

Logical Model of Human Mental Image Processing in Spatiotemporal Language Understanding

MASAO YOKOTA^{*}, ROJANEE KHUMMONGKOL^{**}

Department of Intelligent Information System Engineering, Graduate School
Fukuoka Institute of Technology
3-30-1 Wajirohigashi, Higashi-ku, Fukuoka 811-0295
JAPAN

^{*}yokota@fit.ac.jp, ^{**}bd14002@bene.fit.ac.jp

Abstract: - Natural language is the most convenient means that people use to communicate with each other conventionally. This is also the case for casual and intuitive interaction between ordinary non-expert people and artifacts such as robots. Therefore, it is doubtless that the technology of Natural Language Understanding (NLU) should play a key role at such scenes of humans and machines. In this paper, focusing on spatiotemporal (or 4D) language, NLU by human based on mental image is attempted to simulate so that robots can understand texts in the same way as people. The proposed methodology is quite distinguished from conventional ones and shows a good potential for providing robots with an NLU mechanism guided by humanlike awareness control based on mental image.

Key-Words: - Natural language understanding, Mental image, Temporal logic, Knowledge representation

1 Introduction

In not so far a future, robots will be indispensable partners of humans in various fields. They will be employed not only to assist people in daily life activities but also to save victims' lives at unexpected disasters, as well known as rescue robots. There are several ways to communicate with robots, for example: direct wired control, radio frequency, Wi-Fi, and so on. Anyway, without brains to comprehend natural language, they will not be able to respond to human command correctly and effectively. Thus, Natural Language Understanding (NLU) is a significant field of computer based systems which can understand ordinary human language.

In 1950, A. M. Turing [1] upset the world by his philosophical question 'Can machines think?' and presented his idea, so called 'Turing test', that whether or not a machine is thinking could be judged from its human-likeness in natural language conversations with humans. This test concerns seriously NLU by human, which has been the central theme in such research fields as Artificial Intelligence (AI) and Cognitive Science since then.

Although the Turing test was very influential to AI and Human-Computer Interaction (HCI), these two communities were opposite to each other in the view of the way how human and computer should interact. That is, AI is oriented by knowledge of human intelligence and HCI, by system design for

its implementation. T. Winograd [2] suggested that both of AI and HCI should be needed for human-like intelligence in 'T' shape, namely, T-Shaped model whose depth and width represent AI and HCI discipline, respectively. The famous computer program ELIZA [3] is positioned in a very shallow level of this model, that is, doesn't really understand NL but is designed well enough to pass the Turing test. Then, in order to improve such a defect of this test, several ideas have been proposed, for example, 'Recognizing Textual Entailment (RTE)' by I. Dagan et al [4] and its variant, 'Winograd Schema Challenge' by H.J. Levesque [5]. By the way, it is a hot topic that the Q-A machine 'Watson' [6] developed by IBM won the human champions on a famous American quiz show, Jeopardy, as a real-time challenger but the machine is based on Big Data Analysis employing conventional NLP techniques and does not really understand NL [7], either. However, it is remarkable that the developing team applied UIMA [6, 8, 9], a framework to analyze unstructured data (for example text, voice, etc.) to be structured data (that is, for relational database), resulting in quite great reduction of the execution time, about two hours the system spent to answer a question at the beginning.

As easily imagined from the mention above, it is quite difficult to make the machine understand NL in a human-like way. For example, how can people perceive and distinguish an ambiguous expression

such as S1 so easily? For another example, why can we affirm the question S3 about S2 so immediately?

- (S1) I saw the cloud in my airplane.
- (S2) Tom was with the book in the bus running from Town to University.
- (S3) Did Tom carry the book from Town to University?

An ordinary English speaker can recognize the objects as mental images evoked by S1 - S3, so their spatial relations could be reflected clearly enough to be drawn in pictures. At the present level of NLU technologies, it is extremely difficult to make a robot understand these sentences in the same way as humans do by employing mental images.

M.Yokota [10, 11] has proposed a mental image model and a formal language named 'Language for Mental-Image Language (L_{md})' in Mental Image Directed Semantic Theory (MIDST). Based on MIDST, he and his coworkers [12, 13] have been attempting to simulate such NLU by humans as mental image processing, so called, 'Mental-image Based Understanding (MBU)'.

