

自由会話システムの構成

斉藤良樹 志村正道

(東京工業大学 工学部)

1. まえがき

コンピュータとの会話は質問応答システムとして種々の研究が行なわれている。しかし、ここでは単なる質問応答ではなく、コンピュータとの自由会話を目的として話題を設定し、あるいは話者の話題に応じるようなシステムを対象とする。したがって、会話を構成するデータ量よりもむしろ、いかにして話の筋をつくり、それに必要な知識の枠組みを構成することが重要な課題となる。

本論文ではまず第2節で、概念依存理論 (Conceptual Dependency Theory, CD理論) について述べる。これは、自然言語の意味を表現する方法として、R.C. Shank らによって提案されたもので、認知科学的に大きな意味を持つものとされている。本システムでは基本的にこれを全てこの形で意味表現を扱っている。第3節ではスクリプト (Script) について述べる。スクリプトは人間が常識的に持っている、行為の系列に関する知識のことであり、膨大な背景知識を必要とする談話理解において有用である。第4節では会話システムの構成について述べる。本システムは、相手の発話による情報、およびシステム内の情報処理の状況が作りだすボタンによって駆動されるプロダクション・ルールを核としている。最後に本システムによって実行された会話例を示す。

2. 概念依存理論

概念依存理論 (Conceptual Dependency Theory, 以下CD理論) は、R.C. Shank らによって提案された。CD理論に基づくシステム (MOPs, MARGIE等) は「CD表現」と呼ばれる独自の意味表現の形で、全ての推論その他の意味処理を行う。CD表現では意味表現を比較的少数の基礎的概念に集約化しようとするものである。

2.1 CD理論の基本的仮定

CD理論では与えられた自然言語文の意味を表現するにあたって、次の2つの基本条件を設定している。

- (A) 意味の同じ2つの文に対して、その言語表現の違いにかかわらず、ただ一つの意味表現が存在する。
- (B) たとえ文の字句上に明示されていない情報でも、文中に含まれる情報は全てその文の意味表現において明示されなければならない。

2.2 CD表現の構成要素と構成規則

各概念の演じる役割として次のものを用意する。

- 行為者 (Actor) : 行為を実行する者
- 行為 (ACT) : 対象に対してなされる行為
- 対象 (Object) : 行為を受ける対象
- 受領者 (Recipient) : 行為の結果としての対象を受けとる者
- 方向 (Direction) : 行為の指向する位置
- 手段 (Instrument) : 行為の手段
- 状態 (State) : 対象の状態

概念の型は次のように5つある。

- P P (Picture Producer) 名詞で表現される物理的対象の概念の型
- L O C (Location) 空間位置を表現する概念の型
- T (Time) 時間位置を表現する概念の型
- A A (Action Aider) ACTを修飾する概念の型
- P A (Picture Aider) 対象の属性を表わす概念の型

2.3 基本アクト

ACTを構成する基本アクトは次に挙げる11個である。

- ATRANS : 所有, 所有権, コントロールなど抽象的関係の移動を意味する。
- PTRANS : 対象物の物理的移動を示し、方向・手段を持つ。
- PROPEL : 対象物に物理的な力を加えることを示し、方向を持つ。
- MOVE : 生物がその体の一部を動かすことを示し、方向を持つ。
- GRASP : 対象を握ることを示し、方向を持つ。
- INGEST : 対象を自分の体内に取り入れることを示し、方向を持つ。
- EXPEL : 自分の体内から外界へ対象を排出することを示し、方向を持つ。
- MTRANS : 情報の伝達、移動を示す。
- MBUILD : 旧い情報から新しい情報をつくりだすことを示し、方向、受領者を持つ。
- SPEAK : 音声を発することを示す。
- ATTEND : 注意を向けること、すなわち刺激の方向へ感覚器官の焦点を合わせることを示す。

2・4 因果的概念構造

概念間の構文ルール他に、少し複雑な意味を表現するには、因果的概念依存関係（概念依存構造間の因果関係）を構成する規則が必要である。これは因果タイプ等と呼ばれるが、ここでは省略する。

