PROCEEDINGS OF THE INTERNATIONAL CONFERENCE
ON FIFTH GENERATION COMPUTER SYSTEMS 1992,
edited by ICOT. © ICOT, 1992

Situated Inference of Temporal Information

Satoshi Tojo Hideki Yasukawa†

Mitsubishi Research Institute, Inc.
8-1, Shimomeguro 1, Meguro-ku, Tokyo 153, JAPAN
(phone) +81-3-3536-5813 (e-mail) tojo@mri.co.jp

Institute of New Generation Computer Technology (ICOT)†
Mita Kokusai Bldg. 21F
4-28 Mita 1, Minato-ku, Tokyo 108, JAPAN
(phone) +81-3-3456-3194 (e-mail) yasukawa@icot.or.jp

Abstract

Representations of natural language, describing the same state of affairs may differ between speakers because of their different viewpoints. In this paper we propose the concept of perspectives that are applied to situations, to explain this variety of the representation of an infon, with regard to time. We define the perspective by the relation theory of meaning, namely the relative locations in mind between the described situation, the utterance situation, and the infon. Our aim is to model a situated inference system that infers temporal features of a sentence from the partial temporal information each lexical item carries. For this purpose, we apply our notion of perspectives to actual tense and aspects that are used as lexical temporal features, and in addition we inspect the validity of our formalization for verbs. We show the inference system in the logic programming paradigm, and introduce an ambiguity solver for Japanese *-teiru* that may have multiple meanings, as an experiment of our framework.

1 Introduction

The most common way to represent time is to assume that it is a one dimensional line which extends both to the eternal past and to the eternal future, and a point called 'now' which moves along the line at a fixed speed. The semantics of time in natural languages, so often associated with this physical time parameter 't', has been dealt with. However, the introduction of parameter 't' seems too strong for natural languages, contrary to the case of physical equations. Actually, we cannot always map the temporal property of verbs or temporal anaphora on the time axis correctly.

In opposition to this view of the time, we have been forced to loosen the strongest topology of physical time in some ways. One of the most famous works to represent so-called 'coarse' time is the interval-based theory by Allen [Allen 84]. Kamp [Kamp 79] proposed event calculus where he claimed that 'an instant' was relatively defined by all the known events in his DRT. Our work is to extend this temporal relativity. We will mainly pay attention to the temporal structures of tense, aspects, and verbs within the framework.

From the viewpoint of the history of situation theory, the notion of a spatio-temporal location, or simply a location, was proposed to represent the four-dimensional concept of time and place. In the early stages of situation theory, situations and spatio-temporal locations were distinguished [Barwise 83] as: *In s, at l σ holds.* However the consideration of spatio-temporal location seems to have been rather neglected since then, and we have only found Cooper's work [Cooper 85] [Cooper 86] to give a significant interpretation of locations for time semantics.

The authors have worked for the formalization of a temporal location as a meaning carrier of temporal information [Tojo 90]. In this paper, we aim to show a paradigm of an inference system that merges temporal information carried by each lexical item and resolves any temporal ambiguity that a word may have. First we review the role of the temporal location following Cooper's work. Our position is to regard temporal locations of infons and situations as mental locations. We define a temporal perspective toward a situation, which decides how an infon is verbalized, in terms of relative locations of situations and infons. In the following section, we will give accounts for several important temporal features of tense and aspects by the perspectives, not only to define the basic information for the intended situated inference but also to see the validity of our formalization. In the following section, we discuss the computation system that infers temporal features of a natural language sentence. We have implemented an experimentation system of the ambiguity solver in Japanese *-teiru* with a knowledge representation language $\mathcal{QUIXOTE}$, developed at ICOT (Institute of New Generation Computer Technology, Japan).

2 Situations with perspectives

The temporal information in our mind seems preserved in a quite abstract way, and temporal span or duration are relative to events in the mind. In this section, we will discuss the structure of those subjective views for time, and for the real situation.

2.1 Real situations and perspectives

We often write down an infon in the following way:

$$\ll relation, parameters \gg$$

However none have been concerned with those *labels* for the *relation* in an infon. For example, should we admit such *relations* as those contains tense and aspect? If so, and if the following supporting relations are valid:

$$s \models \ll swim, john \gg$$
$$s' \models \ll swam, john \gg$$
$$s'' \models \ll is\text{-}swimming, john \gg$$

then how should we describe the relation in s, s', and s''? Are we making different expressions for the same real situation?

