A THESIS

entitled

TOWARD A GENERAL TEMPORAL THEORY

by

Ma Jixin

Submitted in partial fulfilment of the requirements for the award of the

DEGREE OF DOCTOR OF PHILOSOPHY
in Computer Science

School of Mathematics, Statistics and Computing
UNIVERSITY OF GREENWICH, LONDON, U.K.
(1994)

Supervised by

Prof. Brian Knight
Dr. Don F. Cowell

Sponsored by
Government of P. R. China
British Council

To My Family

Acknowledgements

I would like to express my sincere thanks to my supervisors, Professor Brian Knight and Dr. Don F. Cowell, for their invaluable guidance, encouragement, and assistance during the course of this research.

I would also like to express my gratitude to the staff and postgraduates at the Faculty of Technology for their affability and support.

The financial support from the government of People's Republic of China and the British Council is gratefully acknowledged.

Abstract

The research work presented herein addresses time representation and temporal reasoning in the domain of artificial intelligence. A general temporal theory, as an extension of Allen and Hayes', Galton's and Vilain's theories, is proposed which treats both time intervals and time points on an equal footing; that is, both intervals and points are taken as primitive time elements in the theory. This means that neither do intervals have to be constructed out of points, nor do points have to be created as some limiting construction of intervals. This approach is different from that of Ladkin, of Van Beek, of Dechter, Meiri and Pearl, and of Maiocchi, which is either to construct intervals out of points, or to treat points and intervals separately.

The theory is presented in terms of a series of axioms which characterise a single temporal relation, "*meets*", over time elements. The axiomatisation allows non-linear time structures such as *branching time* and *parallel time*, and additional axioms specifying the *linearity* and *density* of time are specially presented. A formal characterisation for the *open* and *closed* nature of primitive intervals, which has been a problematic question of time representation in artificial intelligence, is provided in terms of the "meets" relation. It is shown to be consistent with the conventional definitions of *open/closed* intervals which are constructed out of points.

It is also shown that this general theory is powerful enough to subsume some representative temporal theories, such as Allen and Hayes's interval based theory, Bruce's and McDermott's point based theories, and the interval and point based theory of Vilain, and of Galton.

A finite time network based on the theory is specially addressed, where a consistency checker in two different forms is provided for cases with, and without, duration reasoning, respectively.

Utilising the time axiomatisation, the syntax and semantics of a temporal logic for reasoning about propositions whose truth values are associated with particular intervals/points are explicitly defined. It is shown that the logic is more expressive than that of some existing systems, such as Allen's interval-based logic, the revised theory proposed by Galton, Shoham's point-based interval logic, and Haugh's MTA based logic; and the corresponding problems with these systems are satisfactorily solved.

Finally, as an application of the temporal theory, a new architecture for a temporal database system which allows the expression of relative temporal knowledge of data transaction and data validity times is proposed. A general retrieval mechanism is presented for a database with a purely qualitative temporal component which allows queries with temporal constraints in terms of any logical combination of Allen's temporal relations. To reduce the computational complexity of the consistency checking algorithm when quantitative time duration knowledge is added, a class of databases, termed time-limited databases, is introduced. This class allows absolute-time-stamped and relative time information in a form which is suitable for many practical applications, where qualitative temporal information is only occasionally needed, and the efficient retrieval mechanisms for absolute-time-stamped databases may be adopted.

Table of Contents

CHAPTER 1

INTRODUCTION

"What, then is time?

If no one asks me, I know;

but if I want to explain it to a questioner, I don't know."

Augustine of Hippo

(*Confessions XI, XIV*)

1.1 The Roles of Temporal Reasoning

The above quotation reflects the fact that humans have a natural perception of the effects of this cosmic reality but are unable to answer the philosophical question of "what there is". From a computing perspective this question could be transformed to a more tractable formal question of naming and quantification, that is, to the assertion "there is something such that ... " [ThL91].

Since the early 70s, the study of time has increasingly become an important part of research efforts in a variety of strands within computer science. Notably, researchers have seen that reasoning about and with time is a task with wide application in many problems in the domain of both artificial intelligence and software engineering. As Galton points out in [Gal87], this has come about because computer science as a whole is both highly formal and deeply rooted in the practice of everyday life, so that a formalism designed to handle the pervasive feature of time has an important natural role in many fields. In the review paper of temporal logics [Lon89], D. Long categorises areas requiring temporal reasoning as:

(1) temporal database management

(2) predication

(3) planning

(4) explanation

(5) learning new physics

(6) natural language understanding

(7) historical reconstruction

1.2 The Problems

Within the last two decades, many systems have been proposed for capturing the temporal aspects of events and processes in computer based systems. Generally speaking, all temporal systems must rely on an assumed theory which satisfies some intuitive notions of time. For some systems this underlying theoretical basis is formally described, and for others it remains assumed as intuitively agreed. In analyzing the theoretical basis of temporal systems, there are three items which must be related: the theory, the model, and the real world. According to Funk's definition [Fun83], a theory is a collection of statements about a subject domain, and a model of a theory is a structure in which the statements of the theory are interpreted as true. In addition, if the model is to be of use in some practical domain, then the real world must also be taken as a model of the theory.

Hence, when designing a system for temporal reasoning, we are firstly faced with a choice of the underling time structure. The most common theoretical basis is the standard time point system assumed by classical physics. In this theory, the time domain consists of a continuum of time points, isomorphic to the real line. Point-based intervals are constructed from points, and the duration of an interval is the real number difference of its left and right end-points. There is a weight of historical evidence to convince us that most everyday phenomena are models of this theory. However, recent research has shown that, for many applications, particularly those in artificial intelligence and natural language understanding, the time-point system is not ideal for either the expression of temporal facts, or for the storage and organisation of incomplete temporal knowledge, which is strictly relative (e.g., A is before B) and has little relation to absolute time points. For these applications, other theories have been proposed; for example, based on time intervals as primitive rather than time points.

In his series of papers [All81,83,84], Allen has given a compelling argument which leads to the approach that takes time intervals as primitive rather than constructing them out of points. Allen argues that, if intervals are constructed out of points, such as those in the systems of Bruce [Bru72], of Beek [Bee89,92], and of Ladkin [Lad86,92], one must address the annoying question of whether the end-points are in the intervals or not, seemingly without any satisfactory solution: If intervals are all closed then adjacent intervals have end-points in common, which when adjacent intervals correspond to states of truth and falsehood of some property, can lead to situations in which a property is both true and false at an instant. Similarly, if intervals are all open, there will be points at which the truth or falsity of a property will be undefined. The solution in which intervals are all taken as semi-open (e.g., see the definition of intervals in Maiocchi's *TSOS* [Mai92]), so that they sit conveniently next to one another, seems arbitrary in choice of "left-closed/right-open" and "left-open/right-closed", and this arbitrariness is intuitively unsatisfactory.

There are some problems with Allen's theory which explicitly excludes time points, or, later, addresses them at a subsidiary status. As argued by Galton, the major problem with a purely interval-based theory, excluding time points, is that it is inadequate for reasoning correctly about continuous change [Gal90]. Vilain and Kautz [Vil82,ViK86], as well as Galton [Gal90], have proposed revised systems to address both time intervals and time points. However, some problems still exist. For Vilain's system, the inadequacy of Allen's theory for reasoning about continuous change still remain, since the case that a time point standing between two time intervals, that is, *immediately after* one and *immediately before* another one, is not satisfactorily addressed. Galton's argument that time points should be treated on the same footing as time intervals is indeed very suggestive. However, in his revised system [Gal90], although Galton attempts to reject as meaningless the question whether or not a given point is part of a given interval, he retains the idea of there being a point at the place where two intervals meet. This may lead back to the original problem that Allen, and Galton himself, try to avoid: viz do properties ascribed to the intervals apply to the point or not? For example, how do we express the situation that a light is turned on, if one must address the point p at which interval i meets interval j, where i refers to *Light_Off*, and j refers to *Light_On*?

Additionally, for computer based temporal systems, the question of consistency is a major issue that must be concerned. Generally speaking, a database may be taken as a finite set of temporal knowledge, and hence the temporal reasoner needs only deal with a finite number of time elements. However, the inferencing mechanisms which may be used to derive facts from the database must rely on an underlying theoretical basis, insofar that the complete axiomatisation is needed to prove some corresponding consistency algorithms.

Finally, the characterisation for the *open* and *closed* nature of primitive intervals is another problematic question of time representation and temporal reasoning. If such a characterisation is formally given, intuitively, it should be consistent with the conventional definitions of *open/closed* intervals which are constructed out of points.

1.3 Contributions of the Thesis

The main contribution of this thesis is the development of a general temporal theory which may be seen as an extension of Allen and Hayes' and Vilain's corresponding theories. This new theory axiomatises both intervals and points as primitive time elements on an abstractly equal footing, and hence is more expressive than some existing representative temporal theories, such as Bruce's [Bru72] and McDermott's [Mcd82] point-based theories, Allen and Hayes's interval-based theory [All83,Alh85,89], Galton's revised temporal theory [Gal90], and the interval & point-based theory of Vilain and Kautz [Vil82,ViK85], while their appealing characteristics are retained.

An advantage of the new theory is that it optionally allows time structures such as *linear/non-linear*, and *dense/discrete*, etc. Formal characterisations for these issues are given by means of some correspondingly additional axioms. On the other hand, the axiomatisation provides a satisfactory characterisation for the *open* and *closed* nature of primitive intervals, which has been a problematic question of time representation in many incomplete knowledge systems.

For computer based systems, the concept of a finite time network based on the general theory is introduced. A formal graphical representation of a finite time network is given. In term of this graphical representation, the necessary and sufficient condition for the consistency of a time network are provided in two forms for cases with, and without duration reasoning, respectively.

For temporal reasoning about propositions whose truth values are associated with particular intervals/points, a temporal logic is presented based on the new time axiomatisation. The syntax and semantics for the logic are explicitly defined. It is shown that the logic is more expressive than that of some existing systems, such as Allen's interval-based logic [All84], Galton's revised theory [Gal90], Shoham's point-based interval logic [Sho87a,b], and Haugh's MTA based logic [Hau87]; and the corresponding problems with these temporal systems are satisfactorily solved.

Finally, as an application of the temporal theory, a new architecture for a temporal database system which allows the expression of relative temporal knowledge of data transaction and data validity times is proposed. A general retrieval mechanism is presented for a database with a purely qualitative temporal component which allows queries with temporal constraints in terms of any logical combination of Allen's temporal relations. To reduce the computational complexity of the consistency checking algorithm when quantitative time duration knowledge is added, a class of databases, termed time-limited databases, is introduced. This class allows absolute-time-stamped and relative time information in a form which is suitable for many practical applications, where qualitative temporal information is only occasionally needed, and the efficient retrieval mechanisms for absolute-time-stamped databases may be adopted.

1.4 Outline of the Thesis

An outline of the rest of the thesis is as follows. In chapter 2, we address some major issues about the nature of time. A review of some representative temporal systems is given in chapter 3. In chapter 4, a general time theory is proposed. Chapter 5 examines different models of the theory, and shows that some representative temporal systems may be derived from the general theory. As applied to computer based systems, a finite time network based on the theory is specially addressed. In chapter 6, the syntax and semantics for a temporal logic utilising the time axiomatisation are presented; and a categorization of temporal propositions is provided. Chapter 7 introduces a new architecture for a temporal database system, which allows the expression of both qualitative and quantitative temporal knowledge of data transaction and data validity times. Finally, chapter 8 provides a summary and some concluding remarks.

In this thesis, we will be using the first-order predicate calculus with equality throughout, with the following conventions:

$\wedge$	and,
$\vee$	or,
$\triangledown$	exclusive-or
$\Rightarrow$	implication,
$\Leftrightarrow$	equivalence,
$\exists$	existential quantifier,
$\exists 1$	uniquely existential quantifier,
$\forall$	universal quantifier,
$\neg$	negation.

CHAPTER 2

MAJOR ISSUES ABOUT THE NATURE OF TIME

When designing a system for temporal reasoning, we are faced with a choice of the underling time structure. The theoretical nature of time is a question with a long philosophical tradition and the literature is full of disputes and contradictory theories. This contrasts sharply with the commonly held view of time, which allows people to cope easily with time in their everyday life - for different objectives or motivations, different people may have different approaches. In the past two decades, many temporal systems have been proposed to address the problem of modelling human temporal concepts in a natural way. These models are similar in many respects, but there are subtle differences in terminology and basic theory which derive from the differences in approach. Generally speaking, there are several major issues which should be addressed in terms of the theoretical basis of proposed systems.

2.1 The Primitive Nature of Time

This is the issue of what should be taken as the primitive elements of time. Abstractly, there are three known choices: points, intervals, or both. The prevalent mathematical picture of time is that of a set of points without duration. This point view of time is an extremely abstract conception, not to be encountered in ordinary situations. For instance, even expressions such as "the exact time of birth" refer to some small intervals with positive length, rather than to zero-length points. As another tradition, interval structure has been proposed repeatedly in many temporal systems. There are two fundamentally different treatments of interval based systems. In the first, intervals are assumed to be constructed out of points, and hence, the corresponding systems may be considered as models of point-based time theories. An example of this is the *time segments* of Bruce's model for temporal references [Bru72]. However, as mentioned in the introduction (section 1.2), modelling intervals by taking their end-points can lead to the end-points problem. The second treatment takes intervals as primitive objects without any definitions of the "end-point" and "internal-point" structures. In Allen's interval based temporal theory (see [All83,84] and [AlH89]), time intervals are taken as primitive, while points are relegated to a subsidiary status as "meeting places" of intervals. Other theories, e.g. that of Vilain [Vil82], and that addressed in this thesis, treat both intervals and points as primitive on an equal footing.

Although there is something counter-intuitive about treating time as a point-based system, Boyer [Boy59] advances the view that such a departure from primary intuitions is fruitful for many applications and necessary for the advance of science. Hence, by and large, scientists and philosophers of various persuasions have managed to live with this point view of time. However, there are advocates of the use of intervals instead of points as primitive. The justification provided by them is that the interval representation is more suitable because:

- It allows imprecision and uncertainty of temporal information;

- It allows the grain of reasoning to be varied;

- It can be understood more easily by humans.

For instance, linguists are finding that the semantics of temporal discourse is more easily explained in terms of intervals than of points. We may say that:

1) It took him about 15 minutes to cross the river.

2) He read a book for two hours before going to bed yesterday.

Additionally, since neither the starting-time nor the finishing-time of the process of "crossing-river", "reading-book", or "going-to-bed" is explicitly expressed, it is easier to characterise these processes with primitive intervals than points or point-based intervals.

2.2 Order Relations

Whatever primitive time elements are taken, all time systems must adopt axioms defining some sort of order relation. Two fundamental issues are associated with time ordering: the linearity of the time axis and the density of time elements. We address these issues as follows:

2.2.1 Linearity of time

This issue refers to whether the time axis can be always considered as *linear* or *non-linear.*

Linear structure corresponds to the classical physical model of time, where there is total order over time elements. An example of this structure is that of the real line. The majority of time modelling approaches consider the time axis as being linear. However, non-linear time structures have also been proposed, where the fundamental order relation allows topologies such as *branching time*, *parallel time* and *circular time*, etc. Branching time has been proposed as a useful model to handle possible worlds [Mcd82], uncertainty about the past or the future and the effects of alternative actions when planning. Unfortunately, as reviewed by D. Long in [Lon89], branching time does not succeed very well in capturing the fact that of all possible futures or pasts there is precisely one *actual* future and past, while all the others will always remain hypothetical (further discussion will be given in section 3.3). As for the parallel time lines, they are proposed as a way of modelling separate parallel and asynchronous processes, and hence, parallel model can be used in developing logics for reasoning about parallel computation and concurrent processes. Circular time is another interesting possibility in which past, present and future coalesce. It can be used in modelling the behaviour of repetitive, cyclical processes, for example the repetition of cycles in the traffic signals at a road junction [Lon89].

2.2.2 Density of time

The density question is associated with the discreteness versus denseness of time. It depends on the type of primitives assumed for the system. For interval based systems, a dense system is taken to be one where every interval is (infinitely) decomposable. For point based systems, a dense system is one in which between any two points on the same time line, there is a third. As an alternative assumption, some approaches assume that time is discrete, in which each time element (possibly except the first and last) is "sandwiched" between unique previous and next time elements [Gal90a]. There are many applications in which time can be naturally and conveniently considered to be discrete - in reasoning about computation, for example, time can be modelled as CPU clock-ticks. The reason has been summarised by Theodoulidis and Loucopoulos [ThL91] as follow:

- references concerning time in database systems are usually made in discrete terms, e.g. hiring occur daily;

- when references to locations in time are made, their representation must be finite, e.g. in a computer system or on a piece of paper;

- from a modelling point of view time intervals may be considered to be point like in discrete terms. (E.g., in the form of a discrete time-sampling system with variable sampling times [Kni92])

The main argument in favour of density for time is that it corresponds to both the usual intuitive structure for time and also the conventional model of time adopted in classical physics. Dense model of time seems necessary in modelling continuous change since the concept of continuousity itself presupposes a dense time system. However, in the case of finite computation, it will only be needed to identify and reason about a finite set of temporal data. The fact that taking a database as a finite set of temporal knowledge has no bearing on the density question at all, which is a question of the assumed theory only. This theoretical issue impinges upon the inferencing mechanisms which may be used to derive facts from the database, insofar that the density assumption is needed to establish certain consistency proofs.

2.3 Open and Closed Nature of Intervals

There is a conventional way to characterise the *open* and *closed* nature of point-based intervals. If an interval includes its left end-point (right end-point), it will be then considered as left (right) closed. Otherwise, it is left (right) open. However, when intervals are taken as primitive, there are no definitions about their end-points. Hence, to allow successful modelling of the open and closed nature of these primitive intervals, points must taken as primitive as well, on an equal footing to intervals; and, axioms axiomatising the order relation between intervals and points should be properly introduced in the corresponding theory. Additionally, the interpretation of the open and closed nature of primitive intervals should intuitively be in line with the conventional meaning of the open and closed nature for point-based intervals.

2.4 Duration Reasoning

In most applications, it is expected that a temporal system can support duration reasoning. For example, if it is known that interval I_a and interval I_b start together and that the duration of I_a is greater than duration of I_b, we may infer that I_b finishes before I_a. This inference can be made by use of duration knowledge.

The duration assignment to time elements may be characterised by a function from the set of time elements to $\mathbf{R}_0^+$, the set of non-negative real numbers. Intuitively, of course, the duration of time points should be zero, while the durations of time intervals are positive. For point-based intervals, such as Bruce's *time-segment* [Bru72], their durations may be derived from the distance between their left end-point and right end-point. Given a duration assignment over time elements, some corresponding operators, such as *addition*, may be required to be defined, providing consistency of the whole system.

CHAPTER 3

LITERATURE SURVEY

Since the early 70s, many temporal systems have been proposed to address the problem of modelling human temporal concepts in a natural way. These systems are similar in many respects, but there are subtle differences in terminology and basic theory which derive from the differences in approach. In this chapter, we review some representative temporal systems, according to three fundamental considerations:

- *The assumed axiomatic theory*: For all of the systems which we shall consider, there exists an underlying theoretical basis. For some systems this basis is formally described, and for others it remains assumed as intuitively agreed.

- *The expressiveness of the modelling language*: From the point of view of computer databases, it would be impossible/unnecessary to address all times. Hence, the computer based system may be viewed as another model of the theory, in the form of a finite database of temporal facts. Given that the model is incomplete (in terms of a partial knowledge) by reason of storage limitations, there is a drive for efficient storage and retrieval of incomplete temporal knowledge. Expressive modelling languages allow the storage of temporal information which is incomplete in various fashions.

- *The reasoning mechanisms which are provided*: Deductive inference may be performed on the stored data, with reference to the underlying theory, so that any fact which can be proved from the axioms of the theory and the stored temporal database may be assumed true by inference. In this way, the axioms plus database may be viewed as a deductive database from which facts may be retrieved by inference.

3.1 Bruce's *Temporal Model*

An early attempt at mechanizing part of the understanding of time within an artificial intelligence context was Bruce's model for temporal references [Bru72]. In this system a formal framework, based upon first-order logic, is established for the analysis of tenses, time relations, and other references to time in natural language. The axioms of the framework are based on the following definitions: A *time-system* is a pair, (*time*, ≤), where *time* is a set whose elements are called *time-points*, and ≤ is a partial order over *time*. Because there is nothing that has been defined about *time* other than that it is partially ordered by ≤, the theory allows linear time or branching time, discrete time or dense time. The theory is thus more general than that for the standard point-based system, and inferencing mechanisms must be built on weaker axioms.

Bruce then defines point-based intervals, termed *time-segments,* as chains which are convex in the sense that there are no points missing within the chains, where a chain is a totally ordered subset of *time-points*. The related issues about time-segments, such as: density, linearity, boundedness, may hence be derived from the corresponding issues of the time-points which make up the time-segments. The ordering relations between segments are also inherited from the partial order over the time points. Bruce gives seven binary relations between *time-segments,* which can be derived from the ordering relations over their greatest lower bounds and the least upper bounds: *Before, During, Same-time, Overlaps, After, Contains* and *Overlapped.* In terms of these binary relations, a *tense* is defined as a special n-ary relation on time-segments with the following form:

$$R_1\_R_2\_\ldots\_R_{n-1}(S_1,S_2,\ldots,S_n) \equiv R_1(S_1,S_2) \wedge R_2(S_2,S_3) \wedge \ldots \wedge R_{n-1}(S_{n-1},S_n)$$

where each S_i is a time-segment and R_i is a binary relation between S_i and S_{i+1}. S_1 is called the *time of speech*, S_2, ..., S_{n-1} are called the *times of reference*, and S_n is called the *time of event*. For example, the following sentence

- He will have been going to be going to go

has the tense

$$\textit{Before\_After\_Before\_Before}(S_1,S_2,S_3,S_4,S_5) \equiv$$
$$\textit{Before}(S_1,S_2) \wedge \textit{After}(S_2,S_3) \wedge \textit{Before}(S_3,S_4) \wedge \textit{Before}(S_4,S_5)$$

where S_1 is the time of speech, S_2, S_3, S_4 are reference times, and S_5 is the time of event.

Bruce provides a natural language system, termed *CHRONOS*, which consists of a simple English sentence parser, a theorem prover, and a database of facts and events. The system accepts facts about events from the user and the information which is given by tense and time relations can be combined with other facts to allow inferences about the temporal ordering of events.

However, as argued by Galton, there are some limitations with Bruce's *CHRONOS*, deriving in part from the over-simple nature of its translation procedure [Gal87]. Additionally, no consistency checker for the system has been explicitly provided and there are some difficulties in dealing with the treatment of open or closed intervals, that is, the end-points problem (see section 1.2). Mechanisms for duration reasoning are not specified, although these may be defined by introducing a mapping from the time-points to the reals.

3.2 The *Time Specialist* of Kahn and Gorry

In order to store, retrieve, and reason about temporal information, Kahn and Gorry [KaG77] have designed and implemented a module, called *time specialist,* to maintain separate mechanisms for dealing with dated and undated information. The time specialist is endowed with the capacity to order temporal facts in three major ways:

(1) relating events to *dates,*

(2) relating events to special *reference events,*

(3) relating events together into *before-after chains.*

The time specialist can answer different types of questions such as:

- Did event X happen at time expression T?
- When did event X happen?
- What happened at time expression T?

The time specialist can check the consistency of the latest fact with facts previously accepted, and try to resolve inconsistencies through interaction with the user. In such an interaction, the user may withdraw either the new fact, or some old facts whose removal would lead to consistency. However, removing old facts may involve undoing some prior deductions. In order to be able to do this, a deduced fact is marked by those facts used to deduce it.

However, even if the time specialist is able to make deductions and check whether they are consistent with the facts known in the data base, it is weak if the time indications are fuzzy: fuzziness needs to be represented by means of plus/minus error intervals for the dates of events, and for the lengths of times between two events. Additionally, since each of the three methods to organize temporal statements has its own special data structures and routines to work with those structures, for a given set of temporal facts, it is the user, unfortunately, not the time specialist, who has to choose the most appropriate methods.

No formal theory is stated as a basis for the time specialist. The basis for temporal reasoning is contained in the algorithms which make up the system.

3.3 The *Temporal Theory* of McDermott

McDermott [Mcd82] has developed a first-order temporal logic to provide a versatile "common-sense" theory for time: reasoning about causality, reasoning about continuous change, and planning actions. In accordance with the "naive physics" advocated by Hayes [Hay78], McDermott adopts an infinite collection of states as the primitive temporal elements and added several crucial axioms: Every state has a time of occurrence, *d*(s), a real number called its *date*. Time is assumed to be a continuum, with an infinite numbers of states between any two distinct states, where states are partially ordered by the "*no later than*" order relation "≤"; and the future (not the past) is branching, that is, there are many possible futures branching forward in time from the present. Each single branch, called a "*Chronicle*", consists of a connected series of states and is isomorphic to the real line. Developing his theory, Mcdermott examines three major problems that a temporal reasoning system must face: reasoning about causality, reasoning about continuous change, and planning actions.

McDermott's system has formal axioms with time-points (states) and reals as primitives. The theory assumes a partial ordering relation, which gives rise to branching time. Reasoning is via the assumed theory of the real numbers, and no special mechanisms are needed. We can represent a time state, s, as the pair (C_s, t), where t = d(s) and C_s is the set of chronicles that s belongs to. Possible events may be associated with time states.

For illustration, we shall consider the example of a man called John, planning a trip to the theatre. He may go by train or by bus. We may assume that a decision will be made to go by train or bus. If the decision is made to go by train at time state s_{train1}, where $d(s_{train1}) = t_1$, then John will arrive at the theatre at state s_{train2}, and the play will start at state s_{train3}, where $d(s_{train3}) = t_3$. All of these time states lie on a chronicle c_{train}. Alternatively, if the decision is made to go by bus at state s_{bus1}, where $d(s_{bus1}) = t_1$, then he will arrive at the theatre at state s_{bus2}, and the play will start at state s_{bus3}, where $d(s_{bus3}) = t_3$. All of these time states lie on chronicle c_{bus}. These events and states may be represented by the following data:

(decides-to-take-train,	c_{train},	t_1)
(arrives-at-theatre-by-train,	c_{train},	t_2)
(play-starts,	c_{train},	t_3)

(decides-to-take-bus,	c_{bus},	t_1)
(arrives-at-theatre-by-bus,	c_{bus},	t_2')
(play-starts,	c_{bus},	t_3)

Here, s_{train1} has been represented by the pair (c_{train}, t_1), s_{train2} by (c_{train}, t_{train2}) etc.

In this example, illustrated in Figure 3.3(1), we see that time states divide into two separate chronicles c_{train} and c_{bus}, from the state s_0 which may corresponds to finishing supper, as a result of the John's decision. Although it is obviously possible for us to compare times on different chronicles by means of the t component, McDermott uses the "*no later than*" relation over time states which is restricted to states on the same chronicle. This is to prevent us from making "no later than" comparisons for events which cannot both occur in reality. For example, we are not allowed to ask whether he arrives at the theatre by bus before he arrives by train, since he cannot do both. These two events are said to be in different possible worlds (i.e. chronicles).

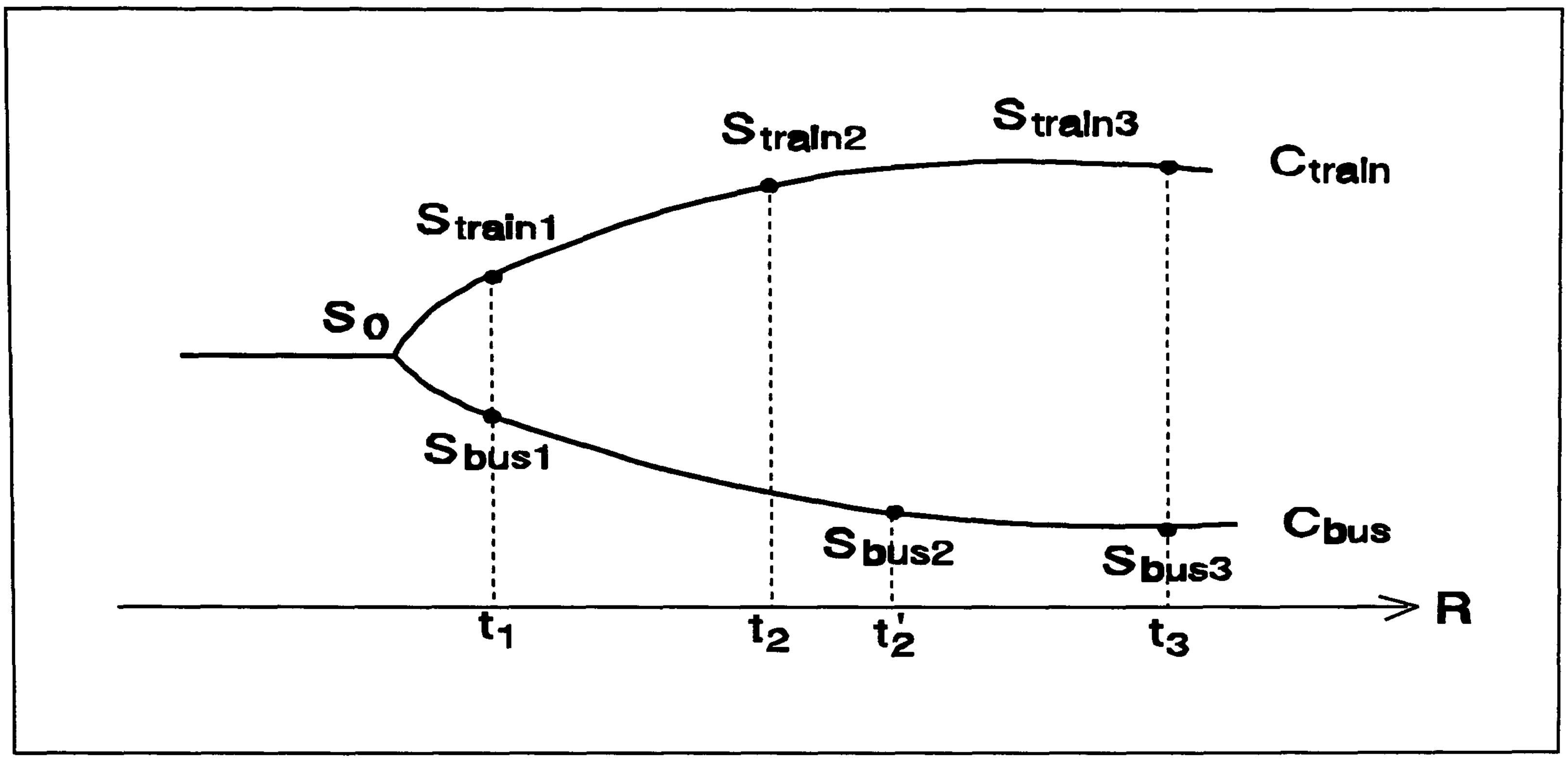

Figure 3.3(1)

McDermott also provides axioms which ensure that chronicles branch only into the future, and this limits the expressiveness of the logic. For, in the example, we have the event "play starts" on two different chronicles which cannot be compared. Using McDermott's logic we must view these as two separate events: "play starts after john's arrival by train", and "play starts after john's arrival by bus". Intuitively, we may judge that the play is independent of John. However, it is not obvious how this independence may be shown in McDermott's system, since we are not allowed to join two chronicles at the state where the play starts.

It is in fact arguable whether we need to consider time as branching in order to model possible worlds. In fact, it is possible to conceptualise the world number, or chronicle, as related to the event data, and not to the time. For example, we can regard the predicate:

(decides-to-take-train, c_{train}, t_1)

as relating:

(event, possible_world, time)

rather than:

(event, time_state)

In this case, time elements are standard linear dense time points, and the axioms for chronicles can be specified independently of those for time.

