An Ontological Consideration of Causal Time in Qualitative Reasoning Systems

Yoshinobu Kitamura, Mitsuru Ikeda, and Riichiro Mizoguchi
I.S.I.R., Osaka University, Osaka 567, Japan
{kita,ikeda,miz}@ei.sanken.osaka-u.ac.jp

Abstract

Human recognition of causal relations is based on recognition of time delay between the cause and the effect. Little, however, is known concerning temporal meaning of causal relations generated by qualitative reasoning systems. Aiming at explicit description of time concepts underlying the models and the causal reasoning engines, this article proposes a causal time ontology which defines a set of general time concepts in qualitative models, called *causal time scales*. Each of them associated with a modeling technique represents a temporal granularity and/or an ontological viewpoint. The causal time scales generalize time concepts in some previous causal reasoning systems. They allow us to specify temporal performances of the engines and to identify a general causal reasoning scheme together with sophisticated feedback analysis. Lastly, we show how modeling schemes and reasoning procedures of causal reasoning systems are determined according to the given temporal requirements.

Introduction

Causality plays a crucial role in human understanding of behavior of physical systems. A lot of research has been carried out on qualitative reasoning systems for causal understanding, which can derive causal relations (i.e., causal ordering) from behavior of the target systems, e.g., (de Kleer and Brown, 1984; Iwasaki and Simon, 1994). Human recognition of causal relations is based on recognition of time delay (i.e., time interval) between the cause and the effect. Little, however, is known concerning temporal meaning of causal relations generated by the reasoning systems, that is, how long (or short) the time intervals in the causal relations in the real physical behavior as discussed in (Iwasaki *et al.*, 1995). There are the following two explanations for it. First, there are many modeling techniques and representations, each of which implies several temporal relations among variables in the models. Secondly, such models are interpreted by the reasoning engines on the basis of their own time concepts behind their reasoning procedures. For example, using the same qualitative differential equations[1], QSIM (Kuipers, 1994) and the causal ordering procedure proposed in (Iwasaki and Simon, 1994) generate different causal relations together with different temporal meanings. The reason for this is that the reasoning engines have different time concepts behind the reasoning procedures. As a consequence of the implicit existence of several time concepts, the temporal meaning of generated causal relations is not clear for the users of the reasoning engines.

The goal of this article is to reveal the structure of causal time underlying the qualitative models and the causal reasoning engines. We propose a set of general time concepts in qualitative models, called *causal time scales*. Each causal time scale associated with a modeling technique represents a temporal granularity and/or an ontological viewpoint. In other words, the set of the causal time scales aims to enumerate all possible temporal meanings of the models, that is, an ontology of time in qualitative models for causal ordering. Ontologies are explicit specifications of concepts (Mars, 1995), which can specify assumptions and performance of problem solving systems (Mizoguchi and Ikeda, 1996).

We have identified 13 causal time scales shown in Table 1. They generalize time concepts in the some previous frameworks (de Kleer and Brown, 1984; Iwasaki and Simon, 1994; Kuipers, 1994; Rose and Kramer, 1991).

In terms of causal time scales, we can explicitly specify the following temporal meanings of causal relations or causal reasoning engines.

- Temporal meanings of causal relations
- Temporal performances of the reasoning engines
- A general causal reasoning scheme
- Sophisticated analysis of feedback
- Design processes of the reasoning engines

First, causal relations generated by the reasoning engines can be categorized into one of the causal time scales. For example, causal relations generated by QSIM are categorized into the causal time scale named $Ta3$ which represents the time concept associated with the mathematical integral operation. On the other hand, some of those generated by the causal ordering

[1]Strictly speaking, the causal ordering procedure (Iwasaki and Simon, 1994) needs additional information.

procedure (Iwasaki and Simon, 1994) are categorized into the time scale $Ta2$ which is a finer-grained time concept than $Ta3$.

Secondly, the causal time scales enable us to specify the performances of the reasoning engines with respect to causal ordering, called *causal time resolutions*. In the example mentioned before, the time resolution of the causal ordering procedure is finer than that of QSIM. The time resolutions of other reasoning methods will be shown in Table 2.

Thirdly, we also identified a general and primitive reasoning scheme which can describe essential parts of the conventional reasoning methods. Fourthly, fine-grained time scales enable sophisticated analysis of causality in feedback loop to obtain less ambiguous causal ordering.

Lastly, we can show how modeling schemes and reasoning procedures of causal reasoning systems are determined according to the given temporal requirements. That represents a part of design rationales of the causal reasoning systems. In the second half of this article, we show an example of the design processes based on the causal time scales.

In this article, we do not discuss formal ontology based on axiomatization, aiming at getting on agreement on the content and the terminology. Next, we concentrate on ontological issues. For the details of model representation, reasoning engine and its evaluation, see other articles(Kitamura *et al.*, 1996a; Kitamura *et al.*, 1996b).

A Causal Time Ontology

Theoretical Foundation

In our causal time ontology, behavior over time generated by the reasoning engine is represented in terms of events and links among the events, in similar way in the history model(Forbus, 1984). An event $e \in E$ represents instantaneous changes of qualitative values of parameters and their resultant values at a time point. Changes of quantitative values are assumed to be continuous and differentiable. A new event e_2 is generated by applying an operators $o \in O$ to an old event e_1 according to the model M. A link $l_1 \in L$ between e_1 and e_2 represents a causal relation according to the model M. There is an open time-interval t_1 between e_1 and e_2, corresponding to the causal relation l_1. The roles of operators $o \in O$ are to propagate changes and to generate new events, time intervals and hence partial temporal relations. Note that the symbol 't' always represents not a time point but a time interval in this article. Although events correspond to time points, we concentrate on time intervals in which changes propagate.

