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 1)

Masao Yokota, Haosheng Li, Hailan Quan, 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 the fundamentals of the formal system and how to formalize mental operations and natural concepts within it.

Keywords: Natural intelligence, Mental image model, Natural language, Formal system.

I. INTRODUCTION

In order for facilitating intuitive and *coherent* human-robot interaction, it is essential to develop a systematically computable knowledge representation language (KRL) as well as representation-free technologies such as neural networks for processing unstructured sensory/motory data. This type of language is indispensable to knowledge-based processing such as understanding sensory events, planning appropriate actions and knowledgeable communication with ordinary people in natural language, and therefore it needs to have at least a good capability of representing spatiotemporal events that correspond to human/robotic sensations and actions in the real world. Yokota, M. has employed the formal language so called 'Mental-image Description Language (L_{md})' proposed in his original semantic theory 'Mental Image Directed Semantic Theory (MIDST)' [1]. This language has already been implemented on several types of computerized intelligent systems including IMAGES-M [2]. The most remarkable feature of L_{md} is its capability of formalizing spatiotemporal matter concepts grounded in human/robotic sensation while the other similar KRLs are designed to describe the logical relations among conceptual primitives represented by lexical tokens [3] with the risk of "predicate drift" [4].

Our final goal is to model natural intelligence as a formal system based on MIDST. A formal system is defined as a pair of a formal language and a deductive system consisting of the axioms and inference rules employed for theorem derivation. L_{md} is a formal language for many-sorted predicate logic with 5 types of terms specific to the mental image model. Therefore, the deductive system intended here is to be based on the deductive apparatus for predicate logic. This paper introduces the formal system and focuses on its fundamentals .

II. FORMAL SYSTEM BASED ON L_{md}

A formal system is defined as a pair of a formal language and a deductive system consisting of the axioms and inference rules employed for theorem derivation. L_{md} is a formal language for many-sorted predicate logic with 5 types of terms specific to the mental image model. Therefore, the deductive system intended here is to be based on the deductive apparatus for predicate logic.

1. Syntax of L_{md}

The symbols of L_{md} for the deductive system are listed as (i)-(ix) below. These symbols are possibly subscripted just like A_{01}, G_t, etc.

(i) logical connectives : ~, ∧, ∨, ⊃, ≡
(ii) quantifiers : ∀, ∃
(iii) auxiliary constants : ., (,)
(iv) sentence variables : χ
(v) predicate variables : ψ
(vi) individual variables
  a) matter variables : x, y, z
  b) attribute variables : a
  c) value variables : p, q, r, s, t
  d) pattern variables : g
  e) standard variables : k
(vii) sentence constants : N
(viii) predicate constants : L, =, ≠, >, < (and others to be introduced where needed)
(ix) individual constants
  a) matter constants : to be introduced where needed
  b) attribute constants : A, B
  c) value constants : to be introduced where needed
  d) pattern constants : G
  e) standard constants : K
(x) function constants : arithmetic operators such as +, -, etc. (and others to be introduced where needed)
(xi) meta-symbols : ⇔, →, ↔ (and others to be introduced where needed)
(xii) others: to be defined by the symbols above.

The system is a many-sorted predicate logic with five kinds of individuals employed for one special predicate constant '*L*' so called 'Atomic Locus'. Except this point, the syntactic rules and the theses of the system are the same as those of the conventional predicate logic.

The predicate 'L' is such a seven-place predicate that is given by expression (1).

$$L(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7) \quad (1)$$

Expression (1) is a well-formed formula (i.e. wff) called 'Atomic locus formula' if and only if the conditions below are satisfied. A well-formed formula consisting of atomic formulas and logical connectives is called simply 'Locus formula'.

(a) w_1 is a matter term (variable or constant)
(b) w_2 is a matter term
(c) w_3 is a value or a matter term
(d) w_4 is a value or a matter term
(e) w_5 is an attribute term
(f) w_6 is a pattern term
(g) w_7 is a standard (or matter) term

2. Semantics of L_{md}

The domain-specificity in the syntax and semantics of L_{md} is exclusively related to atomic locus formulas and the essential part of its semantics is subject to their interpretation controlled by the family of domain-specific constants, namely, Attributes, Values, Patterns and Standards intended to correspond well with natural or artificial sensory systems.