We hypothesize a virtual physical world, or in other words an ontological world which was originally proposed as a real situation [Barwise83]. According to the notion of real situations, we can assume that there is proto-lexicon in the world though there are many different ways to verbalize them. We may call those infons that are not yet verbalized *proto-infons*.[1] We can regard proto-infons as the genotype of infons; to describe a proto-infon to make the phenotype one is to give *rel* and *roles* in natural language with a certain viewpoint. We propose an idea of perspective which gives this notion of view next.

In the scheme of individuation, Barwise regarded all linguistic labels as being encoded in the situation itself already [Barwise89]. We will not discuss the adequacy of this idea in this paper, however it gives us a way to formalize situations with perpsectives as follows. In order to state a formula of the form $s \models \sigma$, we are required to assume a certain observer who has cut out s as a part of the world and has paid attention to information σ, so that the formula must already contain someone's view or perspective. In that meaning, in S or in σ, the basic lexicon must be included as linguistic labels. For example, if the observer is a Japanese, Japanese language labels should be used to describe the information. From this point of view, we assume that in the formula of the support relation between a situation and an infon someone's perspective already exist.

$$s \models \sigma \Leftrightarrow \mathcal{P}(s' \models \sigma')$$

It is an open question whether we can strip off all the external perspectives from a support relation, as below:

$$\mathcal{P}_1(\mathcal{P}_2(\cdots \mathcal{P}_n(s \models \sigma) \cdots))$$

Even if we can, this must not be the only way to choose a sequence of $\mathcal{P}_i$'s.

2.2 Temporal perspective

We concern ourselves with the temporal part of the spatio-temporal location of the situation theory here. The natural way to do this is to assume that there is a support relation:

$$s \models \ll \cdots \gg$$

that is already verbalized even though the *relation* inside the infon do not have tense nor aspects. Our formalization is as follows. There is a *perspective* $\mathcal{P}$ for a support relation that adds tense and aspect.

[1] This is a reinterpretation of the concept of *information* in the real situation.

$$\mathcal{P}(s \models \ll rel, \cdots \gg)$$
$$\Downarrow$$
$$\mathcal{P}(s) \models_t \ll rel\_with\_tense\_aspect, \cdots \gg$$

Here, we assumed that the perspective is decomposable to both sides of the supporting relation, the meaning of which is assumed to be independent of each perspective though we add the subscript 't' to represent if $\mathcal{P}$ is temporal. We may omit the subscript hereafter to avoid confusion.

The next work we need to do is to define the structure of a $\mathcal{P}$.

2.3 Relation theory of meaning with regard to time

Suppose that the mental descripton of the temporal length, or the duration, of an infon can be written as $\|\sigma\|_t$. In the same way, we can assume the temporal size of a situation such as $\|s\|_t$, if we use the situation as some time-space expansion. In this case, which is the temporal location l_t, $\|s\|_t$ or $\|\sigma\|_t$? And also how should we interpret the supporting relation with regard to time? One possibility is the inclusion of intervals:

$$s \models \sigma \quad \Leftrightarrow \quad \|s\|_t \supseteq \|\sigma\|_t$$

while other people may say

$$s \models \sigma \quad \Leftrightarrow \quad \|s\|_t \subseteq \|\sigma\|_t$$

is more felicitous. Actually the authors consider that the plausibility between '$\supseteq$' and '$\subseteq$' depends on the temporal feature of the *rel*-ation of an infon[2].

Although we cannot fix the size and the location of information in the real time scale, we assume we can map them in a relative way with other events inside our minds. In this paper, we do not use the notion of temporal locations l_t on the physical time axis. Instead, we will consider mentally decscribed $\|s\|_t$ and $\|\sigma\|_t$. s was a part of the world cut off by a perspective of a certain observer. In this meaning, we may say that $\|s\|_t$ is the temporal area the observer is paying attention to so that we may name it as the *field of view* of the perspective.

> **Field of view:** To which time part of the event the observer pays attention, that is a mental location of the described situation $\|s\|_t$.

We call $\|\sigma\|_t$ an *in-progress state* [Parsons 90] of σ.

> **In-progress state:** the mental time of the duration of σ, namely from the beginning of σ to the finishing point: $\|\sigma\|_t$.

We need another component for our temporal perspective, that decides tense. Tense should be decided in accordance with the relative position between the described situation and the utterance situation (that offers 'now'), in terms of the 'relation theory of meaning', that is, a natural language sentence 'φ' is interpreted as the relation between the utterance situation u and the described situation e, denoted by $u[\![\varphi]\!]e$.[Barwise 83]. According to this theory, we can say that the standpoint of view is the mental location of the utterance situation.

> **Standpoint of view:** From which time point the observer sees the event, that is the mental location of the utterance situation $\|u\|_t$.

We will discuss the temporal features with regard to the above three parameters of $\|\sigma\|_t$, $\|s\|_t$, and $\|u\|_t$[3], as in fig. 1.

