

英·韓 翻譯시스템에 關한 研究

(A Study on English-Korean Translation System)

金 榮 暹*, 金 漢 宇*, 崔 炳 旭*

(Young Sum Kim, Han Woo Kim and Byung Uk Choi)

要 約

本 論文에서는 構文解析部 意味解析部 그리고 文章生成部로 構成된 英韓翻譯 시스템을 提案한다. 構文解析 過程에서의 中間言語로 Montague 文法에 基礎한 Lambda-範疇 文法을 導入하고, 對象言語의 文法記述에는 擴張文脈自由文法을 導入한다. 一般的인 解析規則의 作成과 對象言語의 意味表現은 User interface utility (文法規則과 辭典의 內容을 修正하고 補完할 수 있는)를 構成함으로써 그 表現의 어려운 問題點을 克服할 수 있다.

Abstract

In this paper, an English-Korean machine translation system which consists of a syntax analyzer, a semantic analyzer and a sentence generator is proposed. The intermediate language in the syntax analysis is described by a lambda categorial grammar that is based on Montague Grammar, and the augmented context free grammar is introduced in the description of object language. Much difficulty in constructing general rules and semantic representations of an object language in the parser can be avoided by using the knowledge from the user interface utility that can modify the contents of dictionary and grammar rules.

I. 序 論

機械翻譯은 컴퓨터 自體가 本質的으로 記號를 處理하는 機械이기 때문에, 文字列에서 文字列로의 置換이 可能하다는 點에 着眼하여 그 可能性이 論議되기 始作하였다. 또한 하드웨어와 소프트웨어의 兩面에 걸친 急激한 發展과, 統語論, 意味論等 翻譯過程에 理論的 基礎를 附與하는 言語理論의 顯著한 進歩는 機械翻譯 研究에 土臺를 提供하였다.^{1,2)}

現在는 多數의 商用 시스템도 發表되고 있으며, manual이나 一定 書式等 制限된 範圍에서는 注目할 만한 成果를 거두고 있다.^{1,3,5)}

本 研究에서는 英語와 韓國語의 文法 및 言語構造를 分析하여 英韓翻譯에 따른 問題點을 解決하여 機械翻譯 시스템의 實現 可能性을 立證하는데 그 目的이 있으며, 英文解析과 中間言語의 變換, 그리고 韓國語生成의 3段階로 시스템을 實現한다.

한편 構文解析에는 擴張文脈自由文法을 導入하고, 英語와 韓國語의 深層構造로 Montague 文法^{1,7)}에 基礎한 Cresswell의 Lambda-範疇言語⁸⁾를 導入하여 翻譯을 實現하였다. 그리고 解析規則과 辭典構成의 未備即 知識 베이스의 未備를 補完하기 爲하여 user interface utility를 構成하였다. 또 曖昧性이 存在하는 文章의 解析은 非決定的으로 順次的인 出力을 附與하고 不適當한 出力에 對해서는 診斷機能에 依해 補完할 수 있도록 한다.

*正會員, 漢陽大學校 電子工學科
(Dept. of Elec. Eng., Han Yang Univ.)

接受日字: 1985年 8月 1日

(*本 研究는 韓國科學財團 1984年度 一般研究 支援에 依하여 이루어진 것임.)

II. 英語와 韓國語의 中間表現

本 英韓翻譯 시스템은 翻譯過程의 媒介를 遂行하는

中間表現으로 Montague 文法에 基礎한 Lambda範疇文法을 導入한다.

Lambda範疇言語는 純範疇 言語에 變數와 Lambda를 附加한 것이다. 純範疇 言語는 2個의 基本範疇(名詞, 文)와 이 基本範疇에서 再歸的으로 生成된 範疇가 存在한다. Lambda-範疇 言語는 여기에 變數와 Lambda를 附加시키는 것에 依해 抽象化된 Lambda-表現의 導入을 可能하게 한다. 本 研究에서는 그림1과 같이 英語와 韓國語를 各各 λ-深層構造로 變換하여 그들의 對應에 依하여 翻譯過程을 遂行하는 中間表現을 利用하고 있다.

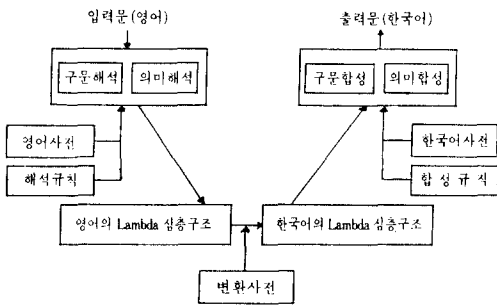


그림 1. 번역과정의 개요
Fig. 1. An outline of translation procedure.