The causal time ontology provides categories of such time intervals, called *causal time scales*. A causal time scale represents a concept of time interval for propagation of effect. The notation $\tau(l) = T$ denotes a time interval t of a causal relation l is categorized into a time scale T. We can say that "the causal relation l is represented on the time scale T".

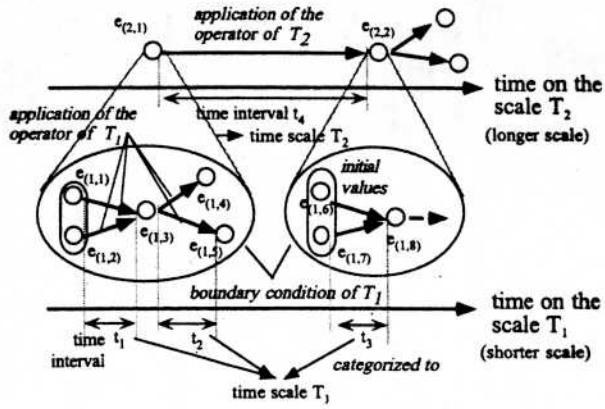

Figure 1: Relation between the two time scales

The ordinal relation $T_1 \prec T_2$ representing a time scale T_1 is shorter (finer-grained) than T_2 is defined as follows;

$$T_1 \prec T_2 \leftrightarrow \forall\, t_1 \in T_1, \forall\, t_2 \in T_2,\ t_1 < t_2$$

In other words, T_1 represents faster events than that T_2 does. This relation is transitive. The relation between T_1 and T_2 where $T_1 \prec T_2$ is shown in Figure 1. When a certain condition becomes true in the reasoning process on a shorter time scale T_1, the reasoning shifts to a neighboring longer time scale T_2. Such a condition is called as a *boundary condition* of T_1 or a *precondition* of T_2. The set of events grouped by the condition $e_{(1,1)}, e_{(1,2)}, ..., e_{(1,5)}$ on T_1 is treated as the instantaneous events $e_{(2,1)}$ on T_2. Then, the *reasoning operator* of T_2 is applied to the event $e_{(2,1)}$. Each time scale has an operator. The resultant values on T_2 can be treated as the initial values on T_1. The same applies to $T_2 \prec T_3$ cases recursively. In summary, a time scale T can be defined by a tuple of three elements, <Pc, Op, Bc>, where these denote a precondition, an operator, a boundary condition, respectively. The elements of T_1 are denoted by T_1:Pc, T_1:Op and T_1:Bc, respectively.

The relations l generated by the reasoning engine do not always make sense from the physical viewpoints. There are such cases where a link l represents an operational order which is not justified by the physical sense. In order to clarify *the physical meaning* of the causal relations, we will discuss two aspects of the physical meaning of each time scale, that is the *interval-meaning* and the *ordinal-meaning*. The former represents a physical justification of existence of the time intervals on the time scale. The latter represents that of order of events on the time scale.

Causal Time Scales

This section defines 13 causal time scales shown in Table 1. The time scales are classified into four categories each of which represents a modeling technique together with particular modeling rationales. The *direct modeling* is to describe models using the mathematical differential equations which directly represent

Table 1: The causal time scales

(a) Direct Modeling
- ca1: changes of parameter values
- *Ta1: Mutual Dependency time scale*
  - ca2: a set of inherently simultaneous equations are satisfied.
- *Ta2: Dependency time scale*
  - ca3: a set of constraints are completely satisfied.
- *Ta3: Integral time scale*
  - Ta3p: Integral-from-equality time scale
  - Ta3i: Integral-to-equality time scale
  - ca4: a set of parameters reaches equilibrium.
- *Ta4: Equilibrium time scale*

(b) Time Constant Modeling
- *Tb1: A Faster Mechanism time scale*
  - cb2: a faster mechanism reaches equilibrium.
- *Tb2: A Slower Mechanism time scale*
  - cb3: a slower mechanism reaches equilibrium.

(c) Component Structure Modeling
- *Tc1: Intra-componnent time scale*
  - cc2: all parameters in a component are satisfied.
- *Tc2: Inter-component time scale*
  - cc3: all parameters in a global structure are satisfied.
- *Tc3: Global time scale*
  - cc4: all parameters in the whole system are satisfied.
- *Tc4: The Whole System time scale*

(d) Modeling of Periods of Interest
- *Td1: Initial Periods time scale*
  - cd1: the first event happens on a time scale.
- *Td2: Intermediate Transitional time scale*
  - cd2: the last event happens on a time scale.
- *Td3: Final Periods time scale*

dynamic behavior over time. The *time constant modeling* is to qualitatively categorize the time constants for modeling of phenomena. The *component structure modeling* is to introduce the concept of "component", aiming at causal relations reflecting the physical structures of the target systems. The *modeling of periods of interest* such as initial responses allows the reasoning engine to neglect changes of no interest. The notation "$Tx\#$" denotes a time scale, where 'x' denotes a category (a,b,c or d) and the number '#' represents ascending order in each category. In Table 1, each condition denoted by a notation "c$x\#$" represents the boundary condition of the time scale listed above and the precondition of that listed below.

