音声言語情報処理 1－2
（1994. 5. 20）

頑健な自然言語処理の研究動向と課題

松本裕治　　今一 修
奈良先端科学技術大学院大学
情報科学研究科
matsu@is.aist-nara.ac.jp

伝統的な自然言語処理技術の多くは，システムがもつ辞書や文法によって定義される適格な文 (well-formed sentence) を対象としてきた．しかし，現実に直面するテキストや話し言葉には様々な非文法的な表現や誤りが含まれる．このような対象に対しては，文法的に適格な入力だけを仮定する訳にはいかない．実用的な自然言語処理システムは，システムの予測を越えるような入力にも対処できる能力を持つ必要がある．

本稿では，頑健な自然言語処理 (robust natural language processing) の動向を解説し，その問題点について考察する．自然言語システムの頑健性はいくつかの観点から考えなければならないことを指摘する．また，頑健な自然言語処理手法に関する著者らの考えについても述べる．

Current Issues in Robust Natural Language Processing

Yuji Matsumoto　　Osamu Imaichi
Graduate School of Information Science
Nara Institute of Science and Technology

Traditional natural language processing assumes an ideal situation where all input sentences are grammatically correct. However, in real life sentences, especially in spoken utterances, this assumption is hardly acceptable. Robust processing of natural language is quite important for any practical setting of natural language systems for coping with grammatically ill-formed inputs.

This article surveys the currently proposed techniques for robust natural language processing and highlights the research issues. We also introduce our research efforts in this field.

1 はじめに

自然言語テキストの電子化が急激な勢いで進み，大規模な自然言語テキストデータ (コーパス) の解析に関心が集まっている．例えば，ARPA が主催となって毎年開催されている Message Understanding Conference (MUC) では，ギガバイトオーダーの英語・日本語の新聞記事を解析して情報抽出を行なう研究の評価が行なわれている．

一方，これまでに提案されたほとんどの文法理論や自然言語解析手法は，文法的に正しい適格な文を対象としている．多くのシステムでは，対象領域を限定するか，前処理等によって入力文を修正しなければ，現実の文を正しく解析することができない．

頑健な自然言語処理が目指すのは，たとえシステムにとって完全な解析が不可能な入力に対しても何らかの処理結果を返すことである．処理の手法および処理結果はシステムによって異なり，入力に含まれる非文法性を修正することを目的とする場合もあれば，解析可能な部分だけを結果として返す場合もある．また，解析自体は可能であるが，曖昧性のためにシステムが正しい結果を判別できない場合も多く，曖昧性の解消を含めて頑健性を考えることが重要である．

ただし，本稿では，曖昧性の問題については，最小限の議論しか行なわない．簡単に言えば，本稿で扱う頑健性の問題は，システムが持っている言語に関する様々な制約を満足しない入力文をいかに理解するかという問題であり，曖昧性の問題は，複数の可能な解釈の中からいかに正しい解釈を得るかと言う知識またはヒューリスティックスの問題である．曖昧性解消のためのヒューリスティックスには，既に様々な提案がある [29][9][23]．また，語に関する曖昧性の解消についても，最近は枠組よりも現実的な方法の提案 (例えば，[24] は現実的に入手可能な様々な情報を利用した語義の曖昧性の解消の方法について述べている) や統計的な手法 (品詞のタガーや単語の共起関係の抽出) が多く提案されている．

次節では本稿が対象と考えている文法的不適格性について述べる．第 3 節では，不適格文に対する処理手法の分類と問題点の検討を行なう．続いて，第 4 節では我々の考え方について説明する．第 5 節で今後の研究の課題について述べ，最後の節でまとめを行なう．

2 文法的不適格性

頑健な自然言語処理システムについて論じる前に，文法的不適格性の定義を明確にしておく必要がある．文献 [21] において，著者は文法的不適格性を絶対的および相対的不適格性 (abusolute and relative ill-formedness) に区別する Weischedel[32] らの考え方を紹介した．人間にとっても文法性は普遍的なものではないが，人間が文法的な制約を逸脱していると認める現象を絶対的不適格性という言葉で表現し，たまたまシステムが持っている言語知識にカバーされていないために処理できない現象を (システムに対して) 相対的な不適格性と呼んで区別することにする．