This paper describes MBU based on MIDST to simulate human mental imagery, focusing on 4D expressions to obtain the results in an acceptable way (like Q&A). To explore more in the detail, its remainder is organized as follows. Section 2 considers mental image processing in human, and section 3 applies L_{md} to meaning representation of NL expressions. Section 4 shows the simulation of NLU by humans and the results, and finally section 5 discusses and concludes the paper.

2 Mental Image Processing

MIDST proposes a mind model imitating human's nervous system where mental images are represented by the formal language L_{md} . This is one kind of knowledge representation language employed for many-sorted predicate logic, consisting of five kinds of term: matter (x and y), value (p and q), attribute (a), pattern (g), and standard (k) to be placed in such a formula as (1) called 'Atomic Locus Formula', often abbreviated as (1'). Then, a logical combination of well-formed atomic locus formulas is simply called 'Locus Formula'.

$$L(x,y,p,q,a,g,k) \tag{1}$$

$$L(x,y,p,q,\alpha), \text{ where } \alpha = (a,g,k) \tag{1'}$$

The atomic locus formula is the unit to articulate the image of an event in space and time, roughly reading that x causes y to change attribute a (e.g., color) monotonically from value p (e.g., red) to q (e.g., blue) during some time interval, where p and q are relative to the standard k , and g is the parameter to indicate whether the change event is in time or in space, called, 'Temporal Change Event' and 'Spatial Change Event', respectively.

The matter terms, x and y can refer to 'Event Causer (EC)' and 'Attribute Carrier (AC)', respectively. An atomic locus formula holds for a time-interval, namely, $[t_i, t_f]$, so p and q are used to indicate the values at t_i and t_f , respectively. The attribute a and standard k can be defined as in [10]. As for the parameter g , consider the following examples, S4 and S5.

- (S4) The bus runs from Town to University.
- (S5) The road runs from Town to University.

The sentences S4 and S5 look similarly, but they are quite different because of the concepts of temporal change event ($g=G_t$) and spatial change event ($g=G_s$). These two change events keep on the relationship between AC and the Focus of the Attention of the Observer (FAO) (through eyes or so). S4 refers to a temporal change event such that FAO is kept on the running bus (AC). On the other hand, S5 describes a spatial change event such that FAO runs along the road (AC) extending spatially. Therefore, these events must be distinctively defined as (2) and (3), respectively, where A_{12} stands for the attribute constant, 'Physical Location'. Therefore, for example, (2) can read that x causes y to stay at p (for $p=q$) or move from p to q (for $p \neq q$).

$$(\exists x,k)L(x,x,Town,Univ.,A_{12},G_t,k) \wedge bus(x) \tag{2}$$

$$(\exists x,k)L(x,x,Town,Univ.,A_{12},G_s,k) \wedge road(x) \tag{3}$$

In MIDST, Tempo-Logical Connectives (TLCs) are used to represent both the logical and the temporal relationships between locus formulas [14], of which the most frequently used are 'Simultaneous AND' and 'Consecutive AND'. They are called 'SAND' and 'CAND' in short, symbolized by Π and \bullet , respectively. For the sake of simplicity, quantifiers (i.e., \forall and \exists) are to be omitted and, moreover, (2) and (3) will be reduced to the forms (2') and (3'), respectively.

$$L(Bus,Bus,Town,Univ.,A_t) \tag{2'}$$

L(Road,Road,Town,Univ.,A_s) (3')

Anyway, to let a machine know how people should be aware of natural language, MBU is provided with a methodology to simulate human mental imagery, but how can we translate that imagery into L_{md} ? Now consider S4 again. If we can take the imaged scene out of our brain, it must be like a motion picture of the bus running from Town to University as shown in Fig.1. However, it remains still difficult to translate this image into logical form as is because this picture has too much information. So, a certain abstract alternative such as Fig.2 is needed. In Fig.2, the circles represent the objects in the sentence, and the broken arrow refers to its movement. Here, note that neither the color nor the shape of each symbol is significant at all. What happens if we apply the principle of highly abstract picture to S2? Yes, the answer must be as in Fig.3, and it is not doubtful whether MBU based on MIDST can return the correct answer to the question S3 depicted in Fig.4.