3.4 The *Interval Logic* of Allen

Allen introduced his temporal logic in order to provide a framework for the naive treatment of two major subareas of artificial intelligence: natural language processing and problem solving. Instead of adopting time points (or states which are associated with time points), he takes intervals as the primitive temporal quantity, as being the natural means of human reference to time. As an example, in [All83], Allen gives the following story:

> *Ernie entered the room and picked up a cup in each hand from the table. He drank from the one in the right hand, put the cups back on the table, and left the room.*

In this account we can identify several time intervals, e.g.: the time Ernie was in the room, the time between entering the room and picking up each cup, the time between putting down the cups and leaving the room, and many others. However, the claim is that intervals are sufficient for modelling all the temporal references in human accounts such as this. Even references to apparent point events, such as the time Ernie entered the room, or the time that he put down a cup, are best modelled as small time intervals. The argument is put forward that all apparently instantaneous events can be decomposed further if we examine them more closely. For example, "entering the room" may be decomposed into: opening the door, moving through the doorway, and closing the door. And again, "opening the door" can be decomposed into turning the handle and pushing the door open. As Allen puts it [All83]:

> *There seems to be a strong intuition that, given an event, we can always "turn up the magnification" and look at its structure.*

In order to express temporal ordering of time intervals, Allen takes as primitive a set of nine (mutually exclusive) basic binary relations between any two intervals [All81], extended later to thirteen [All83]: *Equal, Before, Meets, Overlaps, Starts, Started-by, During, Contains, Finishes, Finished-by, Overlapped-by, Met-by, After.* These are based on Bruce's seven relationships, but whereas Bruce's relations were derived from the order within a point-based theory, Allen's are taken as primitive.

These relationships were later formally defined in terms of the single primitive relation "*meets*" by Allen and Hayes [AlH85,89]. This is done by positing the existence of related intervals. For example:

$$Before(i_1,i_2) \Leftrightarrow \exists i(meets(i_1,i) \wedge meets(i,i_2))$$

The set of axioms that axiomatise the primitive "*meets*" relation over time intervals is proposed first in [AlH85], and then revised in [AlH89], as follows:

<M1> $\forall i,j,k,l \in \mathbf{I}(\mathit{meets}(i,j) \wedge \mathit{meets}(i,k) \wedge \mathit{meets}(l,j) \Rightarrow \mathit{meets}(l,k))$

<M2> $\forall i,j,k,l \in \mathbf{I}(\mathit{meets}(i,j) \wedge \mathit{meets}(k,l) \Rightarrow$

$$\mathit{meets}(i,l)$$

$$\veebar \exists m \in I(\mathit{meets}(i,m) \wedge \mathit{meets}(m,l))$$

$$\veebar \exists n \in I(\mathit{meets}(k,n) \wedge \mathit{meets}(n,j)))$$

(N.B. "$\veebar$" means exclusive disjunction.)

<M3> $\forall i \in \mathbf{I} \exists j,k \in \mathbf{I}(\mathit{meets}(j,i) \wedge \mathit{meets}(i,k))$

<M4> $\forall j,k \in \mathbf{I}(\exists i,l \in \mathbf{I}(\mathit{meets}(i,j) \wedge \mathit{meets}(j,l) \wedge \mathit{meets}(i,k) \wedge \mathit{meets}(k,l)) \Rightarrow$

$$j = k)$$

(N.B. "j = k" means j and k represent the same time element.)

<M5> $\forall i,j \in \mathbf{I}(\mathit{meets}(i,j) \Rightarrow$

$$\exists k \in \mathbf{I} \forall m,n \in \mathbf{I}(\mathit{meets}(m,i) \wedge \mathit{meets}(j,n) \Rightarrow \mathit{meets}(m,k) \wedge \mathit{meets}(k,n))$$

Axiom <M1> states that the "place" where two intervals meet is unique and closely associated with the intervals. The role of <M2> is to ensure that meeting places are totally ordered. <M3> makes every interval have at least one neighbouring interval preceding it, and another succeeding. <M4> simply says that there is only one time interval between any two meeting places. Finally, <M5> states that if two meeting places are separated by a sequence of intervals, then there is an interval which connects these two meeting places. Hence, with axiom <M4> and the definition of equality, for any two adjacent intervals, i and j, the ordered union of i and j, written i ⊕ j, is designed.

A limitation of Allen and Hayes' theory, noted by Tsang [Tsa87], is that the axioms are not primitive enough for extensions. For example, it might be hoped that linearity can be removed from the axiomatisation in order to address the issues such as branching time and parallel time, etc. In fact, Tsang points out that it is difficult to see which of Allen and Hayes' axioms entails linearity. Allen and Hayes conclude that the linearity assumption is characterised by means of axiom <M4> in the revised version of the set of their axioms [AlH89]. However, it is indeed axiom <M2>, rather than <M4>, that entails the linearity of time. In fact, if we remove <M2> from the set of axioms, then the time may be circular, parallel, branching, as shown in Figure 3.4(1). In this graphical representation, the arcs of the graph represent time intervals, and the relation *meets*(i,j) is represented by i being in-arc and j being out-arc to a common node:

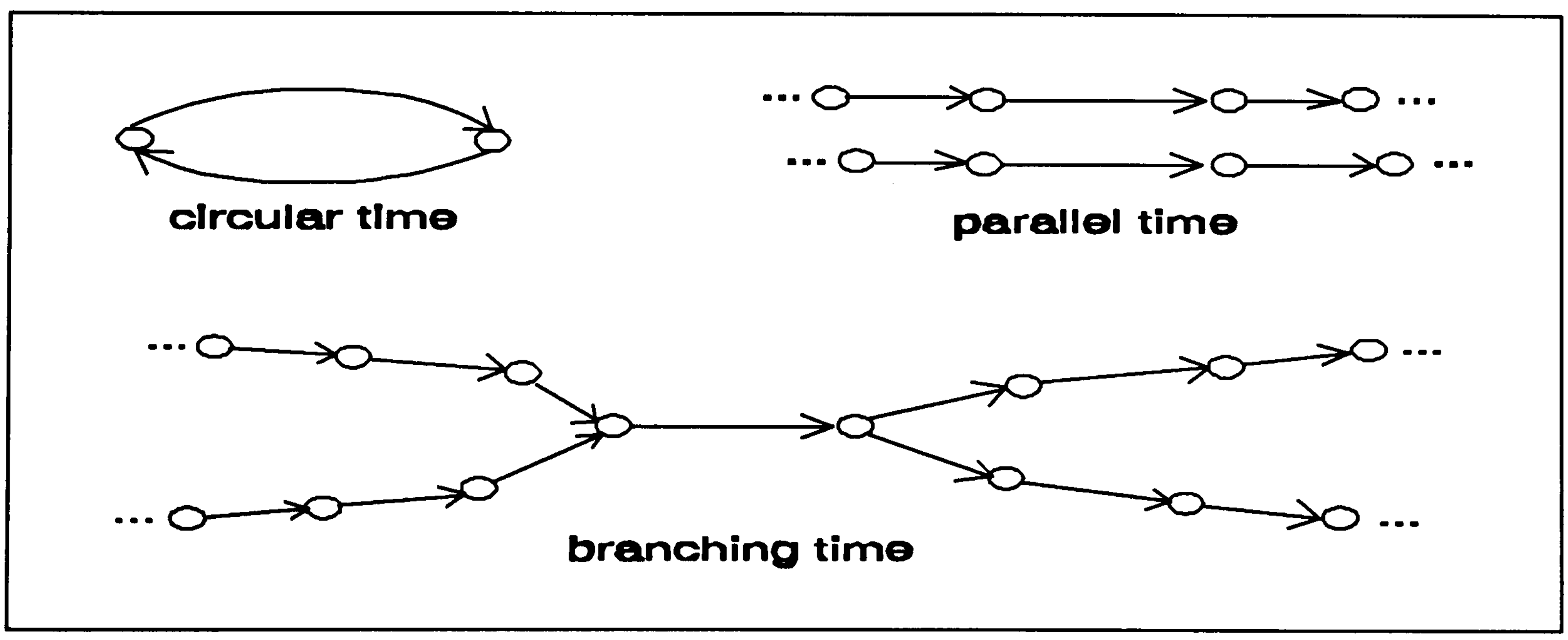

Figure 3.4(1)

Another limitation of Allen and Hayes' time theory is that it takes only intervals, rather than points, as primitive time elements, although points are later introduced as the "*meeting places*" of intervals, or as a maximal set, termed "*nest*", of intervals that share a common intersection, at a subsidiary status within the theory. Allen's original contention is that nothing can be true at a point, for a point is not an entity at which things happen or are true [All83]. However, as Galton shows in his critical examination of Allen's interval logic [Gal90b], Allen's theory of time based on only time intervals is not adequate, as it stands, for reasoning correctly about continuous change.

To characterise the times that some "*instant-like*" events occupy, in [AlH89], Allen and Hayes introduce the idea of very short intervals, called "*moments*". A moment is simply a non-decomposable time interval. The important distinction between moments and points is: although being non-decomposable, moments are defined by having extent and by means of having distinct beginning and ending points (just as for other intervals) [AlH89], while points are defined by having no extent.

However, Allen and Hayes' revised time theory that addresses moments as well is still not adequate for reasoning correctly about continuous change. We may illuminate the problem involved with reference to time points by means of the following example of a ball thrown vertically into the air: The motion may be described qualitatively by the use of two intervals, interval i where the ball is going up, and interval j where the ball is coming down. According to classical physics, there is a point where the ball is stationary. As Allen suggested, in the interval calculus, we may assume that there is a very small interval, that is, a moment, where the ball is stationary. But this does not seem tenable, being inconsistent with the laws of classical physics, no matter how small the interval.

Relating to the "meets" relation, another obvious difference between points and moments is that moments can meet other intervals, and hence stand between them, while points are not treated as primitive objects and cannot meet anything. However, as Allen and Hayes themselves point out, a theory incorporating granularity involves introducing a "*tolerance relation*" that defines when two times are indistinguishable. For example, two intervals, i and j, might be indistinguishable if their beginning points are at most a moment apart, and likewise for their end points. To ensure that the tolerance relation is an equivalence relation, Allen and Hayes propose axiom <M6>, which insists that moments never meet:

<M6> $\forall m,n \in \mathbf{I}(\mathit{moment}(m) \wedge \mathit{moment}(n) \Rightarrow \neg \mathit{meets}(m,n))$

where *moment*(m) is defined by:

$$\forall m \in \mathbf{I}(\mathit{moment}(m) \Leftrightarrow \neg\exists i,j \in \mathbf{I}(m = i \oplus j))$$

Allen and Hayes declare that their formulation permits either discrete or continuous time models, as well as more exotic models that may alternate between continuous and discrete stretches of time. Unfortunately, axiom <M6> leads to another limitation to the primitive time elements: for any interval, either it is non-decomposable, that is, a moment, or it must be infinitely decomposable. For, if it is only finitely decomposable, then it must be the sum of a finite number of moments which would meet one another, contrary to <M6>. This precludes discrete models from the theory containing axiom <M6>. In addition, dense models of the theory, i.e where all intervals are infinitely decomposable, permit no moments at all, so that <M6> is only vacuously true. Hence models of the theory including <M6> which contain moments can be neither solely dense nor solely discrete!

In Allen's system [All81,83], consistency checking is performed by formation of the transitive closure, according to a transitivity table with 144 entries which describes the composition of the thirteen (mutually exclusive) relations. If no conflict is found according to the exclusivity, then the system is consistent. For example, for the system:

Before(a,b), *Before*(b,c)

we may use the transitivity entry:

$$\mathit{Before}(i_1,i_2) \wedge \mathit{Before}(i_2,i_3) \Rightarrow \mathit{Before}(i_1,i_3)$$

to deduce that *Before*(a,c), and no inconsistency arises. However, from:

Before(a,b), *Before*(b,c), *Before*(c,a)

we can deduce *After*(a,b). Hence we have two distinct relations between a and b, *Before*(a,b) and *After*(a,b), which are not allowed due to the exclusivity of these temporal relations. In this way, reasoning in Allen's system relies on the propagation of temporal relations using the transitivity table, in a search for inconsistency.

Allen and Hayes show that the transitivity table in [All83] is a result of the their axioms in [AlH89], following the intuitive reasoning by possible cases which was used to construct the table originally. Additionally, Allen [All83] has suggested that duration reasoning may also be incorporated into the interval-based system by giving examples of rules for duration reasoning. For example:

$$\mathit{During}(\mathrm{a,b}) \vee \mathit{Starts}(\mathrm{a,b}) \vee \mathit{Finishes}(\mathrm{a,b}) \Rightarrow \mathit{duration}(\mathrm{a}) < \mathit{duration}(\mathrm{b})$$

However no comprehensive mechanism has been proposed, and hence the duration reasoning is rather weak.

3.5 Vilain's *Temporal System*

Noting that intervals are not the only mechanism by which human beings understand time, another common construct being that of time points, Vilain and Kautz [Vil82,ViK86]] propose a system which handles time points in much the same way that it handles intervals. The logic of points is arrived at by expanding Allen's logic of intervals: adding new primitive relations and composition rules over them to Allen's interval logic. The new primitive relations may be classified into three groups:

Point-Point: {*Equal, Before, After*}
which relate points to other points;

Interval-Point: {*Before, Started-by, Contains, Finished-by, After*}
which relate intervals to points;

Point-Interval: {*Before, Starts, During, Finishes, After*}
which relate points to intervals.

The mechanism by which Vilain's system makes deductions about points is an extension of that which it uses to make deductions about intervals. In an approach similar to that of Allen, the system maintains a "complete picture" of all relations over intervals and points by means of a transitive closure operation. The operation is performed over the expanded set of composition rules in the newer logic.

However there is a critical omission from the primitive relations between points and intervals in Vilain's system; for the "*Meets*" relation is defined only between intervals and is not allowed between points and intervals. Hence, the problems in modelling continuous change by Allen's system mentioned by Galton in [Gal90b] still exist in Vilain's system. For example, the system is still not capable of modelling the processes of a ball thrown vertically into the air: Let interval i_1 refer to *ball-going-up*, point p refer to *ball-stationary*, and interval i_2 refer to *ball-coming-down*. On the one hand, it is easy to see that p is neither in i_1 nor i_2. On the other hand, according to Vilain's classifications of relations over points and interval, point p is not allowed to meet or be met-by any interval. Hence, we deduce that p is after i_1 and before i_2, that is, there is another time element between i_1 and p, and another time element between p and i_2. This is obviously contrary to our intuition of the processes.

3.6 Kowalski and Sergot's *Event Calculus*

The *event calculus* of Kowalski and Sergot [KoS86] is an approach for representing and reasoning about time and events within a logic programming framework. It is based in part on the situation calculus [McC63,McH69], but focuses on the concept of events as highlighted in semantic network representations of case semantics (see [Kow92]). Its main intended application is the representation of events in updating databases and discourse representation.

Primitives of the theory are events, which are considered to be structureless "points" in time, where "point" is used here only to convey the lack of internal structure. Events start and finish periods of time, during which states are maintained. Events are considered to be after the time periods that they finish and before the time periods that they start, not fully contained within either of these periods.

Using an example about project assignments, Sadri [Sad87] illustrates a number of the general characteristics of the event calculus:

(1) Event descriptions can be assimilated in any order, independent of the order in which events actually take place;

(2) Events can be used for temporal references and need not be associated with absolute times;

(3) Events can be simultaneous;

(4) Events can be partially ordered;

(5) All updates are additive. The effect of deletion is obtained by adding information

about the end of periods;

(6) The event calculus rules are in Horn clause logic augmented with negation by failure;

and

(7) The event calculus allows events to be input with incomplete descriptions.

In [Kow92], Kowalski specially investigates the case of the event calculus connected with database updates. The way in which relational databases, historical databases, modal logic, and the situation calculus deal with database updates is discussed in detail. It is claimed that the event calculus may overcome the computational aspects of the frame problem in the situation calculus, and it is hoped to achieve the efficiency obtainable with "destructive assignment" in relational databases (see [Kow92]). Bernard et. al. [BBG91] have recently presented an adaptation of the event calculus to the problem of determining the temporal structure of operations that must be performed during the realization of some complex objectives. An extension to Kowalski's event calculus model is proposed by Borillo and Gaume [BoG90], by means of the additional spatial component, and the introduction of uncertainty and a general abstract relation among propositions.

The formal theory of Kowalski and Sergot's *event calculus* may be taken as the Horn clause logic plus negation by failure. The event calculus rules can be run as a logic program in *Prolog*. However, the use of negation by failure introduces a procedural element into the axioms. In this respect, the system is thus akin to the time specialist, in that the theory is presented in terms of algorithms.

3.7 Dechter, Meiri and Pearl's *TCSP*

Dechter, Meiri and Pearl [DMP91] have presented a unified approach to temporal reasoning based on constraint-network formalism. In this framework of temporal constraint satisfaction problems (*TCSP*), variables represent time points, and temporal information is represented by a set of unary and binary constraints, each specifying a set of permitted intervals. The unique feature of this framework lies in permitting the processing of metric information, namely, assessments of time differences between events. Algorithms are presented for performing some reasoning tasks, such as finding all feasible times that a given event can occur, finding all possible relationships between two given events, and generating one or more scenarios consistent with the information provided. A *TCSP* involves a set of variables, X_1, ..., X_n, having continuous domains; each variable represents a time point. Each constraint is represented by a set of intervals: $\{I_1, ..., I_n\}$, where these intervals are similar to Bruce's time-segments, that is, they are point-based, may be closed, open, or semi-open. A simple temporal problem (*STP*) is a *TCSP* in which all constraints specify a single interval. The duration of an interval may be defined by the distance between its greatest lower bound and least upper bound. Relations between intervals, such as the thirteen relations defined by Allen, may be derived from the known total order relation among their greatest lower bound and least upper bound. The consistency checking for a *TCSP* is transformed to a corresponding examination of its graphic representation.

The theory is formally stated, with points and real numbers as primitives, and intervals being constructed out of points. It assumes a dense set of time-elements, but time may be branching. Duration reasoning is encompassed by the system, by means of a consistency checking algorithm. The limitation of the *TCSP* model is it's assumption that all the addressed point-based intervals have the same open/closed nature, that is, either interval are all assumed to be closed, or they are all assumed to be open, or all assumed to be semi-open. This assumption can lead to problems: if intervals are all closed then adjacent intervals have ending-points in common, which, when adjacent intervals correspond to states of truth and falsehood of some property, can lead to situations in which a property is both true and false at an instant. Similarly, if intervals are all open, there will be points at which the truth or falsity of a property will be undefined. The solution in which intervals are all taken as semi-open, so that they sit conveniently next to one another, seems arbitrary and unsatisfactory (see [All83,Gal90]).

3.8. Bacchus, Tenenberg and Koomen's *BTK*

Bacchus, Tenenberg and Koomen present a many-sorted temporal logic, termed *BTK* [BTK91], for reasoning about propositions whose truth values might change as a function of time. In order to provide a clear semantics and a well-studied proof theory, they partition both the universe of discourse and the symbols of their language into two sorts, temporal and non-temporal, by which time is given a special syntactic and semantic status without having to resort to reification. In *BTK*, propositions are associated with time objects by including temporal arguments to the functions and predicates, where terms and wffs are defined in the standard fashion, with the only restriction being that arguments of the correct sort must be given for each function and predicate.

Actually, *BTK* is sorted in much the same way as Shoham's *reified logic* [Sho87a,b]. Unlike Shoham's first-order logic in which propositions are expressed just with respect to a pair of time points (denoting a time interval), propositions in *BTK* can be expressed and interpreted with respect to any number of temporal arguments: there is neither a syntactic commitment to the number of temporal objects that any function or predicate may depend upon, nor is there any commitment to interpreting the temporal objects as either intervals or points.

It is interesting to noted that, in their paper [BTK91], Bacchus et. al. have shown that Shoham's logic can in fact be subsumed by *BTK* by defining two transformations, a syntactic transformation, π_{syn}, and a semantic transformation, π_{sem}. π_{syn} maps sentences of Shoham's logic to sentences of *BTK*, while π_{sem} maps models of Shoham's logic to models of *BTK*. Additionally, they argue that Shoham's categorization of propositions over point-based time intervals may also be translated to *BTK*, and the ontology of *BTK* is richer since it allows time intervals to be the primitive temporal objects rather than being defined as pairs of time points.

The major difficulty involved in reasoning in a *BTK* system lies in reasoning with the temporal terms, when the complexity of reasoning is highly dependent on the nature of the temporal domain. However, in *BTK*, there is no axiomatisation characterising the time structure. This question is left open, so that the temporal domain of *BTK* may be defined to be any temporal structure which can be characterised by a set of axioms, for example that of Bruce [Bru72], of Allen and Hayes [AlH89], or of McDermott [Mcd82]. A complete proof theory may then be generated by adding the axioms for the temporal domain to the fundamental axiomatisation of the logic.

3.9 Beek's *Temporal Framework*

In [Bee89,90,92], Beek has separately proposed an interval-based framework, *IA*, and point-based framework, *PA*, for representing and reasoning about incomplete and indefinite qualitative temporal information. Two fundamental reasoning tasks that arise in applications of these frameworks are addressed: Given possible indefinite and incomplete knowledge of the relations between some intervals or between some points,

(i) find a scenario that is consistent with the information provided;

(ii) find the feasible relations between all pairs of the intervals or points.

Following from the approach of Dechter et al. [DMP91], and Ladkin and Maddux [Lad87,92], the reasoning tasks are formalized as binary constraint satisfaction problems. An *IA network* is a network of binary constraints where the variables represent time intervals, the domains of the variables are the set of ordered pairs of rational numbers $\{(s,e) \mid s < e\}$, with s and e representing the starting and ending points of the intervals, respectively, and the binary constraints between variables are represented implicitly by sets of temporal relations over intervals introduced by Allen [All83]. However, these interval relations are induced from the order relation between the starting and ending points of the corresponding intervals. Hence, the interval-based framework *IA* is similar to that of Dechter et al., with intervals being defined in terms of points. Since the rationals are adopted in *IA network* as the underlying representation of time, the time is hence dense, linear, and unbounded. A *PA network* is a network of binary constraints where the variables represent time points, the domains of the variables are the set of rational numbers, and the binary constraints between variables are represented implicitly by sets of the basic point relations proposed in Vilain and Kautz's point algebra [ViK86].

For the point-base framework and the restricted but useful "pointable" version of the interval-based framework, computationally efficient procedures for finding a consistent scenario and for finding the feasible relations are given, which are marked improvements over the previously known algorithms.

It is interesting to note that the frameworks, *IA* and *PA*, deal with temporal relations between intervals, and relations between points separately, that is, the interval-based framework *IA* deals with the thirteen temporal relations (defined by Allen [All83]) between intervals only, while the point-based framework *PA* deals with temporal relations between points only, which are addressed in Vilain and Kautz's point algebra [ViK86]. Relations between intervals and points, such as that proposed in [Vil82], are not addressed at all. Again, like Dechter et al.'s framework, time intervals are not defined as primitive. Indeed, time intervals, and temporal relations between intervals are defined in terms of points (rationals) and the corresponding order relations between points.

3.10 Maiocchi's *TSOS*

TSOS (Temporal Semantic Office Systems) is a system for reasoning about time, presented recently by Maiocchi [Mai92]. In *TSOS*, the temporal domains on which temporal data may be specified in the model are: time points, time intervals, and time extensions. However, only discrete points are taken as primitive time elements, from which other temporal concepts, such as, time intervals and durations are derived. This treatment is quite similar to that of Bruce [Bru72], although in [Bru72] some issues such as durations are not explicitly addressed. In particular, in *TSOS*, time intervals are defined as point-pairs, which are all closed in their *lower end* and open in their *upper end*, and each time interval is connected. However, as mentioned in [All83] and [Lon89], this approach seems arbitrary and unsatisfactory. A time extension denotes a set of consecutive time points at the minimum level of abstraction (quanta of time which is dependent on the application domain). Time concepts such as the distance of a time point from another time point, the duration of a time interval, and dates are then specified in terms of time extensions. For example, "one week" and two days" are time extensions.

In *TSOS*, the concepts of *instantaneous event* and of *proposition* are introduced as the basic elements to which temporal information is associated. Instantaneous events are used to model data to which a single time point is associated, and therefore they are considered instantaneous in the temporal framework of reference for the systems. Propositions model data valid over a time span.

TSOS can be integrated as a time expert in environments designed for broader problem-solving domains. It allows users to infer further information on the temporal data stored in the database through a set of deduction rules handing various aspects of time. To handle imprecise time, *TSOS* supports the concepts of *relative time*, *time granularity*, and *modalities* for propositions, where temporal modalities characterise the possibility of specifying whether a piece of information is always true within a time interval or whether it is only sometimes true, and the capability of answering about the possibility and the necessity of the validity of some information at a given time. The main mechanism for temporal data maintenance supported by *TSOS* is the managements of valid time and transaction time (see [SnA86] and [Sri88]).

CHAPTER 4

A GENERAL TIME THEORY

As discussed in section 3.4, Allen and Hayes' time theory is not primitive enough for extensions [Tsa87], and is not adequate for reasoning correctly about continuous change [Gal90]. Although Vilain's system [Vil82] takes both points and intervals as primitive, it is still not possible to characterise the open and closed nature of intervals, and hence, it is still not adequate for reasoning correctly about continuous change. In this chapter, we propose a general axiomatic framework to serve as an unifying basis for most of representative temporal models in artificial intelligence. The axioms may be seen as an extension of Allen and Hayes' theory [AlH89], to take both intervals and points as primitive objects on an equal footing. This approach is different from that of Vilain and Kautz [ViK86], of Dechter et al. [DMP91], of Ladkin [Lad92], and of Beek [Bee92], which either construct intervals out of points, or treat points and intervals separately.

We present the main body of the axiomatisation in section 4.1. These axioms are independent of the specification of density and linearity. Additional axioms are provided in section 4.2 to specify the linearity and density of time, and, formal definitions are also given for the open and closed nature of primitive intervals. A classification of all possible temporal relations over intervals and points is presented in section 4.3.

4.1 An Axiomatisation of Time based on Intervals and Points

The new general time theory may be seen as an extension of Allen and Hayes' axiomatisation [AlH89], by means of some additional axioms relating to the inclusion of time points as primitive elements, and generalisation of Allen and Hayes' axiomatisation by removing the linearity of time in order to allow non-linear time structures such as branching time, parallel time, etc.

We start the formal theory by posing a nonempty set, **T**, of objects that we shall call **time-elements**, and a function, "*dur*", from **T** to $\mathbf{R}_0^+$, the set of non-negative real numbers. A time-element, t, is called a (time) interval if *dur*(t) > 0, otherwise, t is called a (time) point. According to this classification, the set of time-elements, **T**, may be expressed as $\mathbf{T} = \mathbf{I} \cup \mathbf{P}$, where **I** is the set of intervals, and **P** is the set of points. As in Allen and Hayes' approach, at this early stage we do not make any commitment as to whether all time intervals are decomposable or not. The density question will be addressed by further axioms.

In order to define the primitive order over time elements, we adopt Allen and Hayes' axiomatisation for the single relation "*meets*" between intervals while the axiom characterising the linearity will not be included in the first place. Since the time elements may now be not only intervals but also points, some critical axioms are necessary relating to the treatment of points. The whole set of axioms for the "*meets*" relation over **T** are listed below, where axioms <A1>, <A2>, <A3> and <A4> correspond to Allen and Hayes' <M1>, <M3>, <M4> and <M5> in the above section, respectively:

<A1> $\forall t_1,t_2,t_3,t_4 \in \mathbf{T}(meets(t_1,t_2) \wedge meets(t_1,t_3) \wedge meets(t_4,t_2) \Rightarrow meets(t_4,t_3))$

<A2> $\forall t \in \mathbf{T} \exists t',t'' \in \mathbf{T}(meets(t',t) \wedge meets(t,t''))$

<A3> $\forall t_1,t_2 \in \mathbf{T}(\exists t',t'' \in \mathbf{T}(\ meets(t',t_1) \wedge meets(t_1,t'')$

$\wedge\ meets(t',t_2) \wedge meets(t_2,t'')) \Rightarrow t_1 = t_2)$

<A4> $\forall t_1,t_2 \in \mathbf{T}(meets(t_1,t_2) \Rightarrow$

$\exists t \in \mathbf{T} \forall t',t'' \in \mathbf{T}(\ meets(t',t_1) \wedge meets(t_2,t'')$

$\Rightarrow meets(t',t) \wedge meets(t,t''))$

N.B. For any two time elements, t_1 and t_2, such that $meets(t_1,t_2)$, axioms <A4> and <A3> ensure that there is a unique time element corresponding to the ordered union of t_1 and t_2, which is indicated as $i \oplus j$, and which always implies that $meets(i,j)$.

<A5> $\forall t_1,t_2 \in \mathbf{T}(meets(t_1,t_2) \Rightarrow t_1 \in \mathbf{I} \vee t_2 \in \mathbf{I})$

<A6> $\forall t_1,t_2 \in \mathbf{T}(meets(t_1,t_2) \Rightarrow\ dur(t_1 \oplus t_2) = dur(t_1) + dur(t_2))$

Axiom <A5> is based on the intuition that points will not meet other points, that is, between any two time points, there is a time interval. This is indeed very similar to Allen and Hayes' <M6> which states that moments never meet other moments. However, although <M6> appears to bring little benefit in the form that is presented in [AlH89], dealing with moments, it can be seen that <A5> plays a critical role in the general theory proposed in this chapter, as it is applied to "*time points*". In this case the axiom does not limit the interval structure at all: unlike <M6>, <A5> does not imply the limitation that any decomposable interval must be **infinitely** decomposable. Additionally, axiom <A5> does not affect whether the set of points is dense or not. This issue will be depend on a further assumption ensuring that "within" any time interval, there is a time point (see section 6). Axiom <A6> ensures that the addition operation, "$\oplus$", over time elements is consistent with the function *dur*, which we shall call the **duration assignment** over **T**.

This is the complete fundamental set of axioms concerning the *meets* relation. We denote this set as **A**, and use a pair, (**T**,*meet*), to represent the temporal frame defined by the axiomatisation.

4.2 Some Further Issues

The axiomatisation proposed in the above section defines a general temporal frame based on both intervals and points as primitive objects. In this section, we address some further issues relating to the structure of the frame.

4.2.1 Open and closed nature of intervals

Although intervals are taken in the theory as primitive, that is there are no definitions about the end-points for intervals, the axiomatisation allows the expression of the "open" and "closed" nature of intervals. For example, to represent the process of the ball thrown into the air (see section 3.4), we may relate *ball_going_up*, *ball_stationary*, and *ball_coming_down* to interval i_1, point p, and interval i_2, respectively, where *meets*(i_1,p), *meets*(p,i_2). Intuitively, $t = p \oplus i_2$ relates to *ball_stationary-or-ball_coming_down*. In Figure 4.2.1(1), since i_1 has point p as its immediate successor, we may view i_1 as "right-open" at p, and similarly, i_2 as "left-open" at p. (For clarity, we denote points with bold arcs.) Since interval t (= $p \oplus i_2$) and point p have the same immediate predecessor (i_1) we may view t as "left-closed" at p.

