コーパスからの関係表現の自動抽出

新納浩幸† 井佐原均††

本論文ではコーパスから関係表現を自動抽出する手法について述べる．関係表現とは「に関して」に代表される，助詞相当の働きを持ち，語の挿入や交換が一般に行えない慣用表現の一種である．関係表現は一般に一語として処理するのが有効であるが，その表現を収集することは容易ではない．なぜなら通常の表現と関係表現との違いは不明確であり，つきつめれば，その判定はシステム製作者の主観的な判断によって行われているからである．本論文ではコーパスから関係表現を自動抽出することで，網羅的，かつ統一的な関係表現の収集を目指した．特に本論文では，助詞＋動詞＋付属語（助詞，助動詞）の形を持つ関係表現を抽出することを試みた．本論文は上記の関係表現のもつ２つの特徴に注目する．１つは関係表現中の動詞は，接続的な利用が多く本動詞として利用されることが少ないこと，もう１つは，その動詞に前置する助詞との共起が強いため，動詞に前置する助詞は特異な出現頻度をとるという特徴である．この特徴を利用してまず関係表現中の動詞になりえるものをコーパス中のその語の使われ方の頻度から選出する．次に選出した動詞に前置する助詞をコーパスから収集し，助詞の出現分布を調べることで関係表現を抽出する．

Automatically Extracting Connective Expressions from Corpora

Hiroyuki Shinnou† and Hitoshi Isahara††

Connective expression is a kind of idiomatic phrase corresponding to postposition. Typical example is "NI-KANSI-TE". Generally it is advantageous to handle these expressions as one word, but it is difficult to pick up these expressions. Because it is unclear to distinguish connective expression from normal expression. This paper aims at picking up automatically connective expressions from corpora. We utilize two properties of connective expressions. One is that the verb in connective expression is almost used as connective form and rarely used as main verb. Another is that the verb and postposition in front of the verb are strongly connected. Counting forms of all verbs in corpora, we can pick up verbs which are able to become a part of connective expression. Next, examining the number of postpositions in front of these verbs. Lastly we can extract connective expressions.

1. はじめに

自然言語処理では，入力文を単語列に分解し，各単語を文法規則によって節や句などにまとめあげ，最終的にそれらから全体の構文意味構造を構築するのが普通である．しかしある一連の単語列に対しては，それらを一つの単語として扱う方が処理の効率や精度の面で有利である．

その典型的な例が一語性慣用表現である[1]．一語性慣用表現とは「途方もない」や「に関して」のように，他の語の挿入や語の交換が通常許されない慣用表現の一種であり，そのような表現に対してはその表現を一語として扱うのが妥当である．なぜなら，そのような表現は，個々の構成語の意味からその表現の意味を作り出すことができず[2]，しかもそれらの表現は既存の品詞を持つ単語とほぼ同様な統語的振舞いをするからである．仮にそれらの表現を各単語に分解して処理するとしたら，その表現の統語構造を作り出すことには実質意味がないし，無用な統語的曖昧性も生じることになり，結果的に解析の効率，精度の面で望ましくない．

一語性慣用表現を一語として処理することは妥当ではあるが，その表現を収集するのは容易ではない．なぜなら通常の表現と一語性慣用表現との違いは不明確だからである．例えば「に関して」が「個々の構成語の意味からその表現の意味を作り出せない」かどうか

† 茨城大学工学部システム工学科
Department of Systems Engineering, Faculty of Engineering, Ibaraki University
†† 電子技術総合研究所知能情報部自然言語研究室
Natural Language Section, Machine Understanding Division, Electrotechnical Laboratory

は，そのシステムの持つ「関する」の辞書情報や文法規則に依存している．つきつめればシステム製作者が「に関して」を一語性慣用表現と考えるかどうかの主観的な判断でしか，両者を区別することができない．この主観的な判断の必要性から，慣用表現の収集，整理は人手による多大な労力を要するとともに，収集した表現の網羅性，統一性にも疑問が残る．

この問題に対処する1つの方法として，コーパスから機械的に慣用表現を抽出することが考えられる[3),4)]．コーパスを大規模にすることである程度の網羅性は保証できる．また機械的に取り出すためにその判断は客観的であり，収集したものにある種の統一性も認められる．

機械的に取り出すためには，それら表現が共通して持つ，ある客観的な特徴を発見する必要がある．我々は一語性慣用表現の中で特に(1)の形をした関係表現と呼ばれる助詞相当の語句の場合，

