH モニタ) と文体モニタ (SB モニタ) とから成る。SH モニタは文頭の副詞や前置詞句などを担当するもので、二つの SH ハンドラから成り、各 SH ハンドラは三つの子プロセッサを持つ。SB モニタは文本体の解析を担当するもので、三つの SB ハンドラから成り、各 SB ハンドラは五つの子プロセッサを持つ。SH モニタも SM モニタも SDCG の規則に応じて処理内容が異なってくる。例えば、肯定文に以下の規則 (簡単化している) が適用される場合、

```
sentence_body→ subj_np, vp [3]
sentence_body→ inverted sentence [4]
subj_np→ noun_phrase [5]
```

subj_np→ing_clause

[6]

S Bハンドラの子プロセッサの内、例えば1、2および3は、それぞれ

子プロセッサ1 : noun_phrase, vp
" 2 : ing_clause, vp
" 3 : inverted_sentence

なる解析を担当する。但し子プロセッサ1、2では[3]の右辺の解析のためsubj_npをそれぞれ[5]、[6]の右辺で置き換えてチェックを行っている。必要に応じて各プロセッサは他のマスタと交信する。

2) 名詞句マスタ (NPマスタ)

与えられた系列の最左部分で、文法的に正しいすべての名詞句を同定する。またANマスタ(後述)との交信により、adj_noun_matchに基づく、意味的に正しい名詞句を同定する。

NPマスタは五つのNPハンドラから成る。更に各々のNPハンドラは三つの子プロセッサを持ち、各子プロセッサは以下の解析を担当する。

子プロセッサ1 : 形容詞の修飾なし
" 2 : 一つの形容詞による修飾
" 3 : 二つ以上の形容詞による修飾

3) 形容詞*名詞マスタ (ANマスタ)

与えられた一つの形容詞と一つの名詞に対し、互いに矛盾しない意味(語義)の組み合わせをすべて同定する。

4) 主語*動詞マスタ (SVマスタ)

与えられた一つの動詞と、一つの名詞のある特定の意味とに対し、その動詞のすべての意味をチェックし、互いに矛盾しない意味の組み合わせをすべて同定する。

以上の他、意味辞書マスタと動詞*目的語マスタ(VOマスタ)とがある。

上記マスタによって例文S1が並列的に処理される過程を述べる。並列処理を明確にするため、toughの品詞および語義として以下を追加しよう。

noun(tough,[tough3])

sem(tough3,[head(man)])

与えられた入力系列S1はまずSマスタに引き渡される。SマスタはS1をSHモニタとSBモニタの両方に引き渡す。SHモニタは結果として解析に失敗する(文頭が空であるため)。SBモニタは五つのSBハンドラの内、例えばSBハンドラ1にS1を引き渡す。SBハンドラ1は子プロセッサ1~3にS1を引き渡す。それらの内子プロセッサ1はS1の最左句がnoun_phraseであることを知る必要がある。そこでSマスタはNPマスタを呼び、S1を引き渡す。

NPマスタは五つのNPハンドラの内、例えばNPハンドラ1にS1を引き渡す。NPハンドラ1は三つの子プロセッサにS1を引き渡す。子プロセッサ1は直ちに'The tough'をNPハンドラ、更にNPマスタに送り返す。子プロセッサ3は'The tough coach'をNPハンドラ、更にNPマスタに送り返す。

NPマスタは一方において'The tough'をSマスタに送り返し、他方において'The tough coach'をANマスタに引き渡す。ANマスタは結果的にtough1+coach1とtough2+coach2とをNPマスタに送り返し、NPマスタはこれらをSマスタに送り返す。

Sマスタは送り返された'The tough'と'The tough1 coach1'および'The tough2 coach2'とを順次処理して行く。まず'The tough'を待機中のSBハンドラ1の子プロセッサ1に引き渡す。子プロセッサ1は動詞coachを同定し、引続きsubject_verb_matchを試みるため、SVマスタとの交信に入る。

Sマスタは次に'The tough1 coach1'と'The tough2 coach2'それぞれに関し、SBハンドラ1の子プロセッサの中で受人可能なものがあれば、それらに引き渡し処理を進める。受人可能なものがなければ、SBハンドラ2の子プロ

セッサの中に受人可能なものがあるかどうかをチェックする。
 以上のように並列処理はトップダウンで左から右へと行われる。並列解析器は Intel iSPC/D6 上に組み込まれつつある。このマルチプロセッサは、各プロセッサが隣接する六つのプロセッサと直接通信できる。

