The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Towards natural intelligence modeling as a formal system based on Mental Image Directed Semantic Theory (Part 2)

Masao Yokota, Taiki Tometsuka, Rong Zhao, Kaoru Sugita and Tetsushi Oka

Fukuoka Institute of Technology, 3-30-1 Wajiro-higashi, Higashi-ku, Fukuoka-shi, 811-0295, Japan
(Tel : 81-92-606-5897; Fax : 81-92-606-8923)
(yokota@fit.ac.jp)

Abstract: Yokota, M. has proposed his original semantic theory "Mental Image Directed Semantic Theory (MIDST)" and has been challenging to model natural intelligence as a formal system. This paper presents a brief sketch of the attempt on systematic representation and computation of subjective spatiotemporal knowledge based on certain hypotheses of mental image in human.

Keywords: Natural language, Multimedia understanding, Robotic sensation and action.

I. INTRODUCTION

Another paper of ours for this session [1] presents the fundamentals of the formal system for natural intelligence and how to formalize mental operations and natural concepts within it. The forma l system consists of the formal language $\boldsymbol{L_{md}}$ [1] and the deductive system. The latter is based on the deductive apparatus for predicate logic and is to be provided with postulates concerning human empirical knowledge of space and time as well. This paper focuses on systematic formalization of human empirical knowledge pieces of space and time as postulates for the deductive system and its application to natural language understanding (NLU) for spatiotemporal expressions.

II. PROVISION OF POSTULATES

The deductive system must be as well provided with knowledge pieces in order to solve certain problems in its world or task domain. Such knowledge pieces as called "postulates" here stand for human intuitive laws of the world and are to be treated as equivalents to axioms. Those presented below concern exclusively space and time in order for spatiotemporal language understanding.

1. Fundamental Properties of Locus

The postulates **P1** and **P2** state that *a matter never has different values of an attribute with a standard at a time*. These are called "Postulates of Identity in Assigned Values". **P1** is employed exclusively to detect semantic anomaly in such a sentence as "The red box is black" while **P2** is useful to detect event gaps in such a context as "Tom was in London yesterday and he is in Paris today."

The syntax of $\boldsymbol{L_{md}}$ allows Matter terms to appear at Values and Standard in order to represent their values in each place at the time and over the time-interval, respectively. This rule can be formulated as **P3** and **P4**. The postulate **P3** is to be utilized for such inference as "Mary went to Tom when he was in the garden. Therefore, Mary went to the (same) garden." while **P4** is for such inference as "Jim is taller than Tom. Tom is 2m tall. Therefore, Jim is taller than 2m."

P1. $L(x,y,p_1,q_1,a,g,k)\Pi L(z,y,p_2,q_2,a,g,k)$
$.\supset.\ p_1=p_2 \wedge q_1=q_2$

P2. $L(x,y,p_1,q_1,a,g,k)\bullet L(z,y,p_2,q_2,a,g,k)$
$.\supset.\ q_1=p_2$

P3. $L(x_0,y,z_1,z_2,a,g,k)\Pi L(x_1,z_1,p_1,q_1,a,g,k)\Pi$
$L(x_2,z_2,p_2,q_2,a,g,k).\supset_0.L(x_0,y,p_1,q_2,a,g,k)$

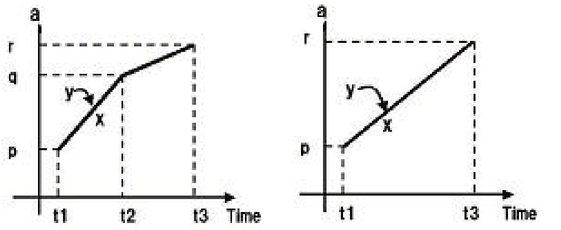
P4. $L(x_0,y,p_1,p_2,a,g,z)\Pi L(x_1,z,q,q,a,g,k)$
$.\supset_0.L(x_0,y,p_1,p_2,a,g,q)$

It is quite subjective how to articulate a locus. For example, whether the point (t_2, q) in Fig.1-a is significant or not so as in Fig.1-b, more generally, locus articulation depends on the precisions or the granularities of these standards, which can be formulated as **P5** and **P6**, so called, '*Postulates of Arbitrariness in Locus Articulation*'.