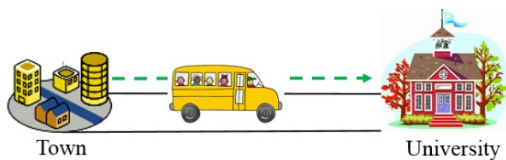


Fig.1. Mental imagery evoked by S4.

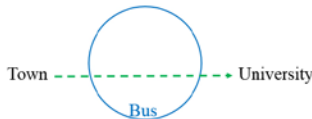


Fig.2. Highly abstract picture of S4.



Fig.3. Highly abstract picture of S2 (= S6).

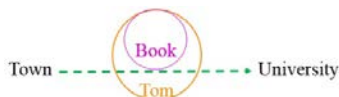


Fig.4. Highly abstract picture of S3.

3 Application of L_{md}

For this work on MBU, L_{md} was applied to three types of stimulus sentences as follows, where SS, PrP, PaP and C denote 'simple sentence', 'present particle construction', 'past particle construction' and 'conjunction', respectively.

[Type I] SS + PrP

For example,

(S6) Tom was with the book in the bus *running* from Town to University. (=S2)

[Type II] SS + PaP

For example,

(S7) Tom was with the book in the car *driven* from Town to University by Mary.

[Type III] SS + C + SS

For example,

(S8) Tom kept the book in a box *before* he drove the car from Town to University with the box.

As easily convinced, S6 - S8 are syntactically ambiguous that may be rather easy for humans to understand, but it is not the case for robots.

For example, consider S6 (or S2). How can the machine know who/what was running from Town to University? —Tom, or book, or bus? Here, to see its syntactic possibilities, Dependency Grammar (DG) is employed to determine the relations between head words and their dependents. In principle, S6 can have twelve possible dependency trees, that is, syntactically ambiguous in twelve ways as shown in Fig.5. This can be formulated by a set of local dependencies such as (4), where each pair of parentheses is for the alternatives causing the syntactic ambiguity.

$$\{D11, D12, (D13|D13a), (D21|D21a|D21b), D22, (D23|D23a)\} \quad (4)$$

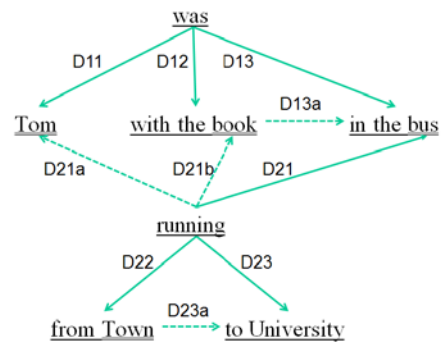


Fig.5. Dependencies possible for S6.

According to our psychological experiment, almost all the human subjects reach very easily the most plausible image (i.e., Fig.3) that corresponds directly to the dependency tree defined by (5) and can be formulated as (6) in L_{md} .

$$\{D11, D12, D13, D21, D22, D23\} \quad (5)$$

$$L(\text{Tom,Book,Tom,Tom},\Lambda_i) \text{IIL}(z,\text{Tom,Bus,Bus},\Lambda_i) \text{IIL}(\text{Bus,Bus,Town,Univ.},\Lambda_i) \quad (6)$$

Quite in the same way, the most plausible interpretations of S7 and S8 are given by (7) and (8), respectively.

$$\begin{aligned} &L(\text{Tom,Book,Tom,Tom},\Lambda_t)\Pi \\ &L(z,\text{Tom,Car,Car},\Lambda_t)\Pi L(\text{Mary,Mary,Car,Car},\Lambda_t) \\ &\Pi L(\text{Mary,Car,Town,Univ.},\Lambda_t) \end{aligned} \quad (7)$$

$$\begin{aligned} &L(\text{Tom,Book,Box,Box},\Lambda_t)\bullet \\ &(L(\text{Tom,Tom,Car,Car},\Lambda_t)\Pi \\ &L(\text{Tom,Car,Town,Univ.},\Lambda_t)\Pi \\ &L(\text{Tom,Box,Tom,Tom},\Lambda_t)) \end{aligned} \quad (8)$$