自然言語処理の過程を，形態素解析，構文解析，意味解析，語用論解析に分けて考えた場合に，現実の不適格文に現れる現象をこれらの処理レベルによって分類してみたものが表 -1である．相対的不適格文は自然言語処理システムの処理能力に依存するので，ここで挙げたのは，あくまで典型的な例である．

並列句や慣用表現のように人間にとっては適格であるが，文法規則として記述するのが困難な言語現象もある．システムがこのような現象に対応できない場合は，これらの現象は相対的不適格性に分類される．また，我々が日常的に目にする文においても，英語の人称や数の不一致などのように誰の目から見ても文法的に正しくないにもかかわらず人間がその誤りに気がつかずに理解できる文もある．

表中で，中止文とは文末が不完全なまま終る文である．連続文とは追い込み文 (run-on sentence) とも呼ばれ，句点などが忘れられて二文が一つになってしまった文のことである．特殊構造には様々なものがあり，箇条書き，章や節番号，数式などのように言語以外の要素を含むような記述が現実の文章には多く存在する．不適格文を処理することの重要性は多くの研究者から指摘されている [31][8]．たとえ人間の推敲や校正を経た文章であっても，不適格文の存在を仮定しない訳にはいかない．

表 -1の中で，本稿が主として議論するのは，構文レベル以降の不適格性についてである．綴誤りや未知語の問題など形態素レベルの不適格性については，他の不適格性と独立に議論されることが多く，手法としても他の部分と関連が強くない．もちろん，このようなエラーを単独で処理するには限界があり，他の不適格性の処理と融合する方向の研究が重要である．

構文レベル以降の不適格性は，何らかの制約の違反として捕らえることができるので，レベルの違いに大きな影響を受けることなく解析手法について論じることができる．

	形態素レベル	構文レベル	意味レベル	語用論レベル
絶対的 不適格性	綴誤り タイプ誤り 区切り誤り 言い淀み	数・人称等の不一致 語順誤り・倒置 中止文，語の欠落 冗長文 (繰り返し) 連続文	必須格の欠落 選択制約違反 比喩・換喩	協調の原則違反 照応のための 情報不足 電報文 (領域依存性)
相対的 不適格性	未知語 省略表現	システムの能力を 越えた構文構造 特殊構造 (箇条書きなど)	語彙知識不足 世界知識不足	(文脈的要因による) 断片文・省略文

表 1: 不適格性の分類

3 文法的不適格文の処理手法

多くの自然言語処理システムでは辞書や文法規則が用意され，システムはそこに記述された制約 (規則) に従って入力を解析する．不適格文の処理が重要になるのは，システムがもつ制約を入力が満足しなかった場合に何の結果も返さないことを避けたいからである．

不適格文を処理するアプローチとして，不適格文の構造を予測し，適格文を記述するのと同じ方法でそれを記述するものもあるが，処理に無駄があり大きなシステムの実現は困難である．また，このような手法では，システムが想定した不適格性しか扱えないと言う欠点がある．不適格文の可能性は与えられた文が適格文として捕らえられない場合にのみ考える方がすぐれている．ただし，上で述べたように文が曖昧性を伴い複数の解釈の可能性を残す場合には，そのうちのどの解釈が最も妥当な解釈であるかを判断する必要がある．不適格文の解析と密接に絡む問題として，比喩の理解の問題があるが，ここでは取り上げない．

システムが予期しない入力に対処できるようにシステムを頑健なものにするための手法を次の二つに分けることができる．

1. 統計的な手法やテンプレートマッチング的な解析を行なうなど画一的な処理のみを行ない，どのような入力に対しても何らかの結果を返すことを保証する．これは従来の規則に基づく自然言語処理に比べると，言語をより浅いレベルで捕らえることによって頑健性をを増す方向の研究であると言える．

2. 解析が失敗するのは，従来の自然言語処理における文法規則や単語における言語情報の記述の詳細度が不充分なためであるとする考え方．より木目の細かい情報を与えることによりシステムの頑健性が増すと考える．ただし，システムが持っている言語的な知識の制約を強めるという意味ではなく，不適格文を処理するためにどの制約を緩和すればよいかをいかに記述し実現するかが興味の対象となる．

本稿で対象とするのは，上の 2 つめの方向である．さらにこの視点からは不適格文の処理のためのアプローチを次のように分類することができる．

パターンマッチング 文中の意味のある断片を発見すること．また断片を用いたパターンマッチングによるアプローチ [2][8]