P5. $(\forall p,q,r,k)(\exists k')L(y,x,p,q,a,g,k)\bullet L(y,x,q,r,a,g,k)$
$.\supset_0.L(y,x,p,r,a,g,k')\wedge k'\neq k$

P6. $(\forall p,r,k)(\exists q,k')L(y,x,p,r,a,g,k).\supset_0.$
$L(y,x,p,q,a,g,k')\bullet L(y,x,q,r,a,g,k')\wedge k'\neq k$

These postulates affect the process of conceptualization on a word based on its referents in the world and moreover they are very useful for spatiotemporal inference in such a context as "Tom flied from Tokyo to Nagoya and consecutively from Nagoya to Osaka. Therefore, he moved from Tokyo to Osaka" or "Tom moved from Tokyo to Osaka. Therefore, he passed somewhere (between the two places)".



(a) (b)

Fig.1. Arbitrariness in locus articulation due to standards: Standard k_1 (a) is finer than k_2 (b).

2. Perception of Time

A perceptual locus can be formulated with atomic locus formulas and temporal conjunctions such as SAND ($\wedge_0$ or Π) and CAND ($\wedge_1$ or $\bullet$). This is not necessarily the case for a conceptual locus corresponding to such a generalized mental image or knowledge piece. For example, people usually interpret the construction '*B* happens *before A* happens' as a general causality, namely, as 'If *A* happens, *B* happens *in advance*'. Whereas this should be formulated with logical connectives other than conjunctions also involved, **D1** [1] is exclusively for perceptual loci so far as it is because there is no interpreting a negated locus formula as a locus with ***a unique time-interval*** necessary to determine a unique temporal relation t_i.

Considering such a definition as '$A\supset B \Leftrightarrow \sim A\vee B$ $(.\equiv.\sim(A\wedge\sim B))$' in standard logic, it is not unnatural to assume the identity of a locus formula with its negative in absolute time-interval, that is, negation-freeness of absolute time passing under a locus referred to by its suppressed absolute time-interval. Therefore, in order to make **D1** valid also for conceptual loci, we introduce a meta-function *d* defined by **D5** and its related postulates **P7** and **P8** as follows, where *d* is to extract the ***suppressed*** absolute interval of a locus formula *c*.

D5. $\delta(\chi)=[t_a,t_b](\in D)$
where $\chi\Pi\varepsilon([t_a,t_b])$.

P7. $\delta(\sim\alpha)=\delta(\alpha)$
where *a* is an atomic locus formula.

P8. $\delta(\chi)=[t_{min}, t_{max}]$
where t_{min} and t_{max} are respectively the *minimum* and the *maximum* time-point included in the absolute time-intervals of the atomic locus formulas, either positive or negative, within *c*.

These postulates lead to **T1** (Theorem of absoluteness of time passing (or negation-freeness of absolute time passing)) below. This theorem can read that absolute time passes during an *objective* event whether it may be perceived *subjectively* as χ or as $\sim\chi$.

T1. $\delta(\sim c)=\delta(c)$

(Proof)

According to **P5** and **P6**, the time-interval of each atomic locus formula involved in *c* is negation-free and therefore so is for $[t_{min}, t_{max}]$ of $d(c)$. **[Q.E.D.]**

The counterpart of the contrapositive in standard logic (i.e. $A\supset B.\equiv.\sim B\supset\sim A$) is given as **T2** (Tempo-logical Contrapositive) whose rough proof is as follows immediately below, where the left hand of ':' refers to the theses (e.g., **PL** is a subset of those in pure predicate logic) employed at the process indicated by the conventional meta-symbol '$\rightarrow$' or '$\leftrightarrow$' for entailment (left-to-right or bi-directional).

T2. $\chi_1\supset_i\chi_2.\equiv.\sim\chi_2\supset_{-i}\sim\chi_1$

(Proof)

D1: $\chi_1\supset_i\chi_2 \leftrightarrow (\chi_1\supset\chi_2)\wedge\tau_i(\chi_1,\chi_2)$
PL: $\leftrightarrow (\sim\chi_2\supset\sim\chi_1)\wedge\tau_i(\chi_1,\chi_2)$
T1: $\leftrightarrow (\sim\chi_2\supset\sim\chi_1)\wedge\tau_i(\sim\chi_1,\sim\chi_2)$
D1: $\leftrightarrow (\sim\chi_2\supset\sim\chi_1)\wedge\tau_{-i}(\sim\chi_2,\sim\chi_1)$
D1: $\leftrightarrow \sim\chi_2 \supset_{-i} \sim\chi_1$ **[Q.E.D.]**

Therefore, S1 and S2 are proved to be paraphrases each other by employing T2 while S3 and S4 are proved so by the definition of tempological conjunctions (i.e. $\wedge_i$).