(a) Direct Modeling In the direct modeling, temporal characteristics of the phenomena are represented directly by the mathematical aspect of the models. The precondition of the time scale $Ta3$ is that a set of parameters are *completely satisfied* where every parameter in the set has values which satisfy all constraints. When the condition holds, the reasoning engine applies the integral operator and hence generates a new event. The integral operator embodies the qualitative mean value theorem $x_{new} = x_{old} + dx/dt$ (de Kleer and Brown, 1984). The time intervals between the old events and the new events are categorized to *Ta3:integral time scale*. The time in QSIM(Kuipers, 1994) corresponds to $Ta3$. Furthermore, $Ta3$ is categorized into two types; $Ta3p$ and $Ta3i$. The former represents the time intervals for integration from the same value as the landmark values to the interval of the landmark values. The latter represents those from the interval to the landmark values. $Ta3p \prec Ta3i$ holds.

On the other hand, until a set of parameters are completely satisfied, the time intervals are categorized to *Ta2:dependency time scale*. The precondition of $Ta2$ is that a set of inherently simultaneous equations[2] are satisfied. The time scale of the causal ordering theory (Iwasaki and Simon, 1994) corresponds to $Ta2$. Until the inherently simultaneous equations are satisfied, the time intervals are categorized to *Ta1:mutual dependency time scale*. Although this time scale has the interval-meaning mentioned before, it has no ordinal-meaning. On the other hand, $Ta2$ and $Ta3$ can have both kinds of the physical meaning. When a set of parameters achieve its equilibrium, the reasoning shifts to *Ta4:equilibrium time scale*.

For example, consider a simple system modeled by the direct modeling, $y = x - z, dz/dt = y$. A variable takes one of the three qualitative values, [+], [0] and [−], where the landmark value is 0. In the initial state, all variables take [0] except for a disturbance $x = [+]$. Figure 2 shows causal relations generated on $Ta2$, $Ta3$ and $Ta4$. In this case, the method of constraint satisfaction is simple propagation of values. First, the value of y becomes greater than 0 (denoted by $[0] \rightarrow [+]$ in the figure) according to $y = x - z$. Next, the value is propagated to the derivative of z (denoted by dz in the figure). At this point, every parameter has a value which satisfies all constraints, that is, the precondition of $Ta3$ becomes true. Then the reasoning shifts to the longer time scale $Ta3$. On the scale $Ta3$, the integral operator is applied to z, then z becomes greater than 0. Next, on $Ta2$, the new value of z is propagated to y and so on. In the case that the system eventually achieves its equilibrium, the event $e_{(3,1)}$ is generated on the time scale $Ta4$. Note that there is no causal relation between y and dz (i.e., both events happen at the same time point) on the time scale $Ta3$, while the change of y causes that of dz after a small time interval t_2 on the time scale $Ta2$.

When model builders describe a phenomenon in terms of differential equations, the modeling rationale is to capture dynamic changes in the transitional behavior in $Ta3$ to its equilibrium. In general, it implies that the time interval to achieve its equilibrium is longer than the other phenomena.

(b) Time Constant Modeling In order to represent differences in time constants, this modeling technique divides the target system into such parame-

[2]This term represents such simultaneous equations which cannot be solved by substitution alone, borrowed from (de Kleer and Brown, 1984).

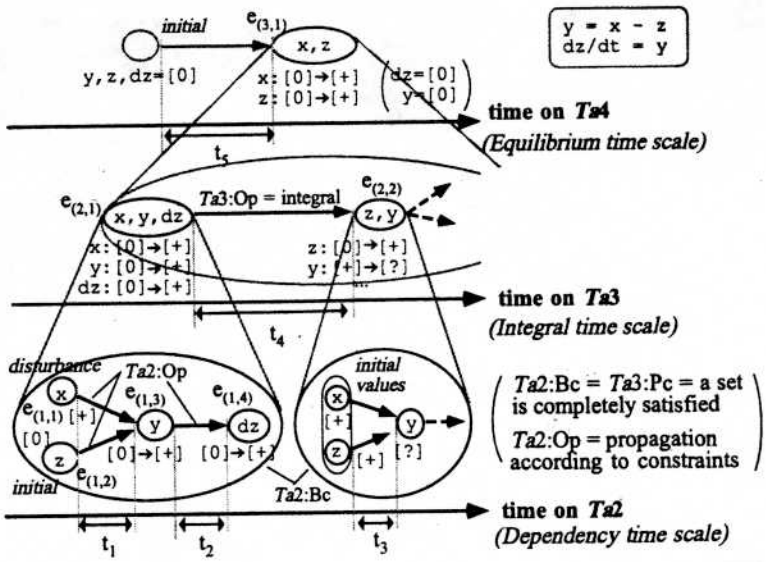

Figure 2: Relation among the time scales $Ta2 \prec Ta3 \prec Ta4$

ter sets where the time intervals to achieve equilibrium $Ta4$ are extremely different from each other. In such a model, the (relatively) faster mechanism firstly reaches the equilibrium on the time scale $Tb1$. Then the reasoning shifts to the time scale $Tb2$. This kind of modeling is found in (Iwasaki and Simon, 1994; Kuipers, 1994). This modeling has an advantage in reasoning efficiency because of separation of the reasoning space.

(c) Component Structure Modeling This modeling is to divide the whole system to subparts according to component structures based on the device ontology (de Kleer and Brown, 1984). In this article, devices in the minimum grain size are called "components". $Tc1$ represents internal behavior in components, while $Tc2$ represents behavior between neighboring components. Interactions between the global structures containing components are represented by $Tc3$. Those between more coarse-grained global structures are also represented by $Tc3$. $Tc4$ represents that the whole system eventually reaches equilibrium. The ordinal relations among these time scales reflect structural distances.

