
Faculty of Mathematical Sciences



University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands

Phone: +31-53-4893400

Fax: +31-53-4893114

Email: memo@math.utwente.nl

www.math.utwente.nl/publications

MEMORANDUM No. 1662

On Yao's method of translation

X. LIU¹ AND C. HOEDE

NOVEMBER, 2002

ISSN 0169-2690

¹Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, P.R. China

On Yao's Method of Translation*

X. Liu

Department of Applied Mathematics
Northwestern Polytechnical University
Xi'an, Shaanxi 710072, P.R.China

C. Hoede

Department of Applied Mathematics
University of Twente
P.O.Box 217
7500AE Enschede, The Netherlands

Abstract

Machine Translation, i.e., translating one kind of natural language to another kind of natural language by using a computer system, is a very important research branch in Artificial Intelligence. Yao developed a method of translation that he called "Lexical-Semantic Driven". In his system he introduced 49 "relation types" including case relations, event relations, semantic relations, and complex relations. The knowledge graph method is a new kind of method to represent an interlingua between natural languages. In this paper, we will give a comparison of these two methods. We will translate one Chinese sentence cited in Yao's book by using these two methods. Finally, we will use the relations in knowledge graph theory to represent the "relations" in Lexical-Semantic Driven, and partition the relations in Lexical-Semantic Driven into groups according to the relations in knowledge graph theory.

Key Words: lexical-semantic driven, knowledge graph, relation type, translation

AMS Subject Classifications: 05C99, 68F99

1 Introduction

Having investigated how knowledge graph can be used to translate AC into MC [Liu & Hoede, to appear (a)], we should now pay attention to attempts to translate MC into English and back. Yao

* Research supported in part by SFAC through grant 01J53079.

has developed a method of translation that he called “Lexical-Semantic Driven” [Yao, 1996]. As in knowledge graph theory semantics are put on the first place too, a comparison seems natural.

Lexical-Semantic Driven (LSD) is a method to translate Chinese to English and vice versa, that has been developed at the Northeast University in China. In order to translate languages, there are 49 “relation types” introduced to represent sentences. Basically, this method is a kind of formal grammar with a complex feature set and complex operations leading to a unified representation.

The “relations” in LSD are not pure semantic relations. They include relationships between the nodes in the sentences, the compound relations between sentences, and something that are really not relations in the traditional sense. In a way they form the extension of case relations. They include compound case relations (including agent, object, implement, ...), event relations (including condition, connection, ...), semantic relations (including part-of,...), and complex relations (including possessor, purpose, ...). There are 49 types of “relations” in all.

The knowledge graph method is a new method to represent natural languages as an interlingua. In this paper, we will give a comparison of these two methods. We will translate one Chinese sentence cited in Yao’s book by using these two methods. Finally, we will use the relations in knowledge graph theory to represent the relations in Lexical-Semantic Driven, and partition the relations in LSD into groups according to the relations in knowledge graph theory.

2 Parsing in Lexical-Semantic Driven

Lexical-Semantic Driven (LSD) is used mainly as a tool to build a machine translation system between English and Chinese. The machine translation system CETRAN (Chinese English TRANslation system) based on LSD consists of two main parts: Parser and Generator. There are the following 6 steps in the parser.

- 1、 Input of the original Chinese sentence.
- 2、 Word segmentation.
- 3、 Lexical analysis.
- 4、 Syntactic analysis
- 5、 Semantic analysis
- 6、 Concept analysis and representation by inter-transmittal language

In these processes, the Machine Dictionary and the Rule Set are the basis. The system CETRAN contains a very complex machine dictionary and a very complex rule set. In his book named Natural Language Understanding [Yao, 1996], Yao gives us the basic ideas of his theory.

Now, we will give an example to show how his theory works.

2.1 Input of the original Chinese sentence

可 怜 的 女 孩 在 门 口 站 了 30 分 钟 。

ke3lian2 de nv3hai2 zai4 men2kou3 zhan4 le 30 fen1zhong1.

pitiful of girl at doorway stand (complete) 30 minute

Note: In order to explain his method, we will add the Chinese spelling and according English words under the original sentence in italics.

2.2 Word segmentation

As we know, there are no division symbols between words or characters in Chinese sentences. So, the first step in CETRAN, or in any other system, is word segmentation. In his book, Yao does not give this step in detail. But he gives the result of the word segmentation as follows.

可 怜 (a, v, a, v) (*ke3lian2, pitiful*)

的 (u, n) (*de, of*)

女 孩 (n) (*nv3hai2, girl*)

在 (p, v, v, v, p, d) (*zai4, at*)

门 口 (s, s) (*men2kou3, doorway*)

站 (v, n, s, v) (*zhan4, stand*)

了 (u, v, v, u, y) (*le, complete*)

30 (m, m) (*30, thirty*)

分 钟 (t) (*fen1zhong1, minute*)

。 (g) (*。 full stop*)

The letters between brackets after each word denote the possible part of speech that word can play. For example, the word “可怜(*ke3lian2, pitiful*)” has four: two kinds of adjective and two verbs.

In the machine dictionary in the system CETRAN, the Chinese word is divided into 22 classes according to the result of this linguist. Moreover, Chinese words are divided into 71 subclasses according to the semantic functions and syntax roles of the word displayed in the sentences. So, the word “可怜(*ke3lian2, pitiful*)” belongs to four subclasses: two adjectives and two verbs.

The letters and their meanings are shown in the following table.

| a | v | u | N | p | d |
|-----------|------|----------------|------|-------------|--------|
| adjective | verb | auxiliary word | noun | preposition | adverb |

| | | | | | |
|------------------|-----------|-------------|-----------|-------------|--|
| s | y | m | T | g | |
| orientation word | mood word | number word | time word | punctuation | |

2.3 Lexical analysis

The kernel of LSD, as its name shows, is to analyze sentences by means of lexical information and semantic information. The lexical analysis consists of two parts: to determine the part of speech and to determine phrases.

2.3.1 To determine the part of speech

In order to determine the part of speech of each word in the sentence, the system has a very complex rule set. For example, in order to determine the word in the sentence, we can use the following rules.

- (1) To determine the word “可怜(*ke3lian2, pitiful*)”

LEX=可怜, (dkind_a), RULE: @head, , ^ (CCAT.a) + ‘的’ +(CCAT.n; CCAT.r) → setmark (a)

(LEX=*pitiful*, (dkind_a), RULE: @head, , ^ (CCAT.a) + ‘de’ +(CCAT.n; CCAT.r) → setmark (a))

The meaning of each symbol is shown in the following list.

| | |
|------------------|--|
| LEX=可怜 | The current word is “可怜(<i>ke3lian2, pitiful</i>)” |
| (dkind_a), RULE | The rule to determine an adjective |
| @head | Detect whether the current word is the first one in the sentence |
| ^ (CCAT.a) | Detect whether the current word is an adjective |
| (CCAT.n; CCAT.r) | Noun or pronoun (where r denotes pronoun) |
| Setmark (a) | Set the current word to be an adjective |

The rule says that if the following word of the current one is ‘de’ and the next is a noun or a pronoun, then the current word is an adjective.

- (2) To determine the word “的(*de, of*)”

LEX=的, (dkind_u), RULE: ^ (‘的’ , n, u) + (—CCAT.g) → setmark (u)

(LEX=*de*, (dkind_u), RULE: ^ (‘de’ , n, u) + (—CCAT.g) → setmark (u))

As above the list of meanings is

| | |
|-----------------|---|
| LEX=的 | The current word is “的(<i>de, of</i>)” |
| (dkind_u), RULE | The rule to determine an auxiliary word |
| ^ (‘的’ , n, u) | Detect whether the current word is noun or auxiliary word |

| | |
|-------------|--|
| (—CCAT.g) | The following is not punctuation. |
| Setmark (u) | Set the current word to be an auxiliary word |

- (3) To determine the word “在(zai4, at)”

LEX=在, (dkind_v), RULE: ^(CCAT.p, CSUBCAT.pp2)+(CCAT.s) → setmark (p)

(LEX=at, (dkind_v), RULE: ^(CCAT.p, CSUBCAT.pp2)+(CCAT.s) → setmark (p))

The meaning of each symbol is:

| | |
|-----------------------|---|
| LEX=在 | The current word is “在(zai4, at)” |
| (dkind_p), RULE | The rule to determine a preposition |
| ^(CCAT.p.CSUBCAT.pp2) | Detect whether the current word is a preposition, and its subcategory is preposition to form a frame. (where pp2 denotes preposition to form a frame) |
| (CCAT.s) | Orientation word (where s denotes orientation word) |
| Setmark (p) | Set the current word to be a preposition |

- (4) To determine the word “站(zhan4, stand)”

LEX=站, (dkind_v), RULE: ^((n, v); (v, d))+ (CCAT.u) → setmark (v)

(LEX=stand, (dkind_v), RULE: ^((n, v); (v, d))+ (CCAT.u) → setmark (v))

The meaning of each symbol is:

| | |
|-------------------|--|
| LEX=站 | The current word is “站(zhan4, stand)” |
| (dkind_v), RULE | The rule to determine a verb |
| ^((n, v); (v, d)) | Detect whether the current word is (noun, verb), or (verb, adverb) |
| (CCAT.u) | auxiliary word |
| Setmark (v) | Set the current word to be a verb |

- (5) To determine the word “了(le, complete)”

LEX=了, (dkind_u), RULE: (CCAT.v; CCAT.a; CCAT.z) + ^((CSUBCAT.uu1; CSUBCAT.uu2), CCAT.v) → setmark (u)

(LEX=了, (dkind_u), RULE: (CCAT.v; CCAT.a; CCAT.z) + ^((CSUBCAT.uu1; CSUBCAT.uu2), CCAT.v) → setmark (u))

The meaning of each symbol is:

| | |
|--------------------------|---|
| LEX=了 | The current word is “了(le, complete)” |
| (dkind_u), RULE | The rule to determine an auxiliary word |
| (CCAT.v; CCAT.a; CCAT.z) | Verb, or adjective, or state word. (where z denotes state word) |

| | |
|---|--|
| $\wedge((\text{CSUBCAT.uu1}; \text{CSUBCAT.uu2}), \text{CCAT.m})$ | Detect whether the subcategory of the current word is a structure auxiliary word or a tense auxiliary word, the following word is number. (where uu1 and uu2 denote structure auxiliary word and tense auxiliary word, respectively) |
| Setmark (u) | Set the current word to be an auxiliary word |

After executing the above rules, we get the following.

| | | | | |
|-----------------|-----------------|----------------|-------------------|------------------|
| 可怜 | 的 | 女孩 | 在 | 门口 |
| <i>ke3lian2</i> | <i>de</i> | <i>nv3hai2</i> | <i>zai4</i> | <i>men2kou3</i> |
| <i>pitiful</i> | <i>of</i> | <i>girl</i> | <i>at</i> | <i>doorway</i> |
| adjective | auxiliary word | noun | preposition | orientation word |
| 站 | 了 | 30 | 分钟 | 。 |
| <i>zhan4</i> | <i>le</i> | <i>30</i> | <i>fen1zhong1</i> | <i>ju4hao4</i> |
| <i>stand</i> | <i>complete</i> | <i>thirty</i> | <i>minute</i> | <i>full stop</i> |
| verb | auxiliary word | number word | time word | punctuation |

2.3.2 To determine phrases

In order to get phrase of the sentence from the words, Yao gave the following rules. The purpose of the phrase processing is only to form phrases by means of deleting some function words, and to combine these discrete phrases (forming a forest) to form a phrase tree.