4 Simulation of MBU

Every semantic interpretation (e.g., (6)) of an NL expression (i.e., S6) is generated by unifying the word meanings according to its corresponding dependency tree (i.e., (5)). In this process, functional words such as verbs and prepositions are employed for structuring the locus formulas. For example, the meanings of ‘with’, ‘in’, ‘run’, etc. are given by (a) - (h) (where $p \neq q$), whose concepts (e.g., $\text{with}(x,y)$) are defined as (9) - (16) in L_{md} based on the authors’ own mental experiences.

$$\begin{aligned} &(a) \ x \text{ (be) with } y: \\ &\text{with}(x,y) \leftrightarrow L(x,y,x,x,\Lambda_t) \end{aligned} \quad (9)$$

$$\begin{aligned} &(b) \ x \text{ (be) in } y: \\ &\text{in}(x,y) \leftrightarrow L(z,x,y,y,\Lambda_t) \end{aligned} \quad (10)$$

$$\begin{aligned} &(c) \ x \text{ run from } p \text{ to } q: \\ &\text{run}(x,p,q) \leftrightarrow L(x,x,p,q,\Lambda_t) \end{aligned} \quad (11)$$

$$\begin{aligned} &(d) \ x \text{ carry } y \text{ from } p \text{ to } q: \\ &\text{carry1}(x,y,p,q) \leftrightarrow L(z,y,x,x,\Lambda_t)\Pi L(x,x,p,q,\Lambda_t) \end{aligned} \quad (12)$$

$$\begin{aligned} &(e) \ x \text{ carry } y \text{ from } p \text{ to } q: \\ &\text{carry2}(x,y,p,q) \leftrightarrow L(z,x,p,q,\Lambda_t)\Pi L(x,y,p,q,\Lambda_t) \end{aligned} \quad (13)$$

$$\begin{aligned} &(f) \ x \text{ drive } y \text{ from } p \text{ to } q: \\ &\text{drive}(x,y,p,q) \leftrightarrow L(x,x,y,y,\Lambda_t)\Pi L(x,y,p,q,\Lambda_t) \end{aligned} \quad (14)$$

$$\begin{aligned} &(g) \ x \text{ move } y \text{ from } p \text{ to } q: \\ &\text{move}(x,y,p,q) \leftrightarrow L(x,y,p,q,\Lambda_t) \end{aligned} \quad (15)$$

$$\begin{aligned} &(h) \ x \text{ keep } y \text{ in } z: \\ &\text{keep}(x,y) \leftrightarrow L(x,y,z,z,\Lambda_t) \end{aligned} \quad (16)$$

On the other hand, entity names such as ‘Tom’, ‘book’ and ‘bus’ are non-functional but utilized for

disambiguation in syntactic dependency. Our psychological experiment revealed that the subjects remembered their own experiences in association with the entity names and that they selected the dependency corresponding to their most familiar experience among all the possibilities. For example, the names in S6 made the people remember the images in the way as formulated by (17) – (19), where $A \approx B$ reads that A evokes B, and + and – denote whether the image is positive (i.e., probable) or negative (i.e., improbable), respectively.

$$\begin{aligned} \text{Tom} \approx &\{ +L(x,\text{Tom,Human,Human},\Theta_t), \\ &+L(\text{Tom,Tom,p,q},\Lambda_t), \dots \} \end{aligned} \quad (17)$$

$$\begin{aligned} \text{Book} \approx &\{ -L(\text{Book,Book,p,q},\Lambda_t), \\ &+L(\text{Human,Book,Human,Human},\Lambda_t), \dots \} \end{aligned} \quad (18)$$

$$\begin{aligned} \text{Bus} \approx &\{ +L(\text{Bus,Bus,p,q},\Lambda_t), +L(\text{Bus,x,p,q},\Lambda_t), \\ &+L(x,\text{Human,Bus,Bus},\Lambda_t), \dots \} \end{aligned} \quad (19)$$