Figure 3 shows an example of causal relations in the local components c_1 and c_2. Given the disturbance $x = [+]$ in c_1, the values within c_1 are changed by the intra-component propagation. Then, on the scale $Tc2$, the value of y is propagated to c_2.

Although $Tc2$ and $Tc3$ have the interval-meaning, the connection information alone cannot give the ordinal-meaning to them. We will discuss additional knowledge for the ordinal-meaning later. On the other hand, $Tc1$ has no physical meaning in any sense. This modeling technique implies such modeling rationales that the causal relations should reflect functioning components and the medium flow along the structures.

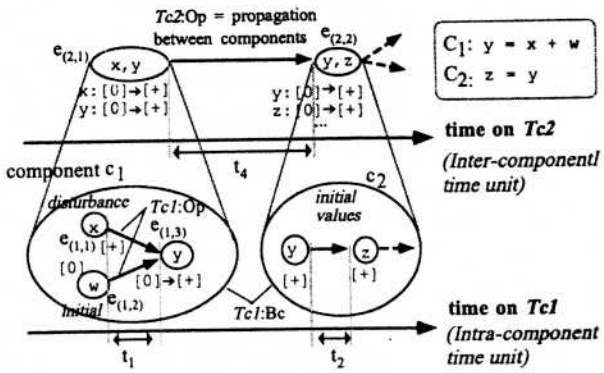

Figure 3: Relation between the time scales $Tc1 \prec Tc2$

(d) Modeling of Periods of Interest This modeling allows the reasoning engine to treat only particular temporal periods of interest such as initial behavior. The time scales constrain not length but the number of time intervals. For example, QUAF (Rose and Kramer, 1991) reasons only the initial changes $Td1$ and the final responses $Td3$ without the intermediate transient behavior. This technique contributes to disambiguation of reasoning results and avoiding reasoning costs.

Causal Time Scales in Reasoning Systems

Let us characterize some of the existing reasoning systems in terms of the causal time scales. In general, a time resolution of a reasoning system is specified by a set of combinations of the primitive time scales discussed thus far. The notation $T_1 : Tx_1 \& Tx_2$ represents that the time scale T_1 consists of Tx_1 and Tx_2. Table 2 shows the time scales which can be treated by some conventional qualitative reasoning systems. For example, QSIM(Kuipers, 1994) can cope with behavior on $Ta3$ and $Ta4$. QSIM uses only mathematical differential equations and adopts a kind of generate-

Table 2: Causal time scales in reasoning systems

QSIM (Kuipers, 1994)	$T1_{qs}$: $Ta3$, $T2_{qs}$: $Ta4$
QSEA (Kuipers, 1994, ch.7)	T_{qa} : $Ta4$
Time-Scale (Kuipers, 1994, ch.12)	$T1_{ts}$: $Ta3$ & $Tb1$, $T2_{ts}$: $Ta4$ & $Tb1$, $T3_{ts}$: $Ta3$ & $Tb2$, $T4_{ts}$: $Ta4$ & $Tb2$
QUAF (Rose and Kramer, 1991)	$T1_{qf}$: $Ta3$ & $Td1$, $T2_{qf}$: $Ta3$ & $Td3$
Mythical Time (de Kleer and Brown, 1984)	$T1_{mt}$: $Ta1$ & $Tc1$, $T2_{mt}$: $Ta1$ & $Tc2$, $T3_{mt}$: $Ta3$ & $Tc4$
Causal Ordering (Iwasaki and Simon, 1994)	$T1_{co}$: $Ta2$, $T2_{co}$: $Ta3$
Abstraction (Iwasaki and Simon, 1994)	$T1_{ab}$: $Ta2$ & $Tb1$, $T2_{ab}$: $Ta3$ & $Tb1$, $T3_{ab}$: $Ta2$ & $Tb2$, $T4_{ab}$: $Ta3$ & $Tb2$

and-test method for constraint satisfaction. Thus, no causal relation among transitional behavior to $Ta3$ is identified. The time of QSIM is corresponds to $Ta3$. QSEA(Kuipers, 1994, ch.7) treats only equilibrium states represented by $Ta4$. The time-scale abstraction(Kuipers, 1994, ch.12) is a kind of the time-scale modeling represented by Tb. QUAF(Rose and Kramer, 1991) reasons only the initial changes $Td1$ and the final responses $Td3$ on the integral time scale $Ta3$.

The method proposed in (de Kleer and Brown, 1984) can generate causal relations among more fine-grained time scale $Ta1$, called "mythical time", on the basis of the concept of device. Causal relations on $T1_{mt}$, however, do not always have the physical meaning because $T1_{mt}$ consists of $Ta1$ and $Tc1$. On the other hand, in order to give the ordinal-meaning to $T2_{mt}$, de Kleer and Brown employ general heuristics representing physical intuitions. Causal relations generated by them, however, are ambiguous due to the arbitrariness of heuristics application.

The causal ordering theory (Iwasaki and Simon, 1994) derives causal relations on $Ta2$, which have the ordinal-meaning representing mathematical dependency. The theory, however, does not try to derive those on $Ta1$. Two abstraction techniques corresponding to Tb are also discussed.

Primitive Reasoning Scheme

The primitive reasoning scheme of a reasoning system can be specified by the set of time scales which the system can cope with. Let TS be such a set and E_c be a current set of events to be carried out. The generic reasoning scheme for a current time scale T_c and neighboring time scales T_1 and T_2 where $T_1 \prec T_c \prec T_2$ is defined as below.