- (1) To determine the role of the word “可怜(*ke3lian2, pitiful*)”

LEX=可怜, (a), RULE: $\wedge\text{CCAT.a ' + ' 的 ' + ' (CCAT.n; CSUBCAT.nn1)} \rightarrow \text{.删. } \wedge 1, \text{定. } \wedge 1, 1 \wedge \text{改. } \downarrow (\text{SEMREAL.EXP0})$

(LEX=*pitiful*, (a), RULE: $\wedge\text{CCAT.a ' + ' de ' + ' (CCAT.n; CSUBCAT.nn1)} \rightarrow \text{.DELETE. } \wedge 1, \text{SET. } \wedge 1, 1 \wedge \text{CHANGE. } \downarrow (\text{SEMREAL.EXP0})$)

The meaning of each symbol is as follows.

| | |
|--|--|
| LEX=可怜 | The current word is “可怜(<i>ke3lian2, pitiful</i>)” |
| (a), RULE | Determine the role of adjective in the sentence |
| $\wedge \text{CCAT.a.}$ | The current word is an adjective |
| (CCAT.n.CSUBCAT,nn1) | The word is a noun, and its subcategory is a general noun (where nn1 denotes general noun) |
| .删. $\wedge 1$ | Delete the first word on the right of the current word |
| 定. $\wedge 1$ | Set the first word on the right of the current word to be the current word |
| $1 \wedge \text{改. } \downarrow (\text{SEMREAL.EXP0})$ | Change the first word on the left of the current one as the son of the current word, and set the semantic relation to be EXPERIENCE, 0 denotes the relation is inverse |

The rule tells us that if the current word (“可怜(ke2lian2, pitiful)”) is an adjective, the following word is “的(de, of)”, and the next (“女孩(nv3hai2, girl)”) is a noun with the subcategory belonging to a general noun, then we have to delete the first word “的(de, of)” on the right of the current word, make the first word (“女孩(nv3hai2, girl)”) on the right of the current word the current word, to change the first word (“可怜(ke2lian2, pitiful)”) on the left of the current one to be the son of the current word, and to set the semantic relation to be inverse EXPERIENCE.

After executing the above rules, we get a phrase (“可怜女孩(ke2lian2 nv3hai2, pitiful girl)”) by deleting the function word “的(de, of)”.

The rules used to determine the roles other words are playing in the sentence are as follows.

- (2) To determine the role of the word “站(zhan4, stand)”

LEX=站, (v), RULE: ^ (v, -vv6) + (CSUBCAT.uu2, .取. \$AASP1) → .赋. ^(^TENSE, \$AASP1), , .删. ^1

(LEX=stand, (v), RULE: ^ (v, -vv6) + (CSUBCAT.uu2, .TAKE. \$AASP1) → .GIVE. ^(^TENSE, \$AASP1), , .DELETE. ^1)

The symbols indicate the following meaning.

| | |
|------------------------|---|
| LEX=站 | The current word is “站(zhan4, stand)” |
| (v), RULE | Determine the role of verb in the sentence |
| ^ (v, -vv6) | The current word is a verb, but not a copula |
| (CSUBCAT.uu2) | The subcategory is tense auxiliary word |
| .取. \$AASP1 | Take the value in the variable \$AASP1 |
| .赋. ^(^TENSE, \$AASP1) | Give the value of the variable \$AASP1 to the tense of the current word |
| .删. ^1 (.DELETE. ^1) | Delete the first word on the right of the current word |

The rule tells that if the current word (“站(zhan4, stand)”) is a verb, but not a copula, the following word (“了(le, complete)”) is a tense auxiliary word, then we have to take the value in the variable \$AASP1 to give it to the tense of the current word, and to delete the first word on the right of the current word (“了(le, complete)”). Here the variable #AASP1 describes the aspect.

After executing the above rules, we delete the word “了(le, complete)”, but we get the tense (perfect) of the verb.

- (3) To determine the role of the word “站(zhan4, stand)” further

LEX=站, (v), RULE: ‘在’ + _ + ^ (CCAT.v), , (1^Rrest, 121) → @cpd(^, 2^), , .删. 2^, 1^ .改. ^ ↓ (SEMREAL.LOC), ^ .增. (PHRASE.vp)

(LEX=stand, (v), RULE: ‘at’ + _ + ^ (CCAT.v), , (1^Rrest, 121) → @cpd(^, 2^), , .DELETE. 2^, 1^ .CHANGE. ^ ↓ (SEMREAL.LOC), ^ .ADD. (PHRASE.vp))

| | |
|-------|---------------------------------------|
| LEX=站 | The current word is “站(zhan4, stand)” |
|-------|---------------------------------------|

| | |
|--------------------------|--|
| (v), RULE | Determine the role of verb in the sentence |
| ‘在’ + _ + ^ (CCAT.v) | The current word is a verb, and foregoing is the word “在(zai4, at)” |
| (1^Rrest, 121) | The semantic classifying code of the foregoing word of the current one is 121 where 121 denotes the space word |
| @cpd(^, 2^) | Put the dynamic feature table of the 2 nd word on the left of the current word to the current word (@cpd is a function) |
| .删. 2^ (.DELETE. ^2) | Delete the 2 nd word on the left of the current word |
| 1^ .改. ^ ↓ (SEMREAL.LOC) | Change the first word on the left of the current one as the son of the current word, and set the semantic relation to be LOCATION |
| ^ .增. (PHRASE.vp) | Add the verb phrase to the current word |

The rule tells that if the current word (“站(zhan4, stand)”) is a verb, the 1st word (“门口(men2kou3, doorway)”) on the left is a space word, and the 2nd word on the left is (“在(zai4, at)”), then we have to put the dynamic feature table of the 2nd word (“在(zai4, at)”) on the left to the current word (“站(zhan4, stand)”), and we have to delete the 2nd word (“在(zai4, at)”) on the left, have to change the 1st word (“门口(men2kou3, doorway)”) on the left as the son of the current word, to set the semantic relation to be LOCATION, and to add the verb phrase to the current word.

After executing the above rules, we delete the word (“在(zai4, at)”), but we get the verb phrase (the location of the act described by the verb stand is doorway).

For the semantic of words, LSD follows Roget’s International Thesaurus [Roget, 1977] to define a semantic classifying code.

- (4) To determine the role of the word “30(30, thirty)”

LEX=30, (m), RULE: ^ (CCAT.m) + (CCAT.n; CCAT.t) → .定. ^1, ,1^ .改. ^ ↓ (SEMREAL.NUM), ,^ .增. (PHRASE.np)

(LEX=30, (m), RULE: ^ (CCAT.m) + (CCAT.n; CCAT.t) → .SET. ^1, ,1^ .CHANGE. ^ ↓ (SEMREAL.NUM), ,^ .ADD. (PHRASE.np))

| | |
|--------------------------|--|
| LEX=30 | The current word is “30(30, thirty)” |
| (m), RULE | Determine the role of a number word in the sentence |
| ^ (CCAT.m) | The current word is a number word |
| (CCAT.n; CCAT.t) | The word is a noun or a time word |
| .定. ^1 | Set the 1 st word on the right to be the current word |
| 1^ .改. ^ ↓ (SEMREAL.NUM) | Change the first word on the left of the current one to be the son of the current word, and set the semantic relation to be NUMBER |
| ^ .增. (PHRASE.np) | Add the noun phrase to the current word |

The rule tells us that if the current word (“30(30, thirty)”) is a number word and the 1st word (“分钟(fen1zhong1, minute)”) on the right is a noun or a time word, then we have to set the 1st word (“分钟(fen1zhong1, minute)”) on the right as the current word, to change the 1st word (“30(30, thirty)”) on the left as the son of the current word, to set the semantic relation to be NUMBER, and to add the noun phrase to the current word.

After executing the above rules, we get the noun phrase (“30 分钟(30 fen1zhong1, thirty minutes)”).

- (5) To determine the role of the word “站(zhan4, stand)” further

LEX=站, (v), RULE: $\wedge(\text{CCAT.v, -vv8, -vv9}) + (\text{CCAT.q; CCAT.t}), \wedge 1 \downarrow \#1 (.取. \$\text{CCAT1}) \rightarrow \wedge 1 .改. \wedge \downarrow (\text{SEMREAL.QNT, SYNREAL.comp})$

(LEX=stand, (v), RULE: $\wedge(\text{CCAT.v, -vv8, -vv9}) + (\text{CCAT.q; CCAT.t}), \wedge 1 \downarrow \#1 (.TAKE. \$\text{CCAT1}) \rightarrow \wedge 1 .CHANGE. \wedge \downarrow (\text{SEMREAL.QNT, SYNREAL.comp})$)

| | |
|--|---|
| LEX=站 | The current word is “站(zhan4, stand)” |
| (v), RULE | Determine the role of a verb in the sentence |
| $\wedge(\text{CCAT.v, -vv8, -vv9})$ | The current word is a verb, but not a necessary auxiliary verb or a possible auxiliary verb (-vv8 and -vv9 denote necessary auxiliary verb and possible auxiliary verb respectively) |
| (CCAT.q; CCAT.t) | The word is a qualifier (quantity word) or a time word |
| $\wedge 1 \downarrow \#1 (.取. \$\text{CCAT1})$ | A son of the 1 st word on the right of the current word takes the value of the variable \$CCAT1 |
| $\wedge 1 .改. \wedge \downarrow (\text{SEMREAL.QNT, SYNREAL. Comp})$ | Change the first word on the right of the current one to be the son of the current word, to set the semantic relation to be QUANTITY and the syntactic relation to be Comp. |

The rule tells that if the current word (“站(zhan4, stand)”) is a verb, but not a necessary auxiliary verb or a possible auxiliary verb, the following word (“分钟(fen1zhong1, minute)”) is qualifier or time word, and a son (“30(30, thirty)”) of the 1st word on the right takes the value of the variable \$CCAT1, then we have to set the 1st word (“分钟(fen1zhong1, minute)”) on the right to be the son of the current word (“站(zhan4, stand)”), to set the semantic relation to be QUANTITY, and the syntax relation to be Comp. The variable \$CCAT1 describes the category.

After executing the above four rules, the remaining words in the sentence, note that words were deleted, are the words “女孩(nv3hai2, girl)”, “门口(men2kou3, doorway)”, “站(zhan4, stand)” and “分钟 (fen1zhong1, minute)”.

Yao now combines the result of the foregoing analyses into a graph structure. This takes place as follows.

The role of “可怜, ke3lian2, pitiful” turned out to be that of an adjective to “女孩, nu3hai2, girl”. These words are the first and third in the sentence. There are 10 words (including punctuation) in all, that can be found in the graph structure via the number 0, 1, 2, ..., 9. By the result mentioned

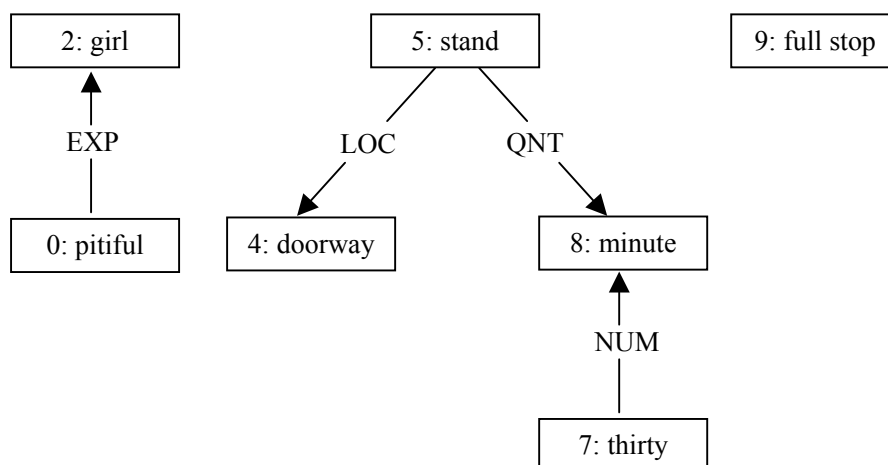
Yao constructs

2: girl
!
! ←EXP→0: pitiful

where we have omitted all further information between parentheses. The exclamation marks are used to copy the word as well as linking the two words. They function like a token. In this way three parts are given.

2: (女孩(nv3hai2), n, nn2, 1111, girl)
!
! ←EXP→0: (可怜(ke3lian2), a, aaa, 2314, pitiful)
!
5: (站(zhan4), v, vv3, 2102, stand, vp, pef)
!
! —LOC→4: (门口(men2kou3), s, sss, 1212, doorway)
!
! —QNT→8: (分钟(fen1zhong1), t, tt1, 122216, minute, np)
!
! —NUM→7: (30, m, mm1, 133503, thirty)
9: (., g, gg1, , .)

The equivalent “graph” representation is as follows.



Note: The numbers before each line are the numbers of the word position in the sentence (the number of the first word is 0). Word 1 was “de”, which is deleted to determine the relation between “girl” and “pitiful”. Word 3 was “zai4” which is deleted to get the LOCATION of “stand”. Word 6 was “le”, which indicates the tense. The numbers in brackets are semantic classifying codes of the words. 1111 denotes biology, 2314 denotes psychology phenomena, 1212 denotes position relation, 2102 denotes leg action, 122216 denotes time period, 133503 denotes name of quantity, nn2 denotes the relation between people, aaa denotes a general adjective, pp1 denote singleness preposition, sss denotes location word, vv3 denotes entirety self-motion verb, tt1 denotes time word, mm1 denotes general number word, and gg1 denotes full stop.