(S1) It gets cloudy *before* it rains.
=If it rains, it gets cloudy *in advance*.
($\equiv$Raining$\supset_{-5}$ Getting_Cloudy)
(S2) It does not rain *after* it does not get cloudy.
=Unless it gets cloudy, it does not rain *later*.
($\equiv\sim$Getting_Cloudy $\supset_5$ ~Raining)
(S3) It got cloudy *before* it rained.
($\equiv$Raining$\wedge_{-5}$Getting_Cloudy)
(S4) It rained *after* it got cloudy.
($\equiv$Getting_Cloudy$\wedge_5$Raining)

3. Reversibility of Spatial Event

As already mentioned in [1], all loci in attribute spaces are assumed to correspond one to one with movements or, more generally, temporal events of the FAO. Therefore, the L_{md} expression of an event is compared to a movie film recorded through a floating camera because it is necessarily grounded in FAO's movement over the event. And this is why S5and S6 can refer to the same scene in spite of their appearances, where what 'sinks' or 'rises' is the FAO as illustrated in Fig.2 and whose conceptual descriptions are given as (1) and (2), respectively, where 'A_{13}', '$\uparrow$' and '$\downarrow$' refer to the attribute 'Direction' and its values 'upward' and 'downward', respectively.

(S5) The path sinks to the brook.
(S6) The path rises from the brook.

$$(\exists y,z,p)L(\_,y,p,z,A_{12},G_s,\_)\Pi\; L(\_,y,\downarrow,\downarrow,A_{13},G_s,\_)\wedge path(y)\wedge brook(z)\wedge z\neq p \quad (1)$$

$$(\exists y,z,p)L(\_,y,z,p,A12,G_s,\_)\Pi\; L(\_,y,\uparrow,\uparrow,A_{13},G_s,\_)\wedge path(y)\wedge brook(z)\wedge z\neq p \quad (2)$$

Such a fact is generalized as **P9** (*Postulate of Reversibility of Spatial Event (PRS)*), where χ_s and χ_s^R are a perceptual locus and its 'reversal' for a certain spatial event, respectively, and they are substitutable with each other because of the property of '$\equiv_0$'. This postulate can be one of the principal inference rules belonging to people's common-sense knowledge about geography.

P9. $\chi_s^R . \equiv_0 \chi_s$

This postulation is also valid for such a pair of S7 and S8 as interpreted approximately into (3) and (4), respectively. These pairs of conceptual descriptions are called equivalent in the PRS, and the paired sentences are treated as paraphrases each other.

(S7) Route A and Route B meet at the city.

(S8) Route A and Route B separate at the city.

$$(\exists p,y,q)L(\_,Route\_A,p,y,A_{12},G_s,\_)\Pi L(\_,Route\_B,q,y,A_{12},G_s,\_)\wedge city(y)\wedge p\neq q \quad (3)$$

$$(\exists p,y,q)L(\_,Route\_A,y,p,A_{12},G_s,\_)\Pi L(\_,Route\_B,y,q,A_{12},G_s,\_)\wedge city(y)\wedge p\neq q \quad (4)$$

Fig.2. Slope as spatial event.

4. Partiality of Matter

Any matter is assumed to consist of its parts in a structure (i.e. spatial event), which is generalized as **P10** (*Postulate of Partiality of Matter*) here. For example, Fig.3 shows that an ISR x_1 can be deemed as a complex of ISRs x_2 and x_3. This postulate, in cooperation with P9, is utilized for translating such a paradoxical sentence as "The Andes Mountains run north and south." into such a plausible interpretation as "One part of the Andes Mountains runs north (from somewhere) and the other part runs south".

P10. $L(y,x_1,p,q,a,G_s,k)\bullet L(y,x_1,q,r,a,G_s,k)$
$.\equiv_0. \ L(y,x_2,p,q,a,G_s,k)\Pi L(y,x_3,q,r,a,G_s,k)$

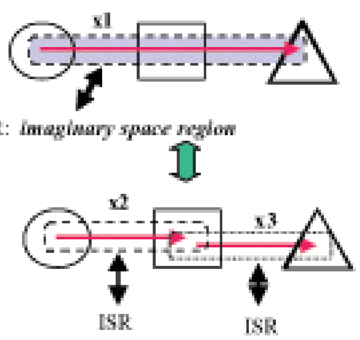


Fig.3. Partiality of ISR.

III. APPLICATION TO NLU

Our intelligent system IMAGES-M [1] can translate systematically natural language and L_{md} expression into each other by utilizing syntactic rules and word meaning descriptions of natural language.