We characterize the notion of a perspective as constraints attached to the described situation, each of which is the temporal relation with the utterance situation or with the information as in Fig. 2 (in that figure, the relations are denoted by '$\sim$' or '$\wr$'), where the verbalized information can be identified with the natural language expression.

2.4 Set-theoretical foundation

The real or physical time space has the strongest topology where any two time points can be separated[4] and all the points are totally ordered; in addition, it is a metric space. However our mental recognition for time is much more vague. We regarded that the temporal recognition of one event is the relative position and the relative length of $\|\sigma\|_t$, $\|s\|_t$, and $\|u\|_t$. Therefore, the mental time space we are considering has a very weak topology.

[2] An instantaneous change of state such as '*understand*' seems to require the temporal vicinity ($\|s\|_t \supset \|understand\|_t$) while *stative* verbs should be valid anytime in s ($\|s\|_t \subset \|is\text{-}running\|_t$).

[3] Reichenbach [Dowty 79] distinguished three kinds of temporal information *time of action*, *time of reference*, and *time of speech*. $\|\sigma\|_t$ and $\|u\|_t$ correspond to *time of action* and *time of speech* respectively, and the $\|\sigma\|_t$ is *time of reference*, though we extend the notions as intervals instead of points.

[4] In the meaning of *Axiom of Separation* in topological spaces.

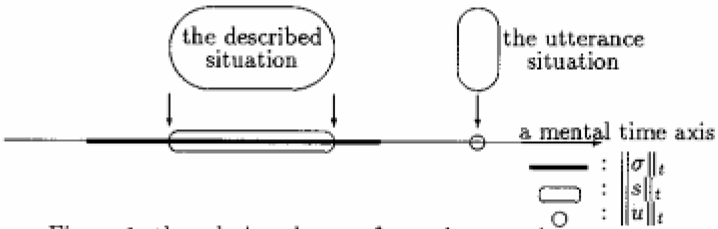

Figure 1: the relation theory of meaning wrt time

$$\begin{array}{ccc} s \models \sigma & & u[\![\phi]\!]s \\ \Downarrow & & \Downarrow \\ \mathcal{P}(s) \models_t \ll \cdots \gg & \equiv & s\begin{bmatrix} \sim & \|u\|_t \\ & \wr \\ \sim & \|\sigma\|_t \end{bmatrix} \models_t \phi \end{array}$$

Figure 2: The denotation of the relation theory of meaning

whose open sets are these $\|s\|_t$'s, $\|u\|_t$'s,[5] and $\|\sigma\|_t$'s. We may call them intervals, sets of points, or even points, however they all should be reinterpreted in terms of open sets in the weak topological space.

We will not mention this topological notion hereafter though we may use some set theoretical notations such as '$\subsetneq$, $\cup$, $\cap$, ϕ'. We adopt the normal subset relation '$\subseteq$' between $\|\sigma\|_t$ and also $\|s\|_t$ to represent the temporal partiality. In addition, we equip ourselves with the normal temporal order, denoting:

$$T_1 \prec_t T_2,$$

for time intervals T_i's, iff all the time points in T_1 chronologically precede those in T_2. Where an instant may be shared between two sets, we use '$\preceq_t$'. We may drop the subscript 't' hereafter as far as there is no confusion.

3 Tense and aspects as temporal perspectives

We have claimed that the temporal perspective depends upon the individual view of the state of affairs. However, as we use natural languages to communicate with others, we may hopefully receive stereotypical views of things. These stereotypes must be tense and aspects. This implies that we can define tense and aspects by situation types that are independent of each σ. In this section, we will introduce inference rules that infers those stereotypes, that are used as ground rules of the system later.

[5] We can assume that a $\|u\|_t$ is a set that does not contain other sets inside in the topological view, if we were to persist with the interval logic.

3.1 Situated inference rules

We defined the role of a perspective $\mathcal{P}$ as follows: if an infon σ is supported by a situation s and if we add a view $\mathcal{P}$, then σ is transformed to another expression σ'.

$$\mathcal{P}(s) \models \sigma' \;\Leftarrow\; s \models \sigma.$$

or:

$$\frac{s \models \sigma}{\mathcal{P}(s) \models \sigma'}.$$

We defined the notion of perspective by the set theoretical relation between $\|s\|$, $\|u\|$, and $\|\sigma\|$. We denote $s[X]$ for the described situation with a perspective, where the contents of $[X]$ is the temporal consraint that observes the following convention.