Note that Yao’s method for attempting phrases in a way “chunks” up the sentence into four chunks: “pitiful girl”, “at doorway”, “stand (perfect)”, and “thirty minutes”. In the complicated notation system also are mentioned syntactic functions and supertypes of the occurring words, like 122216 for “time period” in case of the word “minute”.

2.4 Syntax analysis

Because LSD is based on lexical information and semantic information, syntax analysis is no longer necessary. The reason to keep syntax analysis is to fit the needs of other languages.

After deleting some function words to form phrases, the next step is to determine the role of the words in the sentence, including determining the sentence type. In LSD, sentences are divided into three different types: simple sentences, questions and complex sentences. The discussed sentence is a simple sentence, and there exists the following rule.

LEX=站, (base0), RULE: (1111; 113; 1113) + ^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) → 1^ .改. ^ ↓ (SYNREAL.sub)

^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (121) → 1^ .改. ^ ↓ (SYNREAL.comp)

^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (122) → 1^ .改. ^ ↓ (SYNREAL.obj)

(LEX=stand, (base0), RULE: (1111; 113; 1113) + ^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) → 1^ .Change. ^ ↓ (SYNREAL.sub)

^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (121) → 1^ .CHANGE. ^ ↓ (SYNREAL.comp)

^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (122) → 1^ .CHANGE. ^ ↓ (SYNREAL.obj)

Symbol explanations.

| | |
|---|--|
| LEX=站 | The current word is “站(zhan4, stand)” |
| (base0), RULE | The rule to determine the basic sentence type 0 |
| (1111; 113; 1113) | The semantic classifying code is 1111 (person), or 113 (organization), or 1113 (animal) |
| ^ (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) | The subcategory is a “motion-by-others” verb, general verb, a “self-motion” verb, or a derived verb. |
| 1^ .改. ^ ↓ (SYNREAL.sub) | Change the 1 st word on the left as the son of the current word, and to set the syntax relation is subject |
| (121), (122) | The semantic classifying code is 121 (location), or 122 (time period) |
| ^1 .改. ^ ↓ (SYNREAL.comp) | Change the 1 st word on the right to be the son of the current word, and set the syntactic relation to be comp |
| ^1 .改. ^ ↓ (SYNREAL.obj) | Change the 1 st word on the right of the current one to be the son of the current word, and set the syntactical relation to be object |

The rules tell us that if the current word (“站(*zhan4*, *stand*)”) is a “motion-by-others” verb, a general verb, a “self-motion” verb, or a derived verb, moreover the semantic classifying code of the foregoing word (“女孩(*nv3hai2*, *girl*)”) is 1111(person), or 113 (organization), or 1113 (animal), then we have to change the 1st word “女孩(*nv3hai2*, *girl*)” on the left to be the son of the current word, and we have to set the syntactic relation to be subject. If the semantic classifying code of the word (“门口(*men2kou3*, *doorway*)”, “分钟 (*fen1zhong1*, *minute*)”) following the verb is 121 (location) or 122 (time period), then we have to change the 1st word on the right to be the son of the current word, and set the syntactic relation to be object , or comp respectively.

After executing the above four rules we get that the subject is “女孩(*nv3hai2*, *girl*)”, the comp is “分钟 (*fen1zhong1*, *minute*)”. That is

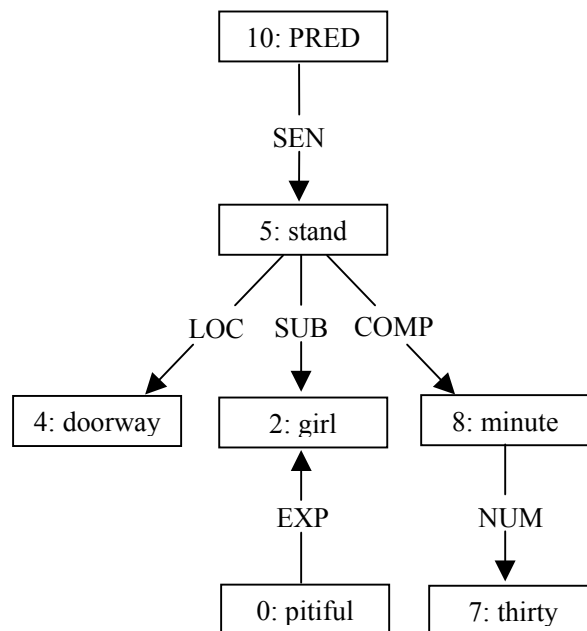
```

10: PRED: (SENTENCE, , , SENTENCE)
!
! —SEN→5: (站(zhan4), v, vv3, 2102, stand, vp, pef)
!
! —LOC→4: (门口(men2kou3), s, sss, 1212, doorway)
!
! —sub→2: (女孩(nv3hai2), n, nn2, 1111, girl)
!
! ←EXP—0: (可怜(ke3lian2), a, aaa, 2314, pitiful)
!
! —comp→8: (分钟(fen1zhong1), t, tt1, 122216, minute, np)
!
! —NUM→7: (30, m, mm1, 133503, thirty)

```

Where PRED denotes the predicate, SEN denotes the relation between sentence node, i.e. the exclamation sign for that, and the predicate of the sentence.

The equivalent “graph” representation is as follows.



Note that in this representation, not for 9, the period, all parts found before are now linked up into one “graph”. The subject role of “girl” and the complement role of “minute” are indicated by link of type SUB and COMP respectively. The links of type LOC and NUM belong to the 49 types that Yao distinguishes. In Yao’s system the facts that “doorway” indicates a location and that “thirty” is a number must be used and therefore be present in the system. They are included in the classifying codes.

2.5 Semantic analysis

In LSD, there exist 49 relation types. The reason is the following. In order to translate a sentence, the first thing is to represent all the information contained in the original sentence. The information includes the combination of words, tense, voice, mood, sentence type and other information. So, Yao chooses for 49 “relation types” (some are not semantic relations at all). So far we have encountered EXPerience, LOCation, QuaNTity and NUMber.

After the syntactic analysis, we can do the semantic analysis. Yao has the following rules.

LEX=站, (base0), RULE: (1111; 113; 1113) + \wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) \rightarrow 1[^].改. \wedge ↓ (SEMREAL.AGT)

\wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (121) \rightarrow 1[^].改. \wedge ↓ (SEMREAL.LOC)

\wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (122) \rightarrow 1[^].改. \wedge ↓ (SEMREAL.QNT)

(LEX=stand, (base0), RULE: (1111; 113; 1113) + \wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) \rightarrow 1[^].Change. \wedge ↓ (SEMREAL.AGT)

\wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (121) \rightarrow 1[^].CHANGE. \wedge ↓ (SEMREAL.LOC)

\wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUCAT.vv3; CSUBCAT.vv4) + (122) \rightarrow 1[^].CHANGE. \wedge ↓ (SEMREAL.QNT))

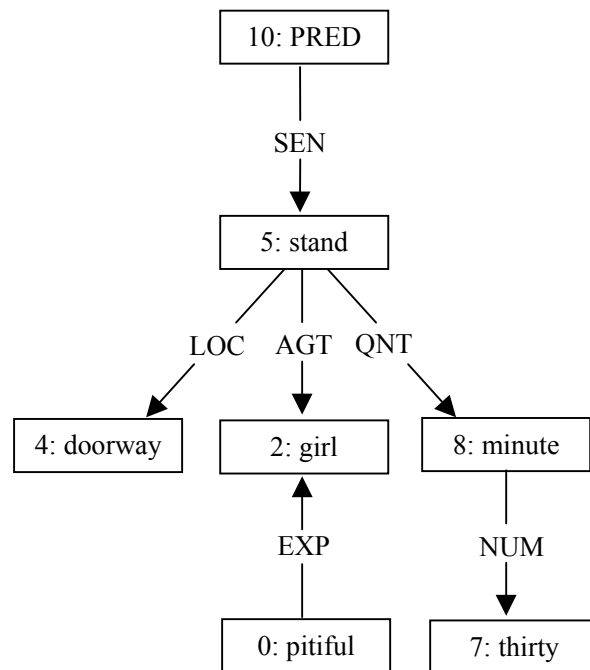
| | |
|---|---|
| LEX=站 | The current word is “站(zhan4, stand)” |
| (base0), RULE | The rule to determine the basic sentence type 0 |
| (1111; 113; 1113) | The semantic classifying code is 1111 (person), or 113 (organization), or 1113 (animal) |
| \wedge (CSUBCAT.vv2; CSUBCAT.vv1; CSUBCAT.vv3; CSUBCAT.vv4) | The subcategory is a “motion-by-others” verb, a general verb, a “self-motion” verb, or a derived verb. |
| 1 [^] .改. \wedge ↓ (SEMREAL.AGT) | Change the 1 st word on the left to be the son of the current word, and set the semantic relation to be agent |
| (121), (122) | The semantic classifying code is 121 (location), (122) (time period) |
| \wedge 1.改. \wedge ↓ (SEMREAL.LOC) | Change the 1 st word on the right to be the son of the current word, and set the semantic relation to be location |
| \wedge 1.改. \wedge ↓ (SEMREAL.QNT) | Change the 1 st word on the right of the current one to be the son of the current word, and set the semantic relation to be quantity |

The meaning of the rule is similar to the meaning given in syntactic analysis.

10: PRED: (SENTENCE, , , SENTENCE)

! —SEN→5: (站(zhan4), v, vv3, 2102, stand, vp, pef)
 ! —LOC→4: (门口(men2kou3), s, sss, 1212, doorway)
 ! —AGT→2: (女孩(nv3hai1), n, nn2, 1111, girl)
 ! ←EXP→0: (可怜(ke3lian2), a, aaa, 2314, pitiful)
 ! —QNT→8: (分钟(fen1zhong1), t, tt1, 122216, minute, np)
 ! —NUM→7: (30, m, mm1, 133503, thirty)

The equivalent “graph” representation is as follows.

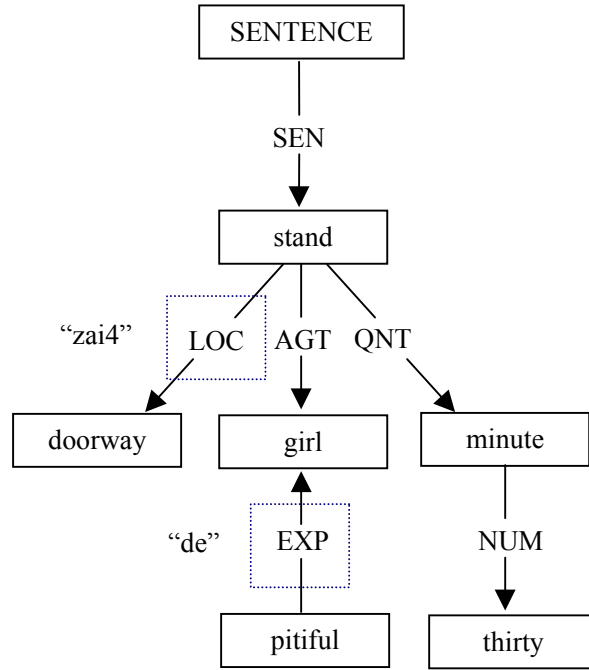


2.6 Representation by “inter-transmittal” by which Yao means “interlingua” language

SENTENCE: (10, , ,)
 ! —SEN→stand: (5, v, vv3, , 2102, prs, pef)
 ! —LOC→doorway: (4, s, sss, 1212, at)
 ! —AGT→girl: (2, n, nn2, 1111)
 ! ←EXP→pitiful: (0, a, aaa, 205e)
 ! —QNT→minute: (8, t, tt1, tp)
 ! —NUM→thirty: (7, m, mm1)

We can see that this structure represents almost all the information of the original sentence.

The equivalent “graph” representation is as follows (there is no corresponding part of “le”).



This is an outline of the parser in Lexical Semantic Driven.

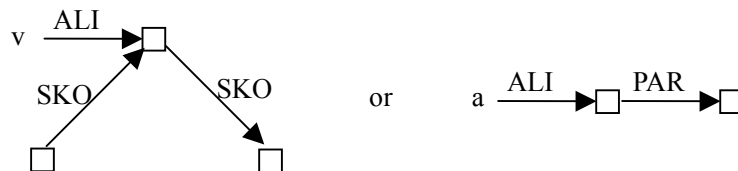
The final result of the parsing process has the structure of a graph. Yao’s method of parsing may therefore also be called a form of structural parsing. The graph formed can be seen as the starting point for the process of uttering the sentence in English. We note that in this example the words had simple counterparts in English, apart from the function word “le”. We will now discuss the method of Yao and make a comparison.