A word meaning description M_w is given by (5) as a pair of 'Concept Part (C_p)' and 'Unification Part (U_p)'.

$$M_w \Leftrightarrow [C_p:U_p] \quad (5)$$

The C_p of a word W is a locus formula about properties and relations of the matters involved such as shapes, colors, functions, potentialities, etc while its U_p is a set of operations for unifying the C_ps of W's syntactic governors or dependents. For example, the meaning of the English verb 'carry' can be given by (6).

$$[(\exists x,y,p_1,p_2,k)\ L(x,x,p_1,p_2,A_{12},Gt,k)\Pi L(x,y,p_1,p_2,A_{12},Gt,k)\wedge x\neq y\wedge p_1\neq p_2:ARG(Dep.1,x); ARG(Dep.2,y);] \quad (6)$$

The U_p above consists of two operations to unify the first dependent (Dep.1) and the second dependent (Dep.2) of the current word with the variables x and y, respectively. Here, Dep.1 and Dep.2 are the 'subject' and the 'object' of 'carry', respectively. Therefore, the surface structure '*Mary carries a book*' is translated into the conceptual structure (7) via the surface dependency structure shown in Fig.4. This process is completely reversible.

$$(\exists y,p_1,p_2,k)L(Mary,Mary,p_1,p_2,A_{12},Gt,k)\Pi L(Mary,y,p_1,p_2,A_{12},Gt,k)\wedge Mary\neq y \wedge p_1\neq p_2\wedge book(y) \quad (7)$$

For another example, the meaning description of the English preposition 'through' is also given by (8).

$$[(\exists x,y,p_1,z,p_3,g,k,p_4,k_0)(\underline{L(x,y,p_1,z,A_{12},g,k)}\bullet L(x,y,z,p_3,A_{12},g,k))\Pi L(x,y,p_4,p_4,A_{13},g,k_0) \wedge p_1\neq z\wedge z\neq p_3:ARG(Dep.1,z); IF(Gov=Verb)\rightarrow PAT(Gov,(1,1)); IF(Gov=Noun)\rightarrow ARG(Gov,y);] \quad (8)$$

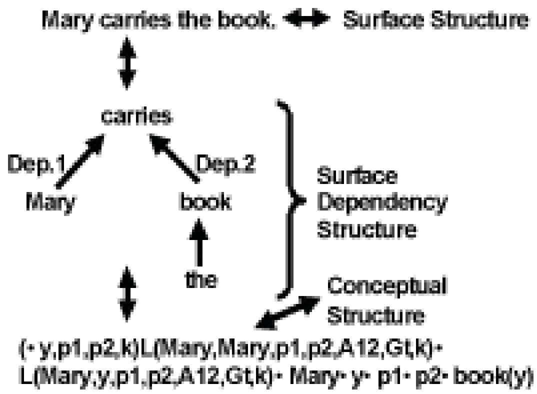


Fig.4. Mutual translation between text and L_{md}

The U_p above is for unifying the C_ps of the very word, its governor (Gov, a verb or a noun) and its dependent (Dep.1, a noun). The second argument (1,1) of the command PAT indicates the underlined part of (9) and in general (i,j) refers to the partial formula covering from the ith to the jth atomic formula of the current C_p.

This part is the pattern common to both the C_ps to be unified. This is called 'Unification Handle (U_h)' and when missing, the C_ps are to be combined simply with '$\wedge$'. Therefore the sentences S9, S10 and S11 are interpreted as (9), (10) and (11), respectively. The underlined parts of these formulas are the results of PAT operations. The expression (12) is the C_p of the adjective 'long' implying 'there is some value greater than some standard of 'Length (A_{02})' which is often simplified as (12').

(S9) The train runs through the tunnel.

$$(\exists x,y,p_1,z,p_3,k,p_4,k_0)(\underline{L(x,y,p_1,z,A_{12},Gt,k)}\bullet L(x,y,z,p_3,A_{12},Gt,k))\Pi L(x,y,p_4,p_4,A_{13},Gt,k_0) \wedge p_1\neq z \wedge z\neq p_3 \wedge train(y) \wedge tunnel(z) \quad (9)$$

(S10) The path runs through the forest.

$$(\exists x,y,p_1,z,p_3,k,p_4,k_0)(\underline{L(x,y,p_1,z,A_{12},Gs,k)}\bullet L(x,y,z,p_3,A_{12},Gs,k))\Pi L(x,y,p_4,p_4,A_{13},Gs,k_0) \wedge p_1\neq z \wedge z\neq p_3 \wedge path(y) \wedge forest(z) \quad (10)$$

(S11) The path through the forest is long.