1. 英語의 Lambda-深層 構造

自然言語의 各 文法範疇에 Lambda-範疇 言語의 統語範疇를 割當함으로써, Lambda-深層構造라고 부른다. 逆으로 自然言語는 Lambda-深層構造 表現에서 logical symbol을 除去하여 얻을 수 있다.

英語의 各 文法範疇에 屬하는 統語範疇의 symbol을 割當한다. 그림 2에 Lambda-範疇 言語의 統語範疇 割當例를 보인다. 여기에서 基本範疇는 0 (文章)과 1 (term)으로 構成된다.

英語의 統語範疇에 割當例	SYNTACTICAL CATEGORY OF CATEGORIAL LANGUAGE
GRAMMATICAL CATEGORY OF ENGLISH SENTENCE	0
term	1
INTRANSITIVE VERB	(0, 1)
TRANSITIVE VERB	(0, 1, 1)
object	(0, 1, 1, 1)
*object	(0, 1)
COMMON NOUN	(0, (0, 1))
NOUN PHRASE	((0, (0, 1)), (0, 1))
DETERMINER	((0, 1), (0, 1))
ADJECTIVE	((0, 1), (0, 1))
ADVERB	((0, 1), (0, 1))
PREPOSITION	((0, 0), 1)
AUXILIARY VERB	((0, 1), (0, 1))
COORDINATE CONJUNCTION	(0, 0, 0)
SUBORDINATE CONJUNCTION	((0, 0), 0)

그림 2. 영어의 통어 범주의 할당
Fig. 2. Syntactic category of English.

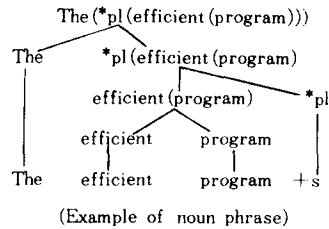
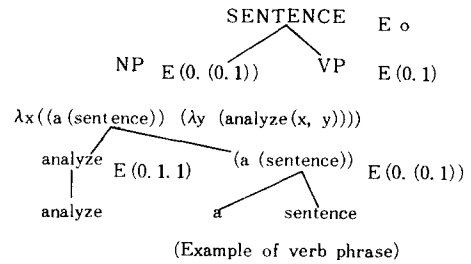


그림 3. 영어의 λ-심층구조의 일례
Fig. 3. An example of English lambda-deep structure.

2. 韓國語의 Lambda-深層構造

英語의 Lambda-深層構造(即, 意味表現 論理式-English-Oriented Formal Representation)에 對應하는 韓國語의 Lambda-深層構造(Korean-Oriented Formal Representation)는 英語와 同一한 方式에 依해서 割當된 韓國語의 統語範疇에 對應되어 있다.

韓國語의 統語範疇의 割當과 Lambda-深層構造의 一例를 그림4, 5에 보인다.

3. 英語와 韓國語의 Lambda-深層構造의 對應

英語와 韓國語의 Lambda-深層構造를 各各 定義했기

한국어의 문법 범주	통어범주의 할당
문 (term)	0
명 사	(0, 1)
동 사	(0, 1, 1)
명 사 구	(0, (0, 1))
완전 형용사	((0, 1), (0, 1))
불완전형용사	((0, 1), (0, 1), (0, 1))
부 사 구	(0, 0)
명사구 연결조사	((0, (0, 1)), (0, (0, 1)), (0, (0, 1)))
접 속 조 사	(0, 0, 0)
감탄사(독립어)	0
조 동 사	((0, 1), (0, 1))
주격, 목적격조사	없 음

그림 4. 한국어 통어 범주의 할당례
Fig. 4. Syntactic category of Korean.

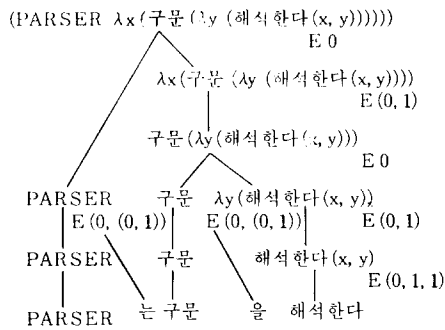


그림 5. 한국어의 λ-심층구조의 일례
 Fig. 5. A example of korean lambda-deep structure.