3 Parsing in Knowledge Graph Theory, a comparison

For parsing the sentence considered in Section 2 we need semantic and syntactic word graphs for all 9 “real” words (without punctuation). We compare with part 6 of Section 2.

可怜, ke3lian2 (a,v,a,v):

The simplest semantic word graph is just $\square \leftarrow \text{ALI} \text{---} \text{ke3lian2}$. If this word can be used only as adjective or as verb, then its syntactic word graph is, see Hoede and Zhang [Hoede & Zhang, 2001],

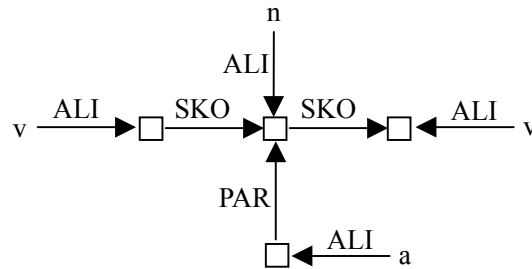


的, de (u, n):

The meaning “of” is expressed by one of the three mereological relationships $\square \text{---SUB} \rightarrow \square$, $\square \text{---PAR} \rightarrow \square$ or $\square \text{---FPAR} \rightarrow \square$ in the ontology of KGT. The use as a noun is most common, and the graphs display clearly a syntactic function.

女孩, nv3hai2 (n):

□←ALI— nv3hai2 is the simplest semantic word graph. The syntactic word graph is, partly



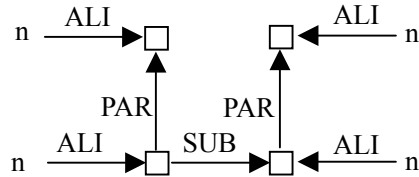
We have now the KGT material for the phrase ke3lian2 de nv3hai2. Yao finds four phrases and we already recommended that these are “chunks” of the sentence. Parsing by chunks, using chunk indicators, was developed by Hoede and Zhang [Hoede & Zhang, 2001].

When we look at the syntactic word graphs, we see that the word “de” must be chosen to have □—PAR→□ as its meaning. This makes the word ke3lian2 an adjective, an adword of the word nv3hai2 and yields the semantic chunk graph ke3lian2—ALI→□—PAR→□←ALI—nv3hai2.

The second phrase, chunk, is 在门口, “zai4 men2kou3”.

在, zai4 (p,v,v,v,p,d):

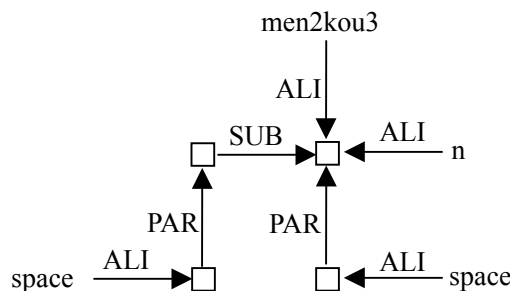
As a preposition Zai4 has a semantic word graph space —ALI→□—SUB→□←ALI— space, as the meaning of “at” is “with space in some other space”. Syntactically the word graph should express that the two spaces (locations) are adwords of two nouns, so



Both space and time are attributes of nouns.

门口, men2kou3 (s,s):

The use of a classification like “orientation word” is partly due to the fact that for Yao the word zai4 as a proposition does not have a well defined meaning. In KGT we just have space—ALI→□—PAR→□←ALI— men2kou3 and □←ALI— n as semantic and syntactic word graph. The only problem is to determine with which token the token for men2kou3 is to be identified. Only the order of zai4 and men2kou3 in the sentence leads to the semantic chunk graph

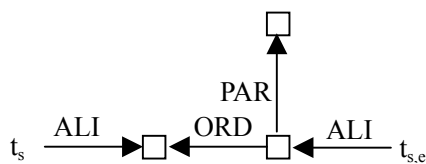


The third phrase, chunk, is 站了, “zhan4le”.

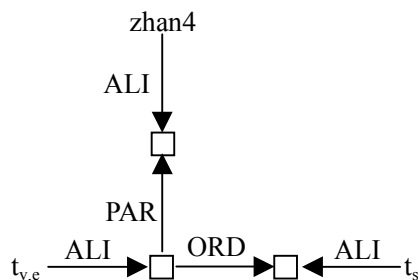
站, zhan4 (v,n,s,v):

There are three different roles possible according to Yao, the choice for zhan4 being a verb coming basically from the auxiliary word le.

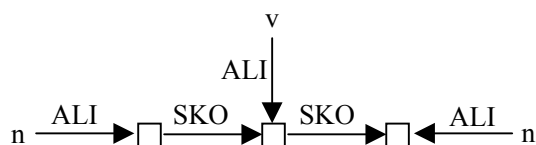
The semantic word graph for le can be taken to be



The unspecified token must be a verb, hence zhan4 must be a verb and we obtain



Syntactically two SKO-arcs indicate the node a verb plays, one of them ties the verb to the subject and the other to the object, both nouns:

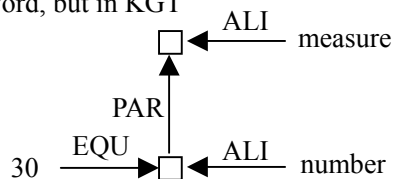


Intransitive verbs have no outgoing SKO-arc in their syntactic word graph.

Finally we have the chunk “30 minutes”.

30 (m,m):

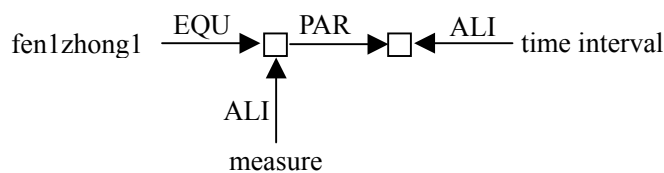
30 is described as a number word, but in KGT



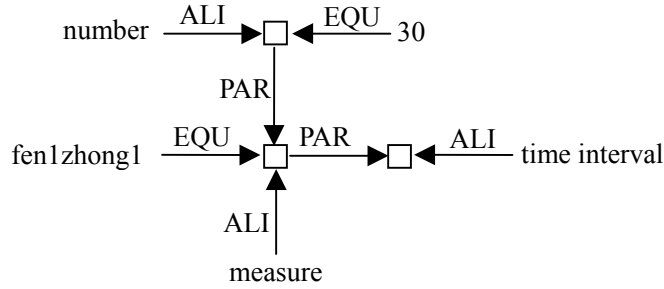
is the semantic word graph. 30 is a value. Values are attributed to measures. The measure is in this case is

分钟, fen1zhong1(t):

The word is called a time word in LSD, but this is not a precise description.

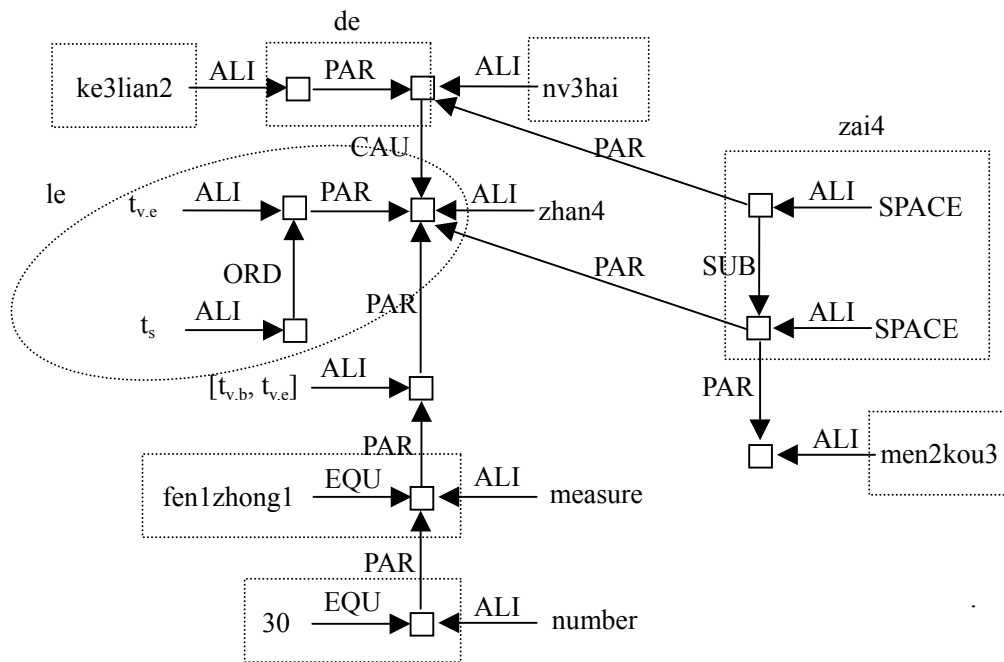


so we get



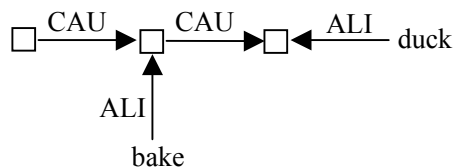
as fourth semantic chunk graph.

Structural parsing in KGT now has to deal with connecting these chunk graphs into a sentence graph, see Hoede and Zhang [Hoede & Zhang, 2001], which for this example leads to the sentence graph



Parsing in KGT is like solving a jigsaw puzzle given semantic and syntactic word graphs for each of the occurring words. The main problem, ambiguity, is present in both LSD and KGT. Let us now compare the resulting graph in Section 2.6 with this sentence graph.

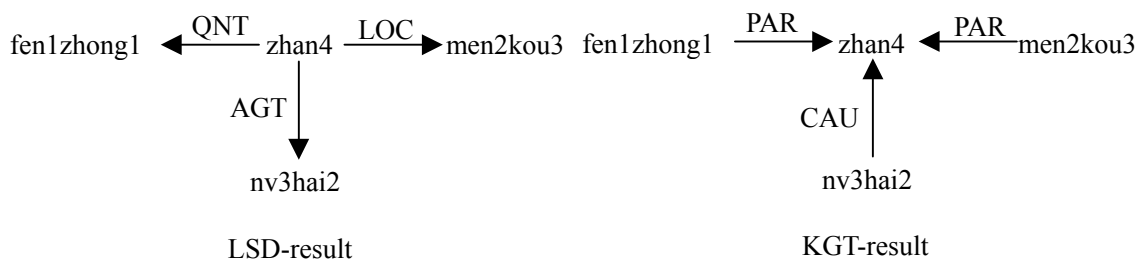
The first striking different is that in LSD representation all kinds of information are added to the words, mainly categorization. For example, nv3hai2 is given with n for “noun”, nn2 for subtype of noun, 1111 for type “person” as well as the English counterpart “girl”. In the KGT representation these are missing. Semantic classifying codes were not necessary in the KGT-parsing. However, in another paper [Hoede & Liu, to appear (b)] we met the sentence “Beijing duck bake well”, and to give the correct sentence graph, we had to know that the basic sentence is



The code for an animal is 1113 and for the verb “bake” an agent with code 1113 may be known to be impossible. This opens the door to huge complexity of the lexicon, but possibly is unavoidable, explaining why LSD is such a complex system. The rules occurring are just solving this kind of ambiguity.

The seemed striking difference is the use of relation types. That “stand” has a LOCATION type of relationship with “doorway” in LSD shows one of the major weaknesses of LSD. As the system does not have the concise ontology of KGT, LSD has to introduce many relationship types. In Yao [Yao, 1996] 49 types are mentioned and we will investigate them in the next section.

The relationship of type EXPERIENCE between “girl” and “pitiful” is an even better example of describing a basically syntactic relationship between a noun and an adjective by a semantic label. AGENT is, by its very meaning, not a relationship. In $\square \leftarrow \text{CAU} \rightarrow \square \leftarrow \text{ALI} \leftarrow$ stand the first token is the agent. Syntactically the corresponding graph is $\square \leftarrow \text{SKO} \rightarrow \square \leftarrow \text{ALI} \leftarrow$ verb and the first token is the SUBJECT. “30” and “minute” are not in a NUMBER-relationship. “30” is a number word used as value of a measure. Finally “stand” certainly does not have a QUANTITY relationship with “minute”. The similarity and the differences come forward most clearly in the subgraphs that contain the four central words 站, zhan4, 门口, men2kou3, 女孩, nv3hai2 and 分钟, fen1zhong1 both of the LSD and the KGT result.