In (17), Θ_t represents ‘Quality’ or ‘Category’ with $g=G_t$, and then $+L(x,\text{Tom,Human,Human},\Theta_t)$ is interpretable as ‘it is positive that Tom is a human’. In the same way, $+L(\text{Tom,Tom,p,q},\Lambda_t)$ as ‘it is positive that Tom moves by himself’, and $-L(\text{Book,Book,p,q},\Lambda_t)$ as ‘it is negative that a book moves by itself’. It is sure that the subjects reached the most plausible interpretation (6) almost unconsciously by using these evoked images for disambiguation. For example, the image for D13 is more probable than that for D13a because of (17) and (19); D21b is improbable because of (18); and the combination of D13 and D21a results in somewhat strange image that Tom was running in the bus, and therefore D21a is seldom selected.

Furthermore, as well as disambiguation, question-answering in MBU was simulated, which is performed by ‘Pattern Matching (PM)’ between the locus formulas of an assertion and a question, for example, (6) of S6 and (20) of S9. Actually, ‘carry’ is defined in the two ways as (12) and (13) but, from now on, only either of them is considered for the sake of simplicity. For example, (20) adopts (12).

(S9) Did Tom carry the book from Town to University?

$$\begin{aligned} &L(z,\text{Book,Tom,Tom},\Lambda_t)\Pi \\ &L(\text{Tom,Tom,Town,Univ.},\Lambda_t) \end{aligned} \quad (20)$$

If (6) includes (20) as is, the answer is positive, but this is not the case. That is, direct trial of PM to the locus formulas (6) – (8) does not always lead to the

desirable outcomes. Therefore, a number of postulates and inference rules must be introduced. The postulates such as P1-P3 are formulas representing pieces of people's commonsense knowledge about the world, where 'A→B' reads 'A implies B' or 'if A then B'.

- (P1) Postulate of Matters as Values:
 $L(z,x,p,q,\Lambda_t) \text{III} L(w,y,x,x,\Lambda_t)$
 $\rightarrow L(z,x,p,q,\Lambda_t) \text{III} L(w,y,p,q,\Lambda_t)$
- (P2) Postulate of Shortcut in Causal Chain:
 $L(z,x,p,q,\Lambda_t) \text{III} L(w,y,x,x,\Lambda_t)$
 $\rightarrow L(z,x,p,q,\Lambda_t) \text{III} L(z,y,p,q,\Lambda_t)$
- (P3) Postulate of conservation of values in time:
 $(L(z,x,p,p,\Lambda_t) \text{III} X_1) \bullet X_2$
 $\rightarrow (L(z,x,p,p,\Lambda_t) \text{III} X_1) \bullet (L(z,x,p,p,\Lambda_t) \text{III} X_2)$

P1 reads that if 'z causes x to move from p to q while w causes y to stay with x' then 'w causes y to move from p to q'. Similarly, P2, so that if 'z causes x to move from p to q while w causes y to stay with x' then 'z causes y to move from p to q as well as x'. Distinguished from these two, P3 is conditional. That is, it is valid only when X_2 does not contradict with ' $L(z,x,p,p,\Lambda_t)$ '.

On the other hand, inference rules such as CS, SS and SC are introduced as follows.

- (CS) Commutativity Law of II:
 $X \text{III} Y \leftrightarrow Y \text{III} X$
- (SS) Simplification Law of II:
 $X \text{III} Y \rightarrow X$
- (SC) Simplification Law of •:
 $X \bullet Y \rightarrow X, X \bullet Y \rightarrow Y$

In order to answer the question S9 to S6, PM is used to compare (6) and (20) as follows.

Apply CS to (6):

$$(6) \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$\underline{L(\text{Bus}, \text{Bus}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}}$$

$$L(z, \text{Tom}, \text{Bus}, \text{Bus}, \Lambda_t) \quad (D1)$$

Apply P1 to D1 (at the underlined part):

$$D1 \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(\text{Bus}, \text{Bus}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D2)$$

Apply SS to D2:

$$D2 \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D3)$$

Apply P2 to D3:

$$D3 \rightarrow L(z, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(\text{Tom}, \text{Tom}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D4)$$

The PM process finds that (20) = D4, and then it is proved that Tom carried the book from Town to University.