때문에, 實際 變換過程에서는 入力 element의 範疇에 對應하는 置換만을 行한다. 한편 變換 後의 Lambda-深層構造는 變換 前의 Lambda-深層構造와 同一한 統語範疇를 가져야 한다.

記號의 置換

book → 책
 F(0, 1)

beautiful → 아름답다 + 바다
 F(0, 1), (0, 1)

Ⅲ. 英語의 構文解析과 深層構造의 生成

1. 對象言語의 記述

英語의 統語, 意味記述은 擴張文脈自由文法과 辭典에 依하여 行한다.^[8,9,12,14] 擴張文脈自由文法은 Pratt가 提案한 LINGOL의^[10] extended 形인 E-LINGOL^[11]에서 使用한 文脈自由文法을 擴張시킨 것이다.

擴張文脈自由文法은 다음과 같은 形式으로 記述한다.^[12]

$$R_n; A \rightarrow \beta_1 \dots \beta_n$$

여기에서 A는 非終端 記號이며, 文, 名詞句等과 같은 對象言語의 文法範疇을 意味한다. 한편 β_i에는 非終端 記號 또는 關係代名詞의 先行詞 削除等の 境遇에 適用되는 (Bi-Ci)와 같은 記號가 位置한다, 各 文法範疇에 對應하는 句構造에는 (Syntactic value) (Semantic value) (Assistant value) 등이 定義된다. 統語值에는 人稱, 數, 性等에 關한 情報을 記述하고, 意味值에는 對應하는 意味表現을 記述하며, 制御值에는 對應하는 句構造의 評價值 또는 解析過程의 制御情報가 記述된다.

英語 辭典의 記述은 LISP의 property 函數로 記述되며 다음과 같은 形式으로 構成된다.

(單語) : ((文法範疇) (統語值) (意味值) (制御值))
 辭典部의 統語, 意味, 制御值은 擴張文脈自由文法

規則에 依하여 句構造의 統語, 意味, 制御값이 再歸的으로 構成될 수 있도록 그 初期值를 附與한다.

多品詞語와 同音異意語에는 同一한 標題를 갖는 複數個의 項目이 存在할 수 있으나, 重要單語가 아니면 可及의 多品詞의 選擇을 排除하였으므로 辭典上에 複數個의 項目을 갖는 單語에는 制御值에 그 情報을 記述하여 處理한다.

그림 6에 辭典構成의 一例을 보인다.

- (PUT 'HAVE 'ENGRWD
- (VT (FORM, ORG) ((SEM, HAVE-) (SEM, HAVE+)) (ASS, 1))
- (PUT 'NEED 'ENGRWD
- ((NOUN (NRB, SGL) (SEM, NEED+) (ASS, 0))
- (VT (FORM, ORG) (SEM, NEED+) (ASS, 1)
- (PUT 'ETHERNET 'ENGRWD (NOUN (NRB, SGL) (SEM, ETHERNET) (ASS, 0))

그림 6. 사전구성의 일례

Fig. 6. An example of dictionary content.

2. 解析規則

對象言語의 言語現象을 規範지우고, 入力文에 對하여 適當한 解析 結果를 誘導하는 解析規則^[9,13,14]은 크게 다음의 4가지로 構成되어 진다.

- 1) EXrule : 對應하는 文法範疇가 goal로서 設定되면 實行한다.
- 2) BORule : Bottom-up parser에서 考慮되어 擴張 文脈自由文法이 그대로 適用된다.
- 3) CORule : (virule) 對應하는 文法範疇의 句構造가 構成되면 實行된다.
- 4) IDrule : 熟語處理等に 使用되며 辭典에 記述된다.

解析規則은 다음과 같은 形式으로 記述된다.

Goal ⇒ (條件 1) - (procedure 1);
 (條件 2) - (procedure 2);
 :
 :
 (條件 n) - (procedure n);

解析規則의 適用은 goal이 設定되었을 때 條件部를 test 해서 true이면 繼續해서 procedure를 實行하는 것으로 한다. 이때 만일 2個 以上の 條件이 true이면 非決定的인 處理를 行한다.

條件部에는 It나 That節等 Lexical element가 豫測되었을 때 實行되는 Lexical test, 그리고 prep(전치사)나 Aux(助動詞)等 word category element가 豫測되었을 때 實行되는 category test 등이 記述된다.

Procedure部에는 任意的인 LISP 프로그램을 記述하는 것이 可能하며 通常 다음과 같은 3個의 sub-procedure가 記述된다.

- 1) Popsem : Scanning된 element에 對한 句構造나 生成된 句構造에 對한 上位의 句構

造가 豫測되었을 때 對應하는 意味情報을 stack에 set시키는 役割을 遂行한다.