$$(\exists x,y,p_1,z,p_3,x_1,k,q,k_1,p_4,k_0)(L(x,y,p_1,z,A_{12},Gs,k)\bullet L(x,y,z,p_3,A_{12},Gs,k))\Pi L(x,y,p_4,p_4,A_{13},Gs,k_0) \wedge L(x_1,y,q,q,A_{02},Gt,k_1) \wedge p_1\neq z \wedge z\neq p_3 \wedge q>k_1 \wedge path(y) \wedge forest(z) \quad (11)$$

$$(\exists x_1,y_1,q,k_1)L(x_1,y_1,q,q,A_{02},Gt,k_1)\wedge q>k_1 \quad (12)$$

$$(\exists x_1,y_1,k_1)L(x_1,y_1,Long,Long,A_{02},Gt,k_1) \quad (12')$$

For simplicity, we have recently employed such a format for text meaning representation as shown in Table 1, so called, 'Discourse Image Tree (DIT)'. This table represents the meaning of S12 where all the formulas are expressed in Polish notation. In general, the leaves of a DIT consist of atomic locus formulas, labeled as 'L_n', and connectives such as CANDs and SANDs. Table 2 shows the DIT for S13.

Table 1*. Discourse Image Tree of S12

Discourse Image	**D_1=Ù S_1S_2**		
Sentence Image	**$S_1=C_1$**		**$S_2=C_2$**
Clause Image	**C_1=P P_1P_2**		**$C_2=P_3$**
Phrase Image	**$P_1=L_1$**	**$P_2=L_2$**	**$P_3=L_3$**
Locus Image	**L_1**	**L_2**	**L_3**
Causer	x	x	x1
Attr_Carrier	road	**P_1**	it
IntVal	Tokyo	west	very. long
FinVal	Osaka	west	very. long
Attribute	A_{12}	A_{13}	A_{02}
Event Type	Gs	Gs	Gt
Standard	k	k_1	k_2

* $\mathbf{P_1}=(\exists x,y,k)L(x,y,Tokyo,Osaka,A_{12},Gs,k)\wedge road(y)$

$\mathbf{P_2}=(\exists x,k_1)L(x,\mathbf{P_1},west,west,A_{13},Gs,k_1)$

Table2. Discourse Image Tree of S13

$D_1=\bullet S_1S_2$					
$S_1=\cdot C_2C_1$					**$S_2=C_3$**
C_1=P P $P_1P_2P_3$				**$C_2=P_4$**	**$C_3=P_5$**
P_1=P P L_1L_2		**$P_2=L_3$**	**$P_3=L_4$**	**$P_4=L_5$**	**$P_5=L_6$**
L_1	**L_2**	**L_3**	**L_4**	**L_5**	**L_6**
tom	tom	x_1	x_2	he	x_3
tom	book	book	**P_1**	he	book
q_1	q_1	red	very.fast	q_2	?q
OSK	OSK	red	very.fast	TKY	?q
A_{12}	A_{12}	A_{32}	A_{16}	A_{12}	A_{12}
Gt	Gt	Gt	Gt	Gt	Gt
k_1	k_1	k_2	k_3	k_1	k_1

A DIT can realize hierarchical representation and computation of text meaning consisting of five levels of image: 1) Locus level image, 2) Phrase level image, 3) Clause level image, 4) Sentence level image and 5) Discourse level image and thereby can cope with higher-order meaning representation as shown just below Table 1.

(S12) The road runs west from Tokyo to Osaka. It is very long.

(S13) Tom carries the red book very fast to Osaka after he reaches Tokyo. Then, where is the book?

IV. CONCLUSION

Most of computations on $\boldsymbol{L_{md}}$ are simply for unifying (or identifying) atomic loci and for evaluating arithmetic expressions such as 'p=q', and therefore we believe that our formalismcan reduce the computational complexities of the traditional ones when applied to the same kinds of problems described here. Moreover, recent employment of DITs has enabled us to program in procedural languages and thereby remarkably reduced computational complexity for $\boldsymbol{L_{md}}$ expressions while the earlier version of IMAGES-M was programmed in PROLOG and therefore inefficient in computation. This advantage of ours comes from the meaning representation scheme normalized by atomic locus formulas, which has come to facilitate higher-order representation and computation as shown in Tables 1 and 2.

Our future work will include further elaboration of the deductive system, improvement of DIT processing algorithms and establishment of learning facilities for automatic acquisition of word concepts from sensory data and more powerful interfaces for human-system communication by natural language under real environments.

REFERENCES

[1] Yokota M et al (2010), Towards natural intelligence modeling as a formal system based on Mental Image Directed Semantic Theory (Part 1), Proc. of AROB2010, Beppu, Oita, Feb. 2010.