For another example, consider the stimulus sentence S7 and the question S10.

(S10) Did Mary carry the car from Town to University?

Adopting (13) for 'carry', the interpretation of S10 can be given by (21).

$$L(z, \text{Mary}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (21)$$

In order to answer the question S10 to S7, PM works as follows, where 'A→B' reads 'B is deduced from A'.

Apply CS to (7):

$$(7) \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Car}, \text{Car}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Mary}, \text{Car}, \text{Car}, \Lambda_t) \quad (D5)$$

Apply P1 to D5:

$$D5 \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Car}, \text{Car}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Mary}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D6)$$

Apply CS to D6:

$$D6 \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Car}, \text{Car}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Mary}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D7)$$

Apply P2 to D7:

$$D7 \rightarrow L(\text{Tom}, \text{Book}, \text{Tom}, \text{Tom}, \Lambda_t) \text{II}$$

$$L(z, \text{Tom}, \text{Car}, \text{Car}, \Lambda_t) \text{II}$$

$$L(z, \text{Mary}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D8)$$

Apply SS to D8:

$$D8 \rightarrow L(z, \text{Mary}, \text{Town}, \text{Univ.}, \Lambda_t) \text{II}$$

$$L(\text{Mary}, \text{Car}, \text{Town}, \text{Univ.}, \Lambda_t) \quad (D9)$$

Hence, the PM proves that (21) = D9 and it is concluded that Mary carried the bus from Town to University.

For the last example, consider the Type III sentence S8, and the question S11 whose interpretation is given by (22).

(S11) Did Tom move the book from Town to University?

$L(\text{Tom, Book, Town, Univ., At})$ (22)

Apply P3 to (8):

(8) \rightarrow $L(\text{Tom, Book, Box, Box, At}) \bullet$
 $(L(\text{Tom, Book, Box, Box, At}) \Pi)$
 $L(\text{Tom, Tom, Car, Car, At}) \Pi$
 $L(\text{Tom, Box, Tom, Tom, At}) \Pi$
 $L(\text{Tom, Car, Town, Univ., At})$ (D10)

Apply CS to D10 several times:

D10 \rightarrow $L(\text{Tom, Book, Box, Box, At}) \bullet$
 $(L(\text{Tom, Car, Town, Univ., At}) \Pi)$
 $L(\text{Tom, Tom, Car, Car, At}) \Pi$
 $L(\text{Tom, Box, Tom, Tom, At}) \Pi$
 $L(\text{Tom, Book, Box, Box, At})$ (D11)

Apply P2 to D11:

D11 \rightarrow $L(\text{Tom, Book, Box, Box, At}) \bullet$
 $(L(\text{Tom, Car, Town, Univ., At}) \Pi)$
 $L(\text{Tom, Tom, Town, Univ., At}) \Pi$
 $L(\text{Tom, Box, Tom, Tom, At}) \Pi$
 $L(\text{Tom, Book, Box, Box, At})$ (D12)

Apply P2 to D12 twice:

D12 \rightarrow $L(\text{Tom, Book, Box, Box, At}) \bullet$
 $(L(\text{Tom, Car, Town, Univ., At}) \Pi)$
 $L(\text{Tom, Tom, Town, Univ., At}) \Pi$
 $L(\text{Tom, Box, Town, Univ., At}) \Pi$
 $L(\text{Tom, Book, Town, Univ., At})$ (D13)

Apply SS and SC to D13:

D13 \rightarrow $L(\text{Tom, Book, Town, Univ., At})$ (D14)

In this way, the system finds that (22) is deduced from (8).