2) Sempart : Stack에 set된 情報을 利用하여 句構造에 對應된 意味를 處理한다.

3) Construct : 最終적으로 豫測된 句構造의 生成을 完了하고 對應하는 句構造를 豫測한다.

以上の 解析規則의 program 記述例를 그림7에 보인다.

```

INPUT 'NP' 'RULE'
  ((DET 'NP') (POP 2)
   (CONSTRUCT NP)))
  ((NP1 ((POP 1) (CONSTRUCT NP))))
INPUT 'NP' 'RULE' (((CONTEST WHICH) (EXP NP) (EXP WHICH) (EXP VP)
  (POP 3)
  (CONSTRUCT NP)))
  ((EXP NP1)
  (POP 1)
  (CONSTRUCT NP)))
INPUT 'VP' 'RULE' ((T (EXP VP) (POP 1) (CONSTRUCT VP)))
INPUT 'VP' 'RULE' ((T (EXP VP) (EXP PRFP)
  (POP 2)
  (CONSTRUCT VP)))
    
```

그림 7. 해석규칙의 program 기술 예
Fig. 7. An example program of grammar rule.

3. 解析規則 Interpreter

解析規則 Interpreter는 擴張文脈自由文法에 依해서 記述된 解析規則과 辭典을 利用해서 入力文에 對한 統語構造와 이에 對應하는 意味記述 即 英語의 Lambda 深層構造를 出力한다.^{11,13}

Interpreter의 概要는 그림 8과 같다.

Interpreter上에서 Sentence → NP·VP라는 擴張文脈自由文法 規則은 Sentence가 goal로서 豫測된 때, 그것을 NP, VP라는 subgoal로 分解한 다음 다시 NP가 goal로서 豫測되면 縮次的으로 procedure를 反復 遂行하여 解析을 行한다.

解析規則의 適用은 非決定的으로 行해지며 解析 tree의 構成은 Bottom-up과 Top-down方式을 適切하게 組合해서 實行한다.

Interpreter上에서 解析規則의 適用은 slot라 부르는 데이터 構造를 利用해서 制御하며 5個의 element로 構成된다.

- 1) b: procedure의 stack
- 2) a: car(b) 現在の goal
- 3) c: 現在 scanning되는 入力 element의 position
- 4) d: partial parse tree에 對應하는 stack.
- 5) e: 省略된 (關係代名詞의 先行詞等) element에 對한 情報를 갖는 stack.

한편 多意語와 多品詞語의 處理는 可及의 排除하였지만 몇몇 主要語에 對해서는 補助 routine을 作成하여 解析을 遂行하였다(그림 9).

多品詞語의 處理는 main routine上에서도 適格한出

```

EROS(nil)=
  (buffer initial setting)
  print - "English Korean machine translation system"
  input - english sentence

PARSER(buffer) = buffer - input sentence
  prog(count - 0);
  if try(cat; 1); (cat - sentence (initial category))
    then print("PARSING SUCCEEDED")

TRY(cat; 1)=
  prog(slot; subslot; a; b; c; d; e);
  slot - list(cons(cat; (FINISH))); position; nil; nil)
  subslot - cons(slot; nil);
  loop; if null(substack)
    then return nil
    else slot - car(slot)
        subslot - cdr(slot)
        tryslot(nil)
    go loop;

TRYSLOT(slot)=
  prog(goal; rules; newslots);
  goal - slot. a;
  newslots - ();
  (gap supplement process, if any)
  if atom(goal)
    then
      (search process to avoid double process)
      rules - invoke (exrule; goal);
      foreach rule rules
        do
          if eval(first (rule))
            then eval(second(rule))
          end
        else eval(goal);
      foreach slot newslots (slot updating)
        do sslot push(slot; sslot)
        end)

EXPAND(a)= a -goal
  prog(goal; rule; borule; newtree; node);
  if and(a -vp ; test(a; c))
    then return nil
    else rule -getp(a; exrule)
        if null(rule)
          then nil
          else expdict(rule)
        getdict(dummy; c)
        if borule -getp(a; borule)
          then
            if rule -assoc(word category; borule)
              then makenode
                rule - invoke(corule; goal)
            (Idiom process routine)
            if node -getdict(goal; c)
              then makenode
                rule -invoke(corule; goal)
            end
    end
    
```

그림 8. Rule interpreter의 개요
Fig. 8. An outline of rule interpreter.