So far, about 50 attributes (e.g. Color, Volume, Taste) related to the physical world have been extracted exclusively from Japanese and English words. The values for each attribute (e.g. Red, Small, Sour) are to be arranged so as to form a structure "Attribute Space", so called.

Correspondingly, six categories of standards (i.e. Rigid, Species, Proportional, Individual, Purposive and Declarative Standards) have been found that are assumed necessary for representing values of each attribute. In general, the attribute values represented by words (e.g. Dark, Large, Bitter) are relative to certain standards. These standards are to be utilized exclusively for coping with vagueness and controlling granularity of attribute values.

As for the pattern term, two constants have been hypothetically provided, that is, G_t (i.e. Temporal event) and G_s (Spatial event) indicating temporal and spatial change in an attribute, respectively.

III. DEFINED CONSTANTS

1. Tempological connective and Empty event

The deductive system employs 'tempo-logical connectives (TLCs)' with which to represent both temporal and logical relations between two loci over certain time-intervals. The definition of a tempo-logical connective C_i is given by **D1**, where t_i, c and C refer to one of ***purely*** temporal relations indexed by an integer 'i', a locus, and an ordinary binary logical connective such as the conjunction '$\wedge$', respectively. The suffix 'i' ($-6 \leq i \leq 6$) indicate 13 types of 1-D topological relation (i.e. 0: EQUALS; +1: MEETS; -1: MET-BY; +2: STARTS; -2: STARTED-BY; +3: DURING; -3: CONTAINS; +4: FINISHES; -4: FINISHED-BY; +5: BEFORE; -5: AFTER; +6: OVERLAPS; -6: OVERLAPPED-BY [5]). The TLCs used most frequently are 'SAND ($\wedge_0$)' and 'CAND ($\wedge_1$)', standing for 'Simultaneous AND' and 'Consecutive AND' and conventionally symbolized as 'Π' and '•', respectively.

D1. $\chi_1 \, C_i \, \chi_2 \Leftrightarrow (\chi_1 \, C \, \chi_2) \wedge \tau_i(\chi_1, \chi_2)$
where
$\tau_{-i}(\chi_2, \chi_1) \equiv \tau_i(\chi_1, \chi_2)$
$(\forall i \in \{0, \pm1, \pm2, \pm3, \pm4, \pm5, \pm6\})$

In order for explicit indication of ***absolute*** time elapsing, 'Empty Event' denoted by 'ε' is introduced as **D2** with the attribute 'Time Point (A_{34})' and the Standard of absolute time 'K_{Ta}', where $\boldsymbol{R}$ and D denote the total sets of real numbers and absolute time intervals, respectively. (Usually people can know only a certain ***relative*** time point by a clock that is seldom exact and that is to be denoted by another Standard in the L_{md}.) According to this scheme, the suppressed absolute time-interval $[t_a, t_b]$ of a locus χ can be indicated as (2).

D2. $\varepsilon([t_i,t_j]) \Leftrightarrow (\exists x,y,g) L(x,y,t_i,t_j,A_{34},g,K_{Ta})$
where
$[t_i,t_j] \in D = \{[t_1, t_2] \mid t_1 < t_2 \; (t_1, t_2 \in \boldsymbol{R})\}$

$$\chi \Pi \varepsilon([t_a,t_b]) \quad (2)$$

2. Mental Operations

People can transform their mental images in several ways such as mental rotation [6]. Here are introduced two kinds of such mental operations, namely, 'reversing' and 'duplicating'.

For example, people can easily imagine the reversal of an event just like 'rise' versus 'sink'. This mental operation is here denoted as a meta-function 'R' and recursively defined as **D3**, where χ_i stands for a perceptual locus. The reversed values p^R and q^R depend on the properties of the attribute values p and q.