助詞＋動詞＋付属語（助動詞，助詞） (1)

(例)「に関して」⟶に＋関する＋て

以下のヒューリスティックス（特徴ともいえる）があることを予想した．

(A) (1)の中の動詞（以下この部分を関係表現の核と呼ぶ）の一般的な使われ方は本動詞（「関しました」）としてよりも，接続的用法である連体形（「関する」）あるいは連用接続形（「関し」「関して」）の形で使われる頻度が圧倒的に多い．

(B) (1)の形をした関係表現は，核と直前の助詞との間に非常に強い結合関係がある．

本論文では，これらのヒューリスティックスを基に(1)の形の関係表現の自動抽出を試みる．またその結果から上記のヒューリスティックスの妥当性も判断する．

関係表現には(1)以外の形をとるものもあり，それらは本手法によって抽出することはできない．しかし(1)の形をした関係表現は特に数も多く，また核が動詞であることから，この表現を単語に分解した場合の処理が複雑になり，関係表現として抽出する効果が大きい．このために本論文では(1)の形をした関係表現をまず対象にした．

本論文の手法は，最初にコーパス中の全動詞に対してその使われ方の頻度を調べ，ヒューリスティックス(A)を用いて，核になりえる動詞を抽出する．次に抽出した核に対応する(1)のパターンの表現をコーパスから取り出し，核の直前の助詞を収集し，その出現頻度から前置する助詞を決定する．この2段階の処理によって関係表現を抽出する．このようにして取り出せる表現は，通常考えられる関係表現を多く含むものと予想する．

本論文では本手法を用いて朝日新聞の記事1か月分（約10 Mbyte）のコーパスから関係表現の抽出実験を行った．この実験結果についても述べる．

2. 関係表現の自動抽出

2.1 抽出する関係表現の形

本論文では以下の4つの形をした関係表現を対象とする．

(a) 助詞＋動詞（連用形）＋助詞「て（で）」
(b) 助詞＋動詞（連用形）
(c) 助詞＋動詞（連体形）
(d) 助詞＋動詞（連用形）＋助動詞「た」（連体形）
(例) 「に関して」…(a)のパターン
「に関し」…(b)のパターン
「に関する」…(c)のパターン
「に関した」…(d)のパターン

上記した例のように，1つのタイプから別のタイプを推測することは可能であるが，ここでは(a)〜(d)の各々を独立して取り出すことにする．

2.2 核と助詞の推定

関係表現を機械的に抽出する際に最も困難な点は，(a)〜(d)の形をした品詞列が関係表現であるか通常の表現であるかの判定である．例えば「を招待して」は，(a)のパターンになっているが，これを関係表現と考えるには無理がある．本論文では，関係表現の核になる動詞の使われ方の頻度，および，動詞の直前の助詞の頻度分布を調べることの2段階の処理によりこの問題に対処する．

2.2.1 核の推定

関係表現は助詞相当の意味を持つことを考えれば，核の動詞には通常の関係子とほとんど同じ，あるいは多少情報を加えた意味，あるいは機能を持つものが使われると考えられる．そのような意味や機能を持つ動詞は，(a)〜(d)の形に代表される接続的な使われ方の頻度が高く，本動詞として利用される頻度は少ないと予想できる．これは1章で述べたヒューリスティックス(A)である．このヒューリスティックスを使って，まず核になりえる動詞を抽出する．

具体的には，まず，コーパス中のすべての動詞に対して，その動詞が

(a′) 接続助詞「て(で)」に前置する連用形
(b′) 助動詞のつかない連用形
(c′) 連体形
(d′) 助動詞「た」の連体形に前置する連用形
(e′) 上記以外

のうちどの使われ方がされているかの回数を計測する．ただし，(a′)の接続助詞「て(で)」の後に「いる」「みる」「おる」などの補助動詞が使われている場合は，(a′)ではなく，(e′)をカウントする．これは補助動詞がついた場合，核の動詞は実質的に本動詞となっているからである．例を示す．**図1**はコーパス中の一部を解析処理した結果である．下線の引いている部分が動詞であり，その下線に付けられている記号が上記(a′)～(e′)の分類を表している．図1の下線部分の情報から出現動詞を**表1**のようにカウントしてゆく．これをコーパス中の全文に対して行う．