4 Expressing the LSD-relationships in terms of the KGT-ontology

The “relations” in LSD are not like the ones in KGT. They are not the pure semantic relations, not the pure syntax relations, and not the pure case relations. They are compound relations including pure semantic relations, pure syntax relations, pure case relations, and even some complex relations. That is to say, the relations introduced in LSD are not the traditional relations. The purpose to introduce these relations is to represent sentences easily. The focus of LSD is not on the representation, but on the tools to translate between languages: Chinese and English. From Yao [Yao, 1996], we get the definitions of the relation types introduced in LSD, and examples. In Appendix, we show these definitions and the examples given by Yao, and, furthermore, we give the representation of each example in terms of knowledge graphs.

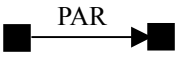
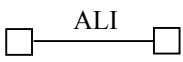
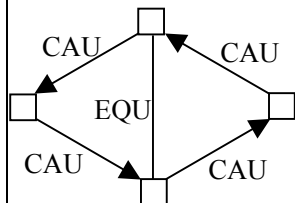
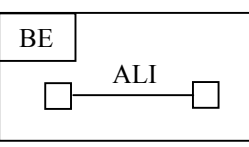
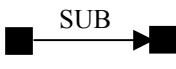
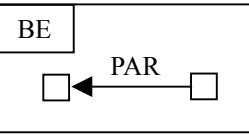
We were indeed able to express each of the 49 types of “relationships” in LSD in terms of the KGT-ontology. We now give the results in the following tables.

4.1 The basic relations (33)

| No. | Relation | Basic KGT relation | No. | Relation | Basic KGT relation |
|------|-----------------------------|--------------------|------|--------------------|--------------------|
| (1) | AGT(agentive) | CAU | (2) | ANL(analogy) | ALI |
| (3) | BAS(basis) | CAU | (4) | BE L(belong) | PAR |
| (5) | BKT(bracket) | EQU | (6) | CAP(capacity) | PAR |
| (8) | CAU(causative) | CAU | (11) | DAT(dative) | ORD |
| (13) | DET(determinate) | PAR | (14) | DIR(direction) | CAU |
| (15) | DST(destination) | ORD | (17) | EXE(execution) | CAU |
| (18) | EXP(experience) | CAU | (19) | EXT(external) | CAU |
| (20) | FCS(focus) | PAR | (21) | FRQ(frequency) | PAR |
| (22) | GOA(goal) | PAR | (23) | INS(instrument) | PAR |
| (24) | LOC(location) | PAR | (25) | MAT(material) | PAR |
| (26) | MAN(manner) | PAR | (27) | MOD(modify) | PAR |
| (29) | NUM(number) | PAR | (30) | OBJ(objective) | CAU |
| (31) | OBS(object besides subject) | EQU | (32) | ORG(original) | ORD |
| (34) | POS(possession) | PAR | (35) | PRT(part of thing) | SUB |
| (36) | QNT(quantity) | PAR | (37) | PDC(product) | CAU |
| (38) | REA(reason) | CAU | (40) | REF(reference) | ORD |
| (41) | RUT(route) | ORD | (42) | SCP(scope) | CAU |
| (43) | SIT(situation) | PAR | (45) | ATT(attribute) | PAR |
| (46) | TES(the example) | EQU | (47) | TIM(time) | PAR |
| (48) | TIT(title in title) | PAR | | | |

4.2 The complex relations (7)

| No. | Relation | Representation | No. | Relation | Representation |
|------|----------------|----------------|------|------------------|----------------|
| (12) | DEG(degree) | | (9) | COM(comrade) | |
| (7) | CAS(condition) | | (10) | CON(conjunction) | |

| | | | | | |
|--|------------------------------|---|------|----------------------|---|
| (16) | DUR(duration) |  | (28) | MOS(modify sentence) |  |
| (33) | PAR(parallel) |  | (39) | RLT(relative thing) |  |
| (44) | SIU(supplementary situation) |  | (49) | VAL(value) |  |
| Note: ■ denotes the sentence frame. | | | | | |

5 Concluding remarks

From the discussion above, we can see that the relations introduced in KGT can represent all 49 types of relations in LSD. There are types of relation including semantic relations, syntactic relations and compound relations. According to the relations in KGT, we divided these 49 types relations into two sub-groups: the basic relations and the complex relations. The basic relations are the relations that have corresponding to basic relations in KGT. The complex relations are the relations for which several concepts are put into relationship.

Inference in Artificial Intelligence systems is very important. Efficient inference lies in the number of relation types introduced. So, the Knowledge Graph Method should turn out be more efficient than the Lexical-Semantic Driven Method.

References

- [Hoede & Li, 1996] C. Hoede and X. Li, *Word Graphs: The First Set*, in *Conceptual Structures: Knowledge Representation as Interlingua*, Auxiliary Proceedings of the 4th International Conference on Conceptual Structures, ICCS '96, (1996) 81-93.
- [Hoede & Zhang, 2001] C. Hoede and L. Zhang, *Structural Parsing*, in: *Conceptual Structures: Extracting and Representing Semantics*, Aux. Proc. of the 9th International Conference on Conceptual Structures(Ed. G. W. Mineau), CA, USA, 75-88, 2001.
- [Liu & Hoede, to appear (a)] X. Liu and C. Hoede, *Translation: an Example from Ancient Chinese to Modern Chinese*, to appear.
- [Liu & Hoede, to appear (b)] X. Liu and C. Hoede, *Passive Sentences and Structural Parsing*, to appear.
- [Roget, 1977] P. Roget, *Roget's International Thesaurus*, 4th Edition, Harper and Row Publishers Inc. 1977.
- [Yao, 1996] TianShun Yao, *Natural language understanding*, Tsinghua University Press, Science and Technology Press in Guangxi Province, 1996. (in Chinese)

Appendix: Representation of Yao’s 49 “relationships” by means of knowledge graph theory

In the appendix, we will represent all Yao’s 49 “relation” types by means of knowledge graphs.

We give sentence graphs for the example sentences, without the details of tense. The relationship is indicated by a frame. In many cases more explicit word graphs for the occurring words are not given in order to avoid too large sentence graphs.

We recall that a relationship between two tokens is described by the knowledge graph containing both concepts. In several cases Yao’s naming of a relationship shows that the explicit semantics of a token is an essential part of the semantics of a relationship, although not explicitly mentioned in the example sentence.

As an example we consider the verb leave. This is done “from” e.g. home town, which can be called the origin. The relationship between the concept “leave” and the concept “home town” is named ORD by Yao (see No. 32). But that means that the specific semantics of “leave”, the ORD arc, turns the concept “home town” into something of type ORIGIN. Such typing are given in capital.

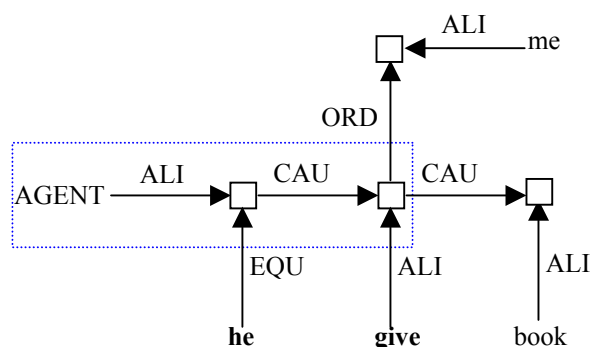
The names of the related concepts are given in bold letters. Sometimes the relationship is named by a linguistic function, as e.g. DAT(ive) in No. 11. The explanations are short versions of those given by Yao. Quite a few examples do not seem to illustrate the type of the relationship well.

1. AGT (agent)

Explanation: The actor of an action.

Example: **He** gives me a book.

Graph:

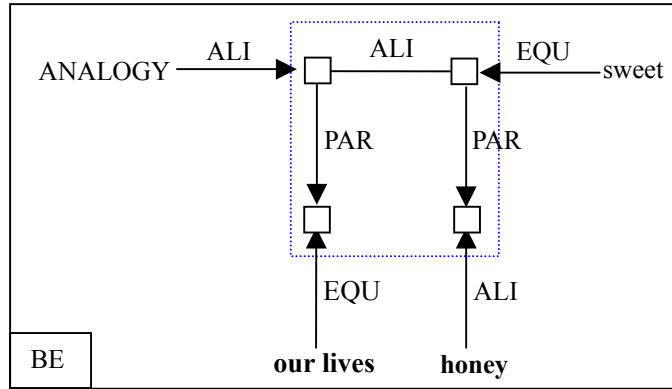


2. ANL (analogy)

Explanation: The similarity or the analogy with the main body in an event.

Example: **Our** lives are like **honey**-sweet.

Graph:

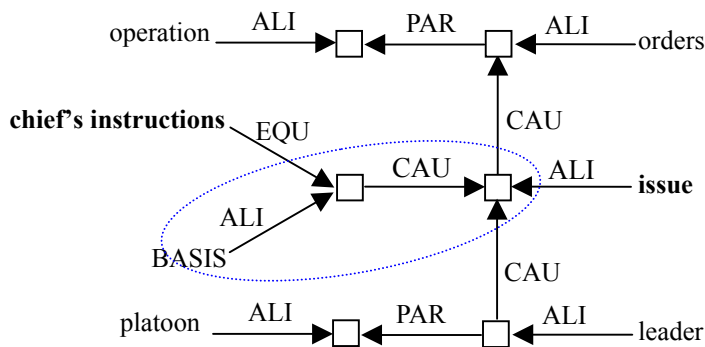


3. BAS (basis)

Explanation: The basis complied to in an event or according to which is acted.

Example: According to the **chief's instructions**, the leader of the platoon **issues** orders of operation.

Graph:

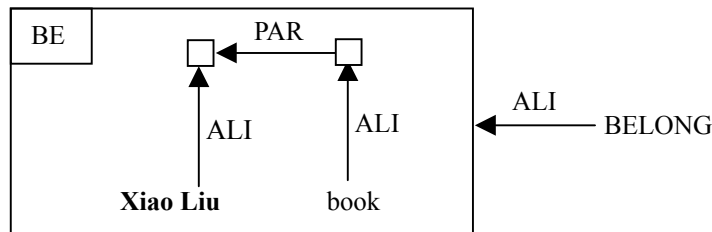


4. BEL (belong)

Explanation: The person or thing belonging to the main body in an event.

Example: **Xiao Liu** has a **book**.

Graph:

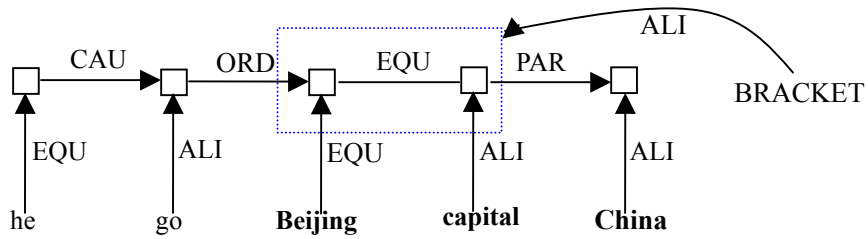


5. BKT (bracket)

Explanation: The relation between the words in the brackets and the foregoing word.

Example: He went to **Beijing (the capital of China)**.

Graph:

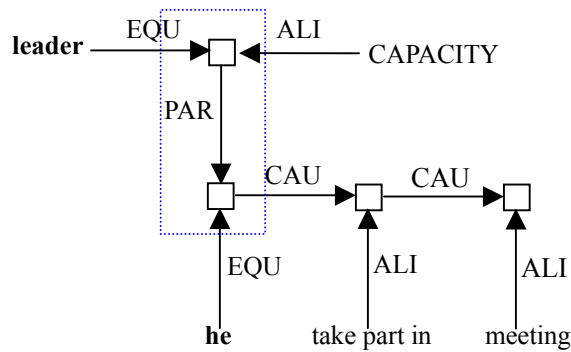


6. CAP (capacity)

Explanation: The capacity or function of some object in an event.

Example: **He as a leader**, takes part in the meeting.