部分解析 全体の解析に失敗したとしても，その時点までに解析された文中の意味のある句や節を捕らえ，それを中心に意味の抽出を試みるもの [13][22][10]

チャート法の拡張 チャート法などの構文解析アルゴリズムの拡張による構文的制約を緩和することによって正しい構文構造を得るアプローチ [7][15][26][30][25]

緩和法 適格文のみを解析する文法にしたがって解析を行ない，失敗時にエラー回復のため制約を緩和する別のメカニズムを用意するもの (緩和法, Relaxation method)[19][32]

段階的緩和 複数のレベルの制約を用意し，段階的に制約を緩和する手法 [3][18]

連続的解釈 意味的また構文的な制約を連続的なものと考え，適格文と不適格文の区別を非連続なものと捕らえない考え方に基づくもの [4][28][6]

意味主導 意味的な情報を構文情報より優先する．意味的情報にガイドされながら構文規則の選択が行なわれる [20][17]

アブダクション 文の解釈のための一般的な枠組を用意し，文の意味理解の過程として不適格文の理解を考えるもの [10][11]

不適格文の処理方法には様々なものが考えられているが，多くの方法ではどの程度の不適格文を処理する能力があるかが明確でないことが多い．IBM の Jensen らは，ともかくすべての文に対して何らかの結果を返す Fitted Parse という考え方を提案している [13]．適格文のみを記述した核文法 (core grammar) を用いて文を上昇型並列に解析し，解析が成功しなかった場合，それまでに得られた構造の中から一番もっともらしい構造の列を選び，それをその後の処理の入力とする．この操作が fitting 処理と呼ばれる．Fitted Parse では，最初に主構成素 (head constituent) を次の優先順位で求める．

1. 主語を伴う時制のある動詞句
2. 主語がなく時制のある動詞句
3. 動詞以外の句 (名詞句，前置詞句など)
4. 不定形の動詞句

これによって入力文全体を覆えない場合は，次の優先順位で残りの構成素を主構成素に付けていく．

1. 動詞句以外の構成素
2. 時制のない動詞句
3. 時制のある動詞句

この方法の特徴は，解析システムを引き続く処理の前処理として位置付け，上のようなヒューリスティックスを使って，ともかく何らかのもっともらしい結果を与えることである．McDonald[22] や Hobbs[10] も，部分的に成功した解析結果を使って，それを中心に文の復元を試みたり，その一部からだけでも有用な情報を抽出することを提案している．

その他のアプローチは，文の不適格性の原因を推定し，それが侵している制約を緩和 (relaxation) することによってその原因の同定と修正を行なう方法を取ることが多い．したがって，適格文を対象とした文法によって文の解析を行ない，それが失敗した段階で，不適格文の同定の処理が行なわれる．

Kwasny ら [19] および Weischedel ら [32] は，Woods の ATN に緩和手法を導入して不適格文の処理を実現している．

緩和法として代表的な Weischedel らの方法を取り上げてみよう．彼らの方法では，ATN によって通常の文法が記述され，不適格文を扱う規則はメタ規則と呼ばれている．メタ規則は，ATN の各状態に関する条件の列をもち，それらが満足されたときに実行すべき動作を記述したプロダクション規則である．ATN による通常の解析が失敗すると，メタ規則が呼び出される．メタ規則は ATN が用いていた適格文のための文法の様々な制約を緩和するための規則である．メタ規則には，この節の最初で分類したほとんどすべての不適格性が考慮に入れられている．代表的なものとして，次のようなメタ規則がある．

- 構文関係のメタ規則
  - 一致条件 (agreement): 数，人称，性の一致
  - 名詞句における冠詞の欠落
  - 置き換わり易い語 (例えば good と well)
  - 関係詞中に現れるはずのない痕跡の要素が現われてしまう誤り
- 意味関係のメタ規則
  - 意味制約の緩和 (擬人化，換喩)
  - 語順条件の違反 (倒置)
  - 格フレームの必須格が表面上現れない

メタ規則として不適格文の処理を統一的に取り扱ったことは優れているが，ATN を対象にしているため，下降型かつ左から右への後戻り解析を行なうことになる．解析が失敗した時にどの部分に対してメタ規則を適用すればよいかが必ずしも明確ではない．また，ATN は後戻りを伴うアルゴリズムによって動くため，同じメタ規則が繰り返して適用されてしまうことがあり得る．