コーパス全文に対して出現動詞を検査した後，表に存在する各々の動詞に対して，その動詞の全体の出現回数 (a′+b′+c′+d′+e′=SUM) に対する接続的な用法の出現回数 (a′+b′+c′+d′) の割合を求める．ここではこの割合を接続割合と呼ぶことにする．この接続割合がある基準値（本論文では 75% とした）よりも大きければ，その動詞を核になりえる動詞と推定する．ただし出現回数がある基準値（本論文では 10 個とした）以下のものは，接続割合にかかわらず，核にはなりえないと判断する．

今は中東危機を頭に描いて作業している.
(a') (e')
もう一歩進んで，カンボジアの問題を考えると
(a') (e')
和平に関する東京会議を開くなど汗を流している.
(c') (e') (e')

図 1 解析例
Fig. 1 Example of analysis.

表 1 動詞の出現回数表
Table 1 Counting table of used verbs.

	a' の形	b' の形	c' の形	d' の形	e' の形	合計 (SUM)
描く	1	0	0	0	0	1
作業する	0	0	0	0	1	1
進む	1	0	0	0	0	1
考える	0	0	0	0	1	1
関する	0	0	1	0	0	1
開く	0	0	0	0	1	1
流す	0	0	0	0	1	1

2.2.2 助詞の推定

日本語は格要素の出現順序が自由なので，核の部分が通常の動詞であれば，直前の助詞の頻度が特定のものに偏るということは少ない．例えば核が通常の動詞「招待して」である場合に，前置する連用句は「先生を」や「我が家に」や「昨日」など様々なものが現れ，その結果として直前の助詞は「を」や「に」あるいは助詞はない，など様々な場合が起こる．一方，関係表現では，構成要素間の結合が強く，語順が不変で，副詞などの挿入も一般に起こらない．これらのことから，(1)の形をした表現が関係表現であれば，その核の直前の助詞を調べると，ある特定のものに偏るということが予想される．例えば「関して」の直前の助詞を調べると「に」が非常に大きな頻度を占めることが予想される．これが1章で述べたヒューリスティックス(B)である．このヒューリスティックスを利用すれば，先に求めた動詞の直前の助詞の分布を調べ，その動詞を核とする関係表現が実際に起こりえるかどうか，また起こりえる場合はその直前に位置する助詞は何かを求めることができる．

具体的には，先の処理によって核になる動詞を推定した後，その動詞を核に持つ(a′)～(d′)の形の表現を各々コーパス中から取り出す．ただし，例えば(a′)の出現回数がその動詞の全体の出現回数 (SUM) に対して，ある基準の割合（本論文では 10% とした）以下であったら，その動詞の(a)の表現は関係表現にならないと判断し，助詞の収集は行わない．(a′)～(d′)に対して各々取り出した直前の助詞の頻度を調べ，最も大きな頻度を持つ助詞がある基準値（本論文では85%とした）を越えた場合，それを核に前置する助詞として，関係表現を生成する．例えば，先の処理によって動詞「対する」が全体の出現回数200回で核になりえると判断されたとする．その200回のうち(a′)の形が100回であったとする．この時，出現回数が基準値(10%)よりも多いので，「対する」の(a′)の形を直前の助詞とともにコーパス中から取り出す．結果は100個取り出せる．その100個のうちその動詞の直前の助詞が「に」であるものが 90 個だとすれば，それは基準値(85%)を越えているので，「に対して」は関係表現であると判断する．

前章も含めて設定した基準値（75%，10%，85%）は，予めいくつかの関係表現を想定しておき，それら

の接続割合や前置する助詞の分布を小規模のコーパス（約4 Mbyteのテキスト）から調べ，想定した表現が関係表現と判定されるように，逆方向から設定した数値である．特に理論面から，あるいは詳細な実験から得られたものではないことを注記しておく．

3. 実　　験

3.1 実験結果

朝日新聞'90年度の記事1か月分のコーパス（約10 Mbyte）から本手法を用いた関係表現の抽出実験を行った．上記したようにコーパス中の動詞各々に対して(a′)～(e′)の頻度を収集し，接続割合の高いものからプリントとした結果の一部が**表2**である．また表2には，比較のために，核になりえない一般的な動詞「招待する」に注目して，その接続割合も示した．核になりえない一般的な動詞は接続割合が低いことが分かる．