Graph:

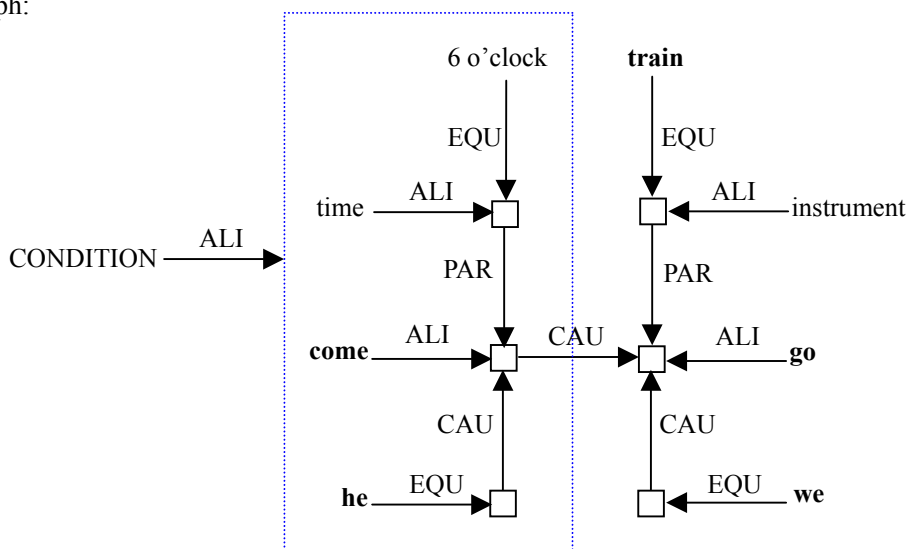


7. CAS (condition)

Explanation: The condition effects that the event happens or develops.

Example: If **he will come at 6 o'clock**, we will go by the next train.

Graph:

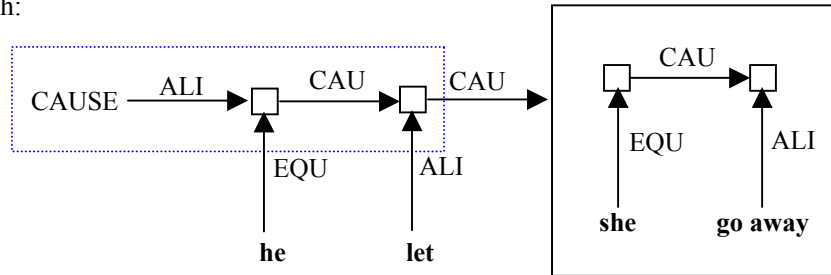


8. CAU (cause)

Explanation: The cause of some action or some status.

Example: **He let her go away**.

Graph:

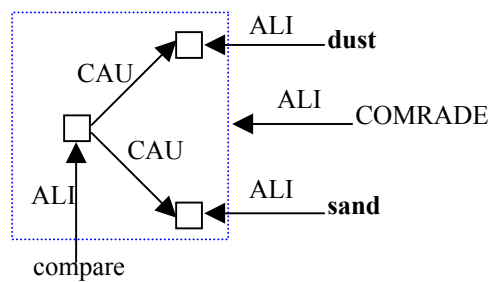


9. COM (comrade)

Explanation: The indirect entity accompanied in or removed from an event.

Example: To compare **dust** and **sand**.

Graph:

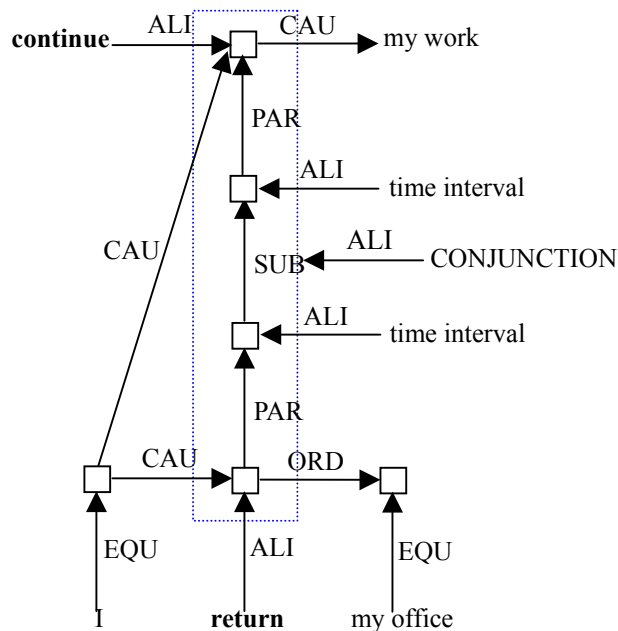


10. CON (conjunction)

Explanation: The conjunction of two events, they occur at the same time or right after each other.

Example: I **return** to my office to **continue** my work.

Graph:

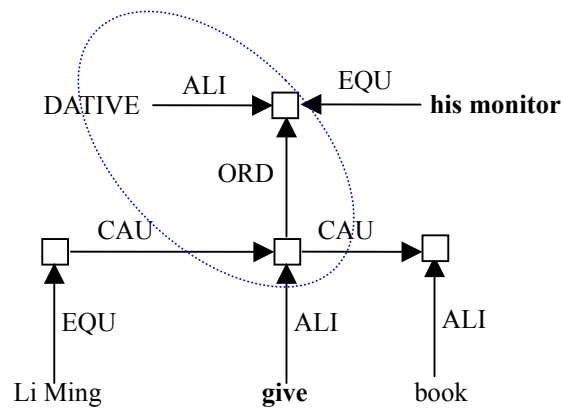


11. DAT (dative)

Explanation: The indirect object with some benefit or loss in an event.

Example: Li Ming **gives his monitor** a book.

Graph:

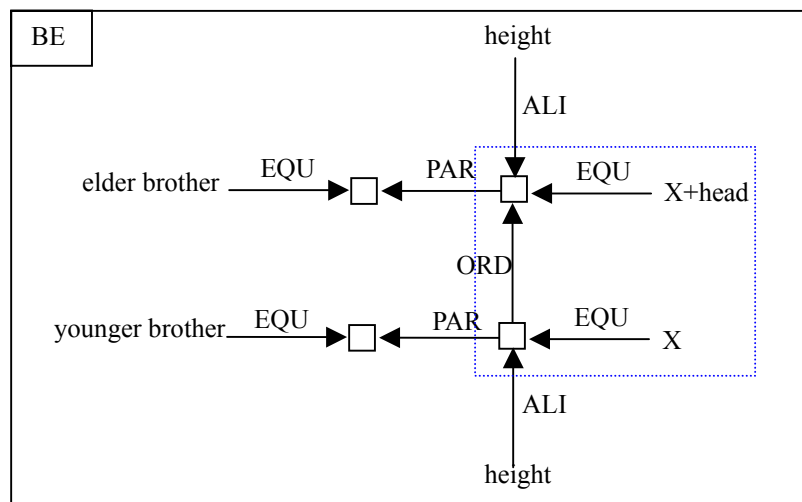


12. DEG (degree)

Explanation: The degree or comparing value.

Example: The elder brother is **one head** higher than his younger brother, (this is not a relationship).

Graph:

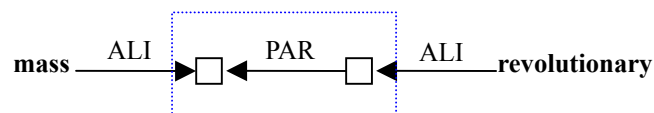


13. DET (determination)

Explanation: The determining of the main body in an event.

Example: **Revolutionary mass**.

Graph:

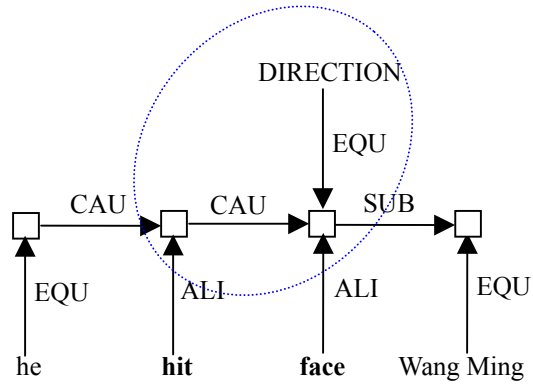


14. DIR (direction)

Explanation: The tendency of an activity in an event or the direction of movement.

Example: He is **hitting** Wang Ming's **face**.

Graph:

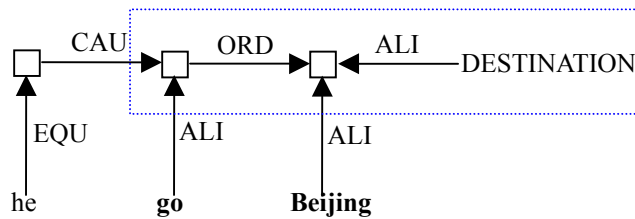


15. DST (destination)

Explanation: The destination of the activity in an event.

Example: He **went to Beijing**.

Graph:

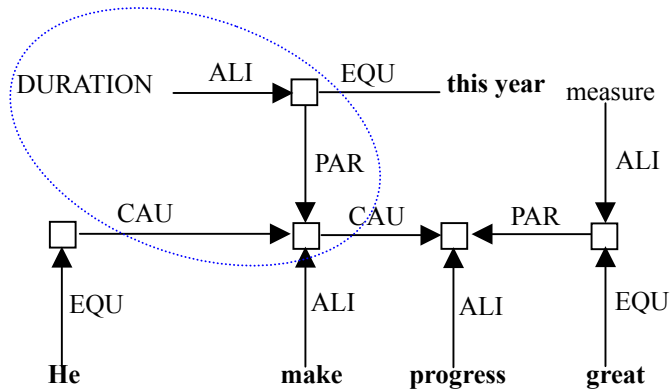


16. DUR (duration)

Explanation: The duration from the beginning to the end of the event.

Example: In **this year**, he **made a great progress**.

Graph:

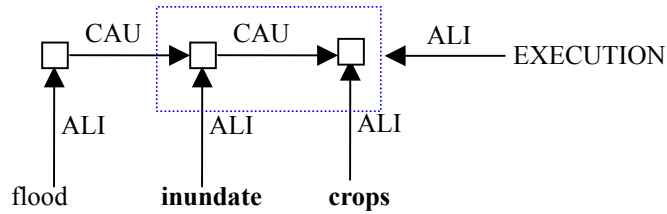


17. EXE (execution)

Explanation: The main body in a natural activity can effect some object.

Example: The flood **inundated the crops**.

Graph:

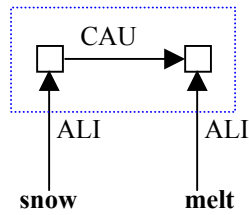


18. EXP (experience)

Explanation: The experience of change.

Example: The **snow melted**.

Graph:

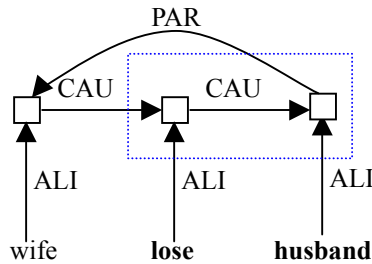


19. EXT (external)

Explanation: The external things that are referred to in the activity in an event, but are not controlled by anything.

Example: The wife **lost** her **husband**.

Graph:

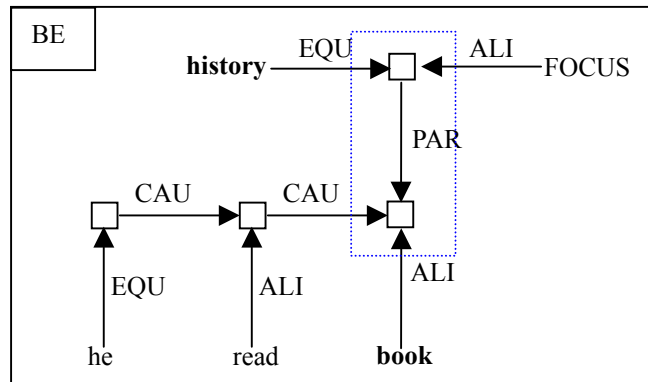


20. FCS (focus)

Explanation: The focus of some object, or the focus of the behavior, status of some object, Emphasis.

Example: It is a **history book** that he read.

Graph:

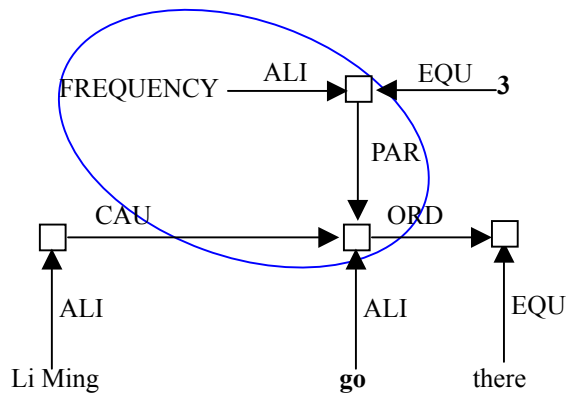


21. FRQ (frequency)

Explanation: The repetition of the action in an event.