緩和法を段階的に用いる方法が，Douglas ら [3] によって提案されている．PATR のような単一化文法を対象にし，そこに記述された制約条件にいくつかのレベルを設けることを提案している．解析が失敗すると，制約を段階的に緩和していくことにより，不適格文の処理と，不適格性の原因の同定を行なうことができる．制約条件を素性構造のような統一的なデータ構造で記述することにより，個別のメタ規則を持つ一般的な緩和法に比べて優れているが，処理可能な不適格性の範囲は広くない．

構文的な制約の緩和については，Mellish[26] によるチャート法 [16] を拡張した不適格文処理がよく知られている．彼の方法が扱っているのは，語の欠落と冗長な追加という構文レベルの不適格性だけである．基本的には

上昇型チャート法によって文の解析が行なわれ，それが成功しない場合に下降型の規則適用と語彙カテゴリの挿入や削除が行なわれる．加藤 [15] は一般の上昇型解析が失敗した場合に，なるべく多くの部分解析結果が上昇型処理によって得られるように，文法規則の左隅以外のカテゴリからの活性弧生成や活性弧同士の結合を行なうことを提案している．得られる情報の増加によって誤りの予測精度は向上するが，同じ文法規則を重複して展開する可能性がある．

適格文を解析する文法だけではなく，色々な不適格性を判断する不適格文用の文法を用意しておくという考え方は素直な考え方であるが，それらをいつどのように用いるかと言うことは自明ではない．工藤ら [18] は，複数のレベルの文法を段階的に用いることにより，不適格性を含む文の解析とその原因の同定を行なうことのできる枠組を提案している．

不適格文の可能性は与えられた文が適格文として捕らえられない場合にのみ考えればよいのだが，曖昧性によって文が複数の解釈の可能性を残す場合には，そのうちのどの解釈が最も妥当な解釈であるかを判断する必要がある．そういう意味で，そもそも適格文，不適格文と文を分類すること自体に問題があると言う考え方も成り立つ．文の不適格性や換喩を優先意味論の枠組で扱い，二元論的な分類を行なわずに頑健な自然言語処理を実現しようという研究が Wilks ら [4] によって行なわれている．

Hobbs らの提唱するアブダクションに基づく文の解釈の研究 [10][11] は，不適格文の解析のための新しい枠組として可能性を秘めた考え方である．文の解釈は，アブダクティブな理由付けによって解釈されて理解されると考えると，この考え方は，上で考えた構文的な不適格文の救済法や緩和法に基づく不適格文処理などの一般的な拡張とも考えることができる．また，彼のアブダクションの枠組ではコストという考え方が採り入れられており，これにより処理の順序がコントロールされることになる．つまり，処理順序のヒューリスティックのための情報が合わせて記述されることになる．コストをどのように割り当てるか，また，コストを動的に決定する仕組みがあまり考えられていないなど，今後詳細化されるべき要素が多く残っている．また，アブタクティブに用いられる規則 (論理式による記述) の妥当性についても論じるべきであろう．

4 少数の原則に基づく不適格文の処理

前節で様々な不適格文処理の手法を説明したが，多くの手法の考え方には，解析の失敗の原因となった制約を緩和することによって不適格文の処理をするという考え方が背景にある．Weischedel ら [32] の方法についていくつかの欠点を指摘したが，メタ規則によってどのような緩和規則も記述できるため，事前に想定できる不適格性に制限はない．しかし，その反面，事前に想定されない不適格性は扱うことができない．また，後戻り処理を伴う ATN 上に実装されているため，メタ規則の適用場所が必ずしも解析失敗部分と一致しない．したがって，正しいメタ規則が正しい位置で適用される保証がない．

制約の段階的な緩和を行なう手法やアブダクションによる手法では，このような欠点が一部解消されているが，緩和処理は，制約の強弱やコストなど，本来の不適格性処理とは関連が強くないかも知れない局所的な情報にコントロールされて処理が進むことになる．部分解析やチャート法の拡張におけるトップダウン予測のような，より目的指向の手法が必要である．