We have implemented our theory of MBU on a PC in Python, high-level programming language, while it is still experimental and evolving. Fig.6 and 7 in APPENDIX show some of the results of question-answering by the MBU system. This can understand User's assertions and answer the questions where the locus formulas were given in Polish notation, for example, as $\bullet \Pi ABC$ for $(A \Pi B) \bullet C$. In the actual implementation, the theorem proving process was simplified as the PM process programmed to apply all the possible

postulates to the locus formula of the assertion in advance and detect any match with the question in the assertion (extended by the postulates) by using the inference rules on the way. During PM, the system is to control its awareness in a top-down way driven by the pair of AC and attribute contained in the question, for example, 'Book' and 'Physical Location (Λ_i)', which is very efficient compared to conventional PM methods without employing any kind of semantic information.

5 Discussion and Conclusion

This paper described an original NLU methodology called MBU (Mental-image Based Understanding), intending so that robots can understand texts in the same way as people. This is quite distinguished from conventional ones and shows a potential good enough to be a very powerful means for realizing awareness in computer and its understanding. To our best knowledge, there is no research similar to ours, namely, NLU based on the model of mental image processing. Therefore, we cannot present any quantitative comparisons with others while we have already commented on our qualitative advantage to conventional methodologies in the previous paper [13]. At conclusion, MIDST could provide the MBU system with an effective methodology to return the correct and satisfied answers in question-answering. The system was designed to disambiguate an input sentence for its most plausible *semantic* interpretation by employing the mental images evoked by the entity names. Disambiguation is the most serious problem for any NLP system. Most of current approaches to it are based on the statistics about certain corpora of texts [15, 16] but they are what lead to the most plausible *syntactic* dependency but not to the most plausible *semantic* interpretation that is most essential to work robots appropriately by words. Our future work will include development of finer measurement of ambiguity for better disambiguation.

References:

- [1] A.M. Turing, *Computing Machinery and Intelligence*, Mind 49, pp. 433 – 460, 1950.
- [2] T. Winograd, *Shifting Viewpoints: Artificial Intelligence and Human-Computer Interaction*, Artificial Intelligence, 2006.
- [3] J. Weizenbaum, *ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine*, Communication of the ACM, 1966.

- [4] I. Dagan, O. Glickman, and B. Magnini, *The PASCAL recognizing textual entailment challenge*, in *Machine Learning Challenges*, Springer Verlag, LNAI 3944, 2006.
- [5] H.J. Levesque, *The Winograd Schema Challenge*, American Association for Artificial Intelligence, 2011.
- [6] D. Ferrucci, et al, *Building Watson: An Overview of the DeepQA Project*, AAAI AI Magazine, Fall 2010.
- [7] C. Thompson, *What Is I.B.M.'s Watson?*, The New York Times Magazine, June 16, 2010.
- [8] D. Ferrucci and A. Lally, *UIMA: An Architectural Approach to Unstructured Information Processing in the Corporate Research Environment*, Journal of Natural Language Engineering, 2004.
- [9] T. Götz and O. Suhre, *Design and Implementation of the UIMA Common Analysis System*, IBM System Journals, 2004.
- [10] M. Yokota, *An Approach to Natural Language Understanding Based on Mental Image Model*, In Sharp, B. (ed.), *Natural Language Understanding and Cognitive Science*, pp. 22 – 31, INSTICC PRESS, 2005.
- [11] M. Yokota, *Aware Computing in Spatial Language Understanding Guided by Cognitively Inspired Knowledge Representation*, Applied Computational Intelligence and Soft Computing, 2012.
- [12] M. Yokota and R. Khummongkol, *Representation and Computation of Human Intuitive Spatiotemporal Concepts as Mental Imagery*, Awareness Science and Technology (iCAST), 2014.
- [13] R. Khummongkol and M. Yokota, *Simulation of Human Awareness Control in Spatiotemporal Language Understanding as Mental Image Processing*, Independent Computing (ISIC), 2014.
- [14] M. Yokota, *A General Theory of Tempological Connectives and Its Application to Spatiotemporal Reasoning*, Soft Computing and Intelligent Systems and Advanced Intelligent Systems, 2008.
- [15] R. Navigli, *Word Sense Disambiguation: A Survey*, ACM Computing Surveys, 41-2, pp.1–69, 2009.
- [16] C.D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.