4 人物の交流および研究会

CRL では人物の交流や研究会の開催に力を入れている。

4.1 コロキアムおよびセミナー

毎週外部から講師を招いてコロキアムが開かれている。1985年11月から1986年2月にかけて招かれた講師と演題の中からは2、3の例を示す。

R. シモンズ (テキサス大) : 機械翻訳のための工学的技術

F. ベレイラ (スタンフォード大) : 文法の性質についての演繹的計算

F. ハラリ (NMSU 客員教授) : 計算科学へのグラフ理論の応用

コロキアムには通常2、30人の研究者が集まる。質疑応答は活発で、講演の中にも盛んに質問が交わられる。講演内容は多少疑問があったり、質問者の意図を汲みきれないときは、しばしば議論が長引く。そのため最後の講演は深まり、講演者も有益な参考意見が得られる。反面、参加者の筆者は1985年12月から二月間、自然言語理解プロジェクトの研究者を対象として以下の特設セミナーを担当した。

題名: 自然言語および図形パターンの総合的意味理解

- 内容:
1. 述語の概念分析
 2. 自然言語概念による図形パターン理解
 3. 物語および動画の文脈解析
 4. 総合的意味理解システム

4.2 AIの基礎に関するワークショップ

1986年2月6~8の間、NMSUの近くのホテルで、「AIの基礎に関するワークショップ」が開催された。後援はNSF, AAAI, CRLおよびNMSUで、企画はCRLが担当した。組織委員長はCRLの、D. パートリッチ博士が努めた。彼は参加者に以下についての意見を求めた。

* AIの哲学的ならびに論理的基礎は何か?

* AIの基礎とプログラムの間の関係は何か?

* AIと他の研究分野の間の関係は何か?

以上に対し、米英の研究者約20名が意見を発表した。以下に幾つかを示す。
 R. ニードハム (ケンブリッジ大, UK) : AIについて何か特別なことがあるのか?

J. フォータ (MIT) : なぜなおかつ思考の言語がなければならないのか?

A. バンディ (エディンバラ大, UK) : AIとはいかなる類の分野か?

J. マッカーシ (スタンフォード大) : AIの推論は拡張性のある論理でなければならぬ。

A. ジョシ (ペンシルバニア大) : AIと言語学

D. ルーメルハート (カリフォルニア大) : AI: 心理学は何を学べるか?

M. ミンスキ (MIT) : AIと認知心理学

R. シャンク (エール大) : 認知心理学は本当に存在するのか?

中心人物の一人であったM. ミンスキが風邪で出席できなかったのは残念であった。以下に1、2の意見の内容を示す。

A. バンディは混迷状態にあるAIを次の三つの分野に分けることを提唱した。

- 1) 数学的AI: 知的行動を与える可能性のある計算技術の研究

2) 技術的 AI : 計算技術の、商業製品構築への応用

3) 科学的 AI : 計算技術の、認知モデルへの応用

彼は例として以下の類比を与えた。

計算科学 : 純粋数学

数学的 AI : 応用数学

技術的 AI : 工学

認知科学 : 理論物理

心理学 : 物理

A. ジョシはまず AI と自然言語処理の関係、および言語学における各分野を整理した。即ち AI において自然言語処理システムは、通常言語的知識、非言語的知識、領域毎の知識、および一般的なる推論戦略を用いる、より大きなシステムの中に組み込まれているもの、と考える。その上で

AI は言語学を有益なものを見たか？

言語学的手法は AI に影響を与えたか？

AI と言語学をより親密にする言語的様相があるか？

などの点を議論した。特に最後の点については言語生成の重要性を指摘した。

5 むすび

CRL の自然言語理解プロジェクトを中心に、米国における AI 研究の一端向を述べた。CRL ではニューメキシコ州の地域社会に立脚した研究課題、即ちビジョン/ロボティクスと人間-機械インタフェイスを設定し、明確な意識を持って AI システムの開発を行っている。また NMSU との関係においても基礎研究にも強い意欲を示している。更に国内および国外の研究者との交流にも力を入れ、かつ国際会議やワークショップなどに積極的に取り組み、この分野でのリーダーシップを取れるよう努力している。

我が国と米国とでは多くの面で事情が異なるが、本稿が、我が国において今後の自然言語理解研究の在り方を議論する際、その一助となることを期待する。

謝辞 お招き頂いた CRL 所長、Y. ウィルクス博士に謝意を表する。

文 献

- (1) A. Bunday: What Kind of Field is AI?, Proc. Workshop of the Foundation of AI, Las Cruces, p.29(Feb. 1986).
- (2) X. Huang: Machine Translation in the SDCG Formalism, Proc. Theoretical and Methodological Issues in MT of Natural Language, pp.135-144, Hamilton(Aug. 1985).
- (3) Y. Wilks: Preference Semantics, in Keenan(ed.), : Formal Semantics of Natural Language, Cambridge University Press(1975).
- (4) X. Huang and L. Guthrie: Parsing in Parallel, Memoranda in Computer and Cognitive Science, NMSU, No.85-40(1985).