1) $s[\prec u]$, $s[\subset \sigma]$, and $s[\supset \sigma]$ are the described situations where $\|s\|_t \prec \|u\|_t$, $\|s\|_t \subset \|\sigma\|_t$, and $\|s\|_t \supset \|\sigma\|_t$ hold, respectively.

2) $s[X, Y]$ means $s[X] \wedge s[Y]$. For example, $s[\subset \sigma, \prec u]$ means s such that $\|s\|_t \subset \|\sigma\|_t$ and $\|s\|_t \prec \|u\|_t$.

3) the relation between σ and u is written as $s[\sigma \prec u]$.

3.2 Tense

We define the tense as the problem of chronological order between $\|s\|_t$ and $\|u\|_t$, independent of $\|\sigma\|_t$.

Let us consider the sentence: 'Ken was running in the park.' There must be a duration of time in which 'Ken runs', which is $\|\sigma\|_t$ where $\sigma = \ll$ run, ken $\gg$ in our definition. However the speaker of the sentence does not necessarily know whether Ken ceased to run at the point this utterance was made. Therefore, whether the verb is past or present does not depend on whether the deed finished. Instead, the only required condition for past is that the area the speaker paid attention to ($= \|s\|_t$) precedes the point at which this utterance was made ($= \|u\|_t$). Fig. 3 depicts this case; in that figure, the deed ($= \|\sigma\|_t$) may or may not finish at the utterance point ($= \|u\|_t$) however both situations can support the same infon: $\ll$ was-running, ken $\gg$.

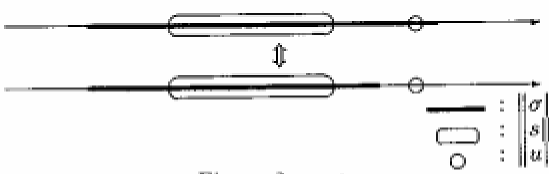

Figure 3: past

The past view for affairs is represented by $\|s\|_t \preceq_t \|u\|_t$. On the contrary, the present tense is represented by $\|u\|_t \subset \|s\|_t$.

We formalize the feature of past and present as follows:

$$s[\preceq u] \models \ll past, \sigma \gg \;\Leftarrow\; s \models \sigma \tag{1}$$

$$s[\supset u] \models \ll present, \sigma \gg \;\Leftarrow\; s \models \sigma \tag{2}$$

3.3 Aspects

The study on how we see the temporal features of events has been done in linguistics and we know the variety of aspects. Among the taxonomy, it seems rather proper to pay attention to the following two important features [Comrie 76] though other features may be omitted, because those distinctions can be found in any language. One is:

- the deed is recognized as a duration of time /a point on the time scale of time (durative/non-durative)

and the other is:

- the finishing of the deed is recognized /not recognized (past or perfective/imperfective)

In summary, perfective/imperfective is distinguished, dependent on if the finishing point of $\|\sigma\|_t$ is before $\|u\|_t$. Durative/non-durative depends on if the field of view wraps up $\|\sigma\|_t$ or not.

Durative The most important feature of our method is the distinction of *durative* and *non-durative* aspects by the relation between $\|\sigma\|_t$ and $\|s\|_t$. We interpret the *progressive* feature in terms of our formalization as follows.

The state of progressive is to see the deed as a durative, and seeing a part of the inside of the deed. Namely the speaker does not pay attention to when the deed began, nor to when it will finish. The state is shown in Fig. 4.

Figure 4: progressive

On the contrary, let us consider the case we do not pay attention to the inside of the deed. When $\|s\|_t$ contains the whole time of the deed, we can conclude that the observer recognized the event as a non-durative one, in which case the event was regarded as a point with no breadth on the mental time axis (Fig. 5), if there is no interaction with other events.[6]

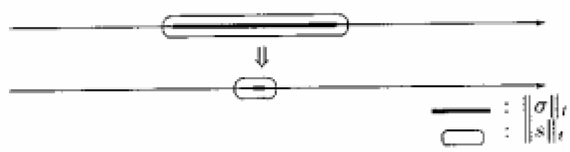

Figure 5: compression to non-durative

A lexical item for the durative view becomes the following:

$$s[\subset \sigma] \models \ll progressive, \sigma \gg \;\Leftarrow\; s \models \sigma \tag{3}$$

[6] From the topological point of view, a set which does not contain other sets inside nor has intersections with other sets is identified with one of the smallest sets of the space, viz. a point.

Perfect and time of reference As Reichenbach claimed in [Dowty 79], present perfect in English refers to the current state. We have shown that present is represented by the relation that $\|s\|_t$ includes $\|u\|_t$. Therefore, to satisfy this issue, our $\|\sigma\|_t$ must precede the $\|u\|_t$ to represent the *present perfect*. In Fig. 6, we have shown the perspective for the present perfect.