Chomsky の生成文法理論が最近の GB 理論のように原理とパラメータの理論として，個別の説明のできない規則を極力廃して来たように，個別のメタ規則に依存する不適格文処理の枠組は好ましいものではない．緩和処理に基づく不適格性の処理もそのような方向に進むべきであろう．このような考え方に基づいて，我々は，少数の原理に基づく頑健な自然言語処理の枠組の研究を行なっている [12]．本節では，これについて説明する．

どの手法もそうであるように，我々の方法も何からの文法的な枠組を想定しなければならない．我々は，下位範疇化情報によって基本的な構文規則を規定した単一化文法を仮定し，構文情報，意味情報とも単語レベルにタイプつき素性構造によって記述されているする．日本語の場合には，用言や付加構造 (adjunction)，複文，埋め込み文などを記述するための最低限の文法規則が存在するが，単語に関する構文的また意味的な情報，特に，その単語がどのような句と結び付いて完成した句になるかを示す下位範疇化情報が個々の単語に与えられているとする．また，タイプ階層により，より一般的な概念が持つ性質は継承するものとする．また，句構造は $\overline{X}$ 構文に従うとしており，すべての句は，同一範疇の単語の投射であると考える．

文の解析は，全体的には文法規則を満たすように解析が進み，文中の個々の単語や句は，自分が完成した (下位範疇化情報が満足された，また，さらに付加句 (adjunct) と結合した) 句になることを目標として緩和処理を行なう．解析中の文のどの部分が緩和処理のライセンスを与えられるかはグローバルな戦略によって決まり，ライセンスが与えられた句は局所的な状況を手がかりとして自

発的に動作する.

我々が想定している文法に関する少数の原則と処理のためのヒューリスティックスに対する原則を簡単にまとめると次のようになる.

文法と語彙についての原則

- 文法は用言に関する規則，複文や埋め込み文などの節を越えた少数の基本的な規則からなる.
- $\overline{X}$ 構文の考え方に従い，あらゆる句や節はある品詞 (名詞，形容詞，動詞など) の投射であると仮定する．また，Head 以外の要素はすべて何らかの語の最大投射である.
- Head となる語や句は，その argument を下位範疇化情報として持つ.
- adjunct に関する情報はタイプ階層を通じて，より一般的な概念から継承される (例えば，動作動詞は時間に関する副詞句を adjunct として持ち得るなどの情報を個々の動作動詞が継承する)

処理のための原則

- 入力文全体を覆う最大投射があれば，それを解析結果の候補とする.
- より広い範囲を覆い，かつ，下位範疇化情報が満たされていない句を処理の対象とする.
- 満たされない下位範疇化情報を持つ句に緩和処理のライセンスを与える．ただし，文を完成するために，満たされない下位範疇化情報を持つ動詞句が最初に高い優先度を与えられる.
  - 隣の句 (または近隣の句) が，下位範疇化情報によって予測される句と同様の投射である場合に，その句に緩和処理のライセンスを与える.
  - その際の (構文的，意味的) 制約の満たされ方に応じて，ペナル ティを設ける.
- 未知語については，下位範疇化情報によって予測される句の lexical head と仮定する.
- タイプ階層を継承して得られる情報は，継承の距離に応じてペナルティを設ける.
- 下位範疇化情報を満たされない未完成な句が、完成した句となるために必要な句をその文内から見つけられない場合、文脈から必要な句を探す。

このような文法および処理における少数の原則を仮定して，次のような例文がどのように解析されるかを見ることにする.

例: この車はよくガソリン食うね.

仮定として，「食う」という語は主語と目的語となる後置詞句 (格助詞を伴う名詞句) を下位範疇化し，それぞれ人間および食物という意味的な制約をもつとする.

1. 通常の解析の結果，「食う」の下位範疇化情報は何も満たされないが，動詞であるため，最初に緩和処理のライセンスを得る.
2. 隣の句として「ガソリン」がある．「食う」によって下位範疇化されるために，その意味制約と後置詞句という条件を目標として緩和処理のライセンスが「ガソリン」に与えられる.
3. 「ガソリン」は，目標を満たすためのペナルティを計算し，自身はその目標を満たす句になる.
4. 「ガソリン食う」という未完成の句が「この車は」という後置詞句に緩和処理のライセンスを与える.

最終的には，主語と目的語のそれぞれの意味的制約条件が緩められ，しかも，目的語については助詞が欠落した文として解析結果が得られる.