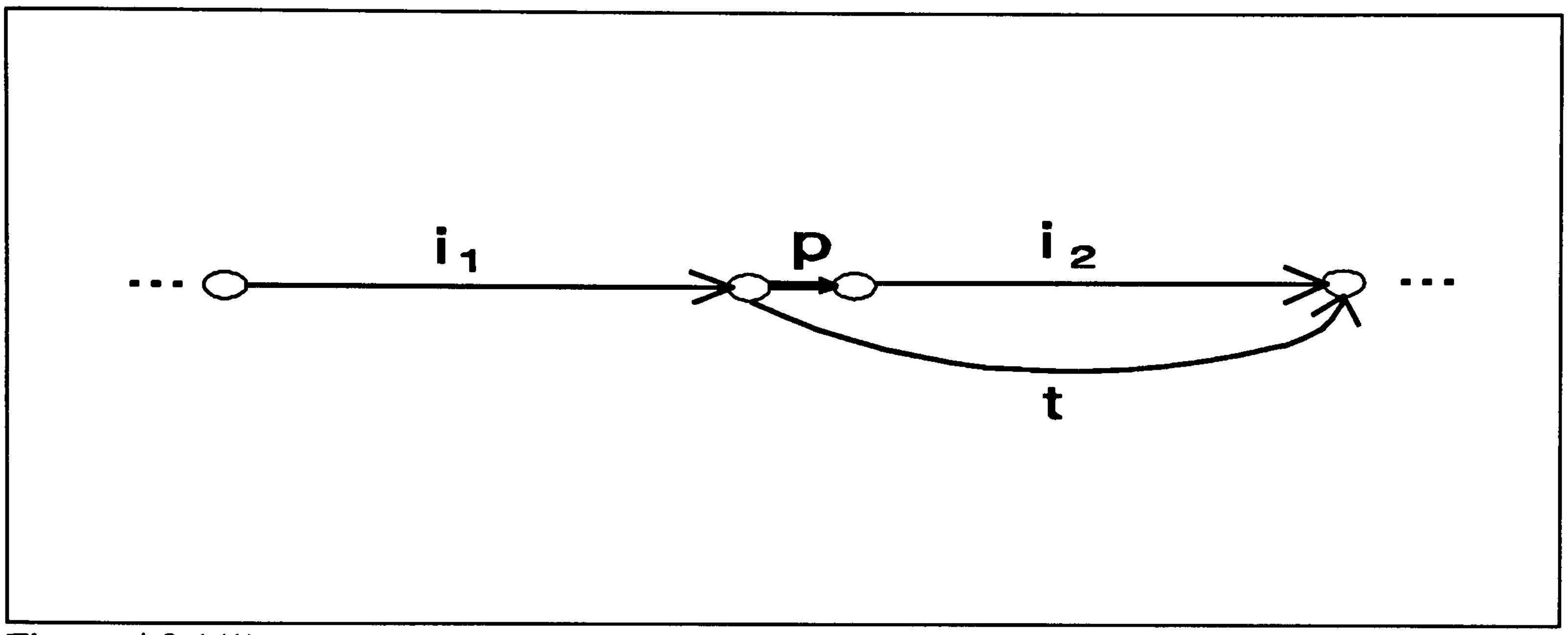

Figure 4.2.1(1)

Formally, the open and closed nature of primitive intervals may be defined as follows:

interval i is **left-open** at point p iff

meets(p, i);

interval i is **right-open** at point p iff

meets(i,p);

interval i is **left-closed** at point p iff

$$\exists i' \in \mathbf{I}(meets(i',i) \wedge meets(i',p));$$

interval i is **right-closed** at point p iff

$$\exists i' \in \mathbf{I}(meets(i,i') \wedge meets(p,i')).$$

It is easy to see that "left-open" and "left-closed" (symmetrically, "right-open" and "right-closed") are exclusive to each other under the axiomatisation. In fact, if interval i is left-open at point p_1, and left-closed at point p_2, then by the above definition, we get:

$$meets(p_1,i) \wedge meets(i',i) \wedge meets(i',p_2), \text{ where } i' \in \mathbf{I}$$

Hence, by axiom <A1> we can infer that $meets(p_1,p_2)$, which contradicts axiom <A5>.

The above interpretation of the "open" and "closed" nature of primitive intervals is in fact in line with the conventional meaning of the open and closed nature for point-based intervals. For instance, point-based interval $(p_1, p_2]$ is "left-open" at point p_1, since intuitively p_1 is an immediate predecessor of interval $(p_1, p_2]$; similarly, $(p_1, p_2]$ is "right-closed" at p_2, since both point p_2 and interval $(p_1, p_2]$ have the same immediate successor, $(p_2, \_\}$.

4.2.2 Linearity of time

Time is usually considered as having a *linear* structure. This corresponds to the classical physical model of time, where the structure is that of the real line, extending indefinitely in both directions.

The **(full) linearity** of a temporal frame (**T**,*meets*) can be characterised by adding an axiom, $\langle A_{\mathrm{Linear}} \rangle$, to **A**, the set of axioms proposed in section 4.1:

$\langle A_{\mathrm{Linear}} \rangle$

$$\begin{aligned}
\forall t_1,t_2,t_3,t_4 \in \mathbf{T}(&\mathit{meets}(t_1,t_2) \wedge \mathit{meets}(t_3,t_4) \Rightarrow \\
&\quad \mathit{meets}(t_1,t_4) \\
&\quad \nabla\ \exists t' \in \mathbf{T}(\mathit{meets}(t_1,t') \wedge \mathit{meets}(t',t_4)) \\
&\quad \nabla\ \exists t'' \in \mathbf{T}(\mathit{meets}(t_3,t'') \wedge \mathit{meets}(t'',t_2)))
\end{aligned}$$

N.B. The axiom $\langle A_{\mathrm{Linear}} \rangle$ is in fact the axiom <M2> (see section 2) for Allen and Hayes' interval-based theory. The "exclusive ors" in this axiom have some quite powerful consequences. In particular, they ensure that there can be no **circular, parallel**, and **branching** times. For instance, the following lemmas are straightforward (see [AlH89]):

$\langle \mathrm{Lemma}^1 \rangle \quad \forall t \in \mathbf{T}(\neg \mathit{meets}(t,t))$

$\langle \mathrm{Lemma}^2 \rangle \quad \forall t_1,t_2 \in \mathbf{T}(\mathit{meets}(t_1,t_2) \Rightarrow \neg \mathit{meets}(t_2,t_1))$

$\langle \mathrm{Lemma}^3 \rangle \quad \forall t \in \mathbf{T} \neg \exists t' \in \mathbf{T}(\mathit{meets}(t,t') \wedge \mathit{meets}(t,t'))$

which ensure that there is no possibility of circular time.

However, without $\langle A_{\mathrm{Linear}} \rangle$, a temporal frame usually allows branching into both the past and the future. Branching temporal frames offer a way to handle possible worlds, uncertainty about the past or the future and the effects of alternative actions when planning. A temporal frame which allows branching into the future but not into the past is called **left-linear** (see Figure 4.2.2(1)). This may be characterised by adding to **A**, the axiom $<A_{L\text{-Linear}}>$, rather than the stronger axiom $<A_{Linear}>$:

$<A_{L\text{-linear}}>$

$$\forall t_1,t_2,t_3,t_4,t\in \mathbf{T}(meets(t_1,t_2) \wedge meets(t_2,t) \wedge meets(t_3,t_4) \wedge meets(t_4,t) \Rightarrow$$

$$meets(t_1,t_4)$$

$$\vee \exists t' \in \mathbf{T}(meets(t_1,t') \wedge meets(t',t_4))$$

$$\vee \exists t''\in \mathbf{T}(meets(t_3,t'') \wedge meets(t'',t_2)))$$

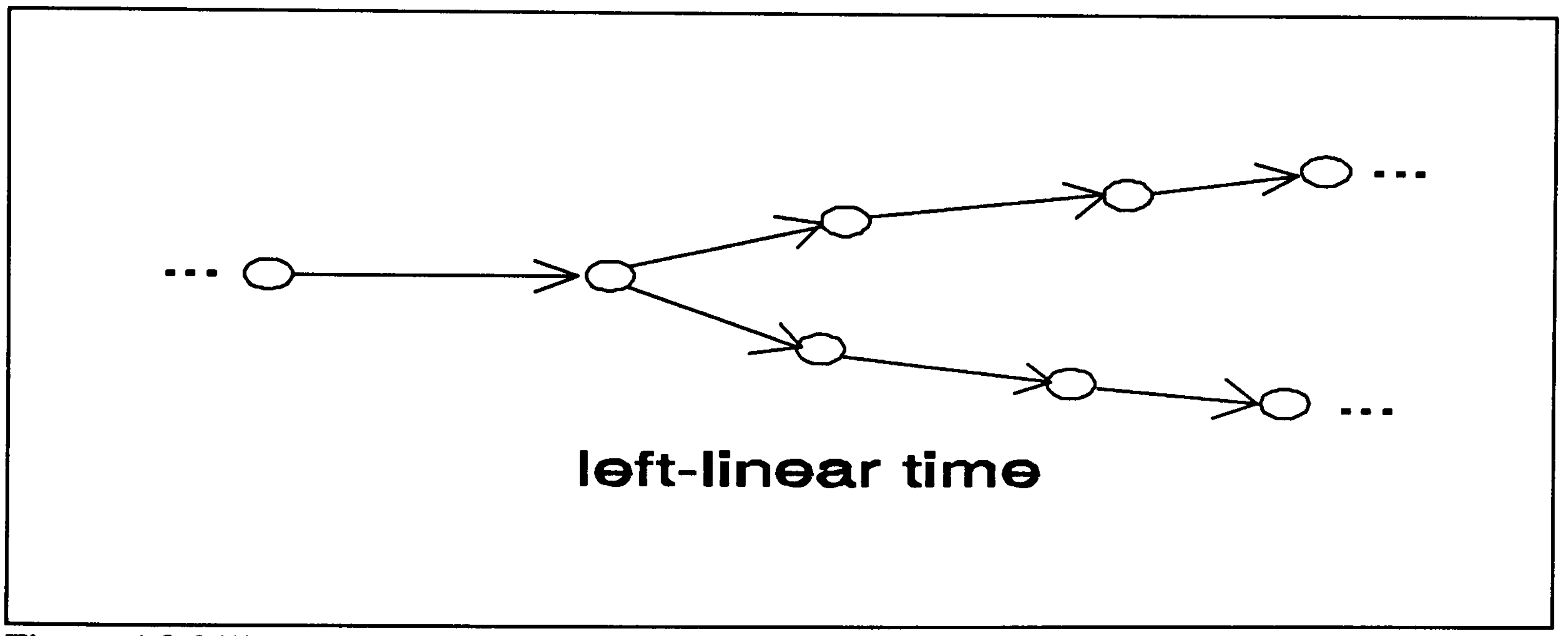

Figure 4.2.2(1)

Analogously, **right-linearity** is defined by means axiom $<A_{R\text{-Linear}}>$:

$<A_{R\text{-}Linear}>$

$$\forall t,t_1,t_2,t_3,t_4 \in \mathbf{T}(meets(t,t_1) \wedge meets(t_1,t_2) \wedge meets(t,t_3) \wedge meets(t_3,t_4) \Rightarrow$$
$$meets(t_1,t_4)$$
$$\vee\ \exists t' \in \mathbf{T}(meets(t_1,t') \wedge meets(t',t_4))$$
$$\vee\ \exists t'' \in \mathbf{T}(meets(t_3,t'') \wedge meets(t'',t_2)))$$

As Galton puts it in [Gal90a], it is interesting to note that **left-linearity** and **right-linearity** together just fail to imply **(full) linearity**, the exception being the case of parallel time lines as shown in Figure 4.2.2(2).

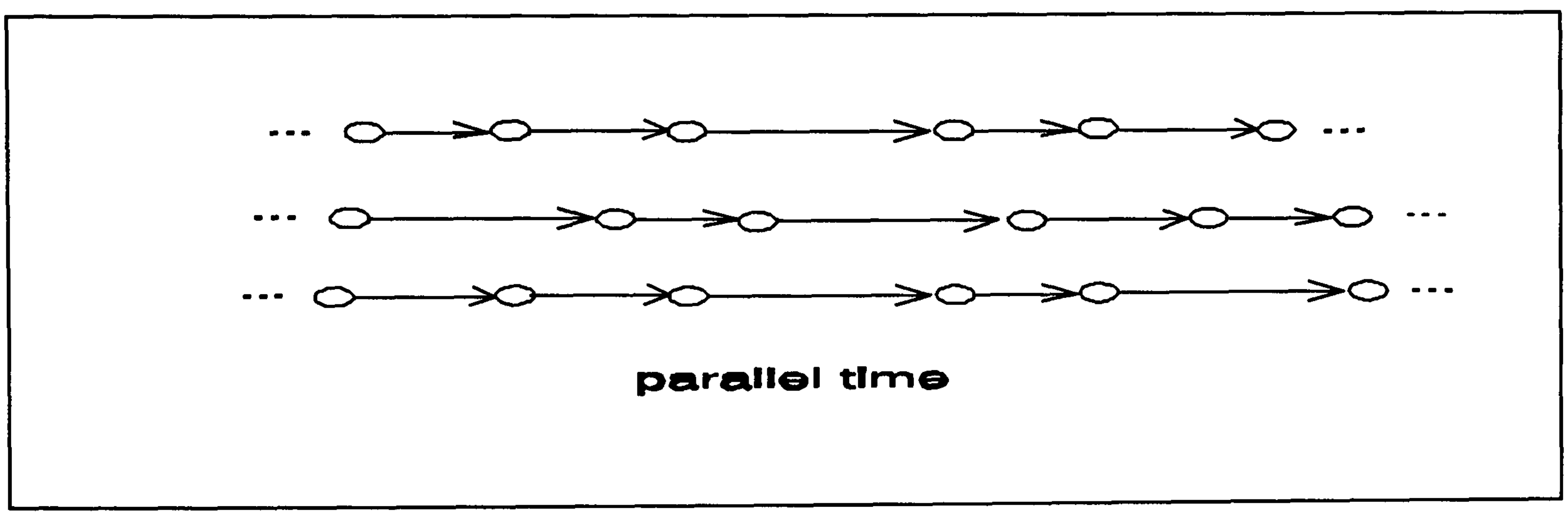

Figure 4.2.2(2)

Parallel temporal frames provide a way of modelling separate and asynchronous processes, and might prove useful in developing logics for reasoning about parallel computation and concurrent processes.

4.2.3 Dense and discrete time

According to Axiom <A2>, for each time-element t, there is a time-element which "*meets*" it, and another one which it "*meets*". Therefore, in particular, axiom <A4> and <A5> additionally ensure that, between any two distinct time points on the same time line, there is always a time interval. However, for time intervals, can we always assume that any interval can be decomposed into two distinct contiguous intervals? If so, we say that the set of time elements forms a dense system.

We may use the following axiom to characterise the density of a temporal frame (**T**,*meets*):

$<A_{Dense}>$

$\forall i \in \mathbf{I} \exists t_1, t_2 \in \mathbf{T}(i = t_1 \oplus t_2)$

We can show that axiom $<A_{Dense}>$ implies that each time interval can be decomposed into two distinct contiguous intervals. In fact, assume interval $i = t_1 \oplus t_2$; if t_1 is a point, then by axiom <A5>, t_2 must be an interval; hence, by $<A_{Dense}>$, $t_2 = t' \oplus t''$. By <A4> and <A3>, we get $i = t_1 \oplus t' \oplus t''$. Since t_1 is a point, axiom <A5> implies that t' must be an interval; hence $t_1 \oplus t'$ is an interval. If t'' is an interval then we have proved that i can be decomposed into two intervals, i_1 and t'', where $i_1 = t_1 \oplus t'$; in the case that t'' is a point, $<A_{Dense}>$ implies that $t' = t_1' \oplus t_2'$, and again, since t_1 and t'' are points, from <A5> we can infer that both t_1' and t_2' must be intervals; hence $i = j_1 \oplus j_2$, where $j_1 = t' \oplus t_1'$, and $j_2 = t_2' \oplus t''$ are two intervals. Similar discussion applies to the case that t_2 is a point which implies that t_1 must be an interval.

In fact, it will turn out that we may need a slightly stronger axiom to characterise a temporal frame in which there is always a time point during any time interval. We introduce it as below:

$<A_{\text{P-Dense}}>$

$\forall i \in \mathbf{I} \exists p \in \mathbf{P} \exists i_1, i_2 \in \mathbf{I}(i = i_1 \oplus p \oplus i_2)$

N.B. By consideration of axioms <A2> and <A5>, we can infer that axiom $<A_{\text{P-Dense}}>$ ensures that between any two distinct time points on the same time line, there is a third. It is easy to show that axiom $<A_{\text{P-Dense}}>$ is stronger than axiom $<A_{\text{Dense}}>$, that is, $<A_{\text{P-Dense}}>$ implies $<A_{\text{Dense}}>$, but not vice versa. (E.g., if we let

$$\mathbf{P} = \varnothing \text{ and } \mathbf{I} = \{ \ [a,b) \mid a,b \in \mathbf{R}, a < b \ \},$$

then we get a time frame which satisfies $<A_{\text{Dense}}>$ but not $<A_{\text{P-Dense}}>$.)

The **discreteness** of a temporal frame (**T**,*meets*) can be characterised by means of adding two axioms, $<A_{\text{L-Discrete}}>$ and $<A_{\text{R-Discrete}}>$ to **A**:

$<A_{\text{L-Discrete}}>$

$\forall t \in \mathbf{T} \exists t_1 \in \mathbf{T}(meets(t_1,t) \land \neg\exists t_2,t_3 \in \mathbf{T}(t_1 = t_2 \oplus t_3))$

$<A_{\text{R-Discrete}}>$

$\forall t \in \mathbf{T} \exists t_1 \in \mathbf{T}(meets(t,t_1) \land \neg\exists t_2,t_3 \in \mathbf{T}(t_1 = t_2 \oplus t_3))$

Axiom $<A_{\text{L-Discrete}}>$ entails the **left-discreteness** of a temporal frame by means of asserting that for each time element, there is a non-decomposable time element which is immediately before it; similarly, Axiom $<A_{\text{R-Discrete}}>$ entails the **right-discreteness** by means of asserting that for each time element, there is a non-decomposable time element which is immediately after it. Consider the case in which the set of time points is empty: by taking t to be a non-decomposable interval (or moment, termed by Allen and Hayes) in the above axioms, since t_1 is by definition a moment we see that <$A_{L\text{-Discrete}}$> or <$A_{R\text{-Discrete}}$> implies that each moment has a predecessor moment or successor moment respectively. Hence, Allen and Hayes' <M6> is inconsistent with the discreteness axioms.

It is interesting to note that there may exist temporal frames in which some intervals are finite sums of moments (see next chapter). This case is axiomatically consistent with our axiom <A5>, but not consistent with Allen and Hayes' <M6>, which implies that each decomposable interval must be infinitely decomposable.

4.3 Derived Temporal Relations over Time Elements

In terms of the primitive relation "*meets*", we may induce the complete set of possible relationships over time elements by means of the following definitions, including the "*meets*" relation itself:

$equal(t_1,t_2) \Leftrightarrow t_1 = t_2,$

$before(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(meets(t_1,t) \wedge meets(t,t_2)),$

$overlaps(t_1,t_2) \Leftrightarrow \exists t,t',t'' \in \mathbf{T}(t_1 = t' \oplus t \wedge t_2 = t \oplus t''),$

starts$(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(t_2 = t_1 \oplus t)$,

during$(t_1,t_2) \Leftrightarrow \exists t',t'' \in \mathbf{T}(t_2 = t' \oplus t_1 \oplus t'')$,

finishes$(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(t_2 = t \oplus t_1)$,

after$(t_1,t_2) \Leftrightarrow$ *before*(t_2,t_1),

overlapped-by$(t_1,t_2) \Leftrightarrow$ *overlaps*(t_2,t_1),

started-by$(t_1,t_2) \Leftrightarrow$ *starts*(t_2,t_1),

contains$(t_1,t_2) \Leftrightarrow$ *during*(t_2,t_1),

finished-by$(t_1,t_2) \Leftrightarrow$ *finishes*(t_2,t_1),

met-by$(t_1,t_2) \Leftrightarrow$ *meets*(t_2,t_1),

It is interesting to note that, since points are allowed now, the above 13 relations have somewhat different "feel" to Allen's 13 temporal relations between intervals. For instance, if i_1 and i_2 are open intervals separated by a point p, then we have *before*(i_1,i_2), although this situation looks very like i_1 "*meets*" i_2 in Allen's system. Again, if i_1 is right-closed, and i_2 is left-closed at point p, respectively, according to the above definitions, we have *overlaps*(i_1,i_2), but again it "looks" like the two intervals meeting. Additionally, from the above definitions, any open interval is "*during*" its closure. What all this means is that, taking both intervals and points as primitive time-elements, we have more than 13 significantly different relationships to considered, because, for example, from almost any point of view, the first case mentioned above (i.e., $meets(i_1,p) \wedge meets(p,i_2)$) is no more similar to the case of two intervals separated by a third interval (a necessary condition of *before* in Allen's system) than it is to the case of two intervals strictly meeting.

On the other hand, as Allen and Hayes show in [AlH89], all the thirteen relations may hold in the case that only intervals are taken as time elements. However, when we examine the general case where elements may also be points, some of these relationships hold and some do not hold.

For example, let $p \in \mathbf{P}$:

$meets(p,t_2)$ may hold for time elements $t_2 \in \mathbf{T}$ according to the axiomatisation.

However, consider the following case:

$overlaps(p,t_2) \Leftrightarrow \exists t,t',t'' \in \mathbf{T}(p = t' \oplus t \wedge t_2 = t \oplus t'')$,

On the one hand, by axiom <A6>, $dur(p) = dur(t') + dur(t)$; and the assumption that p is a point gives:

$$dur(t') + dur(t) = dur(p) = 0 \qquad (1)$$

On the other hand, axiom <A5> ensures that at least one of t' and t is an interval, hence:

$$dur(t') + dur(t) > 0 \qquad (2)$$

(1) and (2) show that *overlaps*(p,t_2) can not hold.

It is straightforward to prove in a similar fashion that all the possible relations over intervals and points may be classified into the following four groups:

Point - Point:

{*equal, before, after*}

which relate points to other points;

Interval - Interval:

{*equal, before, meets, overlaps, starts, during, finishes, finished-by, contains, started-by, overlapped-by, met-by, after*}

which relate intervals to intervals;

Point - Interval:

{*before, meets, starts, during, finishes, met-by, after*}

which relate points to intervals;

Interval - Point:

{*before, meets, finished-by, contains, started-by, met-by, after*}

which relate intervals to points.

N.B. According to the above classification, there are totally 30 possible temporal relations over time-elements which may be both intervals and points. However, in [Vil82,ViK86], Vilain and Kautz have just proposed 26 of these 30 temporal relations. There is a critical omission from the primitive relations between points and intervals in Vilain's system, for the "*meets*" relation is defined only between intervals and is not allowed between points and intervals. This omission leads to some difficulties in modelling the "open" and "closed" nature of intervals, and in reasoning correctly about continuous change. For example, how to express the motion of a ball thrown into the air (see section 3.4)?

CHAPTER 5

MODELS OF THE THEORY

Since the time theory proposed in chapter 4 characterises a very general temporal structure, we may interpret the axiomatisation in various temporal models: dense or discrete, linear or branching, interval-based, point-based, or interval and point-based, etc.

5.1 A Dense Linear Model

As an example of dense and linear models of the axiomatisation, consider an "obvious" interpretation in which the set of time points, **P**, is the set of all real numbers; and the set of time intervals, **I**, is the set of periods which are constructions over all possible point-pairs, $p_1,p_2 \in \mathbf{P}$ such that $p_1 < p_2$, with the following structures:

$$(p_1,p_2,\text{open},\text{open}) =_{def} \{\, r \in \mathbf{R} \mid p_1 < r < p_2 \},$$

$$(p_1,p_2,\text{open},\text{closed}) =_{def} \{\, r \in \mathbf{R} \mid p_1 < r \leq p_2 \},$$

$$(p_1,p_2,\text{closed},\text{open}) =_{def} \{\, r \in \mathbf{R} \mid p_1 \leq r < p_2 \},$$

$(p_1,p_2,\text{closed},\text{closed}) =_{def} \{ r \in \mathbf{R} \mid p_1 \leq r \leq p_2 \}$,

where "<" and "≤" are the ordinary ordering relations over the set, **R**, of real numbers.

N.B. Here, we represent the interval structure by means of the extra primitives: left-type, l, and right-type, r, which take values from a set **Type** $=_{def}$ {open, closed}. There are thus four types of intervals based on points. For convenience of expression, we may identify a point p with (p,p,closed,closed), that is, a special segment whose left end-point and right end-point are identical, with "closed" type for both left-type and right-type.

The duration assignment function, *dur*, can be simply defined by:

$$dur((p_1,p_2,\_,\_)) = p_2 - p_1.$$

We may define the *meets* relation over time elements as following:

$$meets((p_{11},p_{12},l_1,r_1),\ (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$$

$$p_{12} = p_{21} \wedge r_1 = \text{open} \wedge l_2 = \text{closed}$$

$$\vee\ p_{12} = p_{21} \wedge r_1 = \text{closed} \wedge l_2 = \text{open}$$

It is easy to see that this model satisfies axioms <A1> - <A6>. Additionally, the (full) linearity axiom, $<A_{Linear}>$, and the dense axiom, $<A_{Dense}>$, are also satisfied. Hence, the above structure forms a dense and linear temporal model of the theory.

5.2 A Discrete Linear Model

A discrete model satisfying axioms <A1> - <A6>, <A_{Linear}>, <$A_{L\text{-}Discrete}$> and <$A_{R\text{-}Discrete}$> can be constructed by simply limiting all elements of **P** to be integers in the model presented in the above section, although the internal points of intervals are still reals. It is interesting to note that in such a discrete model, although points never meet each other, intervals are not necessarily infinitely decomposable. For instance, according to our axiomatisation, interval (6,8,open,closed) can be only decomposed into at most four (atomic) time elements:

(6,8,open,closed) =

(6,7,open,open)

⊕ (7,7,closed,closed)

⊕ (7,8,open,open)

⊕ (8,8,closed,closed)

However, this model will not be valid for Allen and Hayes' axiomatisation including <M6> (see section 3.4), which implies that if an interval is decomposable then it must be infinitely decomposable. (Otherwise, if it is only finitely decomposable, then it must be the sum of a finite number of moments which would meet one another, contrary to <M6>.)

N.B. As mentioned in section 3.4, in order to interpret Allen and Hayes' axioms in discrete models, their axiom <M6> must be excluded. In another word, axiom <M6> is inconsistent with discrete times. However, the above example shows that the axiom <A5> in our axiomatisation can be satisfied by discrete models.

5.3 Temporal System as Subsumed Models

In what follows, we shall show that our axiomatisation is powerful enough to subsume many representative temporal systems, such as: the point based systems of Bruce, of McDermott, Allen's logic of intervals and Galton's revised theory, and the point and interval based theories of Vilain, of Knight and Ma.

5.3.1 Bruce's point-based system

Bruce's *time-system* is simply a set of time points with a partial order (see section 3.1). In our theory, we may define a partial order, "≤", over the set of points, **P**, as:

$$p_1 \leq p_2 \Leftrightarrow \mathit{Equal}(p_1,p_2) \vee \mathit{Before}(p_1,p_2),$$

where *Equal* and *Before* are introduced as in section 4.3. Hence, the sub-frame, (**P**,≤), of the temporal frame (**T**,*meets*) defined by the axiomatisation, forms a temporal system of Bruce.

In a similar way, we may define Bruce's 7 binary relations over *time-segments* (see [Bru72]), in terms of the temporal relations over intervals introduced in section 4.3.

> N.B. As discussed in the introduction, the temporal theories of Ladkin [Lad86,87,92], of Dechter et al. [DMP91], and of Maiocchi [Mai92] are similar to that of Bruce in the sense that intervals are defined to be constructed out of points. Hence, in a similar way, we may induce the corresponding time model for each of these temporal

frameworks.

5.3.2 McDermott's temporal logic

McDermott's theory assumes a "*no later than*" ordering relation over a dense collection of states (points), which is axiomatised to give rise to a left linear (branching into future) time structure (see section 3.3). Consider the temporal frame axiomatised by axioms <A1>-<A6>, <$A_{L\text{-}linear}$>, and the stronger dense axiom <$A_{P\text{-}Dense}$>. As for Bruce's partial order, we may also define the "no later than" relation over time points in terms of relations *Equal* and *Before*. In this way, we may take McDermott's time structure as a model of the above theory by addressing only time points and the "no later than" relation, while the left-linearity axiom <$A_{L\text{-}linear}$> axiomatises the characteristic that time branches only in future for McDermott's logic.

5.3.3 Allen's interval based model

Since the axiomatisation proposed in this paper may be seen as an extension of Allen and Hayes' interval based temporal theory [AlH89], it is straightforward to subsume Allen and Hayes' theory by taking the set of time points to be empty, and including the linearity axiom <A_{Linear}> in the fundamental axiomatisation. Of course, in this case, axiom <A5> will become vacuous.

N.B. Further examination of Allen's interval based temporal logic and Galton's corresponding revisions will be given later in next chapter (6.3.2).

5.3.4 Vilain's interval & point-based system

Vilain's system based on both intervals and points is arrived at by expanding Allen's 13 temporal relations over intervals to 26, which are primitively defined to relate points to points, intervals to intervals, intervals to points, and points to intervals (see section 4.3). It is interesting to note that Vilain's 26 temporal relations form a subset of the set of 30 relations we introduced in section 4.3. The four relations missing from Vilain's system are: *meets*, *met_by* that relate points to intervals, and *meets, met_by* that relate intervals to points (see N.B. in section 4.3). Hence, if we employ the following more strict axiom instead of <A5>:

$$\forall t_1,t_2 \in \mathbf{T}(meets(t_1,t_2) \Rightarrow t_1 \in \mathbf{I} \wedge t_2 \in \mathbf{I})$$

then we get Vilain's temporal system. The above axiom ensures that if two time elements meet each other, then both of them must be intervals.

5.4 A Time Network for Computer-based Systems

In this section, we concentrate on a special finite model of the theory. The choice of finiteness of time elements in this model is forced by the practicality of the computer based modelling approach. Ordinarily, in computer systems, we have to store information as a discrete (finite) set, and so the semantics of any database of time elements will naturally assume a well-order at some fundamental level. Hence, the computer-based temporal system may be viewed as another model of the theory, in the form of a finite (discrete) set of temporal facts.

5.4.1 Definitions of a finite time network

Assume (**T**,*meets*) is a temporal frame satisfying axioms <A1> - <A6>, A_{Linear}, <$A_{L\text{-}Discrete}$> and <$A_{R\text{-}Discrete}$>. The discreteness property of the temporal frame allows us to form a nonempty finite set $\mathbf{E} \subset \mathbf{T} = \mathbf{I} \cup \mathbf{P}$, such that:

i) $\mathbf{E} = \{t_1, t_2, ..., t_m\}$;

ii) $meets(t_i,t_{i+1})$, i = 1, 2, ..., m-1;

iii) $meets(t_i,t_{i+1}) \Rightarrow t_i \in \mathbf{I} \vee t_{i+1} \in \mathbf{I}$.

These theorems precisely characterise a finite series, **E**, of time elements, which is *similar* to an *initial segment* of the set of natural numbers with the natural order (see [Lip64]), with an *immediate successor* relation. Additionally, it is easy to see that the limitation of axioms <A4>, <A5> and <A6> onto **E** well define the *closure* $\mathbf{C_E}$ of E, under the binary operations of combining adjacent time elements and corresponding addition of duration. For convenience, we call (**E**, M_E, D_E) a *fundamental time network*, and ($\mathbf{C_E}$, M_C, D_E), the *complete time network* corresponding to (**E**, M_E, D_E), where M_E, M_C are the *meets* relations, D_E, D_C are the *duration assignments* over **E** and $\mathbf{C_E}$, respectively. It is clear that the limitation of M_C to **E** is M_E, and the limitation of D_C to **E** is D_E.

The set $\mathbf{C_E}$ includes **E** and all the intervals and points which can be formed from it by means of $\oplus$ and +. However, in an application neither the fundamental set **E** nor the complete set

$\mathbf{C_E}$ may be known. A database of "facts" about $\mathbf{C_E}$ will express knowledge that is incomplete in several ways. For example, the database may contain knowledge of duration assignments for only some of its elements, and may have incomplete knowledge about the *meets* relation. In addition, the database will often contain redundancy, as when facts are known about two elements without the knowledge that they are actually identical. For example, we may know that *meets*(a,b) and *dur*(c) = 1, without knowing that a and c are the same element. To allow for possible duplicate elements, the basic structure of the database is that of a bag, rather than a set.

Accordingly, we use a triad $(\mathbf{K}, M_{\mathbf{K}}, D_{\mathbf{K0}})$ to denote an *(incomplete) time network*, where:

i) $\mathbf{K} = \mathrm{K}_1 \uplus \mathbf{K}_2 \uplus ... \uplus \mathbf{K}_p$, where $\mathbf{K}_i \subseteq \mathbf{K}_{i+1} \subseteq \mathbf{C}_E$, i = 1, ..., p-1; and "$\uplus$" represents the *bag union* (For bag notation adopted here, see [Dil90]);

ii) $M_{\mathrm{K}} = M|_{\mathrm{K1}} \uplus M|_{\mathrm{K2}} \uplus ... \uplus M|_{\mathrm{Kp}}$;

iii) $D_{\mathbf{K0}} \sqsubseteq D_{\mathbf{K}} = D|_{\mathrm{K1}} \uplus D|_{\mathrm{K2}} \uplus ... \uplus D|_{\mathrm{Kp}}$; here, "$\sqsubseteq$" represents the *sub-bag* relation.

Nb. i) expresses our knowledge of what time elements are there;

ii) expresses our knowledge as to how the time elements in $\mathbf{K}$ meet each other;

iii) expresses our knowledge of duration over a sub-bag $\mathbf{K_0}$ of $\mathbf{K}$.

5.4.2 Formal characterisations of the graph of a time network

In this section, we introduce a formal graphical representation of the time network characterised above. The graph is one in which time elements are represented by directed arcs. The *meets* relations are represented by the nodes of the graph: if *meets*(k_1,k_2) then k_1 is the in-arc to a node, and k_2 is the out-arc from the node. All time elements which are known to meet k_2 will be in-arcs to the node, and all time elements which k_1 meets will be out-arcs from the node. Although this representation is intuitively straightforward, the following formal definition of nodes is more involved.