次に接続割合が75%以上である動詞61個各々に対して，(a′)～(d′)の形のそれぞれを直前の助詞とともにコーパスから取り出し，各々の中で最も頻度の高い助詞を見つけ，各々の表現の出現回数に対する割合を調べた．

例として「関連する」を見てみる．「関連する」は接続割合が96.7%であるので，まず核になりえる動詞だと判断する．表2から分かるように，「関連する」の(c′)の形の表現（「関連する」）と(d′)の形の表現（「関連した」）の出現回数は14回と11回であり，両方とも全体の出現回数の151回の10%を越えない．よって(c)(d)の形の関係表現は取り出さない．次に(a′)，(b′)の形をコーパスから取り出すと，直前の助詞の出現回数は**表3**のようになる．表3より「関連する」の(a′)の形の直前の助詞で出現回数の最も大きいものは「に」であり，(a′)の出現回数に対するその割合は94.8%である．基準値（85%）を越えているので，「に関連して」を関係表現とする．同様にして，(b′)も直前の助詞は「に」に定まり（割合100%），「に関連し」を関係表現とする．また**表4**に「招待する」の直前に現れる助詞の分布を示す．核になりえない一般的な動詞の場合，様々な助詞が現れることが分かる．

以上のようにして，77個の関係表現を作成した．その一部を**表5**に示す．

また曖昧性について注記しておく．実験ではコーパスの解析で動詞と他品詞との曖昧性があった場合，動詞と考えることにした．例を示す．

（A）武道を極めて人を知る．（「極めて」は動詞）

（B）この事件は極めて遺憾である．（「極めて」は副

表2 接続割合
Table 2 Rate of connective use.

	接続割合	a'の形	b'の形	c'の形	d'の形	e'の形	SUM
初める	1.000	426	0	0	0	0	426
引き続く	1.000	8	51	0	0	0	59
併せる	1.000	16	19	0	0	0	35
係る	1.000	0	0	31	0	0	31
相手取る	1.000	9	11	0	0	0	20
主催する	1.000	1	3	11	4	0	19
対する	0.991	508	851	854	0	20	2233
関する	0.990	51	34	427	1	5	518
極める	0.987	144	0	0	3	2	149
経る	0.971	122	0	2	10	4	138
関連する	0.967	96	25	14	11	5	151
先立つ	0.956	21	15	7	0	2	45
通じる	0.946	240	29	24	22	18	333
思い切る	0.938	10	1	0	19	2	32
次ぐ	0.923	0	2	22	0	2	26
集う	0.917	0	4	5	2	1	12
連れる	0.909	32	17	0	1	5	55
防止する	0.905	0	4	15	0	2	21
含める	0.903	117	170	12	81	41	421
向ける	0.900	226	72	19	59	42	418
……							
招待する	0.368	0	4	2	1	12	19
……							

表3 選択された動詞に前置する助詞
Table 3 Postpositions in front of a selected verb.

	助詞	出現回数
(a')の場合	に	91
	と	4
	(助詞なし)	1
(b')の場合	に	25

表4 動詞に前置する助詞
Table 4 Postpositions in front of verb.

	助詞	出現回数
「招待する」の直前の助詞	に	3
	を	8
	が	4
	で	1
	(助詞なし)	3

表5 抽出結果
Table 5 Result of extracting.

「に関して」,「に対する」,「に絡む」,「に係る」,
「と題する」,「を相手取って」,「に次ぐ」,「に先立ち」,
「を巡って」,「を通じて」,「を取り巻く」,「に乗り換えて」,
「に基づく」,「に向けた」,「に沿って」,「に伴って」,
「に代わる」,「を交えて」,「を経て」「を踏まえた」,
「を除く」,「と称する」……etc

詞)

例のように「極めて」が出現した場合，それが副詞となるか動詞となるかは形態素解析の段階では判定できない．副詞の場合は，この表現が(a′)の形をしているために，抽出結果への悪影響の恐れもある．しかしここでは動詞として処理することにした．このため「初めて」などの副詞の表現を持つ動詞が核の抽出の部分で取り出されてしまうこともある．しかし副詞の場合，直前の助詞が様々であり，しかも，助詞が現れない場合も多いので，結果的に第2段階の処理で落とされるために実害はない．

3.2 本手法の評価

本手法を評価するために，以下のことを行った．まず，抽出した関係表現の核（39種類）に対する英訳を和英辞典[5]から検索し，抽出した関係表現が使われている用例を調べる．その用例の関係表現部分の英訳が，核の動詞に対応する英語の動詞によって表現されていなければ，その関係表現を妥当なものだと判断することにした．例えば，本手法によって「を巡って」という関係表現が取り出されている．この表現の核の「巡る」を見出しとして和英辞典を調べると，以下の用例を見つけることができる．

(和文) 支払を巡って彼らはけんかをした．

(英文) They had a dispute over the payment.