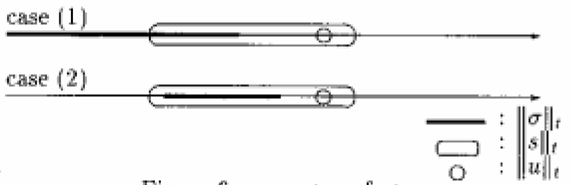

Figure 6: present perfect

In Fig. 6, case (1) shows that perfect is interpreted as a terminative aspect although case (2) shows that perfect is read as an experience.

The perfect view becomes the following:

$$s[\sigma \preceq u] \models \ll perfect, \sigma \gg \;\Leftarrow\; s \models \sigma \qquad (4)$$

4 Inference of temporal information

4.1 Situated inference

In terms of situated inference, we expect the inference with the following rules:

$$S_0 \models \sigma_0 \;\Leftarrow\; S_1 \models \sigma_1, S_2 \models \sigma_2, S_3 \models \sigma_3, \cdots.$$

This sample rule can be interpreted as: if S_1 supports σ_1, S_2 supports σ_2, and so on, then we can infer that S_0 supports σ_0.

This kind of rule can be read as backward chaining from the head, just as with the inference rules of Prolog. We would like to devise a system that computes temporal information, asking questions that corresponds to the head of a rule, and accumulating the temporal information from its body. For example, assume $X, Y, \cdots$ are variables for temporal information.

$$s[X, Y, \cdots] \models \sigma_0 \;\Leftarrow\; s[X] \models \sigma_1, s[Y] \models \sigma_2, \cdots.$$

where all the basic, lexical information such as (1), (2), (3), (4), and so on, defined in the previous section, are considered to be the basic rules. The accumulation of temporal information must not be a mere addition; in our case, it must be the merger of different topologies in $\|s\|_t$'s, $\|\sigma\|_t$'s, and $\|u\|_t$'s. The computation, therefore, must be done in the dual mode; one mode is conventional unification and backward chaining, and the other is the merger of $s[X]$ with consistency. This is the reason why we chose the $\mathcal{QUIXOTE}$ language with its concept of modules (situations, in our case) inside of which features can be defined. We will mention the specification later.

We have developed an ambiguity solver for Japanese *-teiru* that can have three different kinds of meaning that depend on the context.

4.2 The problem of Japanese '*-teiru*'

Prior to introducing the ambiguity solver we have developed, we need to give a short tip on Japanese grammar and the problem we tackled.

In the Japanese language, auxiliary verbs are agglutinated at the tail of the syntactic main verb that is the original meaning carrier. In order to compose a progressive sentence, we need to add an auxiliary verb '*-tei-*', and after that we are required to affix a tense marker. We summarize them below for a Japanese verb '*kiru* (to wear)'.

lexical entry	part of speech	meaning
ki-	verb	to wear
-tei-	aux. verb	be -ing*
-ru	affix	present
-ta	affix	past or perfect
ima	adv.	now
zutto	adv.	all the time
san-nen-mae-ni	adv. phrase	3 years ago
ki-tei-ta	verb phrase	was wearing*

The problem lies in the places marked with * in the table above. The meaning marked by * is not the sole meaning of *-tei*; actually we can interpret the auxiliary verb in three different ways, depending on the context.

We show sample sentences below[7].

[7] This example was shown by the members of the JPSG working group in ICOT

ima ki-tei-ru
(be putting on currently)
zutto ki-tei-ru
(wear all the time)
san-nen-mae-ni ki-tei-ru
(have worn three years ago)

We will build our ambiguity solver, focusing on the area of a deed in JPSG framework that corresponds to our $\|s\|_t$. The partial information that each lexical item carries, as defined in the previous section, is utilized, according to the Japanese lexicon table above. Namely we use the inference rules of past (1) /present (2) and perfect (4), for Japanese '*-ru*/*-ta*'. We use the inference rule of durative (3) for Japanese '*-tei-*'.

The ambiguity of '*-tei-*' is solved as in Fig. 7 where '&' is the merger of information: $X_2 = [\sigma \prec u]$ is incompatible with $[s \subset \sigma]$ in X_5 and $X_6 = [u \subset s]$, so that the value of X_5 necessarily becomes $[\sigma \subset u]$, and this gives '*-tei*' the interpretation of the resultant state.

4.3 Implementation

This section shows an implementation of the treatment of temporal information discussed in this paper. The program is written in the knowledge representation language $\mathcal{QUIXOTE}$[Yasukawa 90],[Yasukawa 92].