1. On the time scale T_c, if an event $e_1 \in E_c$ satisfies the precondition T_c:Pc, the operator T_c:Op is applied to e_1 and then a new event e_2 and a new link l between e_1 and e_2 are generated. $\tau(l) = T_c$ holds.
2. The reasoning process shifts to the shorter time scale T_1. $T_c \leftarrow T_1$ and $E_c \leftarrow e_2$ and go to step 1 recursively[3].
3. If e_2 does not satisfies the boundary condition T_c :Bc, go back to step 1 and $E'_c \leftarrow E_c - \{e_1\} + \{e_2\}$.
4. If e_2 satisfies the boundary condition T_1:Bc, the reasoning process shifts to the longer time scale T_2. All events in T_c are transferred to the event e_3 on T_2. Go to step 1 recursively.

The reasoning process starts with the minimum time scale T_{min} in TS, given the initial value E_c. This reasoning process repeats recursively until the boundary condition of the maximum time scale holds. There are such cases that T_{min} needs a special operator to satisfy the precondition of T_{min}.

The reasoning processes of the conventional systems can be explained by their time scales shown in Table 2. For example, the reasoning method called time-scale abstraction (Kuipers, 1994, ch.12) starts with the minimum time scale $T1_{ts}$. Since $T1_{ts}$ contains $Ta3$, the operator for $T1_{ts}$ is the integration[4]. When the boundary condition of $T1_{ts}$ becomes true, i.e., the faster system reaches equilibrium, the reasoning process shifts to $T2_{ts}$. Because the system is in equilibrium, no reasoning is carried out in $T2_{dx}$. Then, the reasoning process in $T3_{ts}$ starts and then the slower behavior is generated. In principle, the reasoning process at $T3_{ts}$ backs to the shorter time scales $T1_{ts}$ and $T2_{ts}$. In this case, however, because $T2_{ts}$ is in equilibrium and hence has no more events, only checks of values are needed. The primitive scheme of the algorithm shown in (Kuipers, 1994) is identical with this one.

The reasoning result consists of a set of events E and a set of links L each of which has a time scale $T \in TS$ associated with it where $\tau(l) = T$. If there is a (transitive) causal relation between e_1 and e_2, $\tau(e_1, e_2)$ denoting the time scale representing the time interval between e_1 and e_2 is defined as follows;

$$\tau(e_1, e_2) = \max_{l \in Le} \tau(l)$$

where $L_e \subset L$ consists of the links between e_1 and e_2. This implies that a chain of time intervals represented by a time scale can be represented by the same time scale. In other words, time intervals on a time scale T_1 can never become longer enough to be categorized into the longer time scales than T_1. In the cases of no causal relation, if $\tau(e_0, e_1) \prec \tau(e_0, e_2)$ where e_0 represents the last common event (i.e., the junction event), we only can say that e_1 happens before e_2. If not, there is no temporal order between such events.

Feedback and Causal Time Scales

Such phenomena that the effect of an event of a parameter is eventually propagated to the parameter itself are called as feedback. The time delay along the

[3]The symbol '$\leftarrow$' denotes substitution

[4]Strictly speaking, the operator of QSIM is not identical with integration. It represents possible transitions over time for reasoning efficiency.

feedback loop plays a crucial role in human understanding of feedback. For example, in the cases where the time delay along a feedback is very short and then the modeler has no interest in the transitional behavior of the feedback, it is no need to generate causal relations among events in the feedback loop and to trace the changes of parameter values. Therefore, the reasoning engine can treat feedback according to the following heuristics.

Feedback heuristics : Whether or not a phenomenon is recognized as feedback depends on the time delay for the propagation loop according to the pre-defined threshold values T_{s1} and $T_{s2} \in TS$. Let L be a set of the links contained in the propagation loop and T_l be the time scale for the time delay along the loop.

1. If $T_l \preceq T_{s1}$ then the phenomenon is not treated as feedback. The orders of events in L have no physical meaning. If the new value after the feedback is different from the original value, that is viewed as contradiction at the same time point.
2. If $T_l \succ T_{s1}$ and $T_l \preceq T_{s2}$ then the phenomenon is treated as *semi* feedback. The orders of events in L have the physical meaning. If there is a conflict between the old and new values then the new value is neglected.
3. If $T_l \succ T_{s2}$ then the phenomenon is treated as feedback. The orders of events in L have the physical meaning. The values will be changed after the feedback.

The last one corresponds to the usual feedback. The first two are paraphrased as "the feedback is virtual, produced by the sequential operations of the reasoning method" and "there is no feedback which suppresses the original change instantaneously", respectively.

Design of Reasoning Systems

This section discusses design processes of causal reasoning systems based on the causal time ontology. We show how the modeling scheme, the constituents of the models and the reasoning procedure are determined according to the required temporal granularity. In this section, we use our causal reasoning system (Kitamura *et al.*, 1996a; Kitamura *et al.*, 1996b) as an example, which has been developed aiming at finer-grained causal relations than the conventional systems.

Required Time Scales

The system is designed to derive causal relations based on the device ontology. The target systems are fluid-related systems such as power plants. The system is concerned with flow rate, pressure and heat of fluid assuming fluid is incompressible. The requirements are reusability of models and disambiguation of reasoning results. Table 3 shows the requirements for causal relations to be generated and their time scales. For example, in order to cope with global phenomena such as changes in temperatures caused by global heat balances, hierarchical structure ($Tc3$) is needed. There are also such cases where changes in non-neighboring components are simultaneous, called ***globally simultaneous phenomena***. For example, on the assumption that fluid is incompressible, flow rate of such fluid at each component changes at the same time. Thus, such a time scale which is combination of $Tc3$ and $Tb1$ is needed.