Some difficulty is encountered for nodes with only in or out arcs (since in a finite model, there are some time elements in the network that seems to be the "earliest" or the "latest" ones, although in the theory, axiom <A2> assumes that time does not start or stop), but this can be resolved by extending the equivalence relation defined below to include these, by means of the final clause in *Eq_in* and *Eq_out*.

In order to give a proper definition of the nodes of the graph, at first, we define two kinds of equivalence relations over time elements, *Eq_in* and *Eq_out*, in the following meaning:

$$\forall k_1,k_2 \in \mathbf{K}(k_1 \; Eq\_in \; k_2 \Leftrightarrow \exists k \in \mathbf{K}(meets(k_1,k) \wedge meets(k_2,k)) \vee k_1 = k_2)$$

$$\forall k_1,k_2 \in \mathbf{K}(k_1 \; Eq\_out \; k_2 \Leftrightarrow \exists k \in \mathbf{K}(meets(k,k_1) \wedge meets(k,k_2)) \vee k_1 = k_2)$$

Intuitively, *Eq_in* designs a class of time elements known to meet a common element, and

Eq_out designs a class of time elements known to be met by a common element.

According to these two kinds of equivalence relations, we get the equivalence classes of time elements:

$$\mathbf{K}_{Eq\_in,1},\ \mathbf{K}_{Eq\_in,2},\ \ldots\ldots,\ \mathbf{K}_{Eq\_in,s}$$

and

$$\mathbf{K}_{Eq\_out,1},\ \mathbf{K}_{Eq\_out,2},\ \ldots\ldots,\ \mathbf{K}_{Eq\_out,t}$$

We can now define nodes as pairs of equivalence classes:

$$\text{Node}(\mathbf{K}_X,\mathbf{K}_Y) \Leftrightarrow \text{in\_Node}(\mathbf{K}_X,\mathbf{K}_Y) \vee \text{mid\_Node}(\mathbf{K}_X,\mathbf{K}_Y) \vee \text{out\_Node}(\mathbf{K}_X,\mathbf{K}_Y),$$

where

$$\text{in\_Node}(\emptyset,\mathbf{K}_{\text{Eq\_out},j}) \Leftrightarrow \forall k_i \in \mathbf{K}_{\text{Eq\_out},i}(\neg\exists k \in \mathbf{K}(\mathit{meets}(k,k_i)))$$

$$\text{mid\_Node}(\mathbf{K}_{\text{Eq\_in},i},\mathbf{K}_{\text{Eq\_out},j}) \Leftrightarrow \exists k_i \in \mathbf{K}_{\text{Eq\_in},i}\exists \mathbf{k}_j \in \mathbf{K}_{\text{Eq\_out},j}(\mathit{meets}(k_i,k_j))$$

$$\text{out\_Node}(\mathbf{K}_{\text{Eq\_in},i},\emptyset) \Leftrightarrow \forall k_i \in \mathbf{K}_{\text{Eq\_in},i}(\neg\exists k \in \mathbf{K}(\mathit{meets}(k_i,k)))$$

As an example, consider knowledge represented by ($\mathbf{K}$, $M_{\mathbf{K}}$, $D_{\mathbf{K}0}$), where:

$$\mathbf{K} = [\![\ k_{12},\ k_{23},\ k_{24},\ k_{34},\ k_{35},\ k_{45},\ k_{56},\ k_{57}\ ]\!],$$

$$M_{\mathbf{K}} = [\![\ \mathit{meets}(k_{12}, k_{23}),\ \mathit{meets}(k_{12}, k_{24}),\ \mathit{meets}(k_{23}, k_{34}),$$
$$\mathit{meets}(k_{23}, k_{35}),\ \mathit{meets}(k_{24}, k_{45}),\ \mathit{meets}(k_{34}, k_{45}),$$
$$\mathit{meets}(k_{35}, k_{56}),\ \mathit{meets}(k_{35}, k_{57}),\ \mathit{meets}(k_{45}, k_{56})\ ]\!],$$

$$D_{\mathbf{K}0} = [\![\ \mathit{dur}(k_{35})=1,\ \mathit{dur}(k_{45})=1,\ \mathit{dur}(k_{57})=0\ ]\!].$$

According to the equivalence relations, *Eq_in* and *Eq_out*, defined above, we get the equivalence classes of the time elements in $(\mathbf{K}, M_{\mathbf{K}}, D_{\mathbf{K}0})$ as below:

$\mathbf{K}_{Eq\_in,1} = \{ k_{12} \}$,

$\mathbf{K}_{Eq\_in,2} = \{ k_{23} \}$,

$\mathbf{K}_{Eq\_in,3} = \{ k_{24}, k_{34} \}$,

$\mathbf{K}_{Eq\_in,4} = \{ k_{35}, k_{45} \}$,

$\mathbf{K}_{Eq\_in,5} = \{ k_{56} \}$,

$\mathbf{K}_{Eq\_in,6} = \{ k_{57} \}$;

and

$\mathbf{K}_{Eq\_out,1} = \{ k_{12} \}$,

$\mathbf{K}_{Eq\_out,2} = \{ k_{23}, k_{24} \}$,

$\mathbf{K}_{Eq\_out,3} = \{ k_{34}, k_{35} \}$,

$\mathbf{K}_{Eq\_out,4} = \{ k_{45} \}$,

$\mathbf{K}_{Eq\_out,5} = \{ k_{56}, k_{57} \}$.

Hence, we can form seven nodes, n_1, n_2, ..., and n_7, in terms of seven pairs of equivalence classes, $(\varnothing, \mathbf{K}_{Eq\_out,1})$, $(\mathbf{K}_{Eq\_in,1}, \mathbf{K}_{Eq\_out,2})$, $(\mathbf{K}_{Eq\_in,2}, \mathbf{K}_{Eq\_out,3})$, $(\mathbf{K}_{Eq\_in,3}, \mathbf{K}_{Eq\_out,4})$, $(\mathbf{K}_{Eq\_in,4}, \mathbf{K}_{Eq\_out,5})$, $(\mathbf{K}_{Eq\_in,5}, \varnothing)$ and $(\mathbf{K}_{Eq\_in,5}, \varnothing)$, respectively.

Here, n_1 is a *in_Node*, n_2, n_3, n_4 and n_5 are *mid_Nodes*, and n_6 and n_7 are *out_Nodes*. Hence, the network may be represented by a graph as in Figure 5.4.2(1).

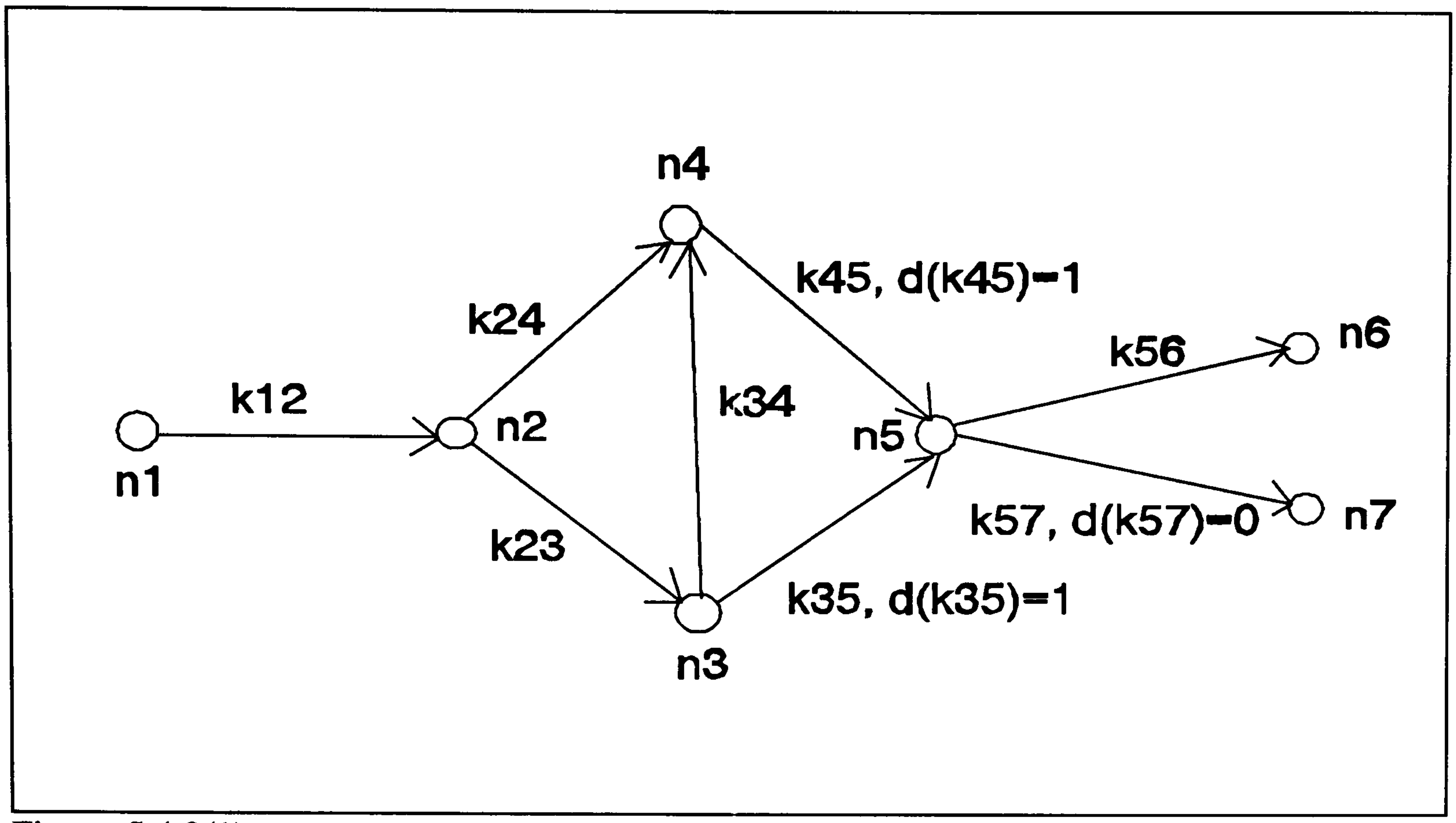

Figure 5.4.2(1)

5.4.3 A necessary and sufficient condition for consistency

To draw inferences from an incomplete time network (**K**, M_K, D_{K0}), we must rely on the assumed properties characterised by the definitions given in section 5.4.1. A consistency checker is needed which will establish whether a triad (**K**, M_K, D_{K0}) is consistent with our basic assumptions about the time network.

In general, a triad (**K**, M_K, D_{K0}) is consistent if we can add to **K** and make any necessary equality assignments, and add to M_K and to D_{K0}, so that the resulting triad (**C**, M_C, D_C) is the closure for some (**E**, M_E, D_E), under the binary operations of combining adjacent time elements and corresponding addition of duration. A necessary and sufficient condition for consistency may be given in terms of the graphical representation introduced in the above section. For convenience, we adopt the notation that k_{ij} represents an arc from node n_i to node n_j, and d_{ij} represents the duration of this arc. We let G be the graph of $(\mathbf{K}, M_{\mathbf{K}}, D_{\mathbf{K}0})$.

Let **Node** = { $n_1, n_2, \ldots\ldots, n_S$ } be the nodes in $G_{\mathbf{K}}$. The network $(\mathbf{K}, M_{\mathbf{K}}, D_{\mathbf{K}0})$ is consistent if and only if:

(I) There is a solution $(x_{i1,j1}, \ldots, x_{iq,jq})$ for unknown durations $(X_{i1,j1}, \ldots, X_{iq,jq})$ which forms a $D_{\mathbf{K}} \sqsupseteq D_{\mathbf{K}0}$, where $x_{i,j} \geq 0$, such that:

(I.1) for each simple circuit in $G_{\mathbf{K}}$, the directed sum of weights is zero.

(I.2) $dur(k_{ij}) + dur(k_{jh}) > 0$, for all i, j, h.

Otherwise, the network is inconsistent.

<u>*Proof of sufficiency:*</u>

a) We first show that if (I) holds, then a function f of **Node** into **R** exists:

n ($\in$ **Node**) -------> f(n) ($\in$ **R**), such that:

(II.1) If $k_{ij} \in \mathbf{K}$, then:

$f(n_i) \leq f(n_j)$,

$f(n_j) - f(n_i) = dur(k_{ij}) \in D_K$;

(II.2) If $k_{ij}, k_{jh} \in \mathbf{K}$, then:

$f(n_h) - f(n_i) > 0$.

Nb. condition (II.2) implies that: $k_{ij} \in \mathbf{I} \vee k_{jh} \in \mathbf{I}$, which is indeed the constraint iii) in section 5.4.1, stating that no two points meet each other.

To show this, we assume G_K to be connected by means of *meets* (the extension to a graph with several connected components is straightforward).

Let y_{ij} denote the duration assignment for $k_{ij} \in \mathbf{K}$, where

$y_{ij} = d_{ij}$, if $d_{ij} \in D_{K0}$; otherwise, $y_{ij} = x_{ij}$.

Now take a directed spanning tree of G_K (i.e. a tree joining all the nodes of G_K, formed by removing some arcs from G_K, where the directed arcs of the spanning tree are as same as those appearing in G_K). Selecting any node n_o as origin, a unique semi-path is determined by the spanning tree between n_o and any other node n (Figure 5.4.3(1)). We may take f(n) as the directed sum of the weighted arcs from n_o to n along this path.

With this assignment, condition (II.1) follows immediately for all arcs on the spanning tree. For any arc k_{ij} not on the spanning tree, we consider the circuit formed by k_{ij} together with

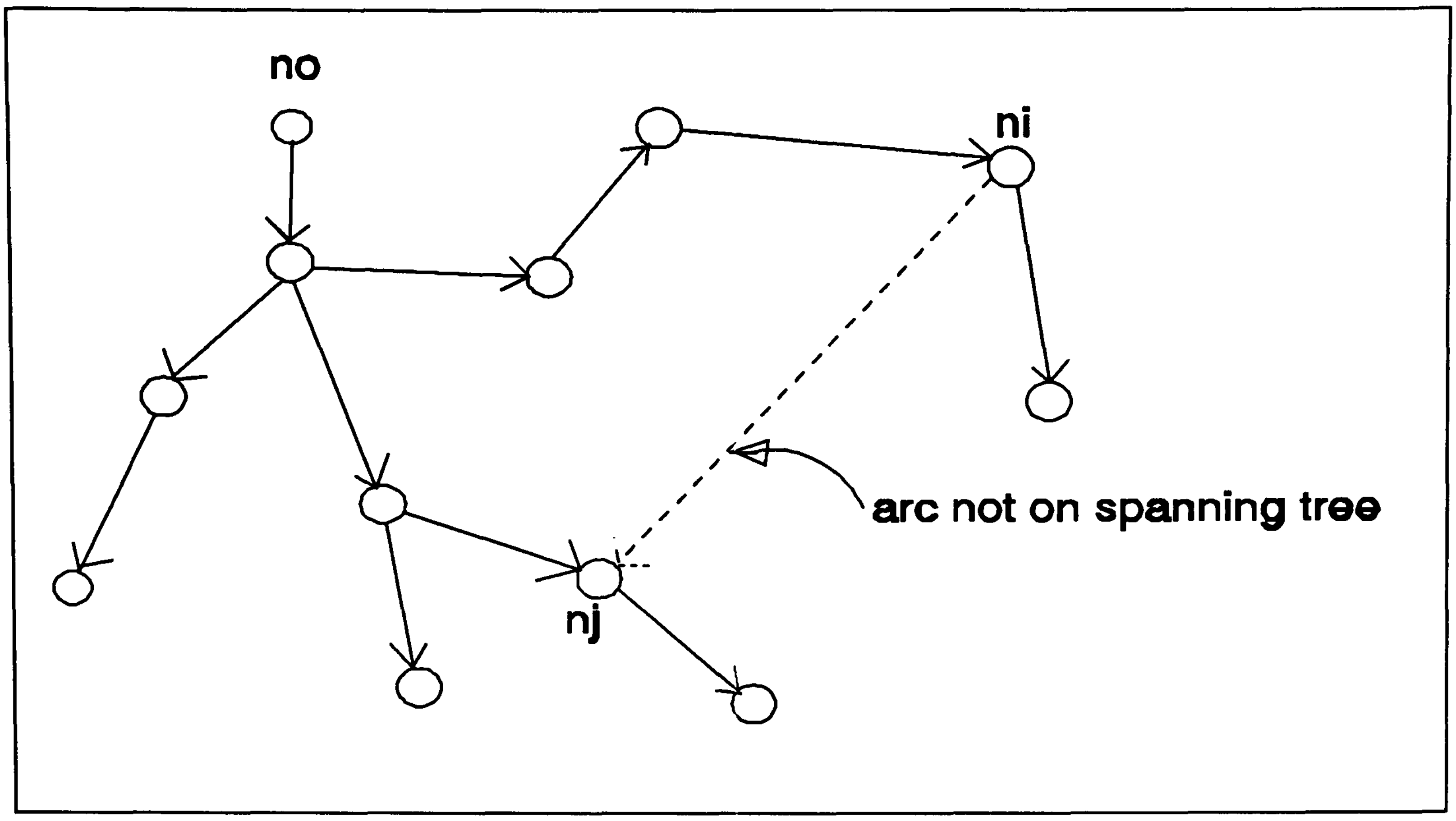

Figure 5.4.3(1)

the spanning tree. Applying condition (I.1), we have:

$$y_{ij} - f(n_j) + f(n_i) = 0,$$

i.e. (II.1) again holds.

Additionally, it is clear that condition (I.2) $\Leftrightarrow$ (II.2).

b) We now show that f(n) may be used to construct (**E**, M_E, D_E). In effect, the function assigns a time measure to the nodes. However, care must be taken to deal with points: if a number of nodes are assigned the same f(n), then we must be sure that we can construct an **E** without two consecutive points. In the procedure that follows, we show how this may be done.

(1) Define equivalent classes N_1, N_2, ... , N_{s1}, $s_1 \leq s$ among **Node** so that:

n_i, n_j belong to the same class $N_r \Leftrightarrow f(n_i) = f(n_j)$;

(2) The nodes within any class N_i are of three types:

(i) those that are in-nodes to zero duration arcs in **K**;

(ii) those that are out-nodes to zero duration arcs in **K**;

(iii) those that are not in- or out- nodes to zero duration arcs in **K**.

Condition I.2 ensures that there are no nodes that are both in-node and out-node to two zero duration arcs. The in-node and out-node to a zero duration arc will be in the same equivalence class, and the in-node must be ordered before the out-node. Accordingly we subdivide each class N_i into two subsets: N_i^1 containing nodes of type (i) and N_i^2 containing nodes of type (ii) and (iii).

(3) The graph of **E** is now formed over the set of subclasses as nodes. The successor relation is defined by the natural ordering of equivalence classes according to f, and by the rule that N_i^2 is the "*successor*" to N_i^1. Duration assignment to **E** is defined by $dur((N_j, N_{j+1})) = f(N_{j+1}) - f(N_j)$, where N_{j+1} is the successor to N_j in G_K.

c) Finally we show that $(\mathbf{K}, M_K, D_{K0})$ is in the closure of $(\mathbf{E}, M_E, D_E)$. We let e_{lm}^{rs} be the arc in the closure of **E** between node N_l^r and N_m^s. We make the following equality assignments over **K**:

$k_{ij} = e_{lm}^{rs}$ if $n_i \in N_l^r$ and $n_j \in N_m^s$, With this assignment, k_{ij} is in the closure of **E**, and

$$dur(e_{lm}^{rs}) = f(\mathbf{N}_m^s) - f(\mathbf{N}_l^r) = d_{ij}.$$

Proof of necessity:

If (**E**, M_E, D_E) exist, and (**K**, M_K, D_{K0}) is in the closure of (**E**, M_E, D_E), then condition (I.1) holds straightforwardly, while constraint iii) implies condition (I.2).

we may use the example given in above section again to illustrate the procedure of establishing the fundamental triad (**E**, M_E, D_E):

There are two elementary circuits in G_K to consider. Setting the directed sum of weights in each of these equal to zero, we get 2 independent constraints:

$$dur(k_{34}) + dur(k_{45}) - dur(k_{35}) = 0,$$

$$dur(k_{23}) + dur(k_{34}) - dur(k_{24}) = 0.$$

By inspection, one consistent solution is:

$$dur(k_{34}) = 0,\ dur(k_{23}) = 1,\ dur(k_{24}) = 1,\ dur(k_{12}) = 0,\ dur(k_{56}) = 0.$$

(N.B. There may be other consistent solutions, for instance:

$dur(k_{34}) = 0$, $dur(k_{23}) = 1.8$, $dur(k_{24}) = 1.8$, $dur(k_{12}) = 3$, $dur(k_{56}) = 10$.)

Correspondingly, let **Node** = $\{n_1, n_2, n_3, n_4, n_5, n_6, n_7\}$ be the nodes in graph (**K**, M_K, D_{K0}), as shown in Figure 5.4.3(2):

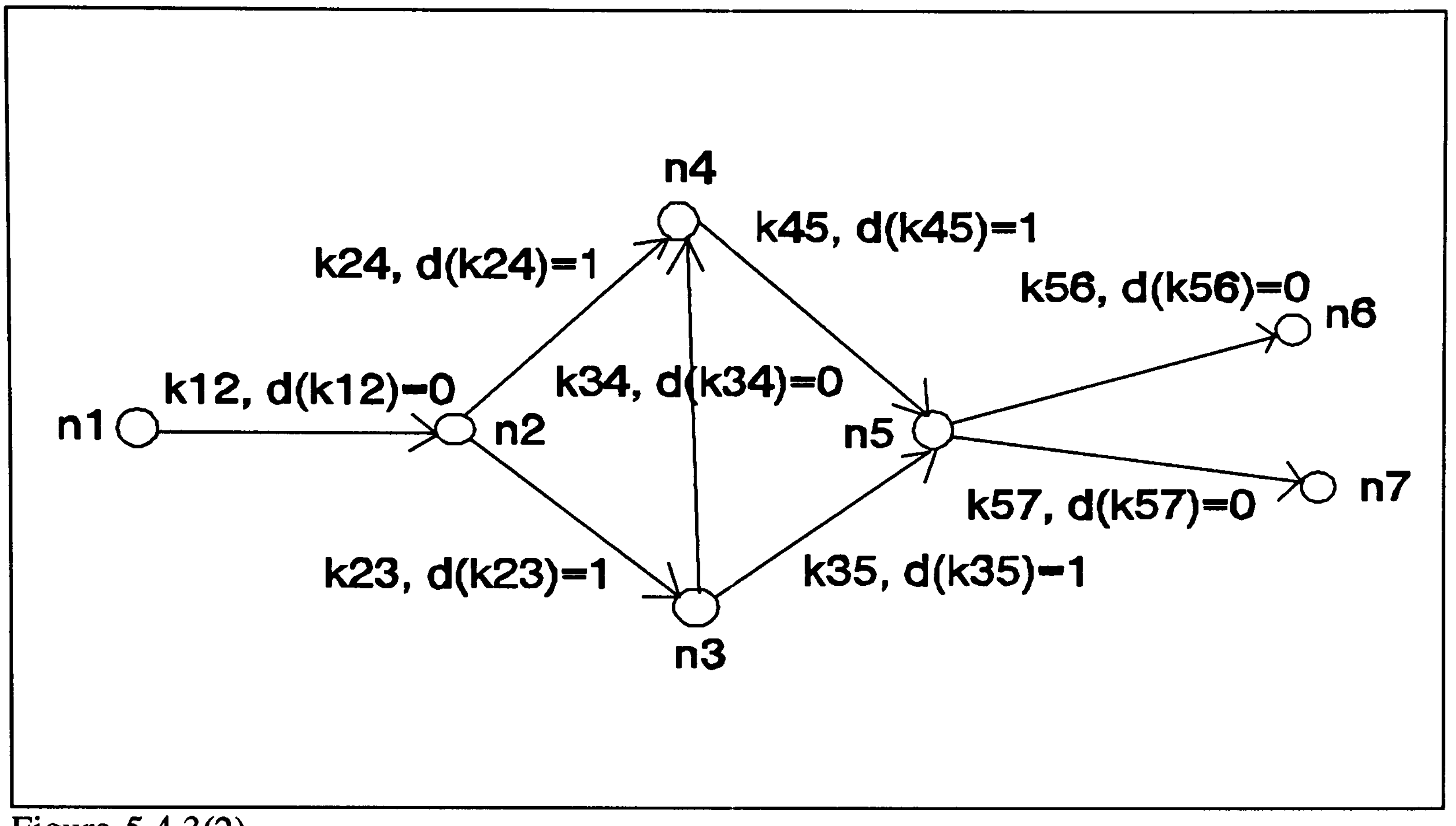

Figure 5.4.3(2)

the function f of **Node** into **R** may be defined as:

$f(n_1) = 0$, $f(n_2) = 0$, $f(n_3) = 1$, $f(n_4) = 1$,

$f(n_5) = 2$, $f(n_6) = 2$, $f(n_7) = 2$.

which satisfies conditions I.1 and I.2.

(1) equivalent classes:

$N_1 = [\![n_1, n_2]\!]$,

$N_2 = [\![n_3, n_4]\!]$,

$N_3 = [\![n_5, n_6, n_7]\!]$;

(2) $N_1^1 = [\![n_1]\!]$, $N_1^2 = [\![n_2]\!]$, $N_2^1 = [\![n_3]\!]$,

$N_2^2 = [\![n_4]\!]$, $N_3^1 = [\![n_5]\!]$, $N_3^2 = [\![n_6, n_7]\!]$

(3) $\mathbf{E} = \{k_{12}, k_{23}, k_{34}, k_{45}, k_{57}\}$,

$M_E = \{meets(k_{12}, k_{23}), meets(k_{23}, k_{34}), meets(k_{34}, k_{45}), meets(k_{45}, k_{57})\}$;

$D_E = \{dur(k_{12})=0, dur(k_{23})=1, dur(k_{34})=0, dur(k_{45})=1, dur(k_{57})=0\}$.

It is easy to see ($\mathbf{E}$, M_E, D_E) satisfies the conditions given in section 5.4.1.

5.4.4 A limited case of the time network

In Allen's interval-based system, no comprehensive mechanism for duration reasoning has been proposed. In the case of modelling a finite set of temporal events, we may take Allen's system as a limited case of a time network defined above, which satisfies axiom $<A_{Linear}>$, but without any actual duration constraints on time intervals. The differentiating property between intervals and points (and also "moments" as termed by Allen and Hayes [AlH89]) is that intervals are allowed to be decomposable, but points are not. We denote this limited model to be a triad ($\mathbf{K}$, M_K, $\varnothing$), or simply as a pair ($\mathbf{K}$, M_K), where $\mathbf{K}$ and M_K are defined as in section 5.4.1, excluding anything related to duration reasoning.

A consistency checker for a limited time network ($\mathbf{K}$, M_K) may be given in terms of its graphical representation: let G_K be the graph of ($\mathbf{K}$, M_K), then ($\mathbf{K}$, M_K) is consistent if and only if:

(I.1) G_K^r is acyclic, where G_K^r is the associated *reduced graph* formed from G_K by merging two nodes connected by a point in G_K and removing the corresponding arc.

(I.2) there are no nodes that are both in-node and out-node to two point-arcs in G_K.

Otherwise, the network is inconsistent.

Proof of sufficiency:

Since G_K^r is not cyclic, by a standard result in graph theory (see Car[79]), we can show that the nodes in G_K can be numbered with integers so that the natural order of the integers is consistent with the relations of "*meets*" over the corresponding time elements. A procedure for this numbering for any acyclic graph G_K^r is:

i) Set variable n = 1

ii) Select any node in the reduced graph G_K^r without in-arc. Such a node exists since G_K^r is acyclic (See [Car79], or any standard text on graph theory).

iii) Number this node n.

iv) Remove this node and associated arcs from G_K^r to form graph $G_K^{r'}$. $G_K^{r'}$ is also acyclic. Set G_K^r to $G_K^{r'}$, increment n by **2** if the deleted node is formed from a pair of nodes in G, otherwise, increment n by **1**.

v) Repeat from ii) until G_K^r is empty.

vi) Form arcs between consecutive integer nodes. In the case that integer n+1 is missed between n and n+2 in the reduced graph G_K^r, then the consecutive integers n and n+1 are associated with the corresponding pair of nodes in G_K, the (K, M_K)-graph.

Then the arcs between consecutive integer nodes form the set **E**, and M_E is formed by the natural order over these integers. Additionally, any element of **K** is an ordered union of some time elements in **E**. Finally, the closure ($\mathbf{C}_E$, M_C) can be formed under the binary operations of combining adjacent time elements. Hence, the network is consistent.

<u>*Proof of necessity:*</u>

The necessity of the consistency condition is straightforward from axioms <A5> and <A_{Linear}>.

Hence the proof of consistency is a test of the graph for the existence of a cycle.

As an example of the consistency checking, we take a case where a network (**K**, M_K) is consistent if an element t_a is not known to be a time point, but inconsistent if it is, where

$\mathbf{K} = \{t_0, t_a, t_b, t_c, t_n\}$,

$\mathbf{M_K} = \{meets(t_0,t_a), meets(t_a,t_n), meets(t_0,t_b), meets(t_b,t_c), meets(t_c, t_n)\}$

If t_a is not known to be a point then the corresponding graph shown in Figure 5.4.4(1) is acyclic, and the network is consistent.

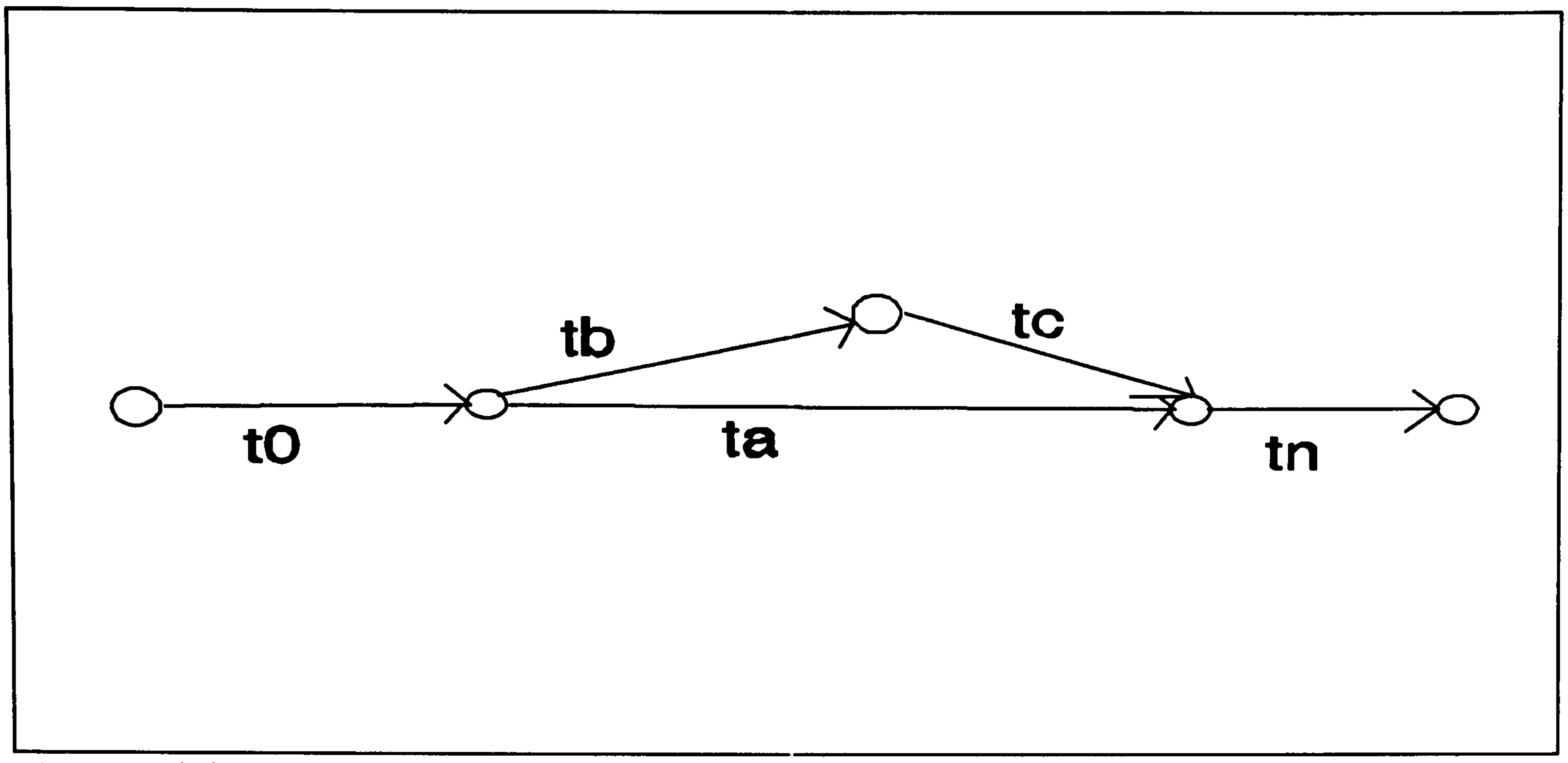

Figure 5.4.4(1)

However, if $t_a \in \mathbf{P}$, then we have the reduced graph in Figure 5.4.4(2), which is cyclic, and we deduce that the network is inconsistent.