4.3.1 $\mathcal{QUIXOTE}$

Terms in $\mathcal{QUIXOTE}$ are extended terms on an order-sorted signature called *object terms*, and written in general as:

$$o[l_1 = o_1, l_2 = o_2 \dots]$$

where o is an atom called **basic object**, l_1, l_2 are atoms called **labels**, and o_1, o_2 possibly be object terms. The domain of atoms (BO) is ordered and constitutes a lattice $(BO, \preceq, \top, \bot)$.

The *subsumption relation*($\sqsubseteq$) is a binary relation over the domain of object terms, and corresponds to so-called isa-relation. Intuitively, $o_1 \sqsubseteq o_2$ (we say o_2 subsumes o_1) holds if o_1 has more arcs than o_2 and the value of a node of o_2 is larger than the value of the corresponding node of o_1 with respect to $\preceq$-ordering. In $\mathcal{QUIXOTE}$, subsumption constraints can be used to specify an object or the relation among objects.

A rule of $\mathcal{QUIXOTE}$ is a prolog-like clause of the form:

$$m :: \sigma <= m_1 : \tau_1, m_2 : \tau_2, \dots, m_n : \tau_n \| C.$$

where $m, m_1, \dots, m_n$ are special extended terms called **module identifiers**, and $\sigma, \tau_1, \dots, \tau_n$ are extended terms, and C is a set of constraints.

4.3.2 Representation in $\mathcal{QUIXOTE}$

There are several points to be explained, that is, how the notions introduced in the preceding sections are represented.

Object terms are used to represent situations, infons, perspective, and so forth.

First, verbalized infons are represented by object terms of the following form:

$$inf[v\_rel = [rel = R, cls = CLS, per = P], \quad args = Args].$$

The CLS takes the symbols $act_1, act_2, act_3, \dots$ as its value which indicates the classifications of verbs on what state, that is, in-progress, target, resultant, each verb can introduce. Here's a list of the classification and the states introduced:

$$act_1 \Rightarrow ip, tar, res$$
$$act_2 \Rightarrow ip, res$$
$$act_3 \Rightarrow tar, res.$$

The relationship among these three classes is given by the following subsumption definition.

$$act_1 \sqsupseteq act_2 ::$$
$$act_1 \sqsupseteq act_3 ::$$

Thus, the verb "*ki-*" can introduce all the three states, while the verbs like "*hashi-*" could not introduce *tar*-state.

For example, the verbalized infon corresponding to the sentence "John is running" is represented as follows:

$$inf[v\_rel = [rel = run, cls = act_1, pers = P], \quad args = [agt = john]].$$

where P is the temporal perspective whose field of view is in-progress and point of view is the *present*. A perspective is also represented by a pair of two object terms as follows:

$$([fov = Fov], [pov = Pov]).$$

where $Fov \in \{ip, tar, res\}$ represents the field of view, and $Pov \in \{pres, past\}$ represents the point of view.

$$\dfrac{\dfrac{s \models \ll\gg}{s[X_2] \models \ll \text{zutto} \gg} \quad \dfrac{\dfrac{s \models \ll\gg}{s[X_4] \models \ll \text{ki} \gg} \quad \dfrac{s \models \ll\gg}{s[X_5] \models \ll \text{tei} \gg} \quad \dfrac{s \models \ll\gg}{s[X_6] \models \ll \text{ru} \gg}}{s[X_3] \models \ll \text{ki\_tei\_ru} \gg}}{s[X_1] \models \ll \text{zutto\_ki\_tei\_ru} \gg}$$

$$\begin{cases} X_1 = X_2 \& X_3 \\ X_2 = \sigma \prec u \\ X_3 = X_4 \& X_5 \& X_6 \\ X_4 = [\,] \\ X_5 = \sigma \prec s \text{ or } s \subset \sigma \\ X_6 = u \subset s \end{cases}$$

Figure 7: the inference tree

The ip and res correspond to in-progress state, target state, and resultant state, respectively.

Among the situations of several kinds, discourse situations are represented by object terms of the following form:

$$dsit[fov = Fov, pov = Pov, src = U]$$

where U is the object term representing the utterance situation.

Thus the propositional content of the sentence "John is running" is represented by

$$\begin{aligned} &dsit[fov = ip, pov = pres, src = u_1] : \\ &\quad inf[v\_rel = [rel = run, cls = act_2, \\ &\qquad pers = [fov = ip, pov = pres]], \\ &\qquad args = [agt = john]]. \end{aligned}$$

For simplicity, the object term $[rel = run, pers = [fov = ip, pov = pres]]$ is written as $is\_running$.