Table 3: Requirements for causal relations

Required causal relations	*Time scales*
Interaction among components	$Tc2$
Global phenomena (heat)	$Tc3$
Globally simultaneous phenomena (flow)	$Tc3$ & $Tb1$
Dependency relations	$Ta1, Ta2$
Dynamic behavior	$Ta3$
Initial and final response	$Td1, Td3$

Table 4: The time scales of our reasoning system

Name of time scale	*Definition*
$T1$:*Intra-Component*	$Ta1/2$ & $Tb1/2$ & $Tc1$
$T2$:*Inter-Component*	$Ta1/2$ & $Tb2$ & $Tc2$
$T3$:*Global*	$Ta1/2$ & $Tb2$ & $Tc3$
$T4$:*Globally Simultaneous*	$Ta1/2$ & $Tb1$ & $Tc3$
$T5$:*Integral*	$Ta3$ & $Tb2$ & $Tc3$
$T6$:*Partial Equilibrium*	$Ta4$ & $Tb2$ & $Tc3$
$T7$:*Complete Equilibrium*	$Ta4$ & $Tb2$ & $Tc4$

$T1 \prec T4 \prec T2 \prec T3 \prec T5 \prec T6 \prec T7$

As a result of design decisions how to cope with the required primitive time scales shown in Table 3, the set of time scales of the system shown in Table 4 are determined. Firstly, it is assumed that there is only one level of faster mechanisms in terms of equilibrium equations. Thus, $T4$:*globally simultaneous time scale* consists of $Ta1$ or $Ta2$ & $Tb1$ & $Tc3$ and the inter-component phenomena ($T2$) and integration ($T5$) are in $Tb2$. Next, as a reflection of no interest in the transient behavior of the inter-component phenomena, the $T2$:*inter-component time scale* is represented by equilibrium equations, i.e., $Ta1$ or $Ta2$. Lastly, although the $T1$:*intra-component time scale* can represent not only a faster mechanisms but also a slower mechanisms, $T1 \prec T4$ holds because $Tc1$ in $T1$ represents finer-grained changes. Because $T4$ represents almost simultaneous phenomena, $T4 \prec T2$ holds.

Moreover, the reasoning engine has some assumptions. Firstly, we assume that the target system has a normal equilibrium state without any perturbation. The behavior represents a response to perturbation to the normal equilibrium. A parameter takes one of the three qualitative values related to the deviation from a normal value which is defined as a permitted range of the parameter in the normal equilibrium. [+]([-]) represents a quantity greater(less) than the normal value. [0] represents a quantity equal to the normal value. Next, it is required that the reasoning engine can infer only the initial changes ($Td1$) and the final responses ($Td3$) and skip the transient behavior in $Td2$. Lastly, we assume that all constraints are continuous.

Table 5: The modeling scheme

Time scales	*Modeling scheme*
$Tc1, Tc2$	Components and connection between ports
$Tb1, Tb2$	Description of time scale of mechanisms
$Td2, Td3$	Stability to equilibrium of parameters
$Ta1\&Tc2$	Causal specification

Required Constituents of the Model

The set of time scales as the performance specification of a reasoning engine governs the constituents of the model. In other words, the model should have such contents as to give the physical meaning to the causal relations to be generated. Table 5 shows the modeling schemes of the system for each time scale. For example, in order to skip the transitional behavior ($Td2$) to the equilibrium state ($Td3\&Ta4$), the knowledge whether or not the behavior will reach to equilibrium, so called stability of equilibrium(Kuipers, 1994), is needed.

Causality of Components A main issue to discuss in order for causal ordering based on the device ontology is what contents of models to give the physical meaning to the causal relations among components ($Tc2$). As discussed thus far, additional knowledge is needed. Considering components have their own causal characteristics, our approach is to explicitly describe inherent causal properties of each parameter in components, called *causal specifications*, context-independently. Such properties, however, prone to dependent on context as discussed in (de Kleer and Brown, 1984). In order to help capture causal properties independent of context, we have identified the three categories of causal relations within a component; *the isolated internal causality, the external causality* and *the combined internal causality.* The isolated internal causality represents such causal relations that a change of a parameter in a component causes that of another parameter in the component through events in the components. The external causality is concerned with direct interactions between connected components. In the case of the combined internal causality, the causal chains between cause and effect include events in other components. Causal specifications from the viewpoint of the isolated internal causality are context-independent.

Model representation

The overall structure of the system is represented by a combination of component models and connections on the basis of the device ontology(de Kleer and Brown, 1984). A component model consists of (1)a set of parameters, (2)constraints over parameters, (3)ports for connections, (4)*causal specifications* representing causal properties of the component, and (5)*time scale* of phenomena.

Constraints are described in terms of qualitative operators and parameters. $D(p)$ represents a derivative of a parameter p with respect to time. It takes one of the three qualitative values [+],[-],[0] which correspond to the sign of derivatives. The integral equation: $p_{(t+1)} = p_{(t)} + D(p)_{(t)}$ holds. Constraints "$\exists t,\ D(p)_{(t)} = [0]$" mean that the parameter converges to the equilibrium state. A parameter can belong to some ports for connections among components. The connection information is represented by relations between the ports. There are global constraints which have connections to local components.

A causal specification of a parameter represents possibility of acting causal roles using the following two flags.