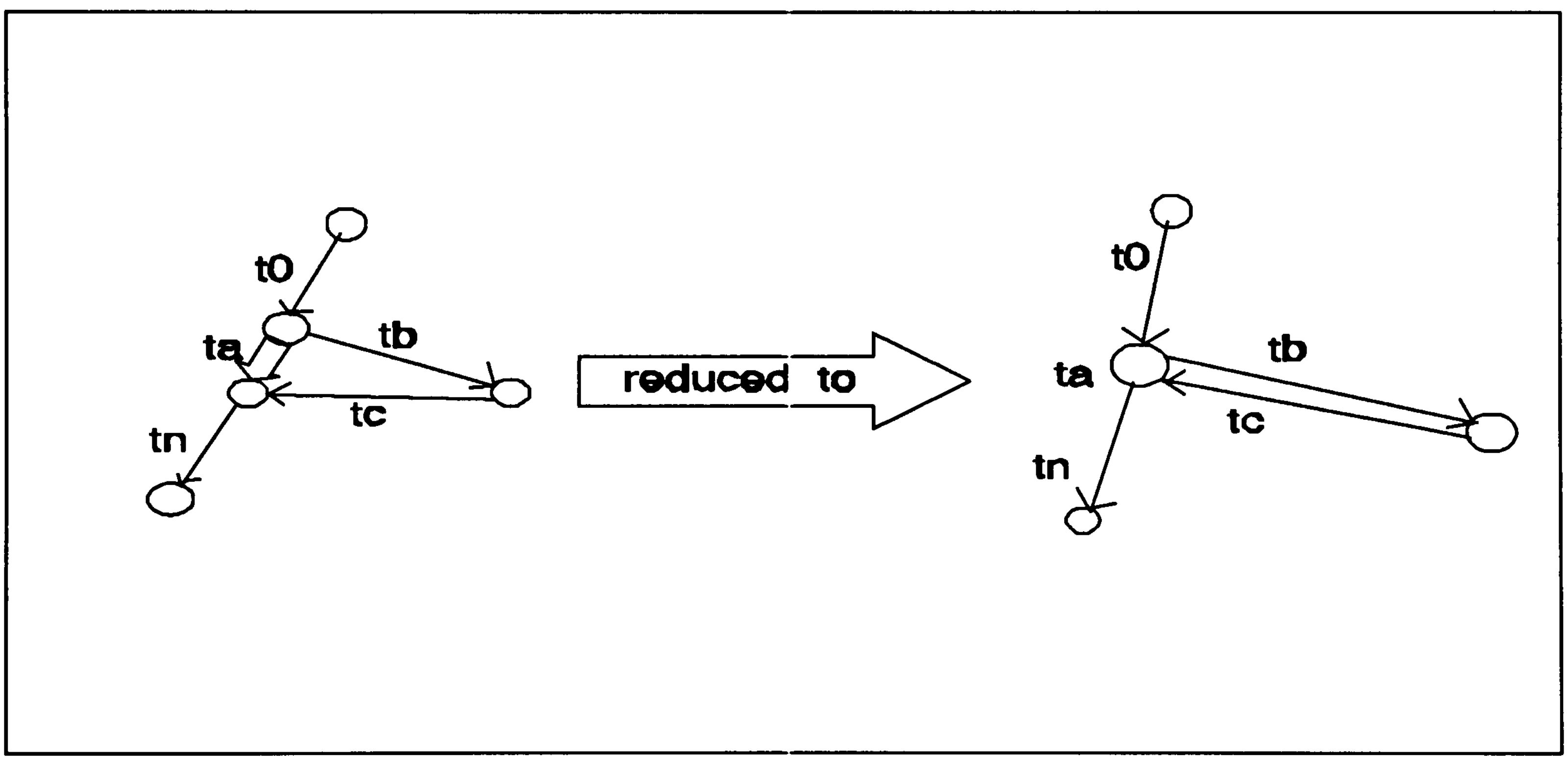

Figure 5.4.4(2)

We can see why this is so intuitively by noticing that in Figure 5.4.4(1):

$t_a = t_b \oplus t_c$

This is consistent until we add the fact that t_a is non-decomposable. Since equation $t_a = t_b \oplus t_c$ states that t_a is decomposable, we reach an inconsistency when $t_a \in \mathbf{P}$.

In Allen's system [All83], consistency checking is performed by formation of the transitive closure, according to a transitivity table with 144 entries which describes the composition of the thirteen (mutually exclusive) relations [All83]. If no conflict is found according to the exclusivity, then the system is consistent. Allen and Hayes show that the transitivity table is a result of their axioms in [AlH89], following the intuitivereasoning by possible cases which was used to construct the table originally. Using the consistency checker given above, we can provide a formal and intuitive proof of the correctness of Allen's transitivity table.

For example, consider the transitivity:

$$before(t_a, t_b),\ during(t_b, t_c).$$

Using the necessary and sufficient condition of consistency in terms of acyclicity of "*meets*", we can prove that the possible relation between t_a and t_c is *before*(t_a, t_c), or *overlaps*(t_a, t_c), or *meets*(t_a, t_c), or *during*(t_a, t_c), or *starts*(t_a, t_c), as follows:

$$before(t_a, t_b) \wedge during(t_b, t_c) \Leftrightarrow$$

$$\exists t', t_1, t_2 \in \mathbf{T}(\ meets(t_a, t') \wedge meets(t', t_b) \wedge t_c = t_1 \oplus t_b \oplus t_2)$$

(1) when $t' = t_0 \oplus t_1$, where $t'' \in \mathbf{T}$, we have:

$meets(t_a, t'') \wedge meets(t'', t_c)$,

i.e: $before(t_a, t_c)$.

From Figure 5.4.4(3), we know this case is consistent since there is no cycle of "*meets*". Hence, we have shown that $before(t_a, t_c)$ is one possible case under the condition $before(t_a, t_b)$ and $during(t_b, t_c)$.

In the similar way, we can show that $overlaps(t_a, t_c)$, $meets(t_a, t_c)$, $during(t_a, t_c)$, $starts(t_a, t_c)$ are also possible cases, as follows:

(2) when $t_a = t'' \oplus t_3$, $t_c = t_3 \oplus t_4$, $t_1 = t_3 \oplus t'$, where $t'', t_3, t_4 \in \mathbf{T}$, we have:

$t_a = t'' \oplus t_3$, $t_c = t_3 \oplus t_4$

i.e: $overlaps(t_a, t_c)$ (see Figure 5.4.4(4)).

(3) when $meets(t_a, t_c)$, we directly get the result (see Figure 5.4.4(5)).

(4) when $t_1 = t_3 \oplus t_a \oplus t'$, where $t_3 \in \mathbf{T}$, we have:

$t_c = t_3 \oplus t_a \oplus t_4$, where $t_4 = t' \oplus t_b \oplus t_2 \in \mathbf{T}$,

i.e: $during(t_a, t_c)$ (see Figure 5.4.4(6)).

(5) when $t_c = t_a \oplus t_4$, $t_1 = t_a \oplus t'$, we have:

$t_c = t_a \oplus t_4$, where $t_4 = t' \oplus t_b \oplus t_2 \in \mathbf{T}$,

i.e: $starts(t_a, t_c)$ (see Figure 5.4.4(7)).

Additionally, we can prove that there is no other possible relation between t_a and t_b as follows:

(6) If *after*(t_a, t_c), then:

$\exists t'' \in \mathbf{T}(meets(t_c, t'') \wedge meets(t'', t_a))$.

However,

meets(t_a, t'), *meets*(t', t_b), *meets*(t_b, t_2), *meets*(t_2, t''), *meets*(t'', t_a)

form a cycle: t_a, t', t_b, t_2, t'', t_a,

which shows inconsistency (see Figure 5.4.4(8)).

Similarly, for other cases:

(7) If *met-by*(t_a, t_c), then *meets*(t_c, t_a),

so that there is a cycle: t_a, t', t_b, t_2, t_a, (see Figure 5.4.4(9)), which shows inconsistency.

(8) If *overlapped-by*(t_a, t_c), then $\exists t_3, t_4, t_5 \in \mathbf{T}$ such that:

$t_a = t_3 \oplus t_5$, $t_c = t_4 \oplus t_3$,

so there is a cycle: t_5, t', t_b, t_2, t_5, (see Figure 5.4.4(10)), which shows inconsistency.

(9) If *started-by*(t_a, t_c), then $\exists t_3 \in \mathbf{T}$ such that:

$t_a = t_c \oplus t_3$,

so there is a cycle: t_3, t', t_b, t_2, t_3, (see Figure 5.4.4(11)), which shows inconsistency.

(10) If *contains*(t_a, t_c), then $\exists t_3, t_4 \in \mathbf{T}$ such that:

$t_a = t_3 \oplus t_c \oplus t_4$,

so there is a cycle: t_4, t', t_b, t_2, t_4, (see Figure 5.4.4(12)), which shows inconsistency.

(11) If *finishes*(t_a, t_c), then $\exists t_3 \in \mathbf{T}$ such that:

$t_c = t_3 \oplus t_a$,

so there is a cycle: t', t_b, t_2, t', (see Figure 5.4.4(13)), which shows inconsistency.

(12) If *finished-by*(t_a, t_c), then $\exists t_3 \in \mathbf{T}$ such that:

$t_a = t_3 \oplus t_c$,

so there is a cycle: t', t_b, t_2, t', (see Figure 5.4.4(14)), which shows inconsistency.

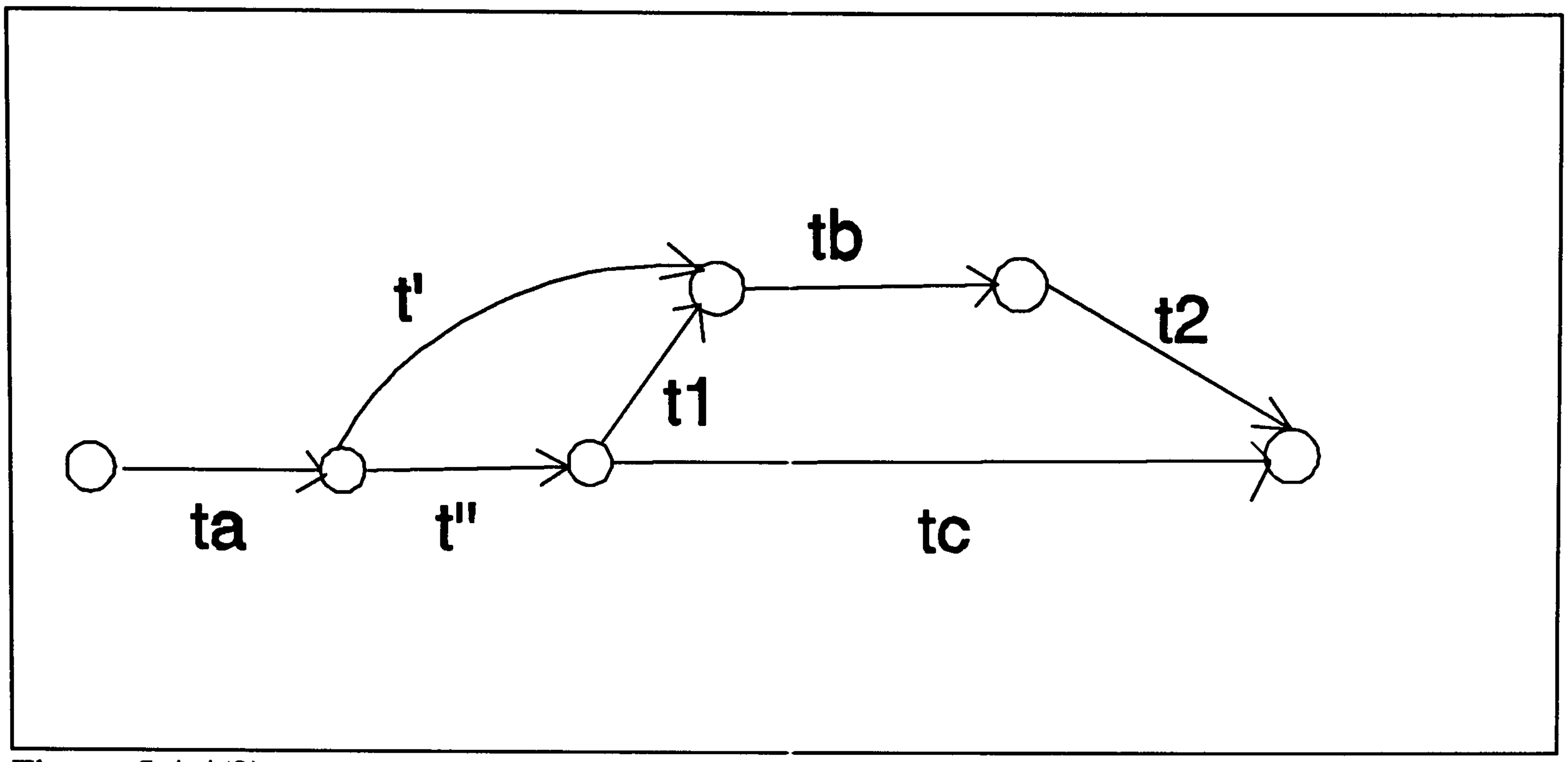

Figure 5.4.4(3)

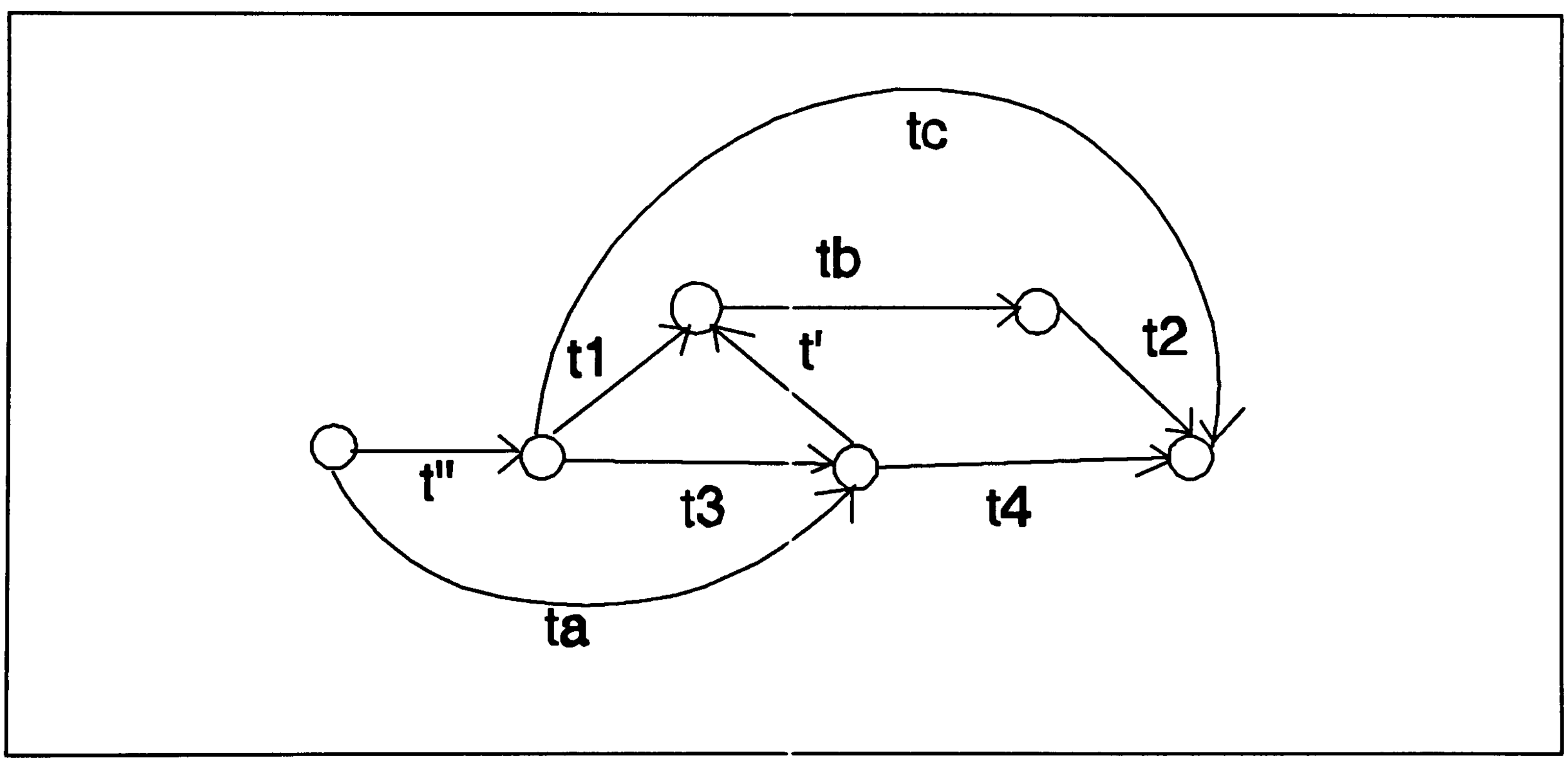

Figure 5.4.4(4)

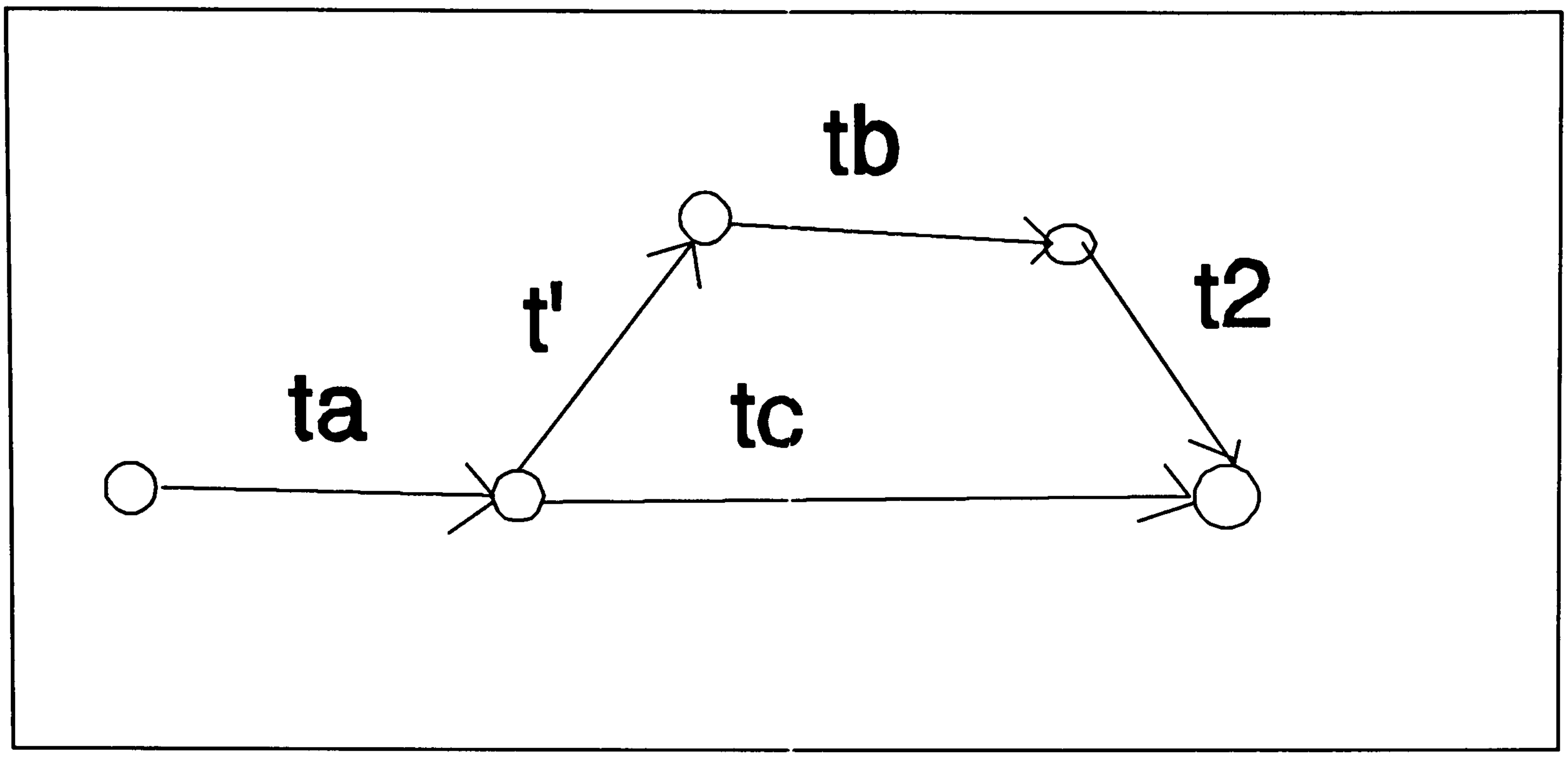

Figure 5.4.4(5)

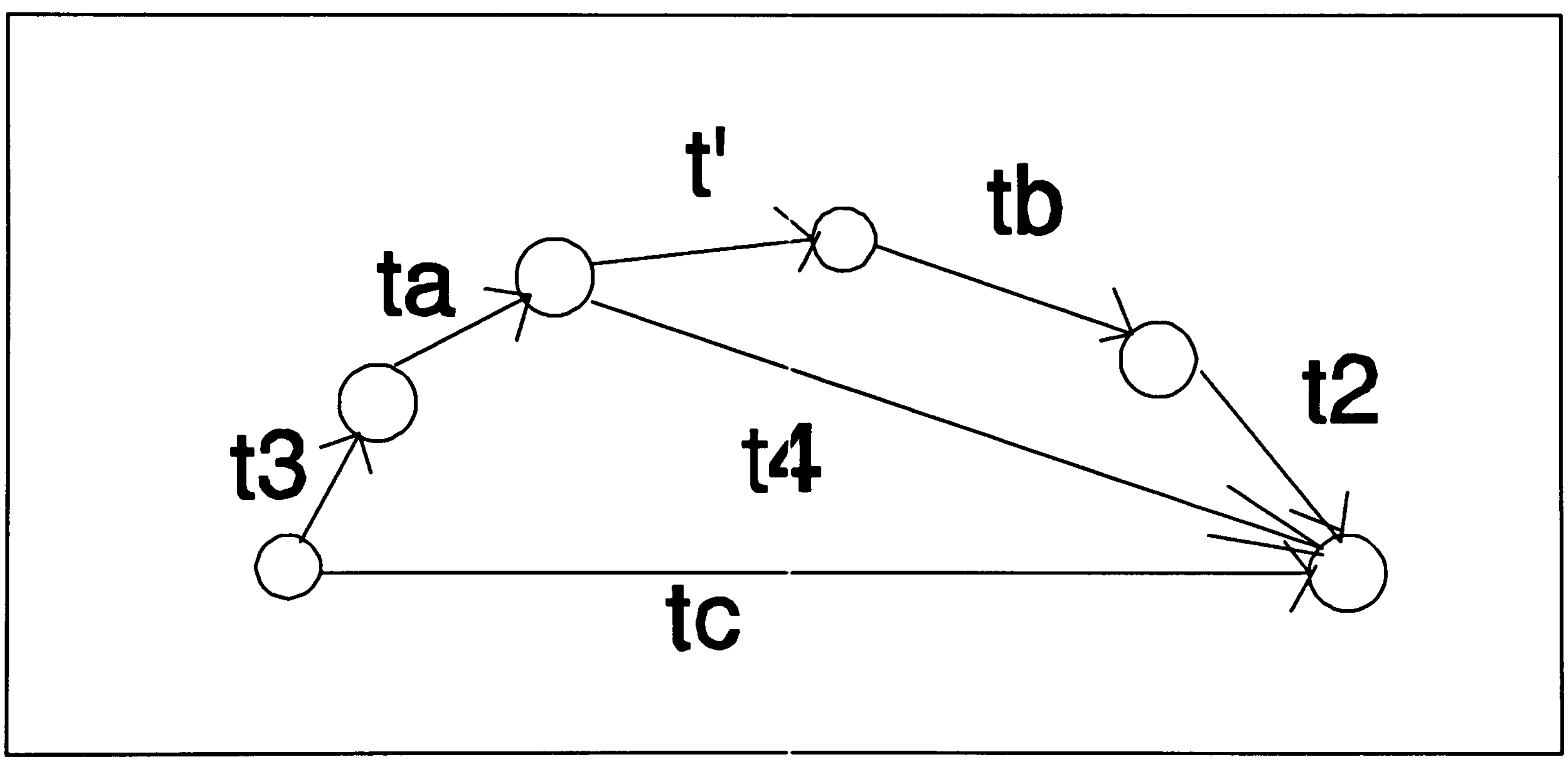

Figure 5.4.4(6)

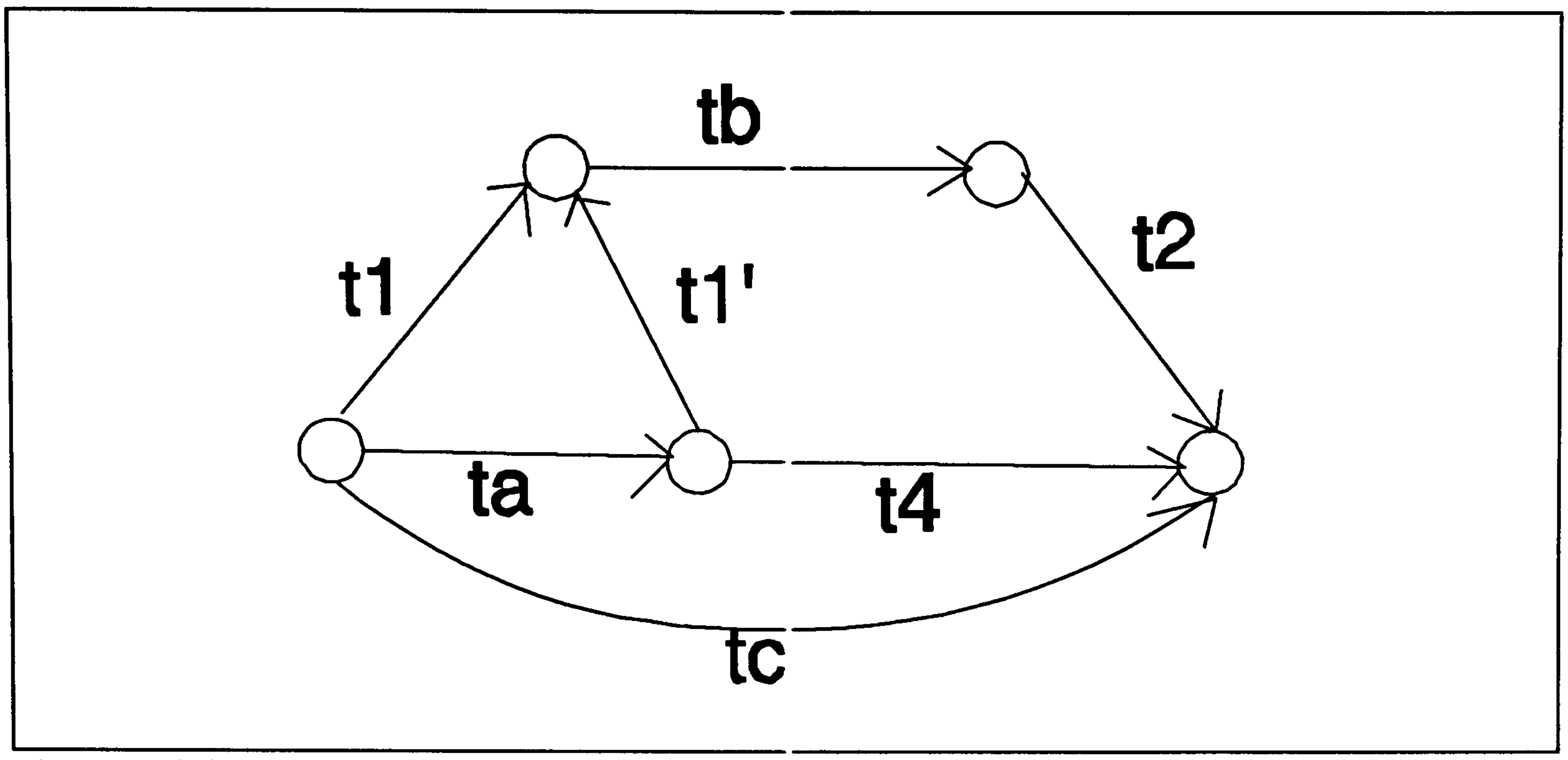

Figure 5.4.4(7)

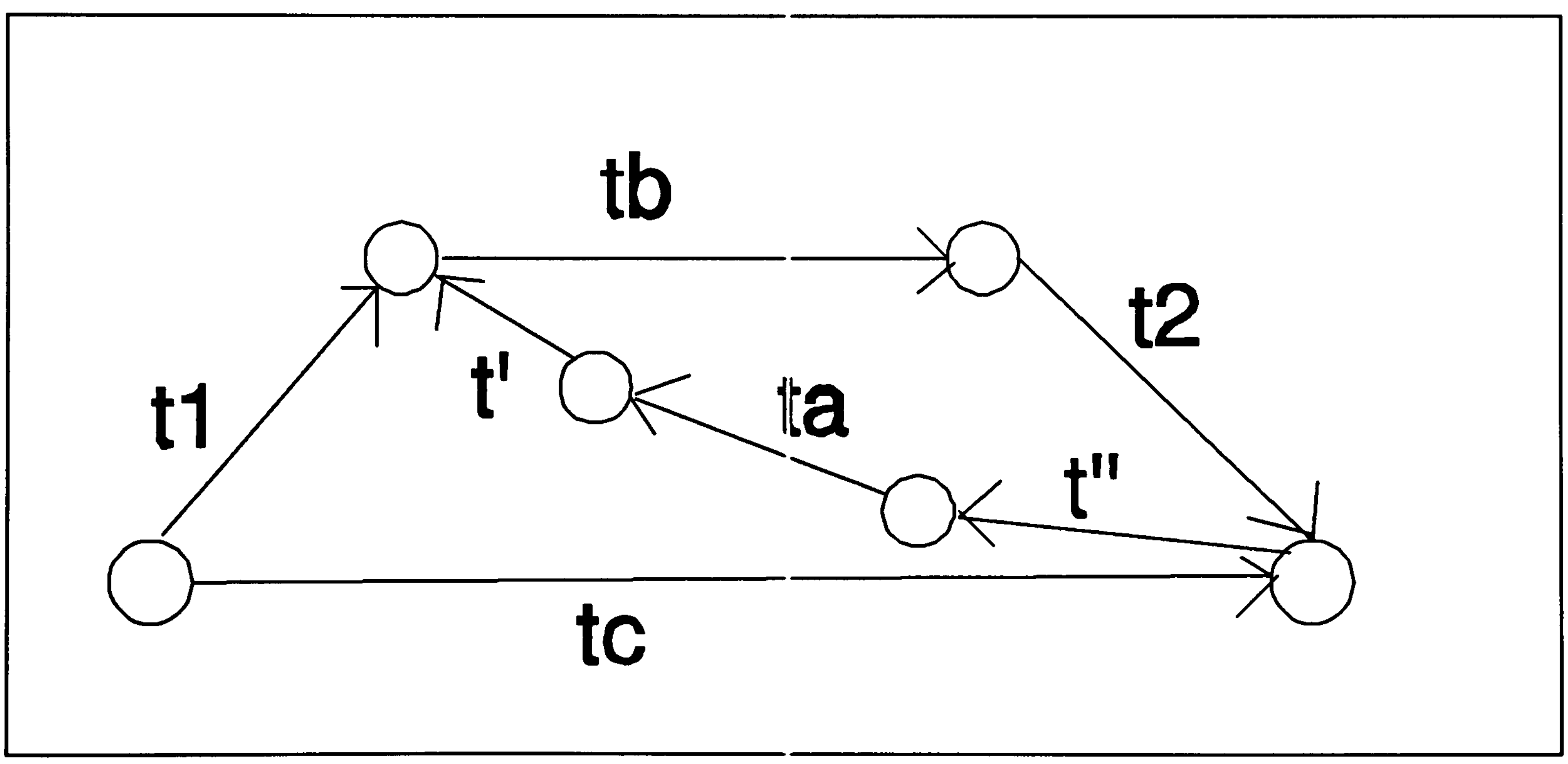

Figure 5.4.4(8)