Next, the lexical entries of Japanese are defined. For example, the Japanese expression "*ki-tei-ru*" consist of three words. The lexical entries are as follows.

$$\begin{aligned} &dict :: v[cls = act_1, rel = put\_on, form = ki] :: \\ &dict :: v[cls = act_2, rel = run, form = hashi] :: \\ &dict :: auxv[asp = state, form = tei] :: \\ &dict :: affix[pov = pres, form = ru] :: \end{aligned}$$

The ambiguity is processed by the mapping from a pair of the class of a verb and the aspect of an auxilially verb:

$$\begin{aligned} (act_1, state) &\rightarrow \{ip, tar, res\}, \\ (act_2, state) &\rightarrow \{ip, tar\}, \\ (act_3, state) &\rightarrow \{tar, res\}. \end{aligned}$$

For example, the expression "*ki-tei-*" has three interpretations, while "*hashi-tei-*" has two interpretations. The target and resultant are the states after an event's having culminated. Thus, it is possible to disambiguate the interpretations if some evidence that the event has culminated or not, for example, the successive utterance of "*mi ni tuke-tei-nai*" (does not wear) makes it clear that the first sentence "*ki-tei-ru*" should be interpreted as having ip as its field of view.

The definition of a tiny interpreter is given in the appendix 5.

A toplevel query corresponds to the definition of the meaning of a sentence in Situation Semantics, and is of the following form:

```
?- mi[u=u1,exp=Exp,e=E,infon=Infon] ||
   {E=dsit[fov=Fov,pov=Pov,src=u1]}.
```

This query says that the meaning of the expression "`Exp`" in an utterance situation "`u1`" is represented by the described (temporal) situation "`E`" and the infon "`Infon`" where the variable "`Fov`" and "`Pov`" represent the temporal perspective.

For example, the following result is given by this interpreter:

```
?- mi[u=u1,exp=[ki,tei,ru],
      e=dsit[fov=Fov,pov=Pov,src=u1],
      infon=Infon].
```

```
Answer:
  Fov = {ip,tar,res}
  Pov = pres
  Infon =
    inf[v_rel=[rel=put_on,cls=act1,pers=P],
        args=_]
  P = [fov=Fov,pov=Pov]
```

This means that the expression "*ki-tei-ru*" has three interpretations depending on which **fov** is applied, because the verb "*ki-*" introduces all the three states[8].

On the contrary, expressions like "*hashi-tei-ru*" and "*waka-tei-ru*" has two interpretations, because those verbs can not introduce all the three states.

```
?- mi[u=u1,exp=[hashi,tei,ru],
      e=dsit[fov=Fov,pov=Pov,src=u1],
      infon=Infon].

Answer:
  Fov = {ip,res}
  Pov = pres
  Infon =
    inf[v_rel=[rel=run,cls=act1,pers=P],
        args=_]
  P = [fov=Fov,pov=Pov]

?- mi[u=u1,exp=[waka,tei,ru],
      e=dsit[fov=Fov,pov=Pov,src=u1],
      infon=Infon].

Answer:
  Fov = {tar,res}
  Pov = pres
  Infon =
    inf[v_rel=[rel=understand,cls=act1,pers=P],
        args=_]
  P = [fov=Fov,pov=Pov]
```

5 Conclusion

We have introduced the idea of temporal perspectives for situations, to explain the variety of language expressions for information in real situations. As a perspective, we assumed a topological relation between the three parameters of the standpoint of view ($\|u\|_t$), the field of view ($\|s\|_t$), and the duration of information ($\|\sigma\|_t$), each of which is the mentally recognized location of the utterance situation, the described situation, and the infon, respectively in terms of the relation theory of meaning.

We have defined tense and several important aspectual distinctions such as perfective /imperfective and durative /non-durative, with the perspective, that should be used as the partial information for the situated inference system. In addition, we tried to define other temporal features of verbs such as telic /atelic and temporal well-/ill-foundedness, to see the validity of our formalization.

Our framework for the situated inference of temporal information is to infer the whole temporal features of phrases or sentences, collecting the partial information that is carried by each lexical item, and to solve the ambiguity partial phrases they may have. We are required to have mechanisms for that system, both Prolog-like backward chaining and maintenance of consistency in modules (situations, in our case), so that we can utilize the knowledge representation language $\mathcal{QUIXOTE}$ in ICOT. We have implemented an inference system to solve the ambiguity of Japanese '*-teiru*', the $\|\sigma\|_t$ of which may refer to different parts of a deed. In that experiment, the problem is solved together with another lexical item which offers the information 'which $\|\sigma\|_t$ coincides with $\|s\|_t$'.