Cause,C: Changes of the value of the parameter can cause those of values of other parameters in the component through events within the component.

Effect,E: Changes of the value of the parameter can be caused by those of values of other parameters in the component through events within the component.

A causal specification takes one of the three values, $C\bar{E}$, $\bar{C}E$ and CE, where "$^-$" is a negation symbol. If there can be a parameter whose change affects the value of the parameter under consideration, then the flag E is associated with the parameter under consideration. And, if there can be a parameter whose value is affected by that of the parameter under consideration, then the flag C is associated with the parameter under consideration. If there is no such parameter, $\bar{C}$ ($\bar{E}$) is associated. Parameters with a constant value, for example, a resistance R in an electric circuit[5], have $C\bar{E}$ as causal specification. The values of such parameters are changed only by influences of other components and/or factors external to the model of the system such as faults. Thus, a parameter is *exogenous*(Iwasaki and Simon, 1994) to the model of the target system if and only if it has $C\bar{E}$ as causal specification and has no connection with other components. The exogenous parameters are candidates of the faults in the diagnostic tasks.

In order to cope with global phenomena, global constraints over local components are described. Such global constraints are justified by general properties of the physical entity such as heat and fluid. Since such properties are specified by the physical laws and the generic topologies of connections among components such as loop, they can be prepared beforehand for each generic topology as a part of the domain ontology. For example, such a general property for the generic loop topology holds in which changes of the temperatures in a loop are caused by the difference between the inflow and the outflow of thermal energy according to heat conservation law. Global constraints are instantiated according to concrete configurations.

Time-scales of phenomena enable the reasoning engine to distinguish *globally simultaneous* phenomena such as changes of flow-rate of incompressible fluid as T4. The global constraints representing such phenomena are called *globally simultaneous constraints*

[5]On the assumption that the resistance R is not changed by heat.

and marked with "simultaneous". Other constraints and local component models are marked with "not-simultaneous".

Reasoning

The reasoning method is based on the primitive reasoning scheme discussed before. Given an initial event, intra-component reasoning at the minimum time scale $T1$ is initiated to determine values of other parameters in the component. In general, the intra-component reasoning is invoked when changes are propagated from other components. The operator in $T1$ is evaluation of constraints according to the causal specification. If a value of a parameter marked with "CÊ" is not determined by other components, it can be assumed to remain an old value in the precedent state. Next, the inter-component reasoning in $T4$ globally propagates events to other components. Then, the reasoning shifts to the longer time scales $T2, T3, T5$, similarly. The reasoning of feedback is based on the feedback heuristics. In our system, since a part of causal relations in $T1$ have no physical meaning and $T4$ represents a very fast mechanism, the threshold values T_{s1} and T_{s2} are set to $T1$ and $T4$, respectively.

Example

The target system of the example shown here is a heat transportation subsystem of a power plant. Figure 4 shows a part of the causal relations generated by the reasoning engine in the case that the valve is closed a bit by hand. Given the component models shown in Figure 4, global constraints for flow-rate and pressure, and `Area = [-]` as the disturbance, the system generates changes of parameters and the causal relations among them together with the causal time scales. When the area available for flow takes `[-]`, the reasoning engine derives the increase of the pressure-drop on T1 (see the sequence No.2 in Figure 4) by introducing an assumption that the flow-rate through the valve specified as CÊ remains the normal value `[0]`. The change of the pressure-drop is propagated to the pressure-difference of the pump simultaneously on T4 through the globally simultaneous constraints concerning pressure-drop (#3). In the pump, the flow-rate decreases because the pump tries to compensate the increase of its load by changing the flow-rate (a type of pumps called pressure-compensative type). The change of the flow-rate of the pump is propagated to the all components simultaneously on T4. Then, flow-rate of the valve Flow = `[-]` is derived (#5) and then the assumed value is dismissed, so-called feedback. Then, value of the pressure-drop becomes ambiguous (#6) because of Area = `[-]` and Flow = `[-]`. Since the time delay along the feedback loop is T4+T1+T4+T1=T4 representing instantaneous phenomena, according to the heuristics, the system obtains Pio = `[+]` which matches reality (#7). In the heat generator and the heat exchanger, the decrease of the flow-rate also causes the changes in the temperatures of the coolant (#6). Since these changes on T2 are not simultaneous, these are propagated to the other components after the simultaneous phenomena (#8).

Furthermore, the reasoning system has been successfully applied to a power plant (Kitamura *et al.*, 1996b). The model of the whole system consists of 27 components, 143 parameters and 102 constraints. All the reasoning results matched those obtained by a domain expert including their ambiguities.

Related Work

The time concept in QSIM is discussed in (Kuipers, 1994) from the mathematical viewpoint, which is categorized into Ta or Tb. Iwasaki and Simon show a causal ordering theory for hierarchical sets of variables and discuss how to generate such hierarchical sets according to time scale and strength of interaction among variables (Iwasaki and Simon, 1994). The causal time ontology allows us to clarify the rationales underlying such sets from the physical viewpoint.

Ontologies of time itself have been discussed elsewhere such as (Allen, 1984) where Allen has identified primitives for representing time itself, and categorized of logical relationship between them. The causal time ontology provides cognitive categories of time intervals from the viewpoint of causal ordering of physical systems.