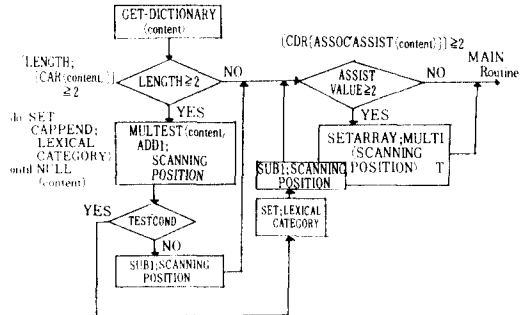


그림 9. 다음사어와 다의어 처리 routine
Fig. 9. Simple flowchart of word class decision and mutivocal word detection routine.

力文을 얻을 수 있지만, 復數個의 品詞中에서 한境遇에 對한 解析을 遂行하여 不適格한 出力이 얻어졌을 때는 backtrack하여야 하므로 process time等 여러面에서 非效率的이다. 그러므로 TEST routine 上에서 scanning position을 하나 增加시켜 單語를 check 하고, 또한 以前의 解析 結果를 參照하여 禁止規則(즉, 冠詞 다음에 動詞, 形容詞 다음에 動詞, 副詞 다음에 名詞等...)을 利用하여 品詞를 決定한다. 그러나 Test가 失敗하면 main routine으로 復歸하여 backtrack에 依한 方式으로 解析過程을 遂行한다.

또한 多意語의 境遇에는 assist value를 check 해서 2 以上이면 對相單語를 多意語 array에 set하고 繼續해서 解析을 遂行한다.

IV. Lambda-深層構造의 變換과 韓國語의 生成

1. Lambda-深層構造의 變換

英語의 入力文을 解析해서 얻어진 Lambda-深層構造 表現을 이에 對應하는 韓國語의 Lambda-深層構造로 變換하는 過程은 對應하는 單語의 置換을 行하는 記號變換過程과 英語와 韓國語와의 表現上의 差異를 補充하기 爲한 構造 變換過程으로 나눌 수 있다.

變換處理는 Lambda-範疇文法 內에서 實行되기 때문에, 어떠한 英語 Lambda-深層構造 表現의 變換도 同一한 範疇內에서 이루어져야 한다.^[14]

本 시스템에서는 先 記號變換만을 念頭에 두고 變換 프로그램을 作成했다.

變換 프로그램은 英語의 構文解析 過程에서 生成된 意味情報 即 Lambda-深層構造 表現과 英語 入力文의 各 element에 對한 辭典情報, 生成된 句構造의 情報, 그리고 句構造의 生成過程을 記述한 情報를 利用하여 實行된다.

變換은 入力 element에 對한 變換辭典을 參照하여 實行하며, 變換辭典에는 生成過程에서 要求가 豫想되는 情報가 記號變換 過程에서 附加될 수 있도록 미리 記述하여 둔다.

그림10에 記號變換의 例를 보인다.

한편 現在의 記號變換만으로도 單純한 文(基本形式文)에서는 適切한 出力을 얻을 수 있으나, 對象文章의 擴張을 爲해서는 構造變換이 試圖되어야 한다. --例로 英語의 no, few, little 등의 否定構文과 形容詞의 名詞化(例: logical formular → 論理式) 등의 境遇에는 Lambda-深層構造 自體의 根本的인 變換이 要求된다.

2. 韓國語의 生成

韓國語文의 生成過程은 變換過程과 同一하게 韓國語의 深層構造를 주어진 順序에 依해 分解하면서 統

일련문 : It consists of logical symbols called Metalanguage #

```
(SEM
  (IT
    (LAMBDA
      X77
      ((A*
        ((F*
          (LAMBDA
            X75
            ((A* METALANGUAGE)
            (LAMBDA X76 (CALL X75 X76))))))
          (LOGICAL SYMBOL)))
          (LAMBDA X78 (CONSIST-OF X77 X78))))))

** SYMBOL TRANSFER **

(SEM
  (E
    (LAMBDA
      X77
      ((A*
        ((F*
          (LAMBDA
            X75
            ((A* METALANGUAGE)
            (LAMBDA X76 (BURDA X75 X76))))))
          (MONRI KIHU))))
          (LAMBDA X78 (KUSUNGDEUDA X77 X78))))))
```

그림10. 기호변환의 예

Fig. 10. An example of symbol transfer.

```
(SEM (THE (COMMAND))
  (LAMBDA X72
    ((ND (OPERAND))
      (LAMBDA X73
        (NEED (X72 X73))))))

(SEM ((GYF T)
  (COMMAND))
  (LAMBDA X72
    (((DR A)
      (OPERAND))
      (LAMBDA X73
        ((PILYO T)
          (X72 X73))))))

(SEM ((GYF T)
  (COMMAND))
  (LAMBDA X72
    ((A* (((GAF T)
      (LAMBDA X115
        ((PILYO T)
          (X72 X73))))
      (OPERAND)))
      (HB S))))
```

그림11. 구조변환의 일례

Fig. 11. An example of structure transfer.