D3. $(\chi_1 \bullet \chi_2)^R \Leftrightarrow \chi_2^R \bullet \chi_1^R$
$(\chi_1 \Pi \chi_2)^R \Leftrightarrow \chi_1^R \Pi \chi_2^R$
$L^R(x,y,p,q,a,g,k) \Leftrightarrow L(x,y,q^R,p^R,a,g,k)$

For another example, people can easily imagine the repetition of an event just like 'visit twice' versus 'visit once'. This operation is also a meta-function recursively defined as **D4**, where 'n' is an integer representing the frequency of a locus formula χ.

D4. $\chi^n \Leftrightarrow \chi \quad (n=1)$
$\chi^n \Leftrightarrow \chi \bullet \chi^{n-1} \quad (n>1)$

3. Natural Concepts

All the natural concepts of physical or metaphysical matters and their relations are to be defined in association with specific loci in specific attribute spaces formalized by the four types of individual constant stated above. These concepts can be introduced as non-logical constants, namely, predicate constants (e.g. carry, snow) or matter constants (e.g. Tokyo, Tom) and defined in context of locus formulas as follows.

A. Event Concepts

An event here, usually referred by a verb, preposition, adjective or so in natural language, is defined as a spatiotemporal relation among certain matters in the world, which is to be conceptualized as a generalization of a perceptual locus, namely, a combination of atomic loci articulated by tempological conjunctions (i.e. $\wedge_i$) with the abstraction operator 'λ'.

For example, the English verb concepts 'carry (=convey)' and 'shuttle' are to be defined as (3) and (4), respectively. In turn, the expression (5) is the definition of the English verb concept 'fetch'. This implies such a temporal event that 'x' goes for 'y' and then comes back with it. In the same way, the English verb concept 'hand' or 'receive' is defined equivalently as (6) or its abbreviation (7). These concepts can be graphically interpreted as Fig.1.

$$(\lambda x,y)\mathrm{carry}(x,y)\Leftrightarrow(\lambda x,y)\mathrm{convey}(x,y) \Leftrightarrow(\lambda x,y)(\exists p,q,k)L(x,x,p,q,A_{12},G_t,k) \Pi L(x,y,p,q,A_{12},G_t,k)\wedge x\neq y\wedge p\neq q \quad (3)$$

$$(\lambda x)\mathrm{shuttle}(x)\Leftrightarrow(\lambda x)(\exists p,q,k) (L(x,x,p,q,A_{12},G_t,k) \bullet L^R(x,x,p,q,A_{12},G_t,k))^n\wedge p\neq q\wedge n\geq 1 \Leftrightarrow(\lambda x)(\exists p,q,k) (L(x,x,p,q,A_{12},G_t,k) \bullet L(x,x,q,p,A_{12},G_t,k))^n\wedge p\neq q\wedge n\geq 1 \quad (4)$$

$$(\lambda x,y)\mathrm{fetch}(x,y)\Leftrightarrow(\lambda x,y)(\exists p_1,p_2,k) L(x,x,p_1,p_2,A_{12},G_t,k) \bullet(\underline{L(x,x,p_2,p_1,A_{12},G_t,k) \Pi L(x,y,p_2,p_1,A_{12},G_t,k)})\wedge x\neq y\wedge p_1\neq p_2 \quad (5)$$

$$(\lambda x,y,z)\mathrm{hand}(x,y,z) \Leftrightarrow(\lambda x,y,z)\mathrm{receive}(z,y,x) \Leftrightarrow(\lambda x,y,z)(\exists k)L(x,y,x,z,A_{12},G_t,k) \Pi L(z,y,x,z,A_{12},G_t,k)\wedge x\neq y\wedge y\neq z\wedge z\neq x \quad (6)$$

$$\Leftrightarrow(\lambda x,y,z)(\exists k)L(\{x,z\},y,x,z,A_{12},G_t,k) \wedge x\neq y\wedge y\neq z\wedge z\neq x \quad (7)$$