Appendix:

```

1 C:\Python27\python.exe C:/Users/20150714yo/PycharmProjects/iWork/first_3.py
2 *****
3 Enter a sentence: Tom was with the book in the bus running from town to university.
4 ***** Your input *****
5 ['tom', 'be', 'with', 'the', 'book', 'in', 'the', 'bus', 'run', 'from', 'town', 'to', 'university']
6 ***** Defined words *****
7 ['PN', 'V', 'P', 'DE', 'N', 'P', 'DE', 'N', 'V', 'P', 'N', 'P', 'N']
8
9 ***** Your turn *****
10 Please enter your question: Did Tom carry the book from town to university?
11 ['Did', 'Tom', 'carry', 'the', 'book', 'from', 'town', 'to', 'university', '?']
12 Locus of Stimulus Sentence = ['SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', ['bus',
', 'bus', 'town', 'university', 'a'], ['bus', 'bus', 'town', 'university', 'a'], ['tom', 'tom', 'town',
', 'university', 'a'], ['tom', 'book', 'tom', 'tom', 'a'], ['tom', 'tom', 'bus', 'bus', 'a'], ['tom',
', 'bus', 'town', 'university', 'a'], ['tom', 'book', 'town', 'university', 'a'], ['tom', 'book', '
bus', 'bus', 'a'], ['tom', 'tom', 'town', 'university', 'a']]
13
14 Locus of Question = [['SAND', ['tom', 'book', 'tom', 'tom', 'a'], ['tom', 'tom', 'town', 'university',
', 'a']], ['SAND', ['tom', 'book', 'town', 'university', 'a'], ['tom', 'tom', 'town', 'university', '
a']]]
15
16 =====
17
18 Your question: Did Tom carry the book from town to university?
19 *****
20 Yes: tom carry the book from town to university
21
22 =====
23
24 Process finished with exit code 0
25

```

Fig.6. Q-A for Type I stimulus sentence.

```

1 C:\Python27\python.exe C:/Users/20150714yo/PycharmProjects/iWork/first_3.py
2 *****
3 Enter a sentence: Tom kept the book in a box before he drove the car from town to university with the
box.
4 ***** Your input *****
5 ['tom', 'keep', 'the', 'book', 'in', 'a', 'box', 'before', 'he', 'drive', 'the', 'car', 'from', 'town',
', 'to', 'university', 'with', 'the', 'box']
6 ***** Defined words *****
7 ['PN', 'V', 'DE', 'N', 'P', 'DE', 'N', 'CJ', 'PN', 'V', 'DE', 'N', 'P', 'N', 'P', 'N', 'P', 'DE', 'N',
']
8
9 ***** Your turn *****
10 Please enter your question: Did Tom move the book from town to university?
11 ['Did', 'Tom', 'move', 'the', 'book', 'from', 'town', 'to', 'university', '?']
12 Locus of Stimulus Sentence = ['CAND', 'SAND', ['tom', 'tom', 'town', 'university', 'a'], ['tom', '
book', 'box', 'box', 'a'], 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', 'SAND', '
SAND', 'SAND', 'SAND', 'SAND', ['tom', 'tom', 'car', 'car', 'a'], ['tom', 'tom', 'town', 'university',
', 'a'], ['tom', 'box', 'tom', 'tom', 'a'], ['car', 'car', 'town', 'university', 'a'], ['box', 'box',
', 'town', 'university', 'a'], ['tom', 'book', 'box', 'box', 'a'], ['tom', 'tom', 'town', 'university',
', 'a'], ['tom', 'box', 'car', 'car', 'a'], ['tom', 'box', 'town', 'university', 'a'], ['tom', 'book',
', 'tom', 'tom', 'a'], ['tom', 'book', 'town', 'university', 'a'], ['tom', 'box', 'town', 'university',
', 'a'], ['tom', 'book', 'car', 'car', 'a'], ['tom', 'book', 'town', 'university', 'a']]
13
14 Locus of Question = [['tom', 'book', 'town', 'university', 'a']]
15
16 =====
17
18 Your question: Did Tom move the book from town to university?
19 *****
20 Yes: tom move the book from town to university
21
22 =====
23
24 Process finished with exit code 0
25

```

Fig.7. Q-A for Type III stimulus sentence.