語構造를 生成하고 다시 終端記號에 對한 形態素 合性을 行하여 最終的인 出力을 얻는다.

英語의 Lambda-深層構造는 對應하는 入力文의 各 lexical element를 一定한 規則에 依해 部分的인 順序關係를 附與해서 tree 構造를 生成한 것으로 看做할 수 있다.

이와같은 意味解析을 前提로 했을 때, 英語의 名詞 表現 pretty girl에 對應하는 Lambda-深層構造는 (pretty)girl로 表現되며, 이것은 pretty라는 形容詞 概念 即 귀여운이라는 附加의 意味를 girl이라는 名詞 概念에 附加하여 名詞 라는 關係로 表現할 수 있다.

形容詞 名詞

以上의 意味解析 過程은 形容詞 概念이 名詞 概念으로 函數 適用되어 새로운 名詞 概念이 導出되며,

最初の形容詞概念에 對應하는名詞의 relationship으로 나타낼 수 있다.

한편 生成過程은 各各의 統語構造와 終端記號에 割當된 情報을 利用하여 助詞處理, 그리고 附加語의 選擇等を 決定한다.

實際 生成過程에서의 助詞處理는 一意的으로 行하기 때문에 生成辭典上에는 各各의 格構造에 對應하는 助詞가 一定하게 定義되어 있다. 그림12는 生成辭典의 一例이다.

```

(FUT 'KYULHAPHADA' GENER '( (NIL 0) ((FHR ACTOR OBJ1 OBJ2)
                                (SEM CONVINCE ACTOR OBJ1 OBJ2)
                                (JOSA (GA, ACTOR))
                                (EUL, OBJ1)
                                (KWR, OBJ2)
                                (E, OBJ2))))))
(FUT 'BAECHIHADA' GENER '( (NIL 0) ((FHR ACTOR OBJ1)
                                (SEM ALLOCATE ACTOR OBJ1)
                                (JOSA (GA, ACTOR))
                                (EUL, OBJ1))))))
(FUT 'HAESUCKHADA' GENER '( (NIL 1) ((FHR ACTOR OBJ1)
                                (SEM ANALYZE ACTOR OBJ1)
                                (JOSA (GA, ACTOR))
                                (EUL, OBJ1)
                                (EBAEHAE, OBJ1))))))
(FUT 'BYUNHWANHADA' GENER '( (NIL 0) ((FHR ACTOR OBJ1 OBJ2)
                                (SEM TRANSDUCE ACTOR OBJ1 OBJ2)
                                (JOSA (GA, ACTOR))
                                (EUL, OBJ1)
                                (E, OBJ1))))))
(FUT 'BOSUCKEUI' GENER '( (ADJ
                            (CHD, 2) (CHANGE KAJINGDEU))))))
(FUT 'SUCKEUI' GENER '( (ADJ
                            (CHOU, 0))))))
(FUT 'MAHJUL' GENER '( (CHOU
                            (CHD, 2) (CHANGE HAEYONG))))))
    
```

그림12. 생성사전의 일례

Fig. 12. An example of generation dictionary.

예를들어 그림12에서 動詞 “결합하다.” (KYULHAPHADA)의 論理表現 즉, Lambda-深層構造上에서의 助詞의 決定은 ACTOR;가, OBJ1;을, OBJ2;과, 에로 이루어진다. OBJ2;[과, 에]의 選擇은 出力結果에 後 編集 概念을 導入하여 實現한다.

한편 多意語의 生成 決定은 Lambda-深層構造의 論理表現만으로는 그 處理가 不可能하므로 生成辭典에 semantic primitive를 割當하여 語意를 決定한다. 그림12의 下段은 semantic primitive M(move)를 갖는 辭典의 一例이다.

生成過程에서 利用하는 辭典은 對應되는 韓國語 單語를 標識로 해서 property list 形式으로, 統語範疇와 格構造에 對應하는 助詞의 pattern, 附加語의 選擇 情報等이 記述된다. 對應하는 附加語의 記述는 獨立된 辭典에 位置한다.^{16, 17)}

그림13에 韓國語 生成過程의 概要를 보인다.