本手法は，現在のところ小規模な実験に留まっている．現実の文章を解析するためにどの程度の原理原則を設定すれば充分であるか，また，実用に耐える程度の解析処理時間を達成できるか，などの問題を解決していかなければならない.

5 頑健な自然言語処理の課題

これまでの節で述べられなかった未解決の問題点についてまとめてみたい．頑健な自然言語処理が何らかの意味で制約の緩和処理であるという立場に立つとしよう．制約をどのような順序で緩和していけばよいか，という問題があるが，これについては，制約に順序を与える，コストを設定する，目標を達成するための緩和を優先する，などの様々な方法がある.

問題は，どこまで制約条件を緩めて良いか，ということである．緩和処理を無制限に許すならば，人間にとって理解不能な文に対してもシステムは何らかの解を与えてしまうだろう．どの程度の逸脱が自然でどの程度が不自然であるかの評価はまだほとんどなされていない.

頑健なシステムを達成するためには，単語や文法に関する情報を詳細にする必要があることを述べた．しかし，

単語が持つ構文的また意味的な記述をいかに行なうべきか，については，共通の認識があるとは言えない．語義に関する記述については，意味マーカやシソーラスだけでなく，KL-ONE[1] のようなより一般的な知識表現が想定される場合もあるが，しかし，それでも充分でないと言う指摘もある．例えば，Pustejovski の Generative Lexicon[27] では，語について，それが意味する概念の概念階層における記述だけでなく，それが存在し得るようになった起源に関する知識や典型的な役割についての知識 (例えば，「本」なら，それが書かれることによって存在し，読まれることによって機能するという知識) が必要であることを指摘している．

どの程度の知識が記述されれば充分かという問題は，永遠に論じられる問題かも知れない．しかし，たとえそれが解決されたとしても，そのような知識を何十万という語に対してどのように記述するかという量的な問題は解決されたとは言えない．語の意味記述を (半) 自動的に獲得 (学習) する方法の研究が極めて重要である．大規模コーパスの利用が現実的に可能になりつつあり，コーパスからの知識獲得の研究は今後盛んになるのは間違いないだろう．

6 おわりに

頑健な自然言語処理について，特に，文の解析レベルにおける不適格文処理の現状を概観した．また，この問題に対する我々の考え方についても述べ，今後重要と思われる点についても述べた．

また，辞書や知識の表現なども自然言語処理に用いる限りはその量の問題を解決しなければならない．コーパスを対象とした言語現象解析には統計的な方法が現在主として用いられている．しかし，今後は，コーパスを対象にしたより学習に関する研究が重要である．

様々な言語現象に対応した解析法や，不適格文の解析法についても優れた手法が数多く考えられている．これらの技術を用いて，いかに見通しよく大規模なシステムを作っていくことができるかが問題である．システムのモジュール性，単調性，などが重要な点であり，我々の提案している少数の原理に基づく処理は，そのための一つの方法であると考えている．

参考文献

[1] Brachman, R.J. and Schmolze, J.G.: An Overview of the KL-ONE Knowledge Representation System, *Cognitive Science*, Vol.9, No.2, pp.171-216, 1985.

[2] Carbonell, J.G. and Hayes, P.J.: Robust Parsing Using Multiple Construction-Specific Strategies, *Natural Language Parsing Systems*, Leonard Bolc (ed.), pp.1-32, Springer-Verlag, 1987.

[3] Douglas, S. and Dale, R.: Towards Robust PATR, *COLING-92*, pp.468-474, 1992.

[4] Fass, D. and Wilks, Y.: Preference Semantics, Ill-Formedness, and Metaphor, *Computational Linguistics*, Vol.9, No.3-4, pp.178-187, 1983.

[5] Fink, P.K. and Biermann, A.W.: The Correction of Ill-Formed Input using History-Based Expectation with Applications to Speech Understanding, *Computational Linguistics*, Vol.12, No.1, pp.13-36, 1986.

[6] Genthial, D., Courtin J. and Kowarski, I.: Contribution of a Category Hierarchy to the Robustness of Syntactic Parsing, *COLING-90*, Vol.2, pp.139-144, 1990.

[7] Goeser, S.: Chart Parsing of Robust Grammars, *COLING-92*, pp.120-126, 1992.