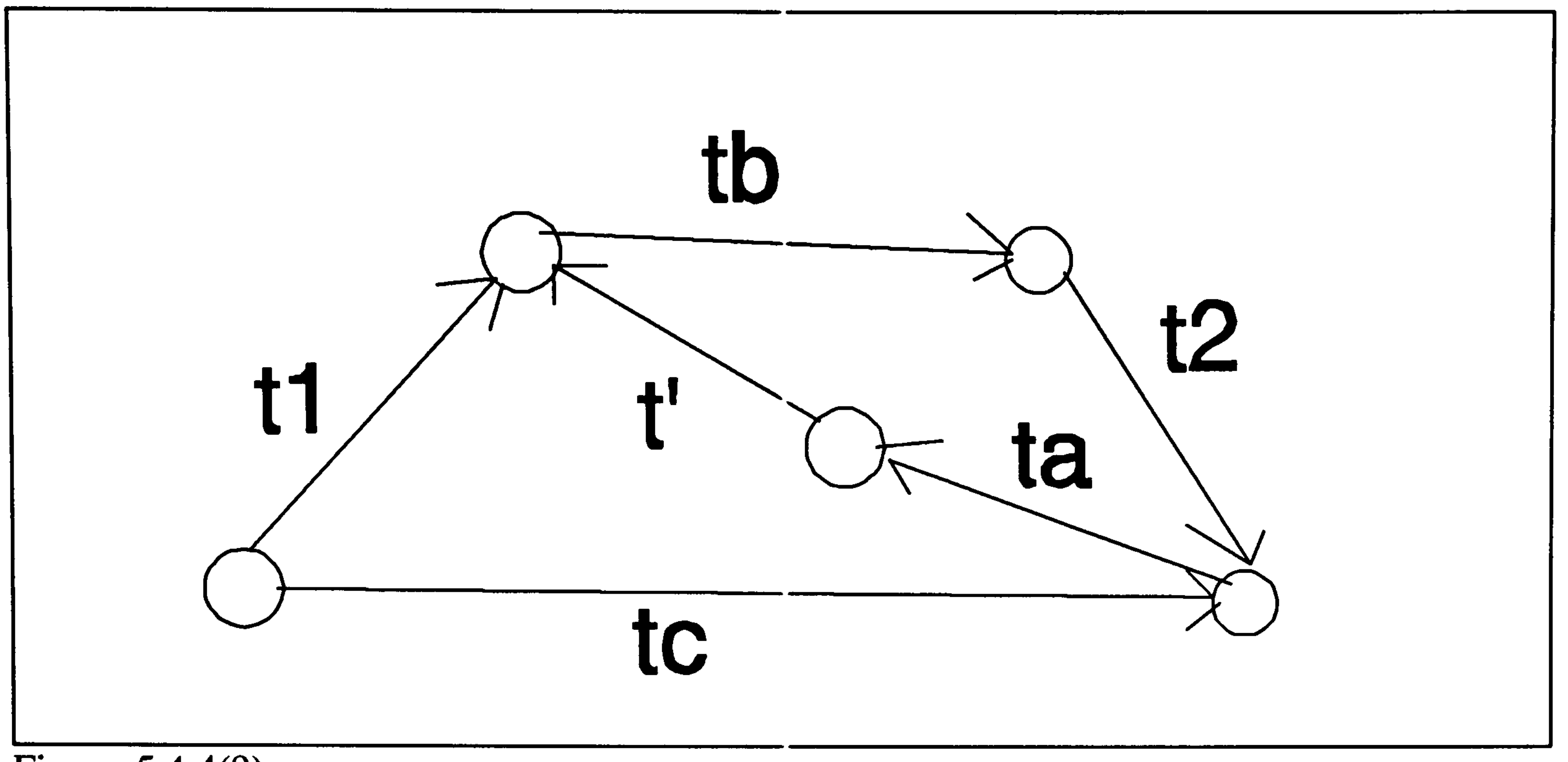

Figure 5.4.4(9)

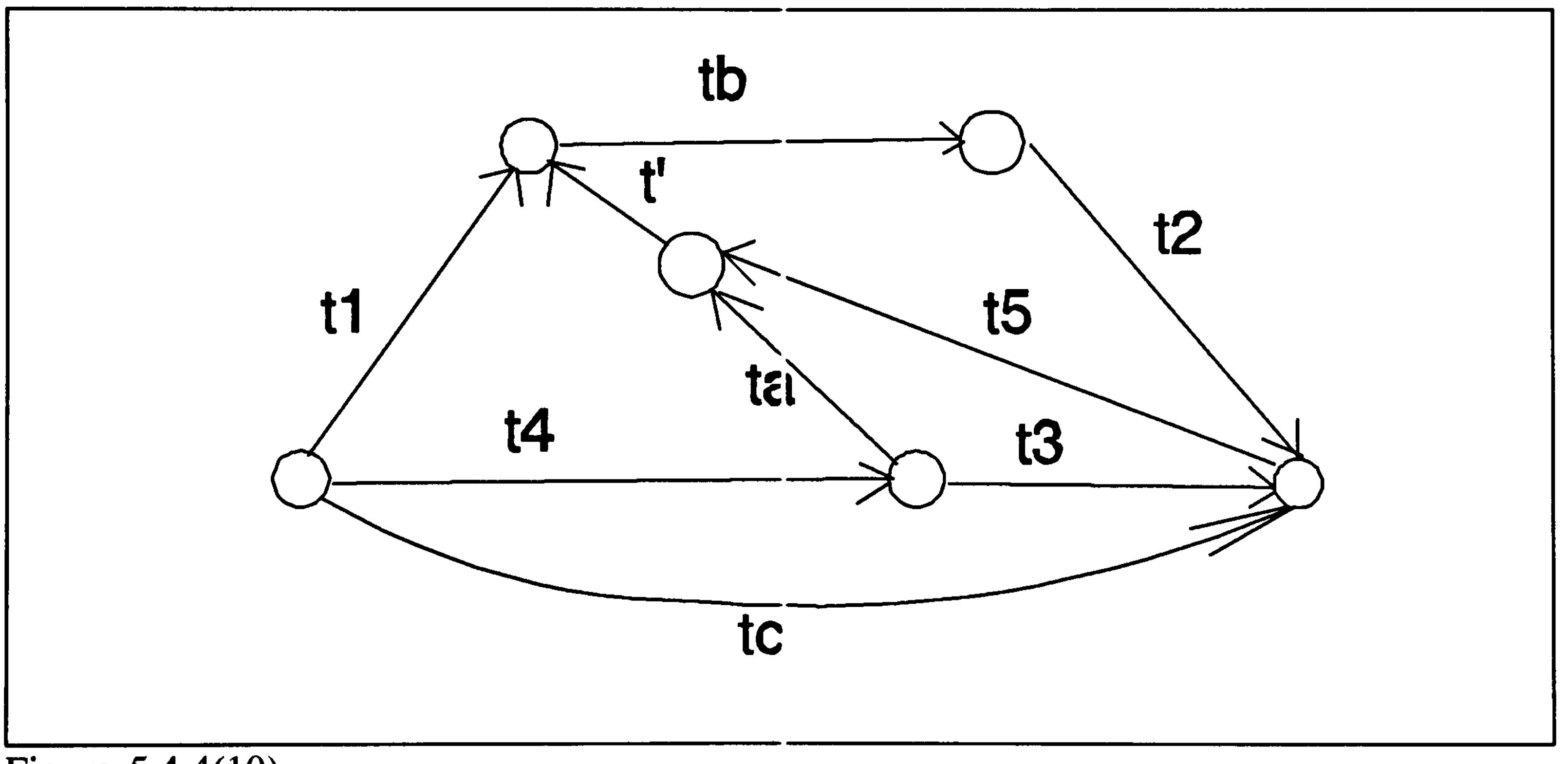

Figure 5.4.4(10)

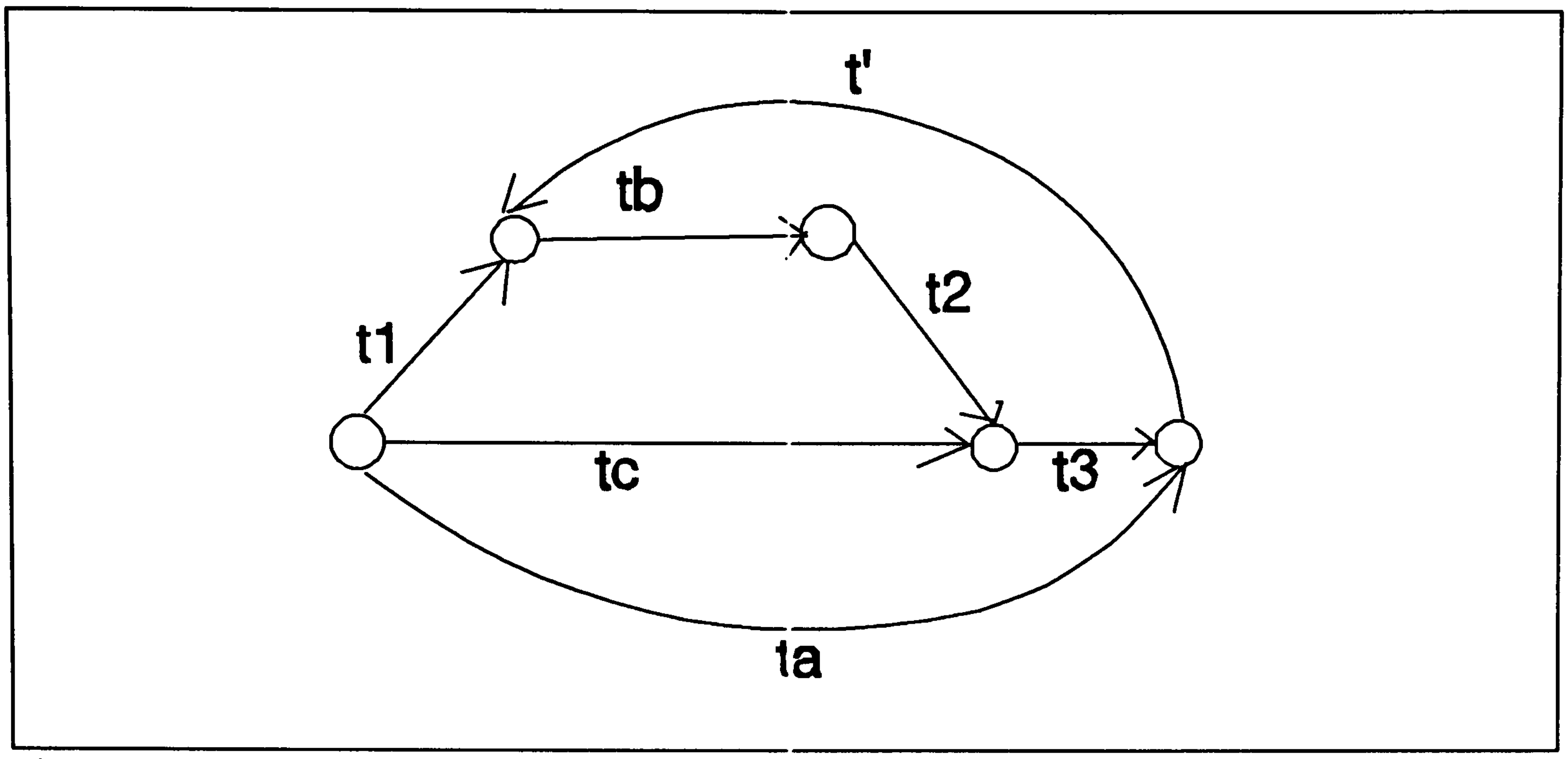

Figure 5.4.4(11)

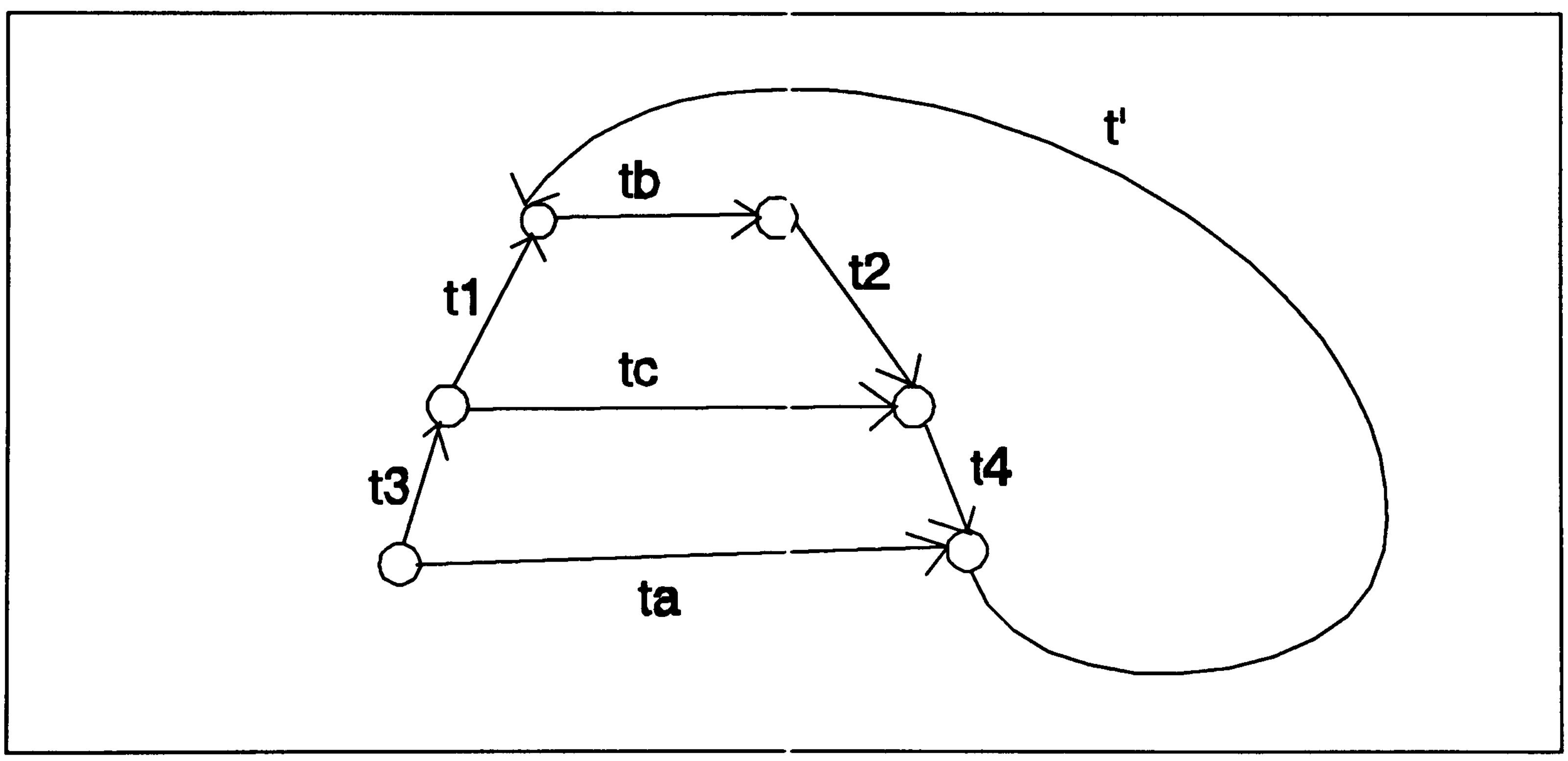

Figure 5.4.4(12)

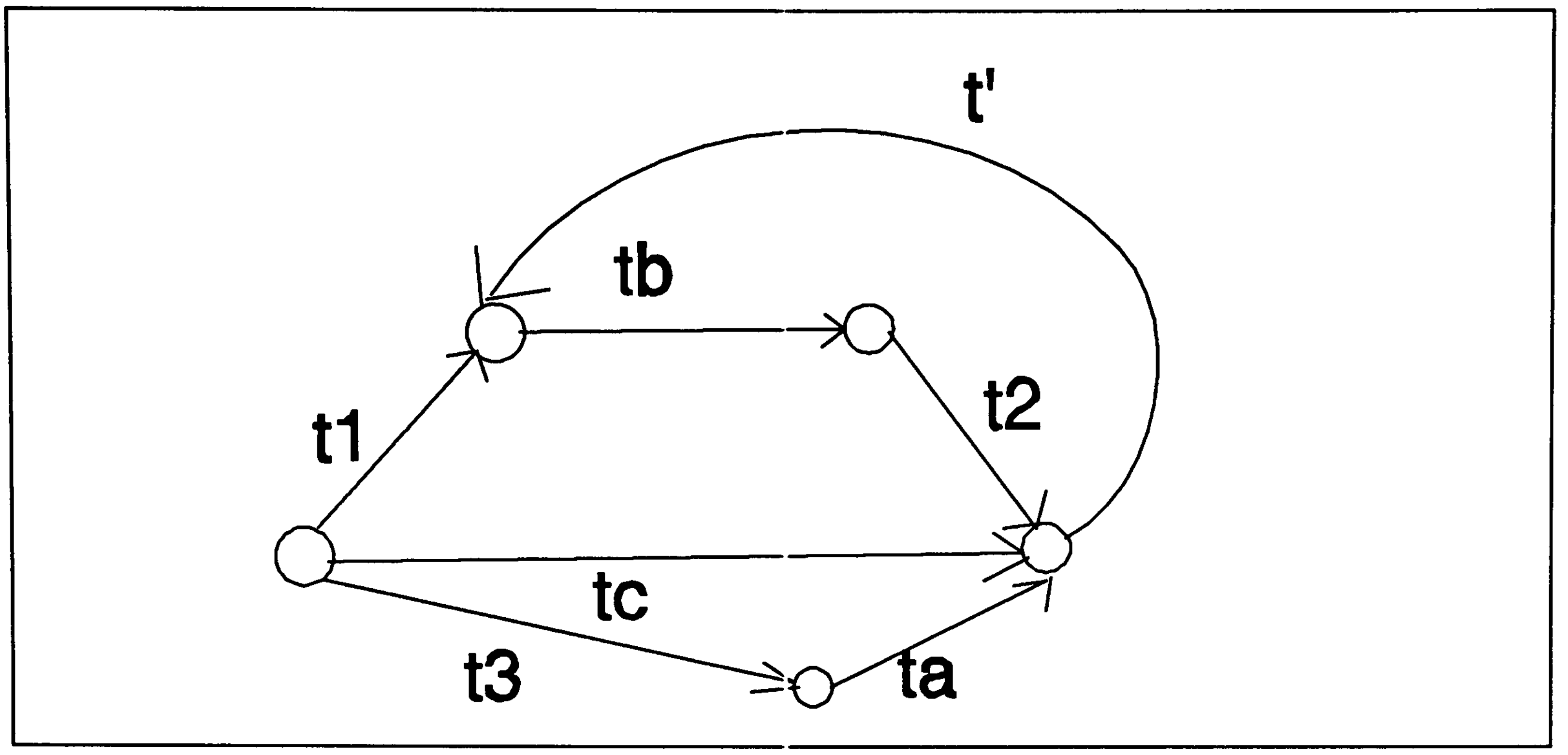

Figure 5.4.4(13)

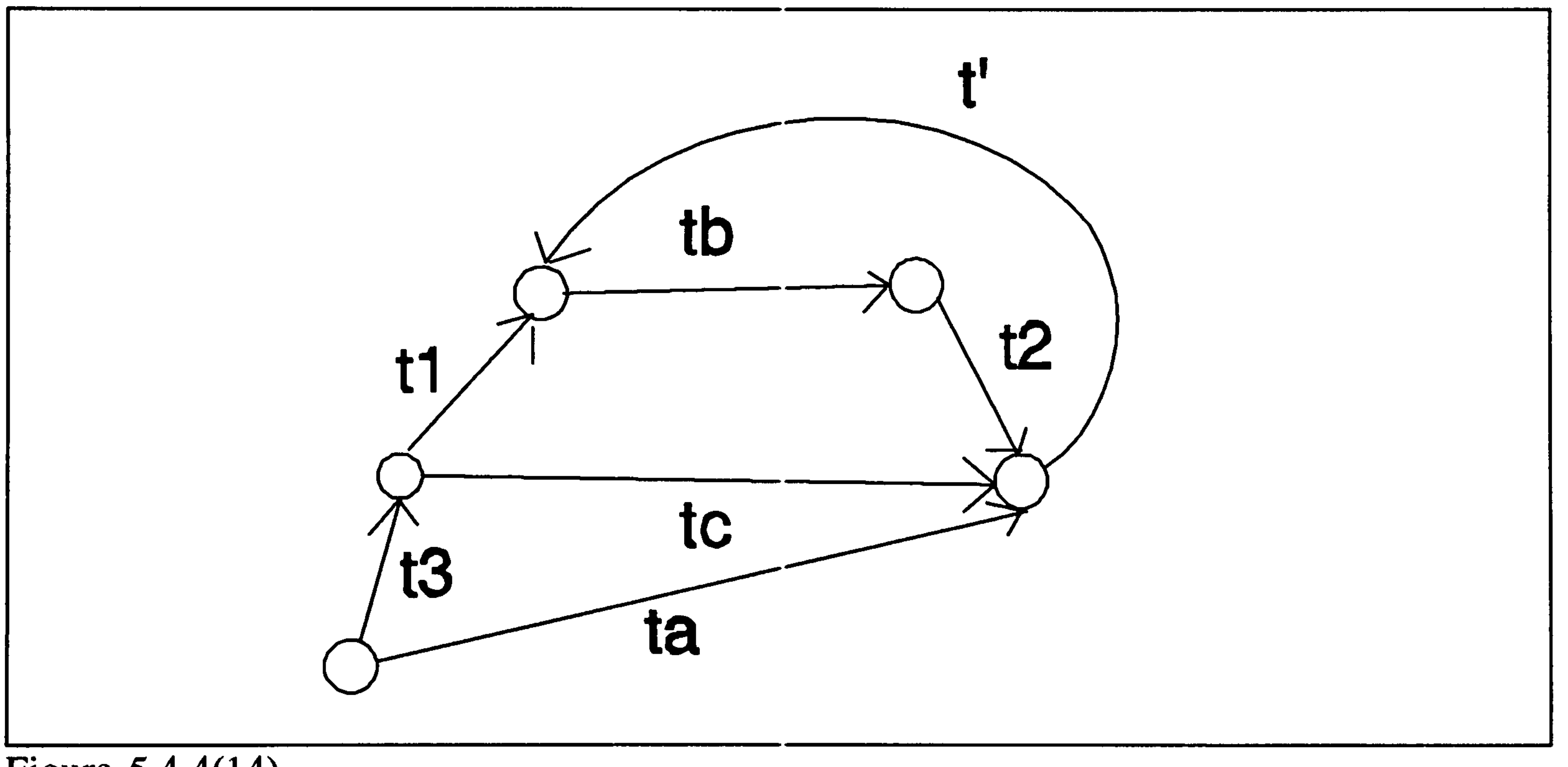

Figure 5.4.4(14)

There is an alternative formal proof of the same result. In fact:

From the definitions about a limited time network given in section 5.4.1, every time element in a consistent network is in the closure of a corresponding fundamental time network, that is, for any time element, it can be written in the form of an ordered union of fundamental time elements. Hence, if t_a, t_b and t_c are elements in a consistent time network (**K**, M_K), we can express them as:

$$t_a = a_0 \oplus a_1 \oplus \ldots \oplus a_{ma},$$

$$t_b = b_0 \oplus b_1 \oplus \ldots \oplus b_{mb},$$

$$t_c = c_0 \oplus c_1 \oplus \ldots \oplus c_{mc},$$

where, a_i, b_j, $c_k \in \mathbf{E}$, $i=0,1,\ldots,m_a$; $j=0,1,\ldots,m_b$; $k=0,1,\ldots,m_c$;

According to the definitions about the fundamental network (**E**, M_E) corresponding to (**K**, M_K), there is a *similar function* between **E** and an initial segment of **N**, the set of natural numbers. Assume that under the *similar function*:

$$a_0 \leftrightarrow n_a,\ b_0 \leftrightarrow n_b,\ c_0 \leftrightarrow n_c, \text{ where } n_a, n_b, n_c \in \mathbf{N}.$$

Hence, $a_i \leftrightarrow n_a + i$, $b_j \leftrightarrow n_b + j$, $c_k \leftrightarrow n_c + k$, where $i=0,1,\ldots,m_a$; $j=0,1,\ldots,m_b$; $k=0,1,\ldots,m_c$;

From *Before*(t_a, t_b) $\wedge$ *During*(t_b, t_c), we get:

Constraint 5.4.4:

$n_a + m_a + 1 < n_b$

$n_c < n_b$,

$n_b + m_b < n_c + m_c$.

It is clear that each of the following additional constraints is consistent with Constraint 5.4.4:

1) $n_c < n_a$,

2) $n_c = n_a$,

3) $n_a < n_c \le n_a + m_a$,

4) $n_c = n_a + m_a + 1$,

5) $n_a + m_a + n < n_c$, $n \in \mathbf{N}$,

and these imply *during*(t_a, t_c), *starts*(t_a, t_c), *overlaps*(t_a, t_c), *meets*(t_a, t_c), *before*(t_a, t_c), respectively.

However, from Constraint 5.4.4, we can infer that,

$n_a < n_a + m_a < n_a + m_a + 1 < n_c + m_c$,

which is contradictory to each one of the following:

after(t_a, t_c),

met-by(t_a, t_c),

overlapped-by(t_a, t_c),

started-by(t_a, t_c),

contains(t_a, t_c),

finishes(t_a, t_c),

finished-by(t_a, t_c).

For instance, *after*(t_a, t_c) implies that

$$n_c + m_c + 1 < n_a,$$

which is obviously contradictory with $n_a < n_a + m_a < n_a + m_a + 1 < n_c + m_c$.

Hence, we have proved that, under the condition *before*(t_a, t_b) $\wedge$ *during*(t_b, t_c), the possible relation between t_a and t_c is one of *before*(t_a, t_c), *overlaps*(t_a, t_c), *meets*(t_a, t_c), *during*(t_a, t_c), and *starts*(t_a, t_c).

All the entries of Allen's transitivity table can be established in either the above two ways. From our assumption we know that a point does not meet or be met-by another point, and from our axioms we have proved that only 3 of the 13 temporal relations between two intervals may hold between two points, and only 7 of them may be hold between a point and an interval, or between an interval and a point. Hence, we can extend Allen's transitivity table to include time points to form the corresponding transitivity mechanisms.

5.5 A Point-based Specialisation of the Time Theory

Generally, in the time theory presented in chapter 4, time intervals and points are both taken as primitive on an equal footing. However, as a specialisation, we may specify time intervals of **I**, to be constructed out of time points which are real numbers, with the following structures, employed in section 5.1:

$$(p_1,p_2,\text{open},\text{open}) =_{def} \{\ p\in \mathbf{P} \mid p_1 < p < p_2\ \},$$

$$(p_1,p_2,\text{open},\text{closed}) =_{def} \{\ p\in \mathbf{P} \mid p_1 < p \leq p_2\ \},$$

$$(p_1,p_2,\text{closed},\text{open}) =_{def} \{\ p\in \mathbf{P} \mid p_1 \leq p < p_2\ \},$$

$$(p_1,p_2,\text{closed},\text{closed}) =_{def} \{\ p\in \mathbf{P} \mid p_1 \leq p \leq p_2\ \},$$

where

$$p_1,\ p_2\in \mathbf{P},\ p_1 \leq p_2.$$

N.B. As mentioned in section 5.1, we may identify **P** with $\{(p,p,\text{closed},\text{closed}) \mid p\in \mathbf{P}\}$, a subset of **I**.

There are three constraints imposed on point-based interval system which can be formally expressed as:

C1): $\forall p_1,p_2\in \mathbf{P}(\ (p_1 < p_2 \Rightarrow \forall l,r\in \mathbf{Type}((p_1,p_2,l,r)\in \mathbf{I}))$

$\wedge(p_1 = p_2 \Rightarrow (p_1,p_2,\text{closed},\text{closed})\in \mathbf{I})\ )$

C2): $\forall i\in \mathbf{I}\exists p_1,p_2\in \mathbf{P}\exists l,r\in \mathbf{Type}((p_1,p_2,l,r) = i)$

C3): $\forall (p_1,p_2,l,r)\in \mathbf{I}\forall p_3,p_4\in \mathbf{P}(p_1 \leq p_3 \leq p_2 \wedge p_1 \leq p_4 \leq p_2 \Rightarrow p_3 \leq p_4 \vee p_4 \leq p_3)$

In the above, constraint C1) states that **I** consists of all the possible intervals constructed from **P** and **Type** = {open,closed}; C2) preserves that all the intervals in **I** are constructed from **P** and **Type** in the fixed way; finally, the role of C3) is to ensure the linear internal-structure of intervals.

The duration assignment function for time elements can be defined by:

$$\forall (p_1,p_2,l,r)\in \mathbf{I}(dur(p_1,p_2,l,r) = p_2 - p_1)$$

Since we employ a specialisation of the structure of time intervals which are constructed out of a set of time elements over which some known temporal order relation have been classically defined, the temporal relations over the whole set of time elements should be given with very careful considerations to ensure that they are consistent with both the time theory itself and the conventional assumptions about time points. The critical primitive relation, *meets*, may be formally defined by:

meet$((p_{11},p_{12},l_1,r_1),\ (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{12} = p_{21} \wedge r_1 = \text{open} \wedge l_2 = \text{closed}$$

$$\vee\ p_{12} = p_{21} \wedge r_1 = \text{closed} \wedge l_2 = \text{open}$$

It is easy to see that the above definitions are consistent with <A1> - <A6>, and <A_{Linear}>. However, there are no constraints about the density of the time. It may be discrete, e.g., when **P** is the set of integers; or dense, e.g., **P** is the set of rationals.