3. 文章のCD表現への変換

CD理論に基づくシステムにおけるパーズングとは、自然言語で表現された文章をCD表現に変換することである。本会話システムでは、パーサとして Shank & Riesbeck が開発した McELI (Micro Version of the English Language Interpreter) を用いている。

3・1 単語の意味定義

McELI の大きな特徴は独立した文法というものを持たず、その代わりに単語の定義自身に「次の単語の予想」を持っている。これは人間の文章処理をモデル化することを目的としたもので、適応性が高いといわれている。単語の基本的な意味定義のしかたは図1ようになる。

3・2 パーズングのアルゴリズム

1. 入力文の先頭に"*START*"をつける。
2. 入力文の先頭の単語の定義(<definition>)をSTACK にpushする。
3. STACKからパケット(<packet>)をひとつ popする。STACK が空ならばSTEP 5へ。
 - (a) そのパケットのリクエスト(<request>) 中にTESTが無い、TESTの結果がTならばSTEP 4へ。
 - (b) TESTの結果がFならばSTEP 3へ（そのパケット全体が捨てられる）。
4. そのパケットの中のASSIGNを（あれば）実行する。そのパケットをスタックTRIGGERED にPUSHしてSTEP 3へ。
5. スタックTRIGGERED が空ならばSTEP 6へ。そうでなければ、スタックTRIGGERED のパケットを一つずつ取り出して、その中にNEXT-PACKET があれば、その定義を再びSTACK にpushしておく。STEP 5へ。
6. 入力文の先頭の単語を一つ取り除きSTEP 2へ。ただし入力文が空になっていればパーズングを終了する。パーズングの実行例を図2に示す。

```
<definition> ::= (<packet>1, <packet>2, ..., <packet>n)
<packet> ::= (<request>1, <request>2, ..., <request>n)
<request> ::= <assign> | <test> | <nextpacket>
<assign> ::= (ASSIGN a1, a2, ..., an)
<test> ::= (TEST S-exp)
<nextpacket> ::= (NEXTPACKET <definition>)
```

図1 辞書部における意味定義のシンタクス

"John went to New York by train."

```
(PTRANS (ACTOR (PERSON (NAME (John)) (SEX (MALE))))
(OBJECT (PERSON (NAME (John)) (SEX (MALE))))
(TO (NewYork))
(FROM (NIL))
(BY (train (S/P (SINGULAR))))
(TIME (PAST)))
```

"John paid one thousand yen."

```
(ATRANS (ACTOR (PERSON (NAME (John)) (SEX (MALE))))
(OBJECT (money (UNIT (yen)) (AMOUNT (1000))))
(FROM (PERSON (NAME (John)) (SEX (MALE))))
(TO (NIL))
(TIME (PAST)))
```

図2 パーズング結果の例

4. 会話システムの構成

4・1 会話システムの基本構成

本システムは次のような手順で行なわれる。

1. 自然言語の文をある種の意味表現に変換する。
2. 明示的に与えられていない情報を推論して付加する
3. 発話方略(Speech Strategy)を選択する。
4. 文章を生成する。

このシステムの基本構成を図3に示す。入力文が与えられるとまずMcELI がこれをCD表現に変換し、CP (Concious Precessor)に付加する。次にプロダクションシステム(以下PS)はこのCPの作り出すパタンによって適用するプロダクション・ルールを決定する。プロダクション・ルールが一つ適用されるたびにCPのパタンは変化する。この適用サイクルは何らかの返答が生成されるか、適当なプロダクション・ルールが尽きる(返答生成に失敗)まで繰り返される。

4・2 スクリプトによる文章理解

スキーマは、知識の構成単位の意味表現形式であり、またそれをいつ・どのように用いるかを記述する手続きまでを包括する。したがってスキーマは受け入れた文章から典型的な事態や関係を成り立たせるのに必要な具体例を探し出して、デフォルト的知識を再認するものである。スクリプトは、与えられた知識に対して”受け皿”的な