Example: Li Ming **went there three times**.

Graph:

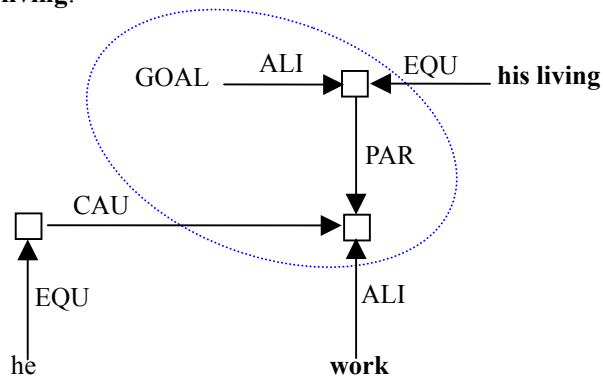


22. GOA (goal)

Explanation: The goal of a consciousness action.

Example: He **works for his living**.

Graph:

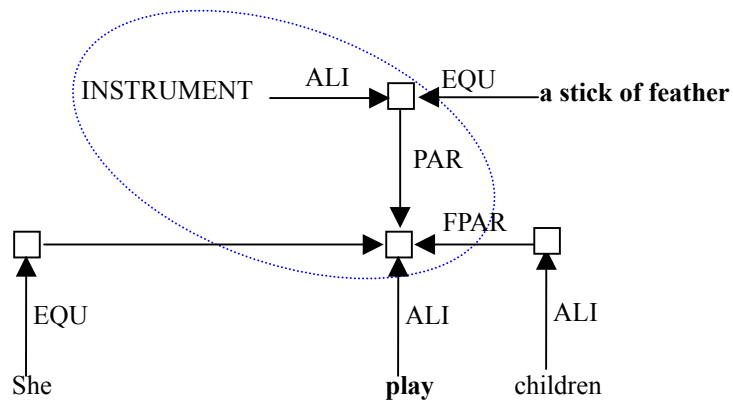


23. INS (instrument)

Explanation: The seeable or hidden instrument or equipment, or the instrument or equipment used by actors.

Example: She **played with the children with a stick of feather**.

Graph:

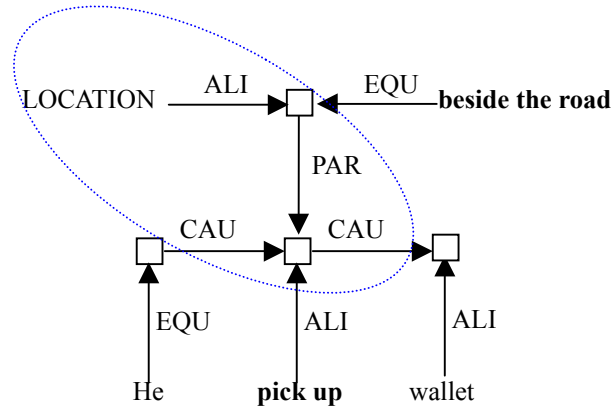


24. LOC (location)

Explanation: The location at which the action or status happened.

Example: He **picked up** a wallet **beside the road**.

Graph:

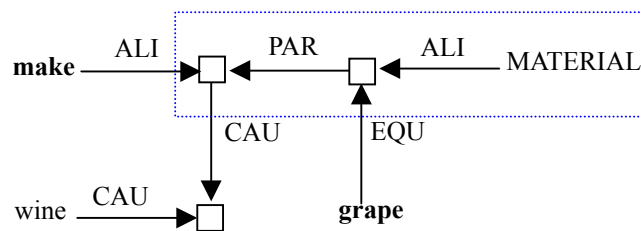


25. MAT (material)

Explanation: The material, energy, or fund which formed or constructed something.

Example: To **make** the wine with **grapes**.

Graph:

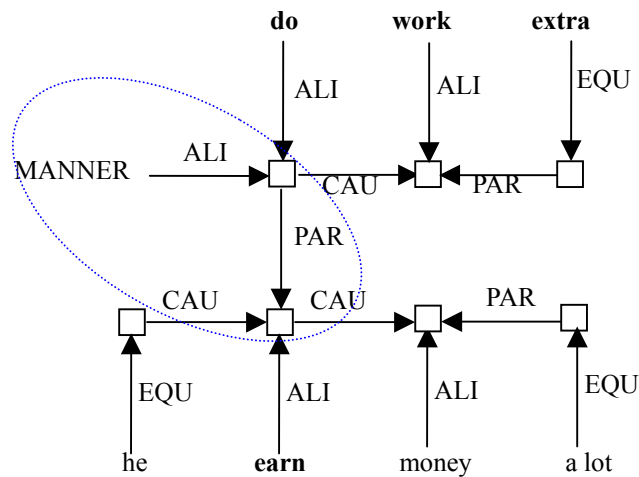


26. MAN (manner)

Explanation: The manner used to obtain some goal.

Example: He **earned** a lot of money by **doing extra work**.

Graph:

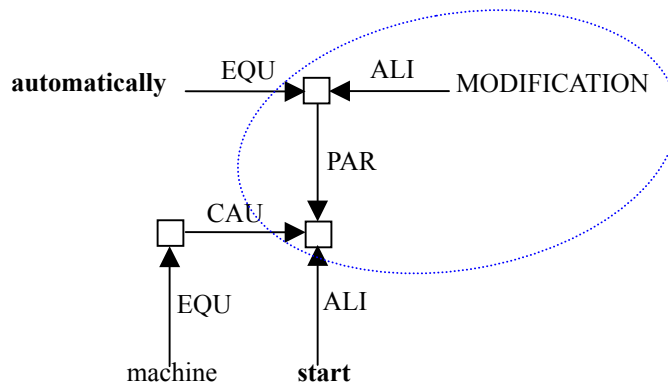


27. MOD (modify)

Explanation: The modification of a verb or a pronoun.

Example: The machine **started automatically**.

Graph:

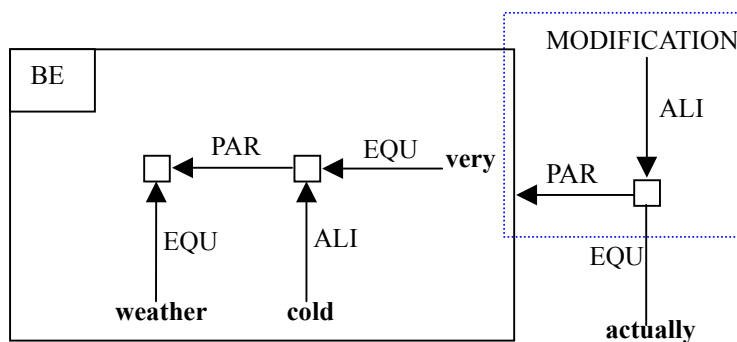


28. MOS (modify sentence)

Explanation: The modifier or phrase of a sentence which the speaker uttered, it is a kind of limitation of the whole proposition.

Example: **Actually, the weather is very cold.**

Graph:

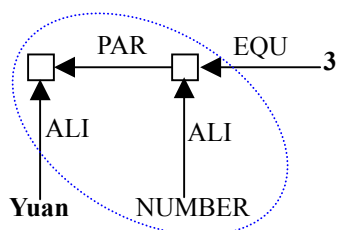


29. NUM (number)

Explanation: The numerical part of some thing.

Example: **three** Chinese **Yuan**.

Graph:

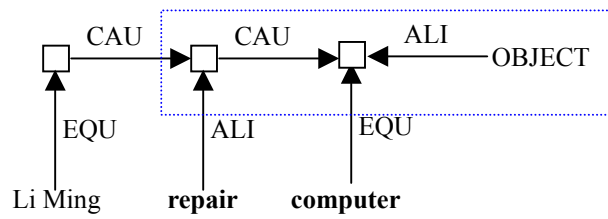


30. OBJ (object)

Explanation: The existing directive object in some event or some activity.

Example: Li Ming **repaired** the **computer**.

Graph:

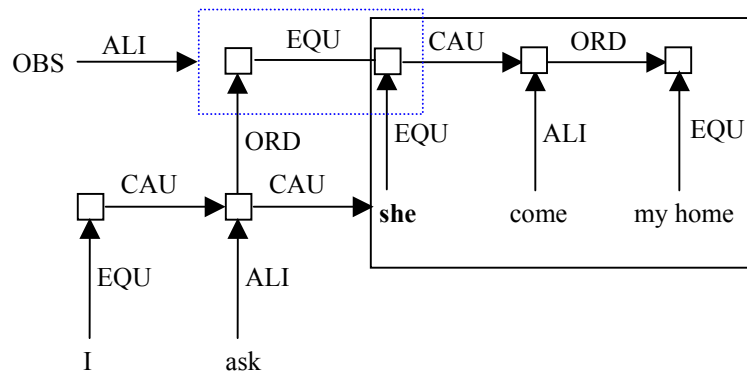


31. OBS (object besides subject)

Explanation: It is used when the subject in the supplement is the same as the object in the sentence.

Example: I asked **her** to come to my home, (this is not a relationship).

Graph:

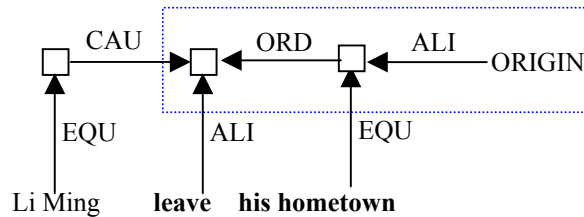


32. ORG (origin)

Explanation: The origination of the activity, or the status before changed in an event.

Example: Li Ming **left his hometown**.

Graph:

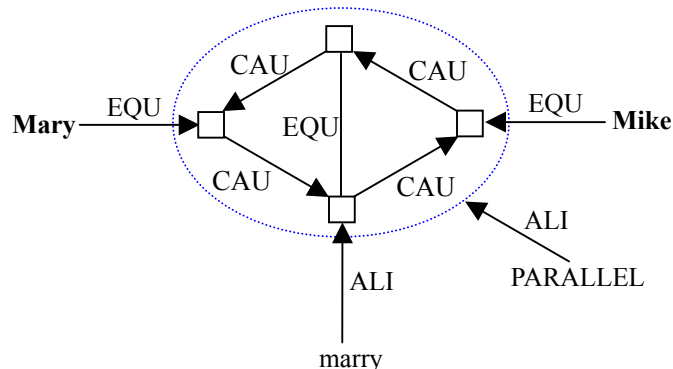


33. PAR (parallel)

Explanation: It is used in a parallel structure sentence.

Example: **Mary and Mike** married.

Graph:

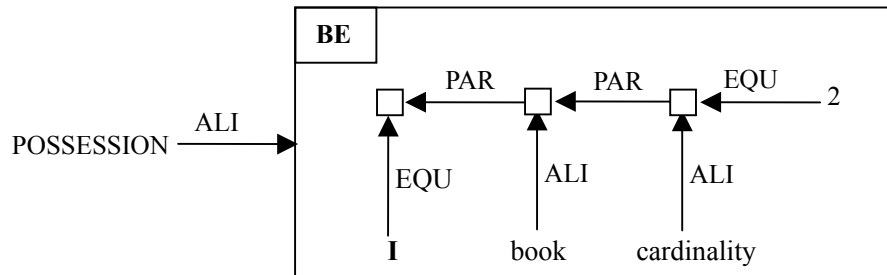


34. POS (possession)

Explanation: The individual in possession.

Example: **I have two books.**

Graph:

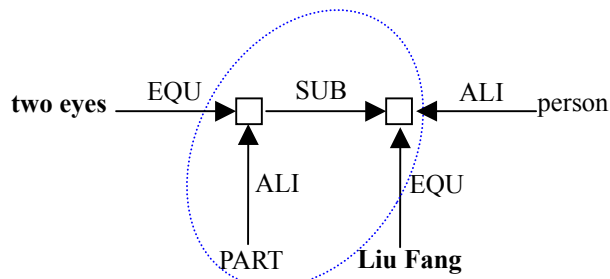


35. PRT (part of thing)

Explanation: The component that formed the main body in an event, being a concrete part.

Example: The big **two eyes** of **Liu Fang**.