[8] Hayes, P.J. and Mouradian, G.V.: Flexible Parsing, *American Journal of Computational Linguistics*, vol.7, no.4, October-December, pp.232-242, 1981.

[9] Hobbs, J.R. and Bear, J.: Two Principles of Parse Preference, *COLING-90*, vol.3, pp.162-166, 1990.

[10] Hobbs, J.R., et al.: Robust Processing of Real-World Natural-Language Texts, *3rd Applied Natural Language Processing*, pp.186-192, Treno, Italy, 1992.

[11] Hobbs, J.R., et al.: Interpretation as Abduction, *Artificial Intelligence*, Vol.63, No.1-2, pp.69-142, Oct. 1993.

[12] 今一 修，松本裕治: 少数の原理に基づく頑健な自然言語処理, 第 8 回人工知能学会全国大会, June 1994.

[13] Jensen, K. et al.: Parse Fitting and Prose Fixing: Getting a Hold on Ill-formedness, *Computational Linguistics*, Vol.9, No.3-4, pp.147-160, 1983.

[14] Jensen, K. and Bonot, J-L.: Disambiguating Prepositional Phrase Attachments by Using On-Line Dictionary Definitions, *Computational Linguistics*, Vol.13, No.3-4, pp.251-260, 1987.

[15] 加藤恒昭: 非文の解析－チャートに基づく新たな手法, 情報処理学会 自然言語処理研究会 83-10, 1991.

[16] Kay, M.: Algorithm Schemata and Data Structures in Syntactic Processing, *XEROX PARC Technical Report*, CSL-80-12, Oct. 1980.

[17] Kirtner, J.D. and Lytinen, S.L.: ULINK: A Sementic-Driven Approach to Understanding Ungrammatical Input, *AAAI-91*, pp.137-142, 1991.

[18] Kudo, I., et al.: Schema Method: A Framework for Correcting Grammatically Ill-Formed Input, *COLING-88*, Vol.1, pp.341-347, 1988.

[19] Kwasny, S.C. and Sondheimer, N.K.: Relaxation Techniques for Parsing Grammatically Ill-Formed Input in Natural Language Understanding Systems, *Computational Linguistics*, Vol.7, No.2, pp.99-109, 1981.

[20] Lytinen, S.L.: Semantic-first Natural Language Processing, *AAAI-91*, pp.111-116, 1991.

[21] 松本裕治: 頑健な自然言語処理へのアプローチ, 情報処理, Vol.33, No.7, pp.757-767, 1992.

[22] McDonald, D.D.: An Efficient Chart-based Algorithm for Partial-Parsing of Unrestricted Texts, *3rd Applied Natural Language*, pp.193-200, 1992.

[23] McRoy, S.W. and Hirst, G.: Race-Based Parsing and Syntactics Disambiguation, *Cognitive Science*, Vol.14, pp.313-353, 1990.

[24] McRoy, S.W.: Using Multiple Knowledge Sources for Word Sense Discrimination, *Computational Linguistics*, Vol.18, No.1, pp.1-30, 1992.

[25] Meknavin, S., Theeramunkong, T. and Tanaka, H.: Parsing Ill-Formed Input with ID/LP Rules, *Fourth International Workshop on Natural Language Understanding and Logic Programming*, pp.158-171, 1993.

[26] Mellish, C.: Some Chart-Based Technique for Parsing Ill-formed Input, *ACL-89*, pp.102-109, June 1989.

[27] Pustejovsky, J.: The Generative Lexicon, *Computational Linguistics*, Vol.17, No.4, pp.409-441, 1991.

[28] Schank, R.C.: *Conceptual Information Processing*, North Holland, 1975.

[29] Schubert, L.K.: On Parsing Preference, *COLING-84*, pp.247-250, 1984.

[30] Stock, O.: Parsing with Flexibility, Dynamic Strategies, and Idioms in Mind, *Computational Linguistics*, Vol.15, No.1, pp.1-18, 1989.

[31] Weischedel, R.M. and Black, J.: Responding Intelligently to Unparsable Inputs, *American Journal of Computational Linguistics*, Vol.6, No.2, pp.97-109, 1980.

[32] Weischedel, R.M. and Sondheimer, N.K.: Metarules as a Basis for Processing Ill-Formed Input, *Computational Linguistics*, Vol.9, No.3-4, pp.161-177, 1983.