この用例で「を巡って」は“over”と訳され，動詞を含んでいないために，関係表現として適切であるとする．

このような処理を抽出した関係表現の核（39種類）に対して行った．結果は以下のようになった．

1. その表現が前置詞（句）に対応している．（16種類）
   （例）：「に関して」→about
2. 適当な用例がないが，英訳する場合，別表現にいい換えると前置詞句の表現になる．（12種類）
   （例）：「に係る」→「に関する」
3. 英訳にはその表現を表す意味が動詞に含まれてしまい表現に出てこない．（2種類）
   （例）：「を率いて（行く）」→take
4. 関係表現部分の英訳が，核の動詞に対応する英語の動詞によって表現されている．（9種類）
   （例）：「に乗り換えて」→change

上記の(1)〜(3)に当たるものは，関係表現と判定して不都合はないと考えると言語的に有効な抽出は76.9%である*．コーパスからの知識獲得に関しては，得られる知識がかなり人間の直観とずれることを考慮すると，比較的良い結果だと考える．また最初に予想したヒューリスティックスもこの数値（76.9%）の度合程度に妥当性があると判断できる．

最後に本手法の欠点について述べる．本手法では1つの核に2つ以上強い共起を持つ助詞がある場合にそこから関係表現を抽出できないという欠点がある．例えば，「含める」の場合，接続割合が90.3%であり，核になりえるが，直前の助詞は「を」と「も」に二分されているために，両方とも取り出せていない．この点については今後，改良する必要がある．

4. 考　　察

4.1 関係表現の定義について

従来，慣用表現に関しては，述語回りを中心として，その定義[6]，処理方法[7),8]，辞書記述の形式[9]などが検討されてきている．しかし，慣用表現の解析にはいくつかの困難な問題が存在し，結局は個別的に情報を記述していく以外ないように思える．このような場合，処理方法から慣用表現を定義，分類してゆくのが有益であると考える．ただしその処理方法をとるべきかどうかを判断する問題は依然として残る．

関係表現に関して言えば，その判断を「一語として登録することで，処理の効率が上がる」という観点から行い[3]，ここに関係表現の定義を求めるのが良いと考える．ただし，一語として扱うことで処理効率が上がるかどうかは複雑な要素が絡み，この判断は難しいが，完全に主観的な言語的定義よりも明確性はある．

また本手法によって得られる関係表現は上記の定義（「一語として登録することで，処理の効率が上がる」）を満たしていると考える．以下の3つの理由からである．1つ目はその絶対数が小さいことである．表現を記憶しておく手法をうまく運用できない原因はその量と検索である．ここで抽出した数程度の表現なら，記憶，検索の無駄は小さい．2つ目は，関係表現は一種の機能語であることである．通常辞書に登録されている単語のほとんどは出現頻度が0に近い．一方機能語は個々の語よりも平均的な出現率が高い．この点を考えれば，出現率の高い表現を記憶しておくことは，少なくとも任意の語彙を増やすことよりも性能向上に貢

---

* (3)を複合動詞の一部と捉える見方もあるが，ここでは接続助詞（「て」）の存在，核の動詞自体も複合動詞である場合があり得ることから，複合動詞の一部とは考えないことにした．

献する．3つ目はここで抽出した関係表現には動詞が含まれていることである．一般に連体修飾，連用接続，どちらの場合にもおいても，その意味や機能を正確に判断するには，さまざまな知識が必要になる．結果的に特異な語に対しては，個々の語にさまざまな情報を持たせると考えられる．つまり，ある動詞がほとんどの場合に連体修飾や連用接続として利用されるとしたら，その動詞に対する辞書情報はその用法に対して通常の動詞とは異なった特別な情報を持ち，処理上はどこかに変換的な処理を起動し，その情報を呼び出して，目的の意味や機能を得ると思われる．これは予め一語と考えておくこととほとんど同じ処理になる．動詞を含んだネストの一段深い統語構造を変換することは比較的処理が重たいので，この点を省くことは大きな効率向上になる．