Our inference system is still small, and needs to be developed to cover many other kinds of lexicon and temporal ambiguity. According to this future work, we might be required to reconsider the structure of perspectives. We are still trying to determine other temporal features of verbs, and, as a task in the near future, we are going to try tto define the temporal perspectives of sentence adverbs.

Acknowledgment

Thanks to the members of STS-WG (Working Group for Situation Theory and Semantics) of ICOT for their stimulating discussions and suggestions. Special thanks to Kuniaki Mukai who gave the authors many important suggestions about the treatment of perspectives in Situation Theory, to Kazumasa Yokota who gave authors

[8] In $\mathcal{QUIXOTE}$, the fact $o[l = \{a, b\}]$ is interpreted as the two facts, $o[l = a]$ and $o[l = b]$.

many important comments, and to Koiti Hasida who explained the JPSG theory for verbs to the authors. Special thanks also to Mitsuo Ikeda of ICOT, who has studied the analysis of Japanese *-teiru*.

References

[Barwise89] J. Barwise. *The Situation in Logic.* CSLI Lecture Notes 17, 1989.

[Barwise83] J. Barwise and J. Perry. *Situations and Attitudes.* MIT Press, 1983.

[Comrie76] B. Comrie. *Aspect.* Cambridge University Press, 1976.

[Cooper85] R. Cooper. Aspectual classes in situation semantics. Technical Report CSLI-84-14C, Center for the Study of Language and Information, 1985.

[Dowty79] D. Dowty. *Word Meaning and Montague Grammar.* D.Reidel, 1979.

[JPSG91] JPSG-WG. The minutes of Japanese phrase structure grammar. 1987-91.

[Allen84] J.F.Allen. Towards a general theory of action and time. *Artificial Intelligence*, 1984.

[Kamp79] H. Kamp. *Events, Instants, and Temporal References*, pages 376–417. Springer Verlag, 1979. in Semantics from Different Points of View.

[Cooper86] R.Cooper. Tense and discourse location in situation semantics. *Linguistics and Philosophy*, 9(1):17–36, February 1986.

[Tojo90] S. Tojo. A temporal representaion by a topology between situations. In *Proc. of SICONLP '90.* University of Seoul, 1990.

[Parsons90] T.Parsons. Events in the Semantics of English. MIT press, 1990.

[Yasukawa90] H. Yasukawa and K. Yokota. "Labeled Graphs as Semantics of Objects", In *Proc. SIGDBS and SIGAI of IPSJ*, Oct., 1990.

[Yasukawa92] H. Yasukawa, K. Yokota, H. Tsuda, Objects, Properties, and Modules in *QUIXOTE*, In *Proc. FGCS'92*, Tokyo, June, 1992.

Appendix - The interpreter

```
%% Subsumption Definition
act1 >=  act2;; act1 >= act3;;

%% Lexical Entry
dict :: v[cls=act_1,rel=put_on,form=ki];;
dict :: v[cls=act_2,rel=run,form=hashi];;
dict :: v[cls=act_3,rel=understand,form=waka];;
dict :: auxv[asp=state,form=tei];;
dict :: affix[pov=pres,form=ru];;
dict :: affix[pov=past,form=ru];;

%% Top level
mi[u=U,exp=[],e=D,infon=Infon];;
mi[u=U,exp=[Exp|R],e=D,infon=Infon] <=
    d_cont[exp=Exp,e=D,infon=Infon],
    mi[u=U,exp=R,e=D,infon=Infon];;

%%%% Interpretation Rules
d_cont[exp=Exp,e=dsit[fov=Fov,pov=Pov,src=U],
       infon=inf[v_rel=V_rel,args=Args]]
  <= dict : v[cls=CLS,rel=Rel,form=Exp] ||
     {V_rel=[rel=Rel,cls=CLS,pers=P]};;
d_cont[exp=Exp,e=dsit[fov=Fov,pov=Pov,src=U],
       infon=inf[v_rel=V_rel,args=Args]]
  <= dict : auxv[asp=ASP,form=Exp],
     map[cls=CLS,asp=ASP,fov=Fov] ||
     {V_rel=[rel=_,cls=CLS,pers=P],
      P = [fov=Fov,pov=_]} ;;
d_cont[exp=Exp,e=dsit[fov=Fov,pov=Pov,src=U],
       infon=inf[v_rel=V_rel,args=Args]]
  <= dict : affix[pov=Pov,form=ru] ||
     {V_rel=[rel=_,cls=_,pers=P],
      P = [fov=_,pov=Pov]} ;;

%% Field of view Mapping
map[cls=act1,asp=state,fov={ip,tar,res}] ;;
map[cls=act2,asp=state,fov={ip,res}] ;;
map[cls=act3,asp=state,fov={tar,res}] ;;
```