Graph:

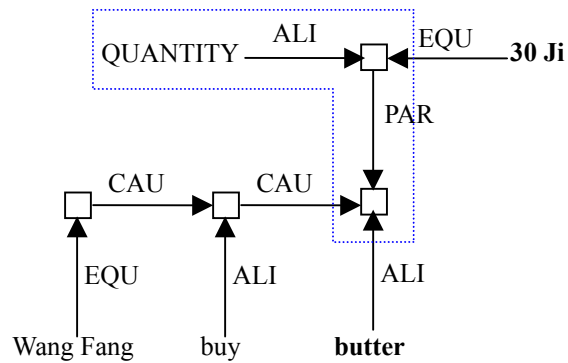


36. QNT (quantity)

Explanation: The relation between thing and quantity.

Example: Wang Fang bought **30 Ji** of **butter**.

Graph:

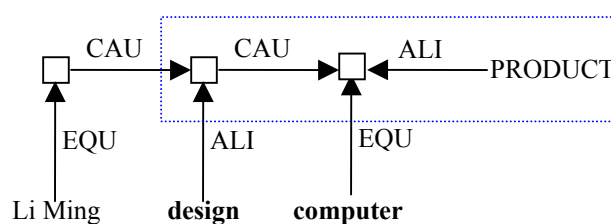


37. PDC (product)

Explanation: The new object created in an activity.

Example: Li Ming **designed a computer**.

Graph:

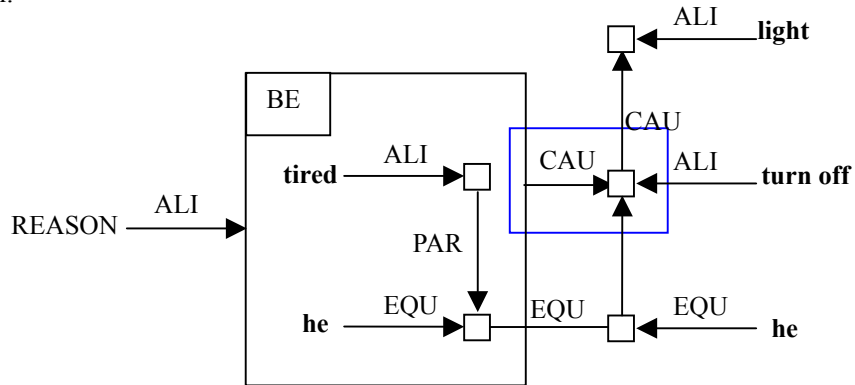


38. REA (reason)

Explanation: The natural reasons of some event.

Example: **He is tired, so he turned off the light.**

Graph:

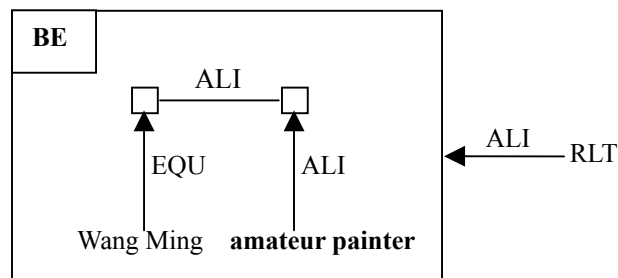


39. RLT (relative thing)

Explanation: The type of some thing.

Example: Mr. Wang is an **amateur painter**.

Graph:

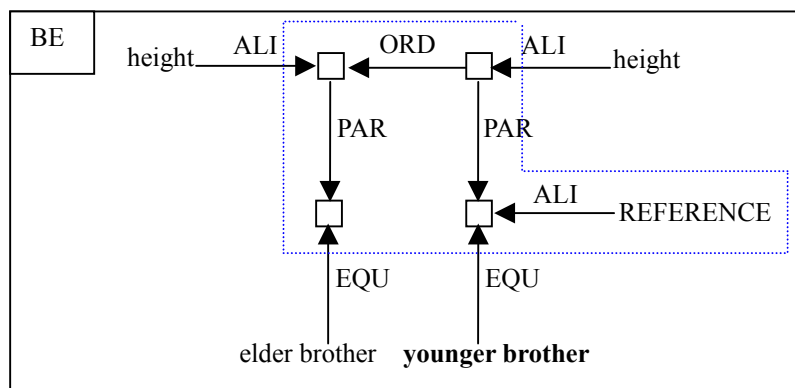


40. REF (reference)

Explanation: The indirect object as the reference used in the comparison or measure in an event.

Example: The elder brother is **higher than** his **younger brother** is.

Graph:

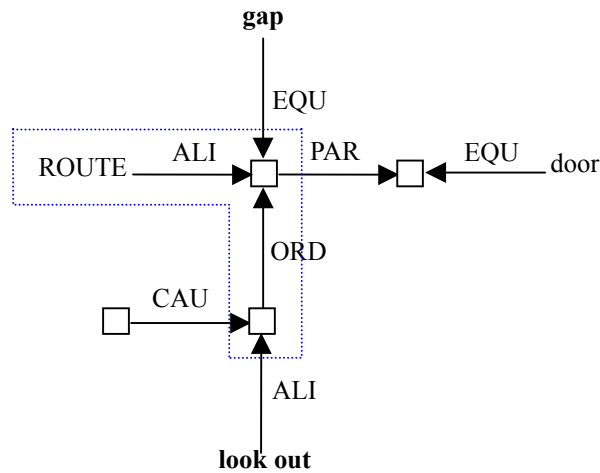


41. RUT (route)

Explanation: The route or process that the event goes through.

Example: To **look out** through the **gap** in the door.

Graph:

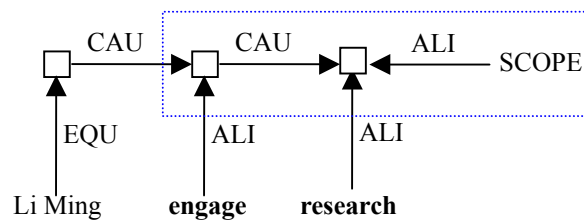


42. SCP (scope)

Explanation: The field or scope that the event touched upon.

Example: Li Ming **engaged** in **scientific research**.

Graph:

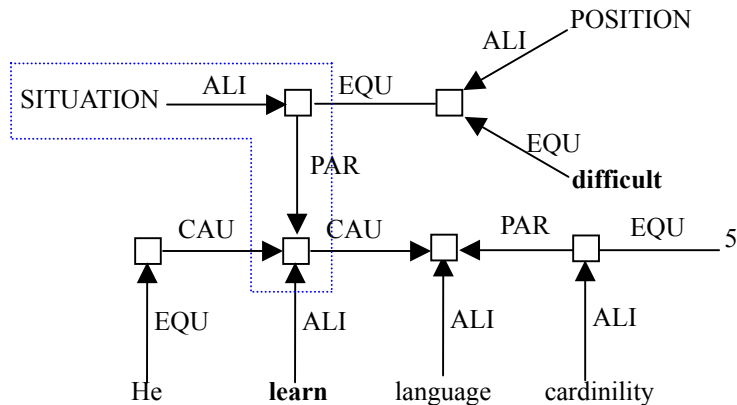


43. SIT (situation)

Explanation: The situation in which the event takes place.

Example: In the **difficult position**, he **learned** 5 languages.

Graph:

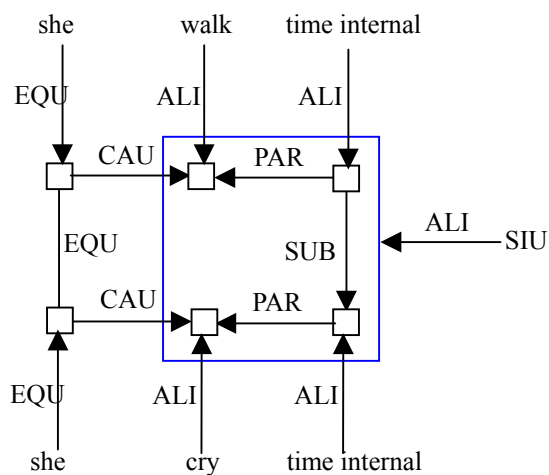


44. SIU (supplementary situation)

Explanation: The modification of a status.

Example: She **cried** while she **walked**.

Graph:

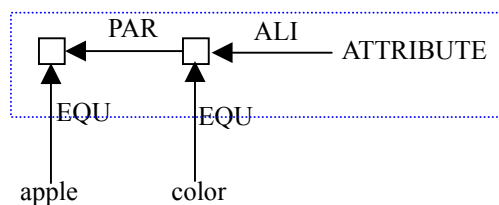


45. ATT (attribute)

Explanation: The attribute of some object.

Example: the **color** of the **apple**.

Graph:

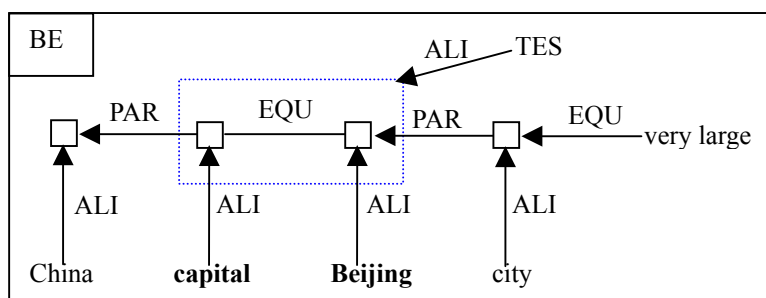


46. TES (the example)

Explanation: The subset or a real instance of the object expressed by the central words.

Example: **Beijing**, *i.e.* **the capital of China**, is a very large city.

Graph:

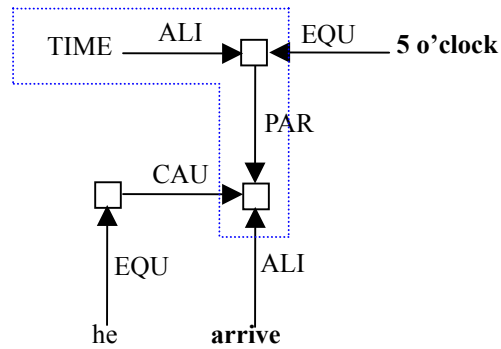


47. TIM (time)

Explanation: The exact time at which an event happens.

Example: He will **arrive** at **5 o'clock**.

Graph:

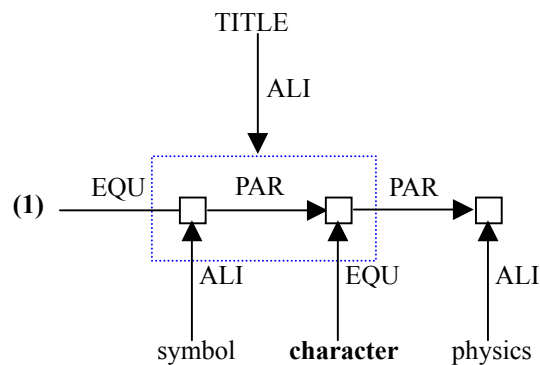


48. TIT (title in title)

Explanation: The relation between the symbol in the first sentence and the article.

Example: **(1)** the **character** of **physics**.

Graph:

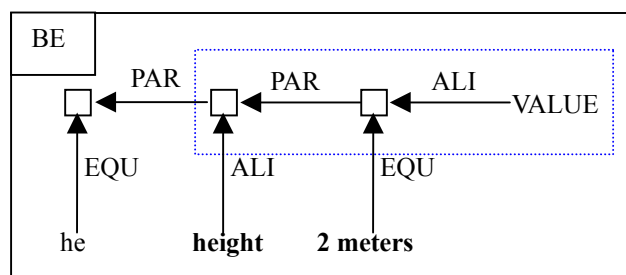


49. VAL (value)

Explanation: the value of ATT (attribute).

Example: His **height** is **two meters**.

Graph:



From the above sentence representations, we see that we represented indeed by knowledge graphs for the 49 relation types as Yao mentioned in his book [Yao, 1996]. We found that the given examples do not always express the exact meaning of the relations.

As we see from the relation types of Yao, indicated by frames in many cases some semantic aspect is taken up into the representing frame. This may be compared with the 15 Chinese words for “in”, see [Hoede & Li, 1996], where also extra semantic differentiation made it possible to indeed find 15 different word graphs.

We must, however, conclude that, although we can understand why semantic information is taken up in the classification of relation types in a system called Lexically Semantic Driven, in knowledge graph theory such information is automatically included on expansion of the occurring words. The large number, 49, of so-called semantic relationships of Yao masks the more basic character of the knowledge graph ontology.