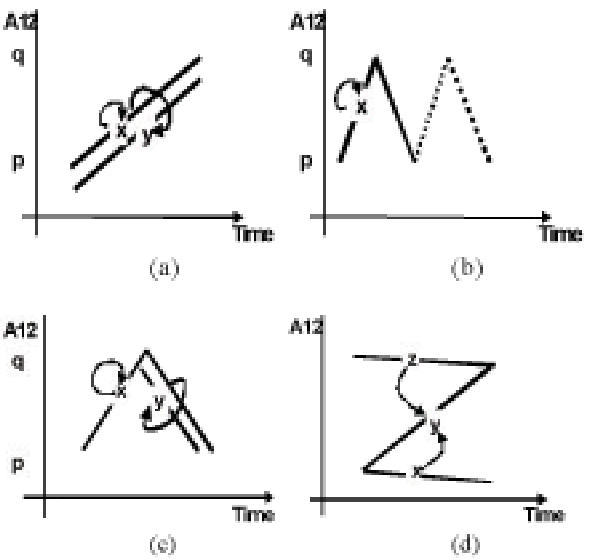

Fig.1. Pictorial interpretation of verb concepts: (a) 'carry', (b) 'shuttle', (c) 'fetch' and (d) 'hand/receive'.

Such locus formulas that correspond with natural event concepts are called 'Event Patterns' and about 40 kinds have been found concerning the attribute 'Physical Location (A_{12})', for example, *move, stay, start, stop, meet, separate, carry, return*, etc.

Employing TLCs, tempo-logical relationships between miscellaneous event concepts can be formulated without explicit indication of time intervals. For example, an event 'fetch(x,y)' is necessarily *finished by* an event 'carry(x,y)' as indicated by the underline at (5). This fact can be formulated as (8), where '$\supset_{-4}$' is the '*implication* ($\supset$)' furnished with the temporal relation '*finished-by* (τ_{-4})'. This kind of formula is not an axiom but a theorem deducible from the definitions of event concepts in the deductive system intended here.

$$(\forall x,y)(\mathrm{fetch}(x,y)\supset_{-4}\mathrm{carry}(x,y)) \quad (8)$$

B .Matter Concepts

A matter, usually referred to by a noun in natural language, is to be conceptualized as a conjunction of the mental images of itself and its relations with others that in turn are to be reduced to certain loci in attribute spaces. In the formal system, a matter concept 'ψ' is defined in such a context as (9), where 'ψ^+' and 'ψ^{++}' are to represent the conceptual images of itself and its relations with others, respectively, and in turn to be reduced to atomic locus formulas of all the attributes.

$$(\lambda z)\psi(z)\Leftrightarrow(\lambda z)\psi^+(z)\wedge\psi^{++}(z) \quad (9)$$

Whereas ψ(z) must be a total description of all the attributes, for simplicity here is to be given only its important part with the symbol '%' representing its abbreviated part. The part $\psi^+(z)$ is given as a combination of atomic locus formulas for the Attribute Carrier 'z' without any other specific matter involved unlike the other part $\psi^{++}(z)$. For example, the matter called 'ice' can be conceptualized as (10). This formula reads that ice is always equal to or less than 0°C cold, is always of no vitality and melts into water (or is something from that H_2O) changes into water)'. In turn, the matter 'snow' can be conceptualized as (11), reading 'Snow is powdered ice attracted from the sky by the earth'. The attributes 'A_{28}', 'A_{39}' and 'A_{41}' refer to 'Temperature', 'Vitality' and 'Quality', respectively. The special symbol '_', defined by (12), is a variable bound by an existential quantifier but does not refer to any specific matter or so in the context while '*' and 'ϕ' represent 'always' and 'no value (or matter)', respectively, defined by (13) and (14).