役割を果たす知識の枠組みとして用いられる。スクリプトによって省略された文章の理解だけではなく、ある状況における「次に起こりそうなこと」の予想、また受け入れられた文章のあるエピソードのどの場面が位置付けることなどを行なう。たとえば「食事スクリプト」には行動の系列として、「レストランへ入る、メニューを見る、注文をする、食事をする、勘定を済ませる、レストランから出てゆく」等が記述されている。またスクリプトにはそのエピソードに必要な登場人物の役割(roles)や、大道具・小道具(props)が設定されている。

本システムではスクリプトによる推論メカニズムとして McSAM〔2〕を用いる。この機構においても意味表現は全てCD表現で表わされる。したがって、すべての行動の系列がリスト構造で書かれたCD表現によってスクリプトに登録されている。このリスト構造は入れ子状の連想リスト(nested-association list)であり、基本的にはフレーム(frame)と同じものである。

このようなスクリプトは一般にはヘッダ(headers)によって呼び出される。呼び出されたスクリプトによって、PROPS、ROLESへ実体が割り当てられ、また文章中の欠けていた情報が補なわれたり(paraphrase)、省略されていた行為そのものが補なわれたりしてあるエピソードの理解が完了する。

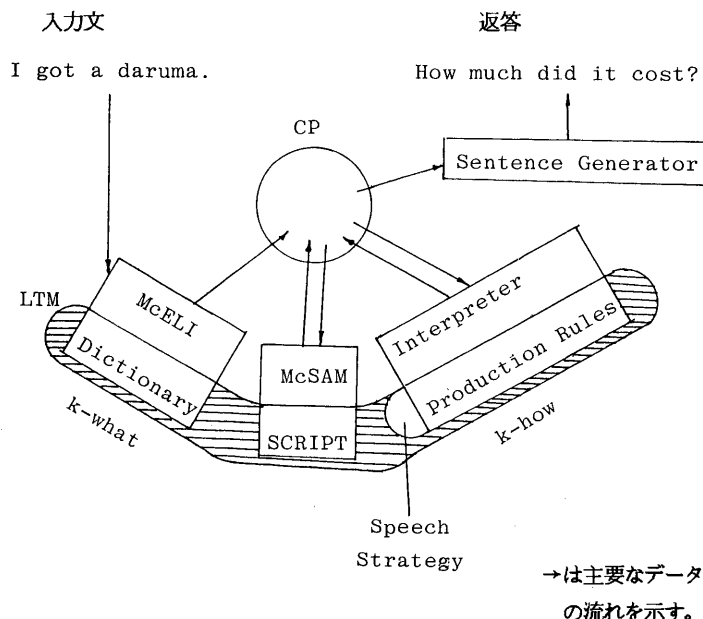


図3 システムの概念図

次にMcSAM のアルゴリズムを示す

1. CD表現に変換された入力文の列から最初の一文を取り出す。文章がなければ終了する。
2. 活性化されたスクリプトがあればSTEP 3へ。なければ現在の文章中のヘッダの有無を調べ、あればそのスクリプトを活性化し、STEP 1へ。なければ、その文章をそのままデータベースに付加し、入力文列から削除する。
3. 活性化されているスクリプト中で、現在の文章がどのイベントにあたるか調べる。すなわち同構造のCD表現を探す。なければSTEP 2へ。あれば、ROLES、PROPSに関する情報のうちで既知のものを用いて、その文章をパラフレーズしてデータベースに付加する。処理の終わった文章を入力文列から削除してSTEP 1へ。

→は主要なデータの
流れを示す。

5. 会話システムの実験

5.1 プロダクション・ルールとCPの形式

本システムは30個弱のプロダクション・ルールを持っている。ここではまずプロダクション・ルールの構文規則および短期記憶の形式について述べる。

<プロダクション>は次のような形のリストである。

(<プロダクション名><条件>₁<条件>₂...<条件>_m
→<アクション>₁<アクション>₂...<アクション>_n)