In (de Kleer and Brown, 1984) and (Top and Akkermans, 1991), although general causal properties of devices have been identified, causal relations generated by their methods are ambiguous in the case of inherently simultaneous equations. The TQ analysis (Williams, 1984) provides heuristics to analyze limited kinds of feedback according to time delay. A part of our causal specification (CÊ) corresponds to the descriptions of "exogenous parameters" (Iwasaki and Simon, 1994) of each component. In (Forbus, 1984) and (Washio, 1989), causal properties of physical processes are described. Our global constraint about heat corresponds to an energy constraint (a global filter) for QSIM (Fouché and Kuipers, 92).

Summary

We have proposed a causal time ontology containing a set of causal time scales shown in Table 1 to reveal time concepts in qualitative models. Some conventional reasoning systems have been characterized with respect to causal ordering using the time scales shown in Table 2. Furthermore, the design processes of causal reasoning systems based on the causal time scales have been discussed.

We confined the topic to continuous changes. A discrete model of a phenomenon is, however, often the result of modeling according to such the rationale that the phenomenon is extremely faster than other phenomena, as discussed in (Iwasaki *et al.*, 1995) and (Nishida and Doshita, 1987). Thus, such discrete models can be viewed as another kind of temporal modeling techniques discussed in this article. Investigation on such discrete changes remains as future work.

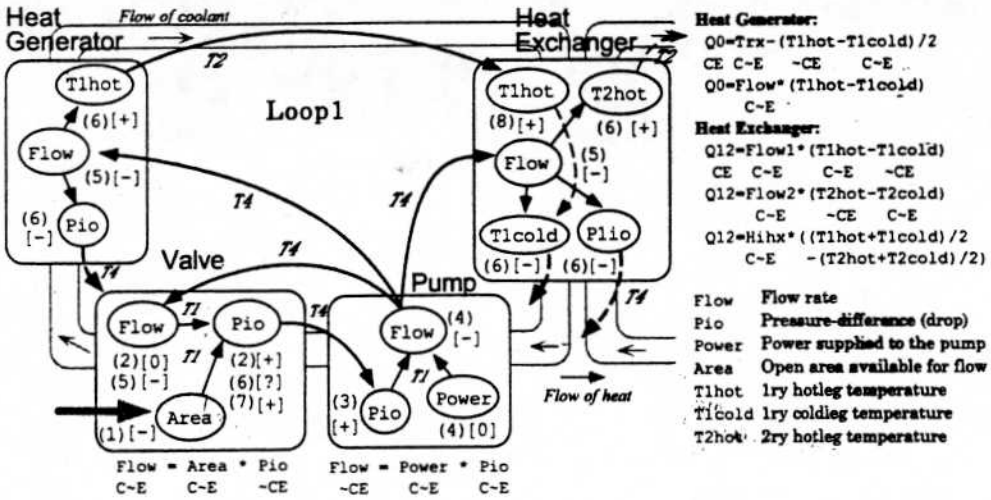

Figure 4: A part of the causal chains in the case of the area of the valve closed a bit

Acknowledgments

The authors would like to thank Shinji Yoshikawa and Kenji Ozawa from Power Reactor and Nuclear Fuel Development Corp. and Munehiko Sasajima from I.S.I.R., Osaka Univ. for their help and comments on building the model of the power plant. The authors are grateful to the anonymous reviewers for valuable comments.

References

Allen, J.F.: Towards a general theory of action and time. *Artificial Intelligence*, Vol.23, pp.123-154 (1984).

de Kleer, J., Brown, J. S.: A Qualitative Physics Based on Confluences. *Artificial Intelligence*, Vol.24, pp.7-83 (1984).

Forbus, K. D.: Qualitative Process Theory, *Artificial Intelligence*, Vol.24, pp.85-168 (1984).

Fouché, P., Kuipers, B. J.: Reasoning about Energy in Qualitative Simulation, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol.22, No.1, pp.47-63 (1992).

Iwasaki, Y., Simon, H. A.: Causality and Model Abstraction, *Artificial Intelligence*, Vol.67, pp.143-194 (1994).

Iwasaki, Y., Farquhar, A., Saraswat, V., Bobrow, D., Gupta, V.: Modeling Time in Hypbrid Systems: How Fast is "Instantaneous"?, *Proc. of IJCAI-95*, pp.1773-1780 (1995).

Kitamura, Y., et.al: A method of qualitative reasoning for model-based problem solving and its application to a nuclear plant. *Expert Systems with Applications*, Vol.10, No.3/4, pp.441-448 (1996).

Kitamura, Y., Ikeda, M., Mizoguchi, R.: A Qualitative Reasoning based on an Ontology of Fluid Systems and Its Evaluation on a Power Plant, *Proc. of PRICAI-96*, pp.288-299, (1996)

Kuipers, B. J.: *Qualitative Reasoning*, MIT Press (1994)

Mars, N.J.I. (Ed.): *Towards Very Large Knowledge Bases*, IOS Press (1995)

Mizoguchi, R., Ikeda, M.: Towards Ontology Engineering, Technical Report AI-TR-96-1, I.S.I.R., Osaka Univ. (1996).

Nishida, T., Doshita, S.: Reasoning about discontinuous change. *Proc. of AAAI-87*, pp.643-648 (1987).

Rose, P., Kramer, M.A.: Qualitative Analysis of Causal Feedback, *Proc. of AAAI-91*, pp.817- 823 (1991).

Top, J., Akkermans, H.: Computational and Physical Causality, *Proc. of the IJCAI'91*, pp.1171-1176 (1991).

Washio, T.: Causal Ordering Methods based on Physical Laws of Plant Systems, MITNRL-033, MIT (1989)

Williams, B. C.: Qualitative Analysis of MOS Circuits. *Artificial Intelligence*, Vol.24, pp.281-346 (1984).