$$(\lambda x)\mathrm{ice}(y)\Leftrightarrow(\lambda x)\mathrm{ice}^+(z)\wedge\mathrm{ice}^{++}(x)$$
$$(\lambda x)\mathrm{ice}^+(x)\Leftrightarrow(\lambda x)((\exists p,q)L(\_,x,p,q,A_{28},G_t,\_) \wedge p\leq 0^\circ C\wedge q\leq 0^\circ C)^*\wedge L^*(\phi,x,\phi,\phi,A_{39},G_t,\phi)\wedge\%$$
$$(\lambda x)\mathrm{ice}^{++}(x)\Leftrightarrow(\exists z,x_1)L(\_,z,x,x_1,A_{41},G_t,\_) \wedge\mathrm{water}(x_1)\,(\wedge H_2O(z))\wedge\% \quad (10)$$

$$(\lambda x)\mathrm{snow}(x) \Leftrightarrow(\lambda x)(\exists x_1)((L(\_,x,x_1,x_1,A_{41},G_t,\_)\Pi L(\mathrm{Earth},x,\mathrm{Sky},\mathrm{Earth},A_{12},G_t,\_)) \wedge\mathrm{powder}(x_1)\wedge\mathrm{ice}(x_1)\wedge\% \quad (11)$$

$$L(\ldots,\omega_i,\_,\omega_j,\ldots)\Leftrightarrow(\exists\omega)L(\ldots,\omega_i,\omega,\omega_j,\ldots) \quad (12)$$

$$\chi^{*}\Leftrightarrow(\forall[p,q])\ \chi\ \Pi\ \varepsilon([p,q]) \quad (13)$$

$$L(\ldots,\omega_i,\phi,\omega_j,\ldots)\Leftrightarrow\sim(\exists p)L(\ldots,\omega_i,p,\omega_j,\ldots) \quad (14)$$

As easily understood, matter concepts include miscellaneous spatiotemporal relations (i.e. events) among matters usually referred by verbs, prepositions or so. Therefore, a matter concept is usually much more complicated than an event concept in definition. By the way, a noun originated or derived from a verb is to be conceptualized so as to include the verb concept. For example, the concept of 'conveyance' is to be introduced as (15).

$$\begin{aligned}&(\lambda z)\mathrm{conveyance}(z)\\ \Leftrightarrow&(\lambda z)\mathrm{conveyance}^{+}(z)\wedge\mathrm{conveyance}^{++}(z)\\ &(\lambda z)\ \mathrm{conveyance}^{+}(z)\\ \Leftrightarrow&(\lambda z)L^{*}(\phi,z,\phi,\phi,A_{32},G_t,\phi)\wedge\%\\ &(\lambda z)\mathrm{conveyance}^{++}(z)\Leftrightarrow(\lambda z)(\exists x,y)\\ &L(\_,\{x,y\},z,z,A_{01},G_t,\_)\Pi\mathrm{convey}(x,y)\end{aligned} \quad (15)$$

It is well known that people perceive more than reality, for example, 'Gestalt' so called in psychology. A psychological matter here is not a real matter but a product of human mental functions, including Gestalt and abstract matters such as 'society', 'information', etc. in a broad sense.

For example, Fig.2 concerns the perception of the formation of multiple objects, where FAO (Focus of Attention of Observer) runs along an imaginary object so called '*Imaginary Space Region (ISR)*'. This spatial event can be verbalized as S3 using the preposition 'between' and formulated as (16) or (17), corresponding also to such concepts as 'row', 'line-up', etc. Employing ISRs and the 9-intersection model [7] all the topological relations between two objects can be formulated in such expressions as (18) or (19) for S4, and (20) for S5, where '*In*', '*Cont*' and '*Dis*' are the values 'inside', 'contains' and 'disjoint' of the attribute 'Topology (A_{44})' with the standard '9-intersection model (K_{9IM})', respectively. Practically, these topological values are given as 3×3 matrices with each element equal to 0 or 1 and therefore, for example, '*In*' and '*Cont*' are transpositional matrices each other.