V. 實驗 및 考察

機械翻譯 시스템의 모든 過程은 VAX-11/780上에서 Interlisp-F4를 利用하여 記述하였다.

解析過程에서 使用한 解析規則은 約220個 程度이며,

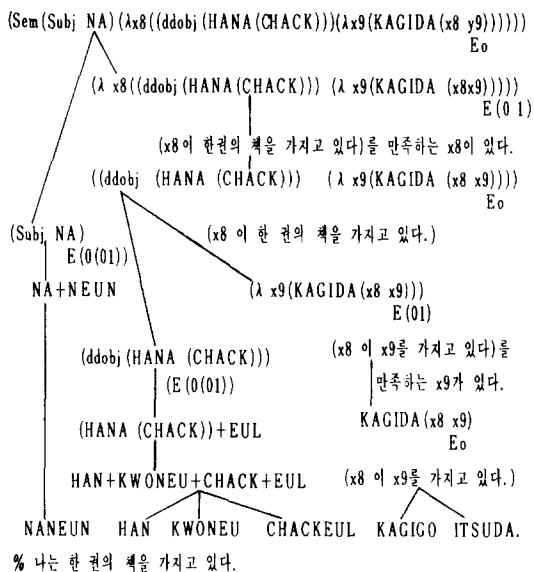


그림13. 생성과정의 개요

Fig. 13. An outline of generation procedure.

解析規則上에서 英語의 動詞記述은 Hornby의 動詞 分類를 參照하였다.

實驗에 使用된 入力文은 中學校 1~2 學年 程度의 教科書와 컴퓨터 manual中에서 比較的 簡單한 文章 (具體的으로는 各 形式分類의 標準例를 中心으로)을 對象으로 實驗하였다. 實驗過程에서 시스템의 作成은 다음과 같은 制約 條件下에서 行한다.

- 1) 辭典에 記述한 英語 單語는 使用頻度가 높은 重要單語가 아니면 可扱의 多品詞語의 選擇을 排除하였다.
- 2) 變換過程에서는 記號變換을 爲主로 한다.
- 3) 韓國語의 格助詞는 一意的으로 決定한다.

實驗過程에서 가장 考慮한 對象은 解析規則의 作成과 그 檢證 그리고 意味處理에 關한 問題이다. 實際, 解析規則과 辭典을 利用하여 構文解析을 行하는 解析過程에서 統語의 曖昧한 文에 對해 그 曖昧性에 對應하는 만큼의 構文解析 結果가 non-deterministic하게 出力한다. 그러므로 初期에 出力된 構文 解析 結果를 check해서 適當하지 않으면 User Interface에 依해서 順次的으로 다음의 出力을 要求하여 實行하였다. 그러나 解析規則의 未備에 依하여 根本的으로 適當한 解析結果를 얻을 수 없는 境遇도 있었다.

變換過程에서는 記號變換을 中心으로 하고 構造變換을 制限하였기 때문에 不自然스런 韓國語 出力을 가져오는 수가 있었으며, 記號變換 自體에서도 專門用語

```

- (EROS)
*** ENGLISH KOREAN MACHINE TRANSLATION SYSTEM (V1) ***
                                --- DO NOT EXCEED THIS POINT * ---
** START INPUT ROUTINE **
INPUT-(IT CONSISTS OF LOGICAL SYMBOLS CALLED METALANGUAGE *)
** INPUT NEW WORD DICTIONARY CONTENT FOR * SYMBOLS **
1 : CATEGORY INPUT-(NOUN)
2 : SYN. VALUE-(NBR. PL)
3 : SEM. VALUE-(SEM. SYMBOL)
4 : ASS. VALUE-(ASS. O)
** PARSING SUCCEEDED **
** SCAN MODE SELECTION G(OOD) R(ETRY) T(OP) D(ISPLAY) **
INPUT-D
SENTENCE
  DCL
    CLAUSE
      NP
        PRONOUN.. IT
      VP
        VP3.. CONSISTS
        PREPP
          PREP.. OF
          NP
            NP
              NP1
                NOUN
                  ADJP.. LOGICAL
                  NOUN.. SYMBOLS
            VPP
              VP23.. CALLED
              NP
                NP1
                  NOUN.. METALANGUAGE
          END.. *
** SCAN MODE SELECTION G(OOD) R(ETRY) T(OP) D(ISPLAY) **
INPUT-G
(SEM
  (IT
    (LAMBDA
      X77
      ((A*
        (P*
          (LAMBDA
            X75
            ((A* METALANGUAGE)
              (LAMBDA X76 (CALL X75 X76) )))
            (LOGICAL SYMBOL) ))
            (LAMBDA X78 (CONSIST-OF X77 X78) )))
** SYMBOL TRANSFER **
(SEM
  (E
    (LAMBDA
      X77
      ((A*
        (P*
          (LAMBDA
            X75
            ((A* METALANGUAGE)
              (LAMBDA X76 (BURDA X75 X76) )))
            (NONRI KIH0) ))
            (LAMBDA X78 (KUSUNGDEUDA X77 X78) )))
** OUTPUT **
ENEUN METALANGUAGERA BULRIUNEUN NONRI KIHORO KUSUNGDEUNDA