ここで<プロダクション名>は任意のアトムであり、

<条件>は任意のリストである。また、<アクション>は任意のLISP関数である。なお、<アクション>中の

"="付きのアトムで表わされる変数には、条件部評価時のCPとのマッチングにおいて、CP中の対応するアトムが値として束縛される。

<CP>は次の様な形式のリストである。

(<パターン>₁<パターン>₂...<パターン>_r)

<プロダクション>の条件部評価時に<CP>内の<パターン>と<プロダクション>の<条件>とのマッチングが調べられ、その結果によって<条件>の真・偽が定まる。図4にCPのパターンとPRの例を示す。

CP : [(B Y) (A (P Q R)) (C S) (A XX)]

プロダクション ルール [Rule 1 (B Y)
(A =X)
(*EQUAL P(*CAR =X)
(C =X)
(*ABSENT(X Y)
→
(*DEPOSIT(U V))
(*DELETE(A =X))
(*DELETE(C =X))]

図4 CDのパターンとプロダクションの例

5.2 スクリプトとその適用プロセス

スクリプトによって、どんな理解が達成されるかを図5に示す。映画スクリプトを例として説明する。

映画スクリプトには、映画を観に行った時に通常起ると考えられる5つの行為がCD表現で記述されている。図中のそれぞれのCD表現の意味は、下の日本語 (i)~(v) で示してある。入力文とスクリプトの一つのCD表現とのマッチングが取られた時に、対応する実体の概念構造が割り当てられる。

(PTRANS (ACTOR (*VAR* AUDIENCE)) (OBJECT (*VAR* AUDIENCE)) (TIME (*VAR* TIME)) (TO (*VAR* MOVIE THEATER)))	(i) ある人が映画館へ行く。
(ATRANS (ACTOR (*VAR* AUDIENCE)) (OBJECT (money (AMOUNT (*VAR* MONEY-AMOUNT)) (UNIT (*VAR* MONEY-UNIT)))) (FROM (*VAR* AUDIENCE)) (TO (cashier))	(ii) 係員に料金を支払う。
(ATRANS (ACTOR (cashier)) (OBJECT (ticket)) (FROM (cashier)) (TO (*VAR* AUDIENCE))	(iii) 係員からチケットを受けとる。
(ATTEND (ACTOR (*VAR* AUDIENCE)) (OBJECT (movie (NAME (*VAR* MOVIE-NAME)))) (TIME (*VAR* TIME)))	(iv) 映画を観る。
(PTRANS (ACTOR (*VAR* AUDIENCE)) (OBJECT (*VAR* AUDIENCE)) (TIME (*VAR* TIME)) (FROM (*VAR* MOVIE THEATER)) (TO (*VAR* ELSEWHERE)))	(v) 映画館から他所へ向う。

ROLES

AUDIENCE 観客
cashier 窓口の係員

PROPS

money お金 MONEY-AMOUNT 金額 MONEY-UNIT お金の単位
movie 映画 MOVIE-NAME 映画名
MOVIE THEATER 映画館

TIME 時制

図5 映画館スクリプト

```

(( PTRANS (ACTOR (PERSON (NAME (John)) (SEX (MALE))))
  (OBJECT (PERSON (NAME (John)) (SEX (MALE))))
  (TO (Movietheater (S/P (SINGULAR))))
  (FROM (NIL))
  (BY (NIL))
  (TIME (PAST)))
  (ATTEND (ACTOR (PERSON (SEX (MALE))))
    (OBJECT (movie (S/P (SINGULAR))))
    (TIME (PAST)))
  (PTRANS (ACTOR (PERSON (SEX (MALE))))
    (OBJECT (PERSON (SEX (MALE))))
    (TO (house))
    (FROM (NIL))
    (BY (NIL))
    (TIME (PAST))))

```

図6 CD表現に変換された“映画ストーリー”

次にスクリプト適用によって、そのような理解が行なわれるかを見ることにする。次の文章a)～c)を“映画ストーリー”とする。またこれらの英文をCD表現に変換すると図6のような3つの行為の列になる。

- (a) John went to the movietheater.
- (b) He saw the movie.
- (c) He went home.