(S3) The square is between the triangle and the circle.

$$\begin{aligned}&(\exists x_1,x_2,x_3,y,p,q)(L(\_,y,x_1,x_2,A_{12},G_s,\_)\Pi\\ &L(\_,y,p,p,A_{13},G_s,\_))\bullet(L(\_,y,x_2,x_3,A_{12},G_s,\_)\Pi\\ &L(\_,y,q,q,A_{13},G_s,\_))\wedge ISR(y)\wedge p=q\ \wedge\mathrm{triangle}(\\ &x_1)\wedge\mathrm{square}(x_2)\wedge\mathrm{circle}(x_3)\end{aligned} \quad (16)$$

$$\begin{aligned}&(\exists x_1,x_2,x_3,y,p)(L(\_,y,x_1,x_2,A_{12},G_s,\_)\bullet\\ &L(\_,y,x_2,x_3,A_{12},G_s,\_))\Pi L(\_,y,p,p,A_{13},G_s,\_)\wedge\\ &ISR(y)\wedge\mathrm{triangle}(x_1)\wedge\mathrm{square}(x_2)\wedge\mathrm{circle}(x_3)\end{aligned} \quad (17)$$

(S4) Tom is in the room.

$$\begin{aligned}&(\exists x,y)L(Tom,x,y,Tom,A_{12},G_s,\_)\Pi\\ &L(Tom,x,In,In,A_{44},G_t,K_{9IM})\\ &\wedge ISR(x)\wedge\mathrm{room}(y)\end{aligned} \quad (18)$$

$$\begin{aligned}&(\exists x,y)L(Tom,x,Tom,y,A_{12},G_s,\_)\Pi\\ &L(Tom,x,Cont,Cont,A_{44},G_t,K_{9IM})\\ &\wedge ISR(x)\wedge\mathrm{room}(y)\end{aligned} \quad (19)$$

(S5) Tom exits the room.

$$\begin{aligned}&(\exists x,y,p,q)L(Tom,Tom,p,q,A_{12},G_t,\_)\Pi\\ &L(Tom,x,y,Tom,A_{12},G_s,\_)\Pi\\ &L(Tom,x,In,Dis,A_{44},Gt,K_{9IM})\wedge ISR(x)\wedge\\ &\mathrm{room}(y)\wedge p\neq q\end{aligned} \quad (20)$$

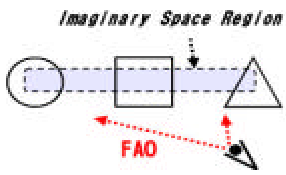

Fig.2. Row as spatial event.

IV. CONCLUSION

The fundamentals of the formal system were presented based on MIDST. The expressive power of the formal language $\boldsymbol{L_{md}}$ was demonstrated with linguistic or pictorial manifestations throughout this paper. Its most remarkable point in comparison with other KRLs resides in that it can provide natural concepts such as *carry*, *snow*, etc. with precise semantic definitions that are normalized by atomic locus formulas and visualized as loci in attribute spaces in both temporal and spatial extents (i.e. temporal and spatial events), which leads to good computability and intuitive readability of $\boldsymbol{L_{md}}$ expressions.

To be continued to another paper of ours [8] for this session.

REFERENCES

[1] Yokota M (2008), In Choi B (Ed.), Humanoid Robots, I-Tech Education and Publishing, 333-362
[2] Yokota M & Capi G (2005), Cross-media Operations between Text and Picture Based on Mental Image Directed Semantic Theory. WSEAS Trans. on Information Science and Applications, 10-2:1541-1550
[3] Dorr B & Bonnie J (1997), Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation, *Machine Translation*, Vol.12, No.4, pp.271-322
[4] Lenat D (1996), CyC: A Large-Scale Investment in Knowledge Infrastructure, Communications of the ACM.
[5] Allen J F (1984), Towards a general theory of action and time. Artificial Intelligence, 23-2: 123-154
[6] Shepard R & Metzler J (1971), Mental rotation of three dimensional objects. Science, 171(972):701-3
[7] Egenhofer M (1991), Point-set topological spatial relations. Geographical Information Systems, 5-2, 161-174
[8] Yokota M et al (2010), Towards natural intelligence modeling as a formal system based on Mental Image Directed Semantic Theory (Part 2), Proc. of AROB2010, Beppu, Oita, Feb. 2010.