```

그림 14. 실험결과와의 一例

Fig. 14. An example of system output.

를 制外 하고는 適切한 對象言語의 選擇에 難點이 있었다.

또한 韓國語 生成時에도 助詞의 選擇을 一意的으로 決定하였기 때문에 助詞處理에 技述의인 未備點(特히 副詞格이나 補格等에)이 存在한다.

意味解析에 있어서는, 入力文 全體의 意味 또는 入力文中 任意的 한 單語上에라도 曖昧性이 存在하면 種種 不適當한 意味를 갖는 出力이 얻어졌다. 勿論 對象 入力文에 限定하여 特別한 辭典記述을 하거나 解析 規則을 作成하면 어느 程度 效果的인 出力이 期待되나, 一般的인 境遇에의 適用에는 不足하다.

그림에 實驗結果의 一例를 보인다.

VI. 結 論

本 論文에서는 Montague 文法에 基礎하여 Lambda-範疇 言語로 英語의 深層意味構造를 生成, 이를 中間 表現으로 해서 韓國語의 出力을 導出하는 方式을 採擇하여 英韓機械翻譯 시스템을 構成하고, 英語의 基本 文型에 對하여 韓國語의 出力을 얻음으로 해서 그 實現 可能性을 보였다.

本 研究에서 提案한 시스템은 아직 翻譯이 可能한 入力文의 pattern이 限定되어 있어, 實用化하기에는 難點이 있으나, 解析規則의 嚴密化와 意味處理를 補完할 수 있는 辭典 記述力의 增大, 그리고 韓國語 生成時 助詞處理等의 諸問題를 研究 補完한다면, manual 이나 一般 書式等과 같은 制限된 範圍에서의 翻譯은 어느 程度 可能하리라고 본다.

參 考 文 獻

[1] A. Barr & E.A. Feigenbaum: The handbook of Artificial Intelligence, *William Kaufmann, Inc.* vol. 1, pp. 233-238.

- [2] 辻井潤一: 人間と機械の自然な對話を目指す自然言語 理解, *Nikkei Electronics*, pp. 185-203, 9. 1982.
- [3] T. Winograd: *Understanding Natural Language*, *Academic Press*, 1972.
- [4] T. Winograd: *Language as a Cognitive Process*, *Addison-Wesley*, 1983.
- [5] David R. Dowty: *A Guide to Montague's PTQ*, *Indiana University Linguistic Club*, 1978.
- [6] M.J. Cresswell: *Logics & Languages*, *METHUN & Co. LTD*, 1978.
- [7] Partee: *Montague Grammar*, *Academic Press*, 1976.
- [8] 長尾眞: 言語工學, pp. 94~125, 昭昇堂, 1983.
- [9] 小坂晃: 現行의 英文解析システムの 現狀と問題點, 京都大學
- [10] Pratt. V. R.: *A Linguistic Oriented Programming Language*, *Proc. IJCAI-73*, pp. 372-381, 1973.
- [11] 田中, 佐藤, 元吉: 自然言語處理のためのプログラミングシステム擴張 LINGOL について, *信學論*, vol. 60,
- [12] 西田, 清野, 堂下: Montague 文法に基づく英文和譯システムの 試作, *情報處理學會論文誌*, vol. 23, no. 2, pp. 107-115, 1982.
- [13] 西田, 堂下: 自然言語解析のためのプログラミングシステム COMPLAN について, *情報處理學會論文誌*, vol. 24, no. 4, pp. 396-405, 1982.
- [14] 堂下修司: 意味表現を 仲介する 科學技術文獻機械翻譯 시스템의 作成に關する研究, 京都大學.
- [15] 長尾眞: 數と式と文の處理, 岩波講座 情報科學.
- [16] 金敏洙: 國語文法論, 일조각, 1978.
- [17] 성광수: 國語助詞에 對한 研究, 螢雪出版社, 1980.