5.3 話題の成立・変化過程とスクリプト間の構造化

会話には必ず何らかの話題が存在する。会話は必ずその会話が“一貫してつながっている(coherent)”という認識に基づいて会話をしている。その一貫性を与えているメインテーマが話題であると考ええる。

話題の成立・変化に対する規則を分析するに当たって Shank は次のように話題を定義している。すなわち「話題は2つの発話の対から成り立ち、新しい発話は古い話題の一部を維持し、関連する新しい話題を導入する。」というものである。

本システムは、5個のスクリプト“休日スクリプト(holiday script)”“買物スクリプト(shopping script)”“映画スクリプト(movie script)”“食事スクリプト(restaurant script)”“電車スクリプト(train script)”を持っていて、それぞれ「休日を過す」「買物をする」「映画を観る」「食事をする」「電車にのる」場合の典型的な(prototypic)な行動列が記述されている。これら

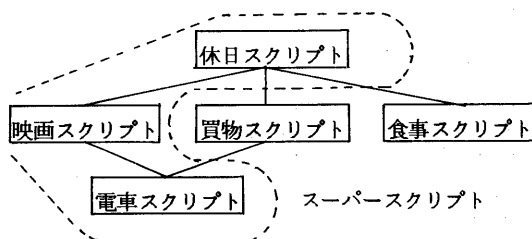


図7 スクリプト階層構造

のスクリプトはそのカテゴリから図7のようなハイアラキーを構成することができる。

このような、上位スクリプト、下位スクリプトによって構成される知識構造によって、本システムは上位話題(super topic)、下位話題(sub-topic)への変化を扱うことができる。

5.4 文章の生成

談話における文章の適格性を考えるにあたって、いわゆる“視点論的アプローチ”が重要である。視点論は最近、従来の認知心理学における知識観・理解観の再検討を迫っている。

人間が何かを発話する時、「どこかに視点を置いた表現」を必ず用いる。従来、スキーマ理論(CD理論も含まれる)ではこのような“視点”にはいっさい注意が払われてこなかった。しかし、人工的に作られた「視点」のふらつく談話が理解度を著しく低下させたり、その談話を再構成させたものは必ず何らかの視点を持っているなどという“視点を無視できない”多くの研究結果が報告されている。視点論的アプローチは今後「理解とは何か」を探るためのキー・コンセプトとなり得ると考えられる。

文章生成の技術的側面は、元の意味表現が計算機内でどんな形式をとっていたかということに依存する。意味表現として、意味ネットワーク(semantic network)、フレーム構造、論理式などが用いられている。

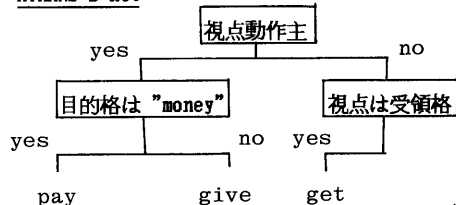
CD表現からの文章生成に必要なものは、弁別ネット(Discrimination Network, 以下D-net とする)と各動詞毎に付加された構文情報である。表層文を生成したいCD表現はまずD-net にかけて、その結果相応の動詞が選出される。その後その動詞に付加された構文情報(格関係、前置詞などの情報)を使用して、表層文が生成される。

弁別のためのファクターはCD表現で目的格の役割を持つPP(Picture Producer)の物理的属性に依存するが、他に主語のとり方、動作の主体と客体の関係(動作主と目的格の関係)等に着目すればより繊細な弁別が可能になる。