Correspondingly, we may specialise the other derived temporal relations over time elements (see section 4.3) as below:

equal$((p_{11},p_{12},l_1,r_1),\ (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{11} = p_{21} \wedge p_{12} = p_{22} \wedge l_1 = l_2 \wedge r_1 = r_2$$

before$((p_{11},p_{12},l_1,r_1),\ (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{12} \leq p_{21} \wedge r_1 = \text{open} \wedge l_2 = \text{open}$$

$$\vee\ p_{12} < p_{21} \wedge r_1 = \text{open} \wedge l_2 = \text{closed}$$

$$\vee\ p_{12} < p_{21} \wedge r_1 = \text{closed} \wedge l_2 = \text{open}$$

$$\vee\ \exists p \in \mathbf{P}(p_{12} < p < p_{21} \wedge r_1 = \text{closed} \wedge l_2 = \text{closed})$$

overlaps$((p_{11},p_{12},l_1,r_1),\ (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{11} < p_{21} < p_{12} < p_{22}$$

$$\vee\ p_{11} < p_{12} = p_{21} < p_{22} \wedge r_1 = \text{closed} \wedge l_2 = \text{closed}$$

starts$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{11} = p_{21} \wedge p_{12} < p_{22} \wedge l_1 = l_2$$

$$\vee\ p_{11} = p_{21} \wedge p_{12} = p_{22} \wedge l_1 = l_2 \wedge r_1 = \text{open} \wedge r_2 = \text{closed}$$

during$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{21} < p_{11} \wedge p_{12} < p_{22}$$

$$\vee\ p_{11} = p_{21} \wedge p_{12} < p_{22} \wedge l_1 = \text{open} \wedge l_2 = \text{closed}$$

$$\vee\ p_{21} < p_{11} \wedge p_{12} = p_{22} \wedge r_1 = \text{open} \wedge r_2 = \text{closed}$$

$$\vee\ p_{11} = p_{21} \wedge p_{12} = p_{22} \wedge l_1 = \text{open} \wedge l_2 = \text{closed} \wedge r_1 = \text{open} \wedge r_2 = \text{closed}$$

finishes$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$

$$p_{11} < p_{21} \wedge p_{12} = p_{22} \wedge r_1 = r_2$$

$$\vee\ p_{11} = p_{21} \wedge p_{12} = p_{22} \wedge r_1 = r_2 \wedge l_1 = \text{open} \wedge l_2 = \text{closed}$$

after$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *before*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

met-by$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *meets*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

overlapped-by$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *overlaps*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

started-by$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *starts*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

contains$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *during*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

finished-by$((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow$ *finishes*$((p_{21},p_{22},l_2,r_2), (p_{11},p_{12},l_1,r_1))$

It is easy to see that for any pair of intervals belong to **I**, one and only one of the above relations will hold. The only case which is not immediately obvious is for the relation between $(p_{11},p_{12},l_1,\text{closed})$ and $(p_{21},p_{22},\text{closed},r_2)$, where $p_{12} < p_{21}$ and $\neg\exists p \in \mathbf{P}(p_{12} < p < p_{21})$; i.e. there are no time elements between $(p_{11},p_{12},l_1,\text{closed})$ and $(p_{21},p_{22},\text{closed},r_2)$. Hence, the relation between them should be semantically defined as:

meets$((p_{11},p_{12},l_1,\text{closed}), (p_{21},p_{22},\text{closed},r_2))$

However, according to the above definitions,

$$\textit{meets}((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2))$$
$$\Rightarrow \{r_1\} \cup \{l_2\} = \{\text{open, closed}\}$$

That is, to ensure *meets*$((p_{11},p_{12},l_1,\text{closed}), (p_{21},p_{22},\text{closed},r_2))$, we must prove that either $(p_{11},p_{12},l_1,\text{closed})$ can be shown to be equal to an interval with an open right-type, or $(p_{21},p_{22},\text{closed},r_2)$ can be shown to be equal to an interval with an open left-type. In fact, the constraint:

$$p_{12} < p_{21} \wedge \neg\exists p \in \mathbf{P}(p_{12} < p < p_{21})$$

enables us to express $(p_{11},p_{12},l_1,\text{closed})$ as $(p_{11},p_{21},l_1,\text{open})$. Similarly, we have $(p_{21},p_{22},\text{closed},r_2) = (p_{12},p_{22},\text{open},r_2)$. Hence, the relation between $(p_{11},p_{12},l_1,\text{closed})$ and $(p_{21},p_{22},\text{closed},r_2)$ is definitely "*meets*".

If the time point system is assumed as dense, then we have:

$$p_{12} < p_{21} \Rightarrow \exists p \in \mathbf{P}(p_{12} < p < p_{21})$$

Hence, the relation between $(p_{11}, p_{12}, l_1, r_1)$ and $(p_{21}, p_{22}, l_2, r_2)$ is definitely "*before*".

CHAPTER 6

A TEMPORAL LOGIC BASED ON THE TIME AXIOMATISATION

The theory of temporal logic concerns reasoning with statements which have some temporal aspects. One way to represent temporal information in artificial intelligence is through what have been termed as *reified* temporal logics, in which nontemporal components of assertions (such as "Light-Is-On"), referring to proposition types [Sho87b], are addressed as arguments to some global "predicates" whose truth values are associated with particular times, which also appear to be arguments to the "predicates". The most influential work in this area is that of McDermott [Med82,DeM87], and of Allen [All83,84]. However, as Shoham points out in [Sho87a,b], both McDermott's and Allen's logics suffer in two respects. Semantically, neither give their sentences a clear meaning, although McDermott does give the semantics of what may be regarded as the propositional theory; and conceptually, Allen's trichotomy of *properties/events/processes*, and McDermott's dichotomy of *facts/events*, are unnecessary at some times and insufficient at others. Additionally, Allen's exclusion of time points from the ontology leads to some awkward formulation, and some inadequacy in reasoning correctly about continuous change (see section 3.4, or [Gal90b]).

Utilising the time axiomatisation, proposed in chapter 4, as the temporal basis, this chapter presents a new structure of temporal logic, termed *TLIP* (Temporal Logic based-on Intervals and Points). We shall give the precise syntax and semantics for the logic, trying to retain the appealing characteristics of other representative logics, such as Allen's interval logic, Shoham's and Haugh's point-based interval logic, etc., but without bearing their corresponding deficiencies discussed in the introduction: unlike those in Shoham's [Sho87a,b] and Reichgelt's [Rei87] reified temporal logic, Halpern and Shoham's propositional modal temporal logic [HaS91], or in Haugh's MTA logic [Hau87], time intervals are taken here as primitive, not necessarily to be constructed out of points; and unlike that of Allen and Hayes's interval-based logic [AlH89], or of Galton's corresponding revised system [Gal90b], time points are also treated as primitive elements, not being relegated to a subsidiary status within the theory by means of defining them as the "*meeting places*" or "*nests*" of time intervals. To allow expressions such as "*valid time*", "*transaction time*", and "*user-defined time*", for instance, in database applications (see [SnA86]), or any other further taxonomy of times, assertions of the logic will in general address an arbitrary number of temporal arguments, associating to the corresponding nontemporal components. We shall show that the proposed logic is more expressive than some representatively existing systems, without bearing their corresponding deficiencies.

6.1 Syntax

The primitive symbols of *TLIP* consist of the elements of the following sets:

T - a set of time element symbols,

TV - a set of time element variables,

C - a set of individual symbols which is disjoint from **T**,

V - a set of individual variables which is disjoint from **TV**,

TF - a singleton set containing a binary temporal function symbol,

TR - a singleton set containing a binary temporal relation symbol,

F - a set of nontemporal function symbols which is disjoint from **TF**,

R - a set of nontemporal relation symbols which is disjoint from **TR**.

Definition 6.1.1 (Temporal terms):

$\mathbf{Term}_{\mathbf{temporal}}$ is defined as the minimal set of temporal terms in *TLIP*, closed under the following rules:

(1) If $t \in \mathbf{T} \cup \mathbf{TV}$ then $t \in \mathbf{Term}_{\mathbf{temporal}}$,

(2) If $t_1, t_2 \in \mathbf{Term}_{\mathbf{temporal}}$, and $f_t \in \mathbf{TF}$ is the binary temporal function symbol, then $f_t(t_1, t_2) \in \mathbf{Term}_{\mathbf{temporal}}$.

Definition 6.1.2 (Nontemporal terms):

$\mathbf{Term}_{\mathbf{nontemporal}}$ is defined as the minimal set of nontemporal terms in *TLIP*, closed under the following rules:

(1) If $s \in \mathbf{C} \cup \mathbf{V}$ then $s \in \mathbf{Term}_{\mathbf{nontemporal}}$,

(2) If $s_1,...,s_n \in \mathbf{Term}_{\mathbf{nontemporal}}$, and $f \in \mathbf{F}$ is an n-ary nontemporal function symbol, then $f(s_1,...,s_n) \in \mathbf{Term}_{\mathbf{nontemporal}}$.

Definition 6.1.3 (Well-formed formulas):

Wff is defined as the minimal set of well-formed formulas in *TLIP*, closed under the following rules:

(1) If $t_1,t_2 \in \mathbf{Term}_{\mathbf{temporal}}$, and $r_t \in \mathbf{TR}$ is the binary temporal relation symbol, then $r_t(t_1,t_2) \in \mathbf{Wff}$,

(2) If $t_1,...,t_m \in \mathbf{Term}_{\mathbf{temporal}}$, $s_1,...,s_n \in \mathbf{Term}_{\mathbf{nontemporal}}$, and $r \in \mathbf{R}$ is an n-ary nontemporal relation symbol, then

$$(t_1,...,t_m;\ r(s_1,...,s_n)) \in \mathbf{Wff},$$

(3) If α, $\beta \in \mathbf{Wff}$, then $\alpha \wedge \beta, \neg\alpha \in \mathbf{Wff}$,

(4) If $\alpha(\mathrm{u}) \in \mathbf{Wff}$, and there is no free occurrence of x in $\alpha(\mathrm{u})$, then $\forall \mathrm{x}(\alpha(\mathrm{x})) \in \mathbf{Wff}$.

N.B. In *TLIP*, we assume the conventional definitions of "$\vee$", "$\Rightarrow$", "$\Leftrightarrow$" and "$\exists$", etc., in terms of "$\wedge$" and "$\neg$".

6.2 Semantics

An interpretation of *TLIP* is a tuple Φ = <**TD**, **D**, $\oplus$, *meets*, **FN**, **RL**, **I**>, where

TD - a nonempty universe of time elements which may be both intervals and points,

D - a nonempty universe of individuals which is disjoint from **TD**,

$\oplus$ - a binary function on **TD**,

meets - a binary temporal relation over the elements of **TD**,

FN - a set of total functions from $2^{\mathbf{D}}$ to **D**,

RL - a set of relations over **D**,

I - a meaning function such that:

I assigns a member of **TD** for each temporal element symbol in **T**;

I assigns a member of **D** for each individual symbol in **C**;

I assigns "$\oplus$" to the binary temporal function symbol of **TF**;

I assigns "*meets*" to the binary temporal relation symbol of **TR**;

for each $f \in$ **F**, and for all $t_1,\ldots,t_m \in$ **TD**, I assigns a member of **FN** to $(t_1,\ldots,t_m, f)$;

for each $r \in$ **R**, and for all $t_1,\ldots,t_m \in$ **TD**, I assigns a member of **RL** to $(t_1,\ldots,t_m, r)$;

where "$\oplus$" and "*meets*" are characterised by the time axiomatisation presented in chapter 4. For convenience of expression, in what follows, we shall identify $\oplus(t_1,t_2)$ with $t_1 \oplus t_2$.

A *variable assignment*, A, is a function assigning each of temporal variables in **TV** some member of **TD**, and each of the individual variables in **V** some member of **D**.

From functions I and A, we can induce an extended meaning, IA, on **T**$\cup$**TV**, (**TD**,...,**TD**,**C**$\cup$**V**) as below:

If $t \in$ **T**, then IA(t) = I(t);

If $t_v \in$ **TV**, then IA(t_v) = A(t_v);

If $c \in \mathbf{C}$, then for all $t_1,\ldots,t_m \in \mathbf{TD}$, $IA(t_1,\ldots,t_m, c) = I(c)$;

If $v \in \mathbf{V}$, then for all $t_1,\ldots,t_m \in \mathbf{TD}$, $IA(t_1,\ldots,t_m, v) = A(v)$;

and on arbitrary terms in the following way:

If $t_1,t_2 \in \mathbf{Term}_{temporal}$, and $f_t \in \mathbf{TF}$ is the binary temporal function symbol, then

$$IA(f_t(t_1,t_2)) = IA(t_1) \oplus IA(t_2);$$

If $s_1,\ldots,s_n \in \mathbf{Term}_{nontemporal}$, and $f \in \mathbf{F}$ is an n-ary nontemporal function, then for all $t_1,\ldots,t_m \in \boldsymbol{TD}$,

$$IA(t_1,\ldots,t_m, f(s_1,\ldots,s_n)) = I(t_1,\ldots,t_m, f)(IA(t_1,\ldots,t_m, s_1),\ldots,IA(t_1,\ldots,t_m, s_n)).$$

The values of wffs under the interpretation, Φ, and the variable assignment, A, are inductively defined as below:

If $t_1,t_2 \in \mathbf{Term}_{temporal}$, and $r_t \in \mathbf{TR}$ is the binary temporal relation symbol, then

$r_t(t_1,t_2)^\Phi = 1$ if $meets(IA(t_1),IA(t_2))$

otherwise $r_t(t_1,t_2)^\Phi = 0$;

If $t_1,\ldots,t_m \in \mathbf{Term}_{temporal}$, $s_1,\ldots,s_n \in \mathbf{Term}_{nontemporal}$, and $r \in \mathbf{R}$ is an n-ary nontemporal relation symbol, then

$(t_1,...,t_m;\ r(s_1,...,s_n))^{\Phi} = 1$

if $(\mathrm{IA}(\mathrm{IA}(t_1),...,\mathrm{IA}(t_m),\ s_1),\ ...,\ \mathrm{IA}(\mathrm{IA}(t_1),...,\mathrm{IA}(t_m),\ s_n))$

$\in \mathrm{I}(\mathrm{IA}(t_1),...,\mathrm{IA}(t_m),\ r)$

otherwise $(t_1,...,t_m;\ r(s_1,...,s_n))^{\Phi} = 0$;

If $\alpha,\beta \in$ **Wff**, then

$(\alpha \wedge \beta)^{\Phi} = 1$ if $\alpha^{\Phi} = 1$ and $\beta^{\Phi} = 1$,

otherwise, $(\alpha \wedge \beta)^{\Phi} = 0$;

$(\neg\alpha)^{\Phi} = 1$ if $\alpha^{\Phi} = 0$, otherwise $(\neg\alpha)^{\Phi} = 0$;

If $\alpha(u) \in$ **Wff**, and there is no free occurrence of x in $\alpha(u)$, then

$\forall x(\alpha(x))^{\Phi} = 1$ if $\alpha(x)^{\Phi} = 1$ for all valuation assignments to x,

otherwise $\forall x(\alpha(x))^{\Phi} = 0$.

6.3 A New Categorisation of Temporal Propositions

In [Sho87a,b], Shoham has proposed a categorisation of *temporal propositions*, which associate a *proposition type* with a point-based time interval. His intention is to replace both

Allen's trichotomy of *properties/events/processes* and McDermott's dichotomy of *facts/events* by a more flexible scheme. In order to do this, Shoham provides the means for distinguishing between fact-like (or property-like) proposition types, event-like proposition types, and so on, in terms of the following series of definitions:

(Sho.1) A proposition type x is *downward-hereditary* if whenever it holds over an interval it holds over all of its subintervals, possibly excluding the two end-points. This can be formally captured by the axiom:

$$\begin{aligned}
&\forall p_1,p_2,p_3,p_4(\ p_1 \le p_3 \le p_4 \le p_2 \\
&\qquad\qquad \wedge\ p_1 \ne p_4 \wedge p_3 \ne p_2 \\
&\qquad\qquad \wedge\ \mathbf{TRUE}(p_1,p_2,\ x)) \\
&\Rightarrow \mathbf{TRUE}(p_3,p_4,\ x)
\end{aligned}$$

(Sho.2) A proposition type x is *upward-hereditary* if whenever it holds for all <u>proper</u> subintervals of some non-point interval (except possibly at its end-points), it also holds over the non-point interval itself. *upward-hereditary* is captured by:

$$
\begin{aligned}
&\forall p_1,p_2(\ p_1 < p_2 \\
&\qquad \wedge (\forall p_3,p_4(\ p_1 \le p_3 \le p_4 \le p_2 \\
&\qquad\qquad \wedge (p_1 \ne p_3 \wedge p_2 \ne p_4) \\
&\qquad\qquad \wedge p_1 \ne p_4 \\
&\qquad\qquad \wedge p_3 \ne p_2 \\
&\qquad \Rightarrow \mathbf{TRUE}(p_3,p_4,\ x))) \\
&\quad \Rightarrow \mathbf{TRUE}(p_1,p_2,\ x)
\end{aligned}
$$

(Sho.3) A proposition type x is *point-downward-hereditary* if whenever it holds over an interval it holds at all of its internal points.

(Sho.4) A proposition type x is *point-upward-hereditary* if whenever it holds at all <u>internal</u> points of some non-point interval it holds also over the non-point interval itself.

(Sho.5) A proposition type x is *interval-downward-hereditary* if whenever it holds over an interval it holds over all of its non-point subintervals.

(Sho.6) A proposition type x is *interval-upward-hereditary* if whenever it holds over all non-point subintervals of some non-point interval it holds also over the non-point interval itself.

(Sho.7) A proposition type x is *liquid* if it is both *upward-hereditary* and *downward-hereditary.*

(Sho.8) A proposition type x is *concatenable* if whenever it holds over two consecutive intervals it holds also over their union.

(Sho.9) A proposition type x is *gestalt* if it never holds over two intervals one of which properly contains the other.

(Sho.10) A proposition type x is *solid* if it never holds over two properly overlapping intervals.

It is interesting to note that Shoham makes some observations in the two versions of his paper, respectively. In the earlier version ([Sho87a],p.189), he points out that:

(1) Allen's properties are exactly the *liquid* propositions;

(2) Allen's events are exactly the *solid* propositions;

and in the later version ([Sho87b],p.101), he revised his "observations" to:

(1') *Point-point-liquid* proposition types coincide with Allen's properties and McDermott's facts, where a point-point-liquid proposition type is both *point-downward-hereditary* and *point-upward-hereditary*. Liquid proposition types coincide with philosophers' *homogeneous* propositions: a homogeneous proposition is true of an interval iff it is true over all its proper subintervals;

(2') Allen's and McDermott's events corresponding either to *gestalt* propositions, or to *solid* ones, or both.

However, through a careful examination, it is found that the above "observations" are in fact not exactly correct. For example, by Allen's axiom (H.1) [All84]:

$$\mathbf{HOLDS}(pro,i) \Leftrightarrow \forall i_1 \in \mathbf{I}(In(i_1,i) \Rightarrow \mathbf{HOLDS}(pro,i_1))$$

N.B. For convenience of expression, Allen introduces the derived temporal relation, "*In*", which summarises the relationships in which one interval, i_1, is a proper subinterval of another interval, i_2, by means of:

$$In(i_1, i_2) \Leftrightarrow During(i_1,i_2) \vee Starts(i_1,i_2) \vee Finishes(i_1,i_2)$$

a property holds over an interval if and only it holds over all its proper subintervals, that is, it coincides with philosophers' homogeneous proposition. However, in Allen's interval-based logic, time points are definitely excluded [All84], or later, introduced as the "*meeting places*" or "*nests*" of intervals at a subsidiary status within the theory [AlH89]. The contention is that nothing can be true at a point, for a point is not an entity at which things happen or are true. Hence, it is obvious that Allen's properties do not, as stated by (i'), coincide with Shoham's point-point-liquid proposition types. It seems that the earlier conclusion, (1), is more acceptable, and also, the earlier version for the axiom capturing downward-hereditary propositions [Sho87a] is more suitable for characterising Allen's properties, since the two end-points of an interval in fact satisfy the condition which is constrained in terms of Allen's relation, "*In*" (see N.B. above).

On the other hand, in contrast to Shoham's observation, (2) or (2'), it seems that Allen's events, which are characterised by

(O.1) $\mathbf{OCCUR}(e,\mathrm{i}) \wedge In(\mathrm{i}',\mathrm{i}) \Rightarrow \neg\mathbf{OCCUR}(e,\mathrm{i}')$

are in fact simply Shoham's gestalt propositions, that is, an event *e* happens over an interval i iff there is no proper subinterval of i over which *e* happens (In fact, we may also easily infer that there is no proper superinterval of i over which the event *e* happens).

The above questions are perhaps not serious, or maybe just need some trivial technical revisions. However, it is the following issue that really make it necessary to revise the fundamental axioms about time itself so as to extend the abstract concept of time elements to include both intervals and points, and the temporal relations between intervals to address points as well.

As we mentioned above, in Shoham's reified temporal logic, time intervals are defined in terms of pairs of time points, and the way to distinguish between different kinds of propositions is by specifying how the truth of the proposition over one interval is related to its truth over the other intervals. For example, in Shoham's definition (D.8), the word "*consecutive*" is used to characterise the temporal relation that a time interval is "*immediately before*" (or "*meets*", in Allen's terminology) another time interval. It seems that, if time intervals are defined to be constructed out of time points, the temporal relations between time intervals, such as those 13 relations introduced by Allen [All83], may be induced from the order relation between the end-points of these time intervals. However, in Shoham's reified temporal logic, there is no formal axiomatisation of time itself, although in the earlier version of his paper, he does simply view time as a total order on time points (see [Sho87a]). In fact, modelling time intervals by means of constructing them out of points may lead to some problematical questions involved with the "open" and "closed" nature of intervals. Consider the "*immediately before*" relation between two intervals, <a,b> and <c,d>:

Intuitively, allowing point-like intervals, we have $a \leq b$ and $c \leq d$. The question is how to decide the order relation between point b and point c. If we assume that interval <a,b> is immediately before <c,d>, that is:

meets(<a,b>, <c,d>)

then the possible relation between b and c is either $b = c$, or $b < c$.

Case 1. $b < c$:

In this case, if we assume that the set of time points is discrete, that is, there may be no other time point between b and c, then we may get *meets*(<a,b>, <c,d>). However, it is interesting to note that, for the representation of facts concerned with continuous change, it must be necessary to take the set of time points to be dense, i.e., between any two distinct time points, there is a third. Hence, in this case, there must be some other time points between point b and point c, and therefore there is at least another time interval between interval <a,b> and interval <c,d>, which contradicts our assumption.

Case 2. b = c.

In this case, as commented by Allen [All83], Haugh [Hau87], and Galton [Gal90b], if we let <a,b> to be a maximal interval throughout which a proposition, *pos*, is true, and <c,d>, a maximal interval throughout which *pos* is false, then one of the following cases must occur:

Case 2.1 b (= c) is part of both <a, b> and <c, d>;

Case 2.2 b (= c) is part of <a, b> but not <c, d>;

Case 2.3 b (= c) is part of <c, d> but not <a, b>;

Case 2.4 b (= c) is part of neither <a, b> nor <c, d>.

In case 2.1, *pos* would have to be both true and false at point b (= c), which is absurd; in case 2.4 *pos* would be neither true nor false at b (= c), which is again absurd; while as for the remaining cases, since there is nothing to chose between them, any decision either way must be arbitrary and unsatisfactory.

The above discussion shows that modelling time intervals by means of constructing them out of points may lead to problems: the annoying question of whether the end-points are in the interval or not must be addressed, seeming without any satisfactory solution.

In what follows, we shall simply revise Shoham's categorization of propositions by means of the time axiomatisation employed in this thesis which addresses both time intervals and time points. However, since time elements may be now intervals or points, we must make sure the corresponding definitions are well-defined for general treatment. First, for convenience of representation, we extend Allen's relation "*In*" (see above) to address both intervals and points. We define:

$$in(t, i) \Leftrightarrow during(t,i) \vee starts(t,i) \vee finishes(t,i)$$

for $t \in \mathbf{T}$ and $i \in \mathbf{I}$, where "*during*", "*starts*" and "*finishes*" belong to the groups of temporal relations classified in chapter 4.

Also, we define a another derived relation, *sub*, as an extension of *in* to include the "*equal*" relation:

$$sub(t_1, t_2) \Leftrightarrow equal(t_1,t_2) \vee during(t_1,t_2) \vee starts(t_1,t_2) \vee finishes(t_1,t_2)$$

where t_1 and t_2 are general time elements which may be intervals or points.

The difference between "*in*" and "*sub*" is that, while $in(i_1, i_2)$ summarises the relationship that interval i_1 is a proper subinterval of interval i_2, the notation $sub(t_1, t_2)$ allows time element t_1 to be t_2 itself.

For the sake of simple expression, in what follows, we shall only consider a simple case of wffs in which the number of temporal arguments is limited to 1, that is, we only address the time over which the nontemporal component, or the *proposition type* in Shoham's terminology (see [Sho87a,b]), x, is believed to be true; and also, following Shoham's notation, we shall represent it as: **TRUE**(t; x).

<Def.1> A proposition type x is *downward-hereditary* if whenever it holds over a time element it holds over all of its sub-elements, that is:

$$\mathbf{TRUE}(t; x) \wedge sub(t', t) \Rightarrow \mathbf{TRUE}(t'; x)$$

<Def.2> A proposition type x is *upward-hereditary* if whenever it holds for all proper sub-elements of some *decomposable time interval*, it also holds over the time element itself, that is:

$$\forall t \in T(in(t; i) \Rightarrow \mathbf{TRUE}(t; x)) \Rightarrow \mathbf{TRUE}(i; x)$$

where $i \in \mathbf{I}$.

N.B. By the above definitions, our downward-hereditary proposition types may deal with both decomposable/nondecomposable time intervals, and time points which are definitely nondecomposable (see axioms <A5> and <A6> chapter 4), while the upward-hereditary proposition types are still defined only for decomposable time intervals.

<Def.3> A proposition type x is *liquid* if it is both *upward-hereditary* and *downward-hereditary*.

<Def.4> A proposition type x is *concatenable* if whenever it holds over two consecutive time elements it holds also over their ordered union, that is:

$$\mathbf{TRUE}(t'; x) \wedge \mathbf{TRUE}(t''; x) \wedge \mathit{meets}(t'; t'') \Rightarrow \mathbf{TRUE}(t' \oplus t''; x)$$

<Def.5> A proposition type x is *gestalt* if it never holds over two intervals one of which <u>properly</u> contains the other, that is:

$$\mathbf{TRUE}(t; x) \wedge \mathit{in}(t', t) \Rightarrow \neg\mathbf{TRUE}(t'; x)$$

<Def.6> A proposition type x is *solid* if it never holds over two <u>properly</u> overlapping intervals, that is:

$$\mathbf{TRUE}(t; x) \wedge (\mathit{overlaps}(t', t) \vee \mathit{overlapped\text{-}by}(t', t)) \Rightarrow \neg\mathbf{TRUE}(t'; x)$$

Of course, we can also give definitions which address time intervals and time points separately, such as *point-down-hereditary*, *point-upward-hereditary*, *interval-down-hereditary* and *interval-upward-hereditary*, and so on. However, all these may be taken as the special cases of the corresponding ones which address time elements in the general form. In fact, as Shoham puts in his paper [Sho87b], the categories of proposition type may be further devised. We prefer to stop our revised categorisation here.

It is interesting to note that, since the above categorisation addresses both time intervals and time points as primitive, Allen's trichotomy of *properties/events/processes* and McDermott's dichotomy of *facts/events* may be subsumed satisfactorily, e.g., Allen's *properties* may be characterised as the *liquid* propositions, while his *events* may be characterised as the *gestalt* propositions, defined above.

6.4 Toward Allen's and Galton's Properties

As Allen points out [All84], one of the most important predicates in his typed first order logic is **HOLDS**, which asserts that a property holds (i.e., is true) throughout a time interval. However, there are some problems with Allen's definition of properties. One limitation to the definition of properties by means of axioms (H.1) (see above section) is that it does not treat time *moments* satisfactorily, because any property *pro*, can be shown to hold over any time moment unconditionally from axiom (H.1), since there is not any proper subinterval within any given time moment. On the other hand, if all time elements are taken as infinitely decomposable intervals, Galton has shown in [Gal90b] that this will lead to inadequacies for reasoning about continuous change.

Another limitation of Allen's predicate, **HOLDS**, is that it characterises only one way of ascribing properties to times, namely to assert that a property holds throughout an interval, which seems too restrictive, representing only one category in the three types of statement introduced by Galton (see below). Additionally, it is interesting to note that, as shown by Galton in terms of his two different formulations, there are some problems with Allen's property-negation which is characterised by the following axiom (H.4):

$$\mathbf{HOLDS}(\mathrm{not}(pro),\mathrm{i}) \Leftrightarrow \forall \mathrm{i}_1 \in \mathbf{I}(In(\mathrm{i}_1,\mathrm{i}) \Rightarrow \neg(\mathbf{HOLDS}(pro,\mathrm{i}_1)))$$

In order to overcome the inadequacy of Allen's theory of action and time, Galton proposes a series of revisions which address time points in the theory as well as time intervals, and diversify the range of predicates assigning temporal locations to properties.

Galton's theory of points and intervals is built up by means of adding two extra relationships between points and intervals, as the extension to Allen's temporal relations between intervals. Rejecting the question whether or not a given point is part of, or a member of a given interval, while retaining the idea of there being a point at the meeting place of two intervals, Galton introduces two additional temporal relations to Allen's time theory: First, the point where two intervals meet each other is said to fall "*within*" the ordered union of these two intervals, and second, the same point is said to "*limit*" both of these two intervals, the former at its end, the latter at its beginning. Galton uses notions, *Within*(p,i) and *Limits*(p,i), to represent that a point p falls <u>within</u>, and <u>limits</u> an interval i, respectively.

Whereas Allen recognises only one way of ascribing properties to times, namely to assert that a property holds <u>throughout</u> an interval, Galton introduces the following three forms:

HOLDS-ON(*pro*, i),

HOLDS-IN(*pro*, i),

HOLDS-AT(*pro*, p),

where $i \in \mathbf{I}$ and $p \in \mathbf{P}$, for three types of statement: a property *pro* holds throughout an interval i, holds during i (i.e., at some time during an interval, not necessary throughout all of it), holds at a point p, respectively.

Commenting that the problems with Allen' system can all be traced to the assumption that all properties should receive a uniform treatment with respect to the logic of their temporal incidence, Galton proposes one of his revisions by distinguishing between two kinds of properties, namely *states of position* and *states of motion*, which have different temporal logics: States of position can hold at isolated points; and if a state of position holds throughout an interval, then it must hold at the limits of that interval, e.g., a body's being in particular position, or moving at a particular speed or in a particular direction. States of motion cannot hold at isolated points; if a state of motion holds at a point then it must hold throughout some interval within which that point falls, e.g., a body's being at rest or in motion (see [Gal90b], p.169). In terms of the above classes of properties, Galton characterises the formal constraints imposed on states of position (SP) and states of motion (SM) by the following axioms:

(SP) $\forall i \in \mathbf{I}(Within(p,i) \Rightarrow \mathbf{HOLDS\text{-}IN}(pro,i)) \Rightarrow \mathbf{HOLDS\text{-}AT}(pro,p)$,

(SM) $\mathbf{HOLDS\text{-}AT}(pro,p) \Rightarrow \exists i \in \mathbf{I}(Within(p,i) \wedge \mathbf{HOLDS\text{-}ON}(pro,i))$,

respectively.

Additionally, Galton lists a series of theorems which can be derived from the above axiomatisation. Some of them, i.e., (T1)-(T.10) ([Gal90b],pp.171-172) hold for general properties, regardless of whether they are states of position or states of motion; others, i.e., (T.11P)-(T.15P) ([Gal90b],pp.173-174), and (T.11M)-(T.15.M) ([Gal90b],pp.174-175) hold so long as *pro* is a state of position, and a state of motion, respectively.

Since the general temporal theory utilised in *TLIP* allows both intervals and points, it is straightforward to subsume Galton's corresponding terminologies. In fact,

a) *Within*(p, i) can be subsumed by *during*(p, i), and

b) *Limits*(p, i) can be subsumed by *meets*(p, i) $\vee$ *met-by*(p, i) $\vee$ *starts*(p, i) $\vee$ *finishes*(p, i),

where $p \in \mathbf{P}$ and $i \in \mathbf{I}$, and *during*, *meets*, *met-by*, *starts* and *finishes* belong to the classification of temporal relations presented in section 4.3.

In Addition, the axiom $<A_{P\text{-}Dense}>$ proposed in chapter 4 is indeed equivalent to the following required rule for Galton's temporal logic:

(I.1) $\forall i \in \mathbf{I} \exists p \in \mathbf{P}(\mathit{Within}(p, i))$

while other required rules for Galton's revised system, i.e.:

(I.2) $Within(p, i) \wedge In(i, j) \Rightarrow Within(p, j)$,

(I.3) $Within(p, i) \wedge Within(p, j) \Rightarrow \exists k \in \mathbf{I}(In(k, i) \wedge In(k, j))$,

(I.4) $Within(p, i) \wedge Limits(p, j) \Rightarrow \exists k \in \mathbf{I}(In(k, i) \wedge In(k, j))$

can be straightforwardly proved as lemmas from the time axiomatisation.

N.B. 1) In Galton's notation, borrowing from [All84], "*In*(i, j)" means that the interval i is a proper subinterval of j.

2) Interpreting *Within* and *Limits* as in a) and b), it is easy to see that these two relations are mutually exclusive to each other.

It is significant to note that, in the interval&point based logic presented in this chapter, the global predicate **TRUE** does not assume homogeneity or any other connection between a property holding for a time element and its holding for any substructure of the time element. In terms of "**TRUE**", for Allen's system, if we limit the set of time elements, **T**, to the set of time intervals **I**, we may simply define Allen's **HOLDS** by means of:

$$\mathbf{HOLDS}(pro, i) \Leftrightarrow \forall i' \in \mathbf{I}(sub(i', i) \Rightarrow \mathbf{TRUE}(i'; pro)),$$

Similarly, let p be a point, and i an interval. Then we may characterise Galton's form **HOLDS-AT**(*pro*, p) as:

HOLDS-AT(*pro*, p) $\Leftrightarrow$ **TRUE**(p; *pro*),

and replace his **HOLDS-IN** and **HOLDS-ON** by means of:

HOLDS-IN(*pro*, i) $\Leftrightarrow$ $\exists$p$\in$**P**(*during*(p,i) $\wedge$ **TRUE**(p; *pro*)),

HOLDS-ON(*pro*, i) $\Leftrightarrow$ $\forall$p$\in$**P**(*during*(p,i) $\Rightarrow$ **TRUE**(p; *pro*)),

respectively.