4.2 本手法の実施容易性について

本手法の特徴としてコーパスの解析が非常に浅いものでよいことが挙げられる．

コーパスからの知識獲得の研究は近年盛んに行われているが[10]，そこにはタグ付けされたコーパスや多大な訓練データが必要であったり，精巧なパーサや大規模な辞書が実際には必要であったりとその手法を別の環境で再現したり，大規模な実験に拡張したりするのには困難な場合が多い．一方，本手法は基本的にプレーンなコーパスから動詞とその活用形を抽出するだけでよい．付属語の部分は種類が限られているので，予め想定されるパターンを持っておくことが可能である．しかも，動詞を抽出した後で，本質的に問題となるのは出現頻度であり，100% 完全に動詞を取り出す必要はなく，ある程度の誤りを許容できる．このため，本手法は動詞の抽出法を工夫することにより，更に単純な実現法が可能になると予想できる．例えば，付属語と動詞だけの辞書（この情報は統語情報だけで良い）を使っても本手法は実現できる．この場合，サ変動詞に対しては漢字列部分を語幹と考え，漢字列に続く平仮名パターンからその原型を予想できる．

5. おわりに

コーパスからの関係表現の自動抽出法とその実験結果について述べた．本手法で抽出する関係表現は，通常考えられる関係表現の多くを含み，一語として扱う効果も大きいことから，本手法の有効な利用法が期待できる．また本手法は基本的にプレーンなコーパスと簡易な辞書だけで実現できるため，実験の再現，拡大が容易であるという特徴ももつ．

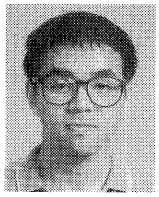
今後は，本手法で用いた考え方を利用して，設定した形以外の関係表現の自動抽出も行いたい．またこれを一般の述語回りの慣用表現にも拡大したい．

謝辞 査読者の方々から有益なコメントを頂きました．感謝いたします．

参考文献

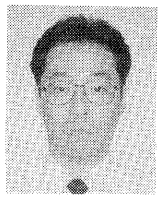
1) 首藤公昭，吉村賢治，武内美津乃，津田健蔵：日本語の慣用表現について，情報処理学会自然言語処理研究会，66-1, pp. 1-7 (1988).
2) 慣用句，日本語学，Vol. 4 (1985).
3) 北 研二，小倉健太郎，森元 逞，矢野米雄：仕事量基準を用いたコーパスからの定型表現の自動抽出，情報処理学会論文誌，Vol. 34, No. 9, pp. 1937-1943 (1993).
4) Church, K. W. and Hanks, P.: Word Association Norms, Mutual Information, and Lexicography, *ACL-89*, pp. 76-83 (1989).
5) ライトハウス英和辞典，研究社 (1990).
6) 野村直之，高橋一裕：3軸モデルによる慣用表現の分類，第41回情報処理学会全国大会論文集，3-75 (1990).
7) 奥 雅博：日本語文解析における熟語相当の慣用的表現の扱い，情報処理学会論文誌，Vol. 31, No. 12, pp. 1727-1734 (1990).
8) 鈴木克志，太田 孝：日英機械翻訳における共起表現の扱い，情報処理学会自然言語処理研究会，82-9 (1991).
9) 末松 博，杉浦真弓，有岡昌子，J. Timothy：電子化辞書における英語連語の記述・表現法，情報処理学会自然言語処理研究会，86-3 (1991).
10) 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).

(平成5年9月 1 日受付)
(平成6年7月 14日採録)



新納　浩幸（正会員）

1961年生．1985年東京工業大学理学部情報科学科卒業．1987年同大学大学院理工学研究科情報科学専攻修士課程修了．同年富士ゼロックス，翌年松下電器を経て，1993年4月より茨城大学工学部システム工学科助手，現在に至る．自然言語処理の研究に従事．人工知能学会，ACL各会員．



井佐原　均（正会員）

1954年生．1978年京都大学工学部電気工学第2学科卒業．1980年同大学大学院工学研究科電気工学専攻修士課程修了．同年通商産業省電子技術総合研究所入所．現在同所知能情報部自然言語研究室主任研究官．主なる研究テーマは，自然言語処理，知識表現，機械翻訳など．日本認知科学会，人工知能学会，ACLなどの会員．