図8に基本アクトPTRANS、ATRANSに関するD-netの例を示す。またこれを用いた表層文の生成例を図9に示す。

次に質問文を生成する場合について述べる。質問文の生成は、計算機内につくられた意味表現のどこに着目しどんなレベルで質問文を生成するかということ特定してやれば、前述のD-netを用いた文章の生成と同様に可能である。しかし、着目点やレベルの決定が、相手の発話内容やシステム内の意味表現によってどのように導かれるかをルールに記述することは非常に困難である。

ATRANS D-net



PTRANS D-net

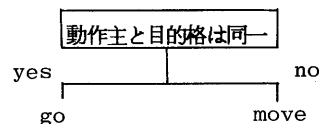


図8 D-net の例

```

((PTRANS (ACTOR (you))
  (OBJECT (you))
  (TIME (PAST))
  (BY (walk))
  (FROM (NIL))
  (TO (store (S/P (SINGULAR))))))
(PTRANS (ACTOR (you))
  (OBJECT (daruma (S/P (SINGULAR))
    (AKO (TOY))
    (COLOR (red))))
  (TIME (PAST))
  (TO (you)))
(ATRANS (ACTOR (store (S/P (SINGULAR)))
  (OBJECT (daruma (S/P (SINGULAR))
    (AKO (TOY))
    (COLOR (red))))
  (FROM (store (S/P (SINGULAR))))
  (TIME (PAST))
  (TO (you)))
(ATRANS (ACTOR (you))
  (OBJECT (money (AMOUNT (1000)) (UNIT (yen))))
  (FROM (you))
  (TIME (PAST))
  (TO (store (S/P (SINGULAR))))))
(PTRANS (ACTOR (you))
  (OBJECT (you))
  (FROM (store (S/P (SINGULAR))))
  (TIME (PAST))
  (BY NIL))
  (TO (movietheater (S/P (SINGULAR))))))
  
```

図9 図8のD-net を用いた表層文生成例

視点 "you"
を与えた場合の生成例

You went to the store by walk.
You moved the daruma to you.
You got the daruma from the store.
You paid a thousand yen to the store.
You went to the movietheater from the store.

視点 "store"
を与えた場合の生成例

You went to the store by walk.
You moved the daruma to you.
The store gave you the daruma.
The store got the money from you.
You went to the movietheater from the store.

Why don't you converse with me? Please tell me something.

>> I WENT ON FOOT. **FIRE* R1, R2 (i)

(1) Where did you go?

>> I WENT TO THE STORE. **FIRE* R3 (ii)

(2) Oh, did you?

>> I WENT TO THE STORE. **FIRE* R4

(3) Did you go shopping?

>> YES, I DID. **FIRE* R5, R6, R61 (iii)

(4) What did you buy at the store?

>> I GOT A RED DARUMA. **FIRE* R62, R65, R61 (iv)

(5) How much did you pay for thr daruma?

>> I PAID A THOUSAND YEN. **FIRE* R62, R65, R61

(6) Where did you go after shopping?

>> I WENT TO THE MOVIE THEATER. **FIRE* R62, R65, R64, R4 (v)

(7) Did you go to see a movie?

>> YES, I DID. **FIRE* R5, R6, R61

(8) What did you see?

>> I SAW TRON. **FIRE* R62, R65, R61

(9) Where did you go after the movie?

>> I WENT HOME. **FIRE* R62, R65, R63, R8 (vi)

(10) All you said is your holiday story?

>> YES. **FIRE* R5, R6, R61

(11) Where did you go for lunch?

>> I WENT TO THE RESTAURAN. **FIRE* R621, R6, R61

(12) What did you eat?

>> I ATE A BIG SANDWICH. **FIRE* R62, R65, R61

(13) How much did the sandwich cost?

>> I PAID A THOUSAND YEN. **FIRE* R62, R65, R61

(14) Where did you go after meal?

>> I WENT TO THE MOVIE. **FIRE* R62, R65, R63, R81, R9 (vii)

(15) Previously, you said you went to the store by walk.

>> YES. **FIRE* R92

(16) And you said you got the daruma from the store.

>> YES. **FIRE* R92

図10 会話例