In fact, we may also give definitions characterising that a property holds in a time element, and holds throughout a time element. However, since time elements may now be intervals or points, we must make sure these definitions are well-defined for general treatment.

The following axiom defines what it is for a property to hold in a time element, namely that there is at least one "sub-element" of the time element for which the property holds.

<d.1> **hold-in**(t; *pro*) $\Leftrightarrow$ $\exists$t'$\in$**T**(*sub*(t', t) $\wedge$ **TRUE**(t'; *pro*)),

Similarly, for a property to hold throughout a time element, the property must hold for any sub-element of the time element, including the whole element itself.

<d.2> **hold-on**(t; *pro*) $\Leftrightarrow$ $\forall$t'$\in$**T**(*sub*(t', t) $\Rightarrow$ **TRUE**(t'; *pro*)).

N.B. The above definitions overcome the problem with Allen's axiom (H.1) for the case in which the addressed time element is non-decomposable, i.e. a *moment* (see section 3.4), and therefore is well-defined for both intervals/moments and points. Additionally, time intervals and points are addressed here on the same footing: there is no necessary connection between a property holding for intervals and its holding for points, while this connection is definitely axiomatised in Galton's theory.

Given the above definitions of **hold-in** and **hold-on**, by making use of our classification of temporal relations over intervals and points, we can readily prove the following theorems, which are in fact very similar to those (T.1 - T.5) given by Galton in [Gal90b]:

<t.1> **hold-in**(t; *pro*) $\wedge$ *sub*(t, t') $\Rightarrow$ **hold-in**(*t'*; *pro*)

which says that if a property holds in some time element t, then it holds in any time element of which t is a sub-element.

<t.2> **hold-on**(t; *pro*) $\wedge$ *sub*(t', t)$\Rightarrow$ **hold-on**(*t'*; *pro*)

which says that if a property holds throughout t, then it holds throughout every sub-element of t.

<t.3> **hold-on**(*t*; *pro*) $\Rightarrow$ **hold-in**(t; *pro*)

which says that if a property holds throughout t, then it holds in t.

<t.4> $\mathbf{hold\text{-}on}(t; \mathit{pro}) \Rightarrow \forall t' \in \mathbf{T}(\mathit{sub}(t', t) \Rightarrow \mathbf{hold\text{-}in}(t'; \mathit{pro}))$

which says that if a property holds throughout t, then it holds in every sub-element of t.

<t.5> $\exists t \in \mathbf{T}(\mathit{sub}(t, t') \wedge \mathbf{hold\text{-}on}(t; \mathit{pro})) \Rightarrow \mathbf{hold\text{-}in}(t'; \mathit{pro})$

which says that if a property holds throughout some sub-element of t, then it holds in t.

N.B. Whereas Galton gives more theorems relating to his **HOLDS-AT** for points, that is (T.6) and (T.7), in our revised system, they are the same as <t.5> and <t.2>, respectively, since points are treated now on the same footing as intervals. Examples relating to these theorems may be found in Galton's paper ([Gal90b], pp.171-172).

Similarly, for general treatment of properties, we can introduce their negation by means of:

<d.3> $\mathbf{TRUE}(t; \mathrm{not}(\mathit{pro})) \Leftrightarrow \neg\mathbf{TRUE}(t; \mathit{pro})$,

and easily prove the following theorems:

<t.6> $\mathbf{hold\text{-}on}(t; \mathrm{not}(\mathit{pro})) \Leftrightarrow \neg\mathbf{hold\text{-}in}(t; \mathit{pro})$,

<t.7> $\mathbf{hold\text{-}in}(t; \mathrm{not}(\mathit{pro})) \Leftrightarrow \neg\mathbf{hold\text{-}on}(t; \mathit{pro})$,

<t.8> $\mathbf{TRUE}(t; \mathrm{not}(\mathrm{not}(\mathit{pro}))) \Leftrightarrow \mathbf{TRUE}(t; \mathit{pro})$.

N.B. We omit the proofs of the above theorems here since they are very straightforward and indeed will be very similar to the corresponding proofs given in the appendix in Galton's paper [Gal90b].

Also, we may formally axiomatise the characteristic of a state of position s_p by:

$$
\begin{aligned}
\forall i \in \mathbf{I} \forall p \in \mathbf{P}(\quad & \mathbf{hold\text{-}on}(i;\ s_p) \\
& \wedge\ (\mathit{met\text{-}by}(i,p) \vee \\
& \quad \mathit{meets}(i,p) \vee \\
& \quad \mathit{started\text{-}by}(i,p) \vee \\
& \quad \mathit{finished\text{-}by}(i,p)) \\
& \Rightarrow \mathbf{hold\text{-}in}(p;\ s_p)\ )
\end{aligned}
$$

and a state of motion s_m by:

$$
\forall p \in \mathbf{P}(\mathbf{hold\text{-}at}(p;\ s_m) \Rightarrow \exists i \in \mathbf{I}(\mathit{during}(p,i) \wedge \mathbf{hold\text{-}on}(i;\ s_m)))
$$

It is interesting to note that, the definitions relating to the open and closed nature of intervals given in section 4.2.1 provide another formal and intuitive characterisation for the distinction between *states of position* and *states of motion*: States of position can hold at isolated points; and if a states of position holds on an interval, then it must hold on the closure of that interval. States of motion hold only on open intervals. For instance, in the example of a ball thrown vertically into the air described in section 3.4, the property *ball_stationary* may be taken as a state of position, while *ball_going_up* and *ball_coming_down* may be taken as states of motion.

However, by some further examination, it is found that there are still some problems with Galton's revised treatments of general properties. Galton's argument that addressing both time points and time intervals on the same footing is necessary for accommodating the representation of facts concerning continuous change is indeed very suggestive. However, as in Allen and Hayes's approach [AlH85,89], Galton defines time points as the meeting places of time intervals. Hence, from the view of the abstract axiomatisation about time itself, time points are still relegated to a subsidiary status, not really treated on the same footing as time intervals. Additionally, to develop his revised logic, Galton imposes a very strict rule, that is, (I.1), which states that for any time interval, there exists a point which falls within this interval. It is easy to see from Galton's definition of time points, rule (I.1) implies that, any time interval i can be decomposed to two proper subintervals i_1 and i_2, such that *meets*(i_1,i_2). Further, it is straightforward to infer that any time interval is required to be <u>infinitely</u> decomposable. Hence, Galton's revised axiomatisation definitely excludes the special time intervals that are non-decomposable, namely moments, in Allen and Hayes' theory. This limitation is perhaps not too serious, since Allen and Hayes' conception of time moments are in fact introduced to characterised the times that some "instant-like" events occupy (although time moments still have positive duration). We may simply utilise time points to play the role of moments.

However, there are some problems with Galton's revisions which we shall show require revisions to the fundamental axioms about time itself in order to extend the abstract concept of time elements to include both intervals and points, and the temporal relations between intervals in order to address points as well.

Let *at-rest* and *in-motion* refer to the properties of a body being at rest, and being in motion; and i_1 and i_2 represent two time intervals **throughout** which *at-rest* and *in-motion* hold, respectively. Intuitively, we may assume that *meets*(i_1,i_2), and if we use p to denote the point at which i_1 "*meets*" i_2, from the definition of "*Limits*" we get:

Limits(p,i_1) $\wedge$ *Limits*(p,i_2).

According to the distinction between the two kinds of properties, Galton claims both *at-rest* and *in-motion* are states of motion (see [Gal90b], p.169).

Assuming that the state of motion *at-rest* holds at point p: by Galton's definition, it must hold throughout some interval i' such that *Within*(p,i'). Hence, together with *meets*(i_1,i_2), *Limits*(p,i_1) and *Limits*(p,i_2), we can infer that *overlaps*(i',i_2). Hence, both properties *at-rest* and *in-motion* will hold throughout an interval which is a common subinterval of both i' and i_2. This is unsatisfactory.

Similarly, if the state of motion, *in-motion*, holds at point p, then it must hold throughout some interval i'' such that *Within*(p,i''). Hence, together with *meets*(i_1,i_2), *Limits*(p,i_1) and *Limits*(p,i_2), we can infer that *overlaps*(i_1,i''). Hence, both *at-rest* and *in-motion* will hold throughout the common subinterval of i_1 and i''.

Hence, the above proof shows that at the point p, the body is neither at rest nor in motion.

Generally, if *$state_1$* and *$state_2$* are two <u>opposite</u> states (i.e., *$state_1$* $\Leftrightarrow$ not(*$state_2$*)), that hold throughout intervals i_1 and i_2, respectively, where *meets*(i_1,i_2), we can prove in the same way as above that the following case:

(a) both *$state_1$* and *$state_2$* are states of motion

is not allowable in Galton's revised logic.

Similarly, consider case

(b) both *$state_1$* and *$state_2$* are states of position:

If we again use p to denote the point at which i_1 meets i_2, then by Galton's definition of a state of position, *$state_1$* must hold at point p, which is one of i_1's limits, since *$state_1$* holds throughout interval i_1; similarly, since *$state_2$* holds throughout interval i_2, it must holds at p as well, which is also one of i_2's limits. Hence both *$state_1$* and *$state_2$* hold at point p. Again, this is absurd.

Hence, according to Galton's classification of states, there are only two possible cases:

(c) *$state_1$* is a state of position and *$state_2$* is a state of motion,

(d) *$state_1$* is a state of motion and *$state_2$* is a state of position.

However, both (c) and (d) seem arbitrary and unsatisfactory since there may be nothing to choose between them. For example, how do we decide which of the following two states

$state_{yes}$: the car is John's

$state_{not}$: the car is no longer John's

should be considered as a state of position, and which must be addressed as a state of motion?

In fact, it is interesting to note that, in Galton's paper, it is not explicitly expressed whether states of position and states of motion are all the possible kinds of properties or not. However, it is not obvious what other kinds of property will be needed to avoid the problem outlined above.

A further problem arises in connection with assignment of properties to time intervals. Noticing that it is necessary to extend Allen's single way of ascribing properties to times, namely to assert that a property holds throughout an interval, Galton introduces three different ways. For the initial, general treatment, he takes the locution **HOLDS-AT** (relating to time points) as primitive, and defines the other two, **HOLDS-IN** and **HOLDS-ON**, in terms of **HOLDS-AT** as below:

(D.1) $\mathbf{HOLDS\text{-}IN}(pro, \mathrm{i}) \Leftrightarrow \exists \mathrm{p} \in \mathbf{P}(Within(\mathrm{p,i}) \wedge \mathbf{HOLDS\text{-}AT}(pro, \mathrm{p}))$,

(D.2) $\mathbf{HOLDS\text{-}ON}(pro, \mathrm{i}) \Leftrightarrow \forall \mathrm{p} \in \mathbf{P}(Within(\mathrm{p,i}) \Rightarrow \mathbf{HOLDS\text{-}AT}(pro, \mathrm{p}))$.

Later, Galton shows that ([Gal90b],pp.174-175), for states of position, **HOLDS-IN** may be taken as primitive instead of **HOLDS-AT** by means of the following theorems for states of position:

(T.14P) $\quad \textbf{HOLDS-AT}(\textit{pro}, \text{p}) \Leftrightarrow \forall \text{i} \in \textbf{I}(\textit{Within}(\text{p,i}) \Rightarrow \textbf{HOLDS-IN}(\textit{pro}, \text{i}))$

(T.15P) $\quad \textbf{HOLDS-ON}(\textit{pro}, \text{i}) \Leftrightarrow \forall \text{i'} \in \textbf{I}(\textit{In}(\text{i',i}) \Rightarrow \textbf{HOLDS-IN}(\textit{pro}, \text{i'}))$

and, for states of motion, **HOLDS-ON** may be taken as primitive instead of **HOLDS-AT**, or **HOLD-IN**, by means of:

(T.14M) $\quad \textbf{HOLDS-AT}(\textit{pro}, \text{p}) \Leftrightarrow \exists \text{i} \in \textbf{I}(\textit{Within}(\text{p,i}) \wedge \textbf{HOLDS-ON}(\textit{pro}, \text{i}))$

(T.15M) $\quad \textbf{HOLDS-IN}(\textit{pro}, \text{i}) \Leftrightarrow \exists \text{i'} \in \textbf{I}(\textit{In}(\text{i',i}) \wedge \textbf{HOLDS-ON}(\textit{pro}, \text{i'}))$

However, consider the following example: let $\textit{pro}_0$ represent the property of *having zero duration*, and $\textit{pro}_+$ represent the property of *having positive duration*, then

$$\forall \text{p} \in \textbf{P}(\textbf{HOLDS-AT}(\textit{pro}_0, \text{p}))$$

hence, by (D.2), we get that for any time interval,

$$\textbf{HOLDS-ON}(\textit{pro}_0, \text{i})$$

That is, the property, *having zero duration*, will hold **throughout** any time interval which is assumed to have a positive duration. This seems contrary to both human intuition and the corresponding axiomatisation. On the other hand, since each interval i, as well as any subinterval of i, has a positive duration, we intuitively have that the property of *having positive duration* holds during interval i, that is:

HOLD-IN(pro_{+}, i)

However, by (D.1) we will get that there exists a point p within the interval i, such that

HOLDS-AT(pro_{+}, p)

that is, this point has a positive duration, which is again contrary to our assumptions.

The source of the above problems is indeed in the determination to define time points in terms of the meeting places of time intervals, and define the corresponding types of predicates ascribing properties to times, either according points conceptual priority over intervals, or regarding intervals as conceptually prior to points.

6.5 Toward Allen's and Galton's Events

In addition to properties, in Allen's interval based system, *processes* and *events*, generally termed as *occurrences*, are addressed as well. However, as Galton argues in his corresponding examination [Gal90b], by locating the distinction between *broad sense* and *narrow sense* [Gal84] in processes rather than the time, Allen's processes may be in fact subsumed from properties and events. That is, it is unnecessary to introduce a category of processes separate from properties and events. Hence, in this section we shall only consider some special issues about events.

As mentioned in section 6.3, Allen's *events* which are characterised by his axiom (O.1):

$$\mathbf{OCCUR}(e,\mathrm{i}) \wedge In(\mathrm{i}',\mathrm{i}) \Rightarrow \neg\mathbf{OCCUR}(e,\mathrm{i}')$$

are in fact coincident with *gestalt propositions*. On the one hand, of course, in Allen's system, instantaneous events are definitely excluded, although they do occur in reality (Examples are given in [Gal90b], p.178). On the other hand, it is interesting to note that, in addition to his three different predicates **HOLDS-AT**, **HOLDS-IN** and **HOLDS-ON** for properties, Galton has gone on to replace Allen's **OCCUR** by means of three predicates **OCCURS-AT**, **OCCURS-IN** and **OCCURS-ON**. The first of these is for locating an *instantaneous* (termed *punctual*) event at the point at which it occurs, the second for locating an event (punctual or *durative*) in an interval within which it occurs, and the third is for locating a durative event which takes time on an interval over which it occurs (i.e., **OCCURS-ON** corresponds to Allen's **OCCUR**).

In *TLIP*, we may also specially define a primitive predicate, **occur**, which locates an event, that is, *gestalt proposition*, over a time element on which it happens. The definition for **occur** is as below:

<d.4> **occur**(t; e) $\Leftrightarrow$ **TRUE**(t; e) $\wedge$ $\neg\exists t_1 \in$ **T**($in(t_1,t)$ $\wedge$ **TRUE**(t_1; e))

which, although it addresses time points as well, is very similar to Allen's (O.1) and Galton's (O.2) (see [Gal90b], p.179). In fact, Allen's (O.1) may be replaced by means of simply limiting the set of time elements, **T**, to the set of time intervals **I**.

Similarly, we can also define another predicate **occur-in**, for locating an event over a time element in which it occurs:

<d.5> **occur-in**(t; e) $\Leftrightarrow$ $\exists t_1 \in$ **T**($sub(t_1,t)$ $\wedge$ **occur**(t_1; e))

N.B. In the extreme case where t is a point, p, since the "*sub*" relation includes "*equal*" relation, we get that: **occur-in**(p; e) $\Leftrightarrow$ **occur**(p; e).

Note that our predicates **occur** and **occur-in** address both instantaneous and durative events: for an instantaneous event e, if we let $t \in P$, Galton's **OCCURS-AT**(e, p) can be simply taken as our **occur**(p; e); for a durative event e, if we let $i \in I$, again, Galton's **OCCURS-ON**(e, i) can be simply taken as our **occur**(p; e); and, for a general event e (instantaneous or durative), if we let $i \in I$, Galton's **OCCURS-IN**(e, i) can be simply taken as our **occur-in**(i; e).

From the definitions of **occur** and **occur-in**, we can straightforwardly prove the following theorems:

<t.9> **occur**(i; e) $\wedge$ $sub(i,i_1)$ $\Rightarrow$ **occur-in**(i_1; e)

<t.10> **occur-in**(i; *e*) $\wedge$ *sub*(i, i_1) $\Rightarrow$ **occur-in**(i_1; *e*)

It is interesting to note that, from <t.9>, <t.10> along with our <d.4> and <d.5>, we can subsume Galton's (O.1)-(O.4), (O.5D), (O.5P), (T.16D) and (t.16P) (see [Gal90b], p.179).

6.6 The Expressive Power of the New Logic

One important intuition which leads to Allen's interval-based logic is that most of human temporal knowledge, especially in the field of AI, is expressed without explicit reference to time points. As Allen repeatedly argues [All81,83,84], if one insists on addressing the ending-points of time intervals, one must consider what knowledge one has at them about properties which are naturally associated only with the intervals. Allen's idea is therefore to take intervals as primitive, excluding the concept of points explicitly from the fundamental theory, and to maintain that only knowledge about properties associated with intervals is necessary. For instance, again consider the example of switching on a light: this case may be expressed by using an interval i to denote the time over which the light is off, and another interval j to denote the time over which the light is on, where interval j is *immediately after* interval i. In a *pure* interval-based system, that is where intervals are treated as primitive time elements, not constructed out of points, this may be conveniently expressed, for example in Allen's notation [All84], as:

HOLDS(*Light_Off*, i),

HOLDS(*Light_On*, j),

meets(i, j).

Knowledge about whether the light is on or off at the "switching point" is not to be represented in this system. But this is to be considered an advantage in this example, since we do not have any firm knowledge about the state of the system at this point.

However, as Galton [Gal90b] has shown, excluding time points may lead to inadequacy in reasoning about continuous change. For instance, in section 3.4, we have illuminated the problem involved with reference to time points by means of the example of a ball thrown vertically into the air, and show that this situation cannot be satisfactorily expressed in terms of Allen's interval-based logic.

In order to overcome this inadequacy, Galton proposes his series of revisions to Allen' system to accommodate the representation of facts concerning continuous change. In terms of Galton's terminology, one can now express the situation of a ball thrown vertically into the air as:

HOLDS-ON(*Ball_Going_Up*, i),

HOLDS-AT(*Ball_Stationary*, p),

HOLDS-ON(*Ball_Going_Down*, j),

meets(i, j),

Limits(p, i),

Limits(p, j),

where both *Ball_Going_Up* and *Ball_Going_Down* are *states of motion*, while *Ball_Stationary* is a *state of position* (Galton's revised terminology, see [Gal90b]).

However, Galton's revisions are in fact achieved by insisting on the existence of a point between any two intervals that meet, which leads back to the very problem that Allen tries to avoid and Galton tries to reject as meaningless: viz do properties ascribed to intervals apply to points or not? For example, how will the situation that a light is turned on be expressed in Galton's revised system?

An advantage of the interval- and point- based logic proposed in this paper is that it does not suffer from these problems. In fact, on the one hand, since intervals and points are now treated as primitive time elements of equal standing, it is not necessary to insist on the existence of a point between any two time intervals which meet each other. Hence, our logic still retains Allen's solution which allows the expression of knowledge of properties over time intervals only. On the other hand, since points are now addressed as time elements on the same footing as intervals, the logic also allows for correct reasoning about continuous change.

For instance, in the new logic, the situation that a light is turned on may be simply expressed as:

hold-on(i; *Light_Off*),

hold-on(j; *Light_On*),

meets(i, j).

Here we have expressed exactly the knowledge about the light being on or off, without being forced into expressing disputable knowledge about the property at any "switching point". However, if such knowledge happened to be available then we could include it. For instance if on some grounds we were to have the knowledge that the light was on and not off at the switching point p between i and j, we would have:

hold-on(i; *Light_Off*),

hold-on(p; *Light_On*),

hold-on(j; *Light_On*),

meets(i, p),

meets(p, j).

Similarly, the situation that a ball thrown vertically into the air may be conveniently expressed as:

hold-on(i; *Ball_Going_Up*),

hold-on(p; *Ball_Stationary*),

hold-on(j; *Ball_Going_Down*),

meets(i, p),

meets(p, j).

N.B. In the above, for the time point p, **hold-on**(p; *x*) may be replaced by **hold-in**(p; *x*).

CHAPTER 7

A NEW ARCHITECTURE FOR TEMPORAL DATABASES

In this chapter, we shall present an architecture for a temporal database system which allows the expression of relative temporal knowledge of data transaction and data validity times. The system is founded on the time axiomatisation given in Chapter 4, which allows both time intervals and time points. A general retrieval mechanism is presented for a database with a purely relative temporal knowledge which allows queries with temporal constraints in terms of any logical combination of possible temporal relations classified in section 4.3. However, when absolute time duration knowledge is added, the consistency checking algorithm upon which the inference mechanism is based becomes a linear programming question. Much work has gone into algorithms for linear programming. As Cormen et al examine in [CLR89], linear programs can in practice be solved very quickly by means of the *simplex algorithm* (see p.539, [CLR89]). However, with some carefully contrived inputs, the simplex method can lead to exponential complexity. General linear programs can be solved in polynomial time by either *Karmarkar's algorithm*, which in practice is often competitive with the simplex method, or the *ellipsoid algorithm* which, however, runs slowly in practice [CLR89]. For the sake of reducing the computational complexity, a class of databases, termed *time-limited* databases, is introduced as a practical solution for the inference mechanism. This class allows absolute-time-stamped and relative time information in a form which is suitable for many practical applications, where relative temporal information is only occasionally needed. The architecture of such a system is given, and it is shown that efficient retrieval mechanisms for absolute-time-stamped databases may be adapted to time-limited databases.

7.1 The Need for the New Architecture

Designing an information system means modelling some portion of the real world in a suitable way, such that the model can be represented on a computer (data definition), providing mechanisms for maintenance and evaluation of that representation (data manipulation). For many applications in database systems (Dbs), it is necessary to retain complete information about objects. The current value of their attributes, as well as their histories, should be stored and managed by the DB. In this case, it is not appropriate to discard old information. Time values are necessary to be associated with data to indicate their periods of validity. However, a conventional database is updated from time to time; old data is deleted and new data inserted. Thus only current information resides in the database so that conventional database management systems (DBMSs) lack the capability to record and process time-varying aspects of the real world. With increasing sophistication of DB applications, the lack of temporal support raises serious problems in many cases. For example, conventional DBMSs cannot support historical queries about past status, nor trend analysis on a series of versions (essential for applications like decision support systems). There is no way to represent retroactive or proactive changes, while support for error correction or audit trail necessitates costly maintenance of backups, checkpoints, or transaction logs to preserve past states. There is a growing interest in applying database methods for version control and design management in computer-aided design, requiring capabilities to store and process time-dependent data. "Without temporal support from the system, many applications have been forced to manage temporal information in an ad-hoc manner" [SnA86].

The incorporation of time into conceptual database models has been an active area for research over the past decade. In [Ari86,91], Ariav has examined the nature of temporally oriented data definition, delineating the problems involved and established a platform for their discussion. Several approaches to the problem have used a relational model which includes temporal attributes. For example, Jones and Mason [JoM80], and Sarda [Sar90], address starting and finishing times as attributes for each whole tuple which define its *validity*; Gadia [Gad88] proposes a system in which each attribute value is stamped with a *temporal element* which is a finite union of point-based intervals; and Clifford [Cli85,87] includes functions from valid times to attribute values for *time-varying attributes*. Additionally, it has been recognised that there may be more temporal attributes required for temporal database management systems. In fact, in Ben-Zvi's time relational model [Ben82], five implicit time attributes have been addressed: *effective-time-start* and *effective-time-stop* are respectively the left and right end points of the time interval for the existence of the real-world phenomenon being modelled; *registration-time-start* is the time at which the effective-time-start was stored; *registration-time-stop* is the time at which the effective-time-stop was stored; and *deletion-time* records the time when erroneously entered tuples are logically deleted. Subsequently, Snodgrass and Ahn [SnA86,Sno87], McKenzie [Mck88,McS91], and Clifford [Cli93] have proposed systems with temporal attributes, which are semantically similar to Ben-Zvi's. E.g., in [Sno87], four implicit times, *valid-from*, *valid-to*, *transaction-start* and *transaction-stop* are addressed. Common to these later approaches is that the attributes encode both valid time and transaction time for each tuple. In the glossary of temporal database concepts [JCG92], Jensen et al. have given definitions of valid time and transaction time, as below:

- *Valid Time*: The valid time of a fact is the time when the fact is true in the modelled reality.

- *Transaction Time*: A database fact is stored in a database at some point in time, and after it is stored, it may be retrieved. The transaction time of a database fact is *the time when the fact is stored in* the database. Transaction times are consistent with the serialisation order of the transactions. Transaction time values cannot be after the current time.

Hence, here, we take the transaction time of a database fact as the time over which the fact is (was) taken as part of the current state of the database. It is interesting to note that the time, or the "point in time", at which a database fact is stored in the database is sometimes termed its *recording time*: in fact, recording time can be taken as the *transaction-start* for the transaction time attribute in Snodgrass' system [Sno87].

Some systems, such as those of Jones and Mason [JoM80] and of McKenzie and Snodgrass [McS91,Sno87] take the tuple as the fundamental transaction unit, so that whenever anything changes within the tuple, the whole tuple is regarded as being renewed. Other systems, such as that of Gadia [Gad88], take individual attribute values as the fundamental transaction unit. Ahn [Ahn86] has shown that these two views are entirely equivalent, and Ling and Bell

[LiB92] have pointed out that the appropriate transaction units for any application are to do with practical questions of storage and retrieval efficiency. We shall adopt the tuple view here, since it makes for clearer presentation.

Most temporal database systems require absolute time values for the temporal attributes. However, there may be temporal knowledge about a valid/transaction interval even though precise starting and finishing times are unknown. For example, we may know that event A happened *before* event B, without knowing the absolute times when they actually started/finished. Relative temporal knowledge such as this is typically derived from humans, where absolute times are rarely remembered, but relative temporal relationships are often remembered.

In order to be capable of representing not only absolute, but also relative and imprecise temporal information, some new temporal database systems have been introduced, examples are those of Chaudhuri [Cha88], and of Koubarakis [Kou93]. Chaudhuri's graph model is proposed as a tool in identifying generic temporal queries, and in describing the process of deduction of temporal relationships. The model represents time elements as nodes, and temporal relationships between time elements as arcs of the graph. Queries to the graph are processed by propagation of temporal relationships along arcs of the graph, according to Allen's transitivity table [All83]. The complexity of this process is a major problem of the deductive method, and Chaudhuri discusses how heuristics may be used in some cases to solve this problem. However, the model is restricted to a subset of binary temporal relationships only; it cannot represent mixed relative time and absolute time duration information. Additionally, it cannot represent disjunctive constraints, and doesn't address issues such as transaction time/valid time explicitly. Consistency checking, which has been a problematic question in most temporal systems, is not addressed in Chaudhuri's model.

Koubarakis's system is based on relation-like representations which can contain variables constrained by the formulas of his temporal theory [Kou93]. The underlying time theory for the system is point-based: points are identified with the rationals while intervals are considered as pairs of points. Hence, the time basis for this system is similar to the point-based constraint network of Dechter et al. [DMP91], where temporal predicates over intervals must be expressed in terms of the temporal order over the interval end-points. However, Vilain [Vil82,ViK86] and Van Beek [Bee89,92] have examined the complexity issues relating to interval/point algebra, showing that the computational complexity of the constraint evaluation algorithm may be prohibitive for practical systems. Additionally, temporal relationships between intervals must be expressible in terms of the order relation over the interval end-points, and the annoying question of end-points must be artificially addressed.

The need for temporal databases of the kind discussed here occurs in applications where one might possess some temporal knowledge in relative form. We give an illustrative example of such a system, from the medical field, to assist the presentation of the ideas.

Consider a patient, who we shall call Lee, attending a clinic: On his first attendance, Lee described that several weeks before, he felt some stomach pain and took a pain-killer (drug a) for some days. However this did no good, and he visited the doctor. A doctor (Dr. Major) cancelled drug a, and prescribed drug b and drug c for three days, asking Lee to come to the clinic after finishing the drugs. Unfortunately, Dr. Major forgot to write down the prescription of drug c on Lee's medical record. The next time that Lee attended the clinic, he was examined by another doctor (Dr. Long). Since Lee's condition had changed for the better, and his medical record wrongly indicated that the treatment had been just drug b, Dr. Long simply prescribed drug b for another three days. Thereupon Lee's condition deteriorated. Lee came back to the clinic just after finishing drug b. This time, Dr. Major discovered the mistake, and corrected Lee's medical record. In addition, Dr. Major prescribed drug b and drug c for another seven days.

This example involves very little data in the traditional sense, but involves many different temporal references. The data essentially describes a treatment (the list of drugs prescribed). The valid time of the treatment may be characterised as an interval with a positive length. The exact left-end-point and right-end-point of each of these periods are not always specified. However, some temporal facts are known relating to the periods. For example, the period of treatment with drug a comes immediately before treatment with b and c, although we do not know its start time, or duration.

Another interesting aspect of temporal databases illustrated by this example is the importance of the transaction time to the inference mechanism. According to the view of the database at the time when Lee was examined by Dr. Long, the database wrongly indicated that Lee's condition changed for the better because of the effect of just drug b, although actually drug b and c together were taken. However, it is not appropriate to delete this wrong information from the database, since if we do so we cannot at a later date tell why doctor Long didn't prescribe drug c at that time. Instead of over-writing the wrong information, the corrected version must be added to the database, but with its own transaction time.

The main problem involved with relative temporal databases is that of temporal inference since for relative temporal data one often needs deeper reasoning to infer facts. For example, if we have a database with the following facts:

Interval i_1:	starting time = 8.00	finishing time = 9.00,
Interval i_4:	starting time = 11.15	finishing time = 11.30,

then we only need to retrieve i_1 and i_4 to find directly that i_1 is before i_4. However, if we have the following relative information:

Interval i_1 before Interval i_2, Interval i_2 before Interval i_3, Interval i_3 before Interval i_4,

then we need to retrieve facts from the database other than those relating to i_1 and i_4 (viz i_2 and i_3), and then we must deduce that i_1 is before i_4. This problem is dealt with in section 7.3 below, where a general mechanism for query evaluation is presented. The mechanism is based on the refutation principle, with a general consistency checker for interval and point based system given previously in section 5.4.4 (or see Appendix A).

The time theory used in the architecture allows both time points and time intervals as primitive, so as to handle human temporal information in a natural way, consistent with Allen's approach. This is different from the approach of Koubarakis, which in fact reduces to a point-based constraint network problem. It is also more comprehensive than the approach of Chaudhuri since it allows mixed relative and absolute time. Also a general retrieval mechanism is given based on refutation by means of a consistency checker, supporting any conventional data conditions, and both conjunctive and disjunctive temporal constraints.

7.2 The Architecture of the Temporal Database

In this section, we present the architecture of the system in terms of the relational model. Following the conventional concepts, a nontemporal relation R with nontemporal attributes A_1, ..., A_n defines a nontemporal relation scheme, denoted as R(A_1, ..., A_2), where each attribute A_i takes values from its domain, D_i, a set of data. A nontemporal relation is a subset of the Cartesian product of one or more domains. Conventionally, a relation is envisioned as a table of data values, where the rows of such a table are termed tuples, and values of an attribute associated with column i of the relation are taken from domain D_i. In the architecture presented here, temporal reference of the system is made by assigning two time-elements to each nontemporal tuple, which denote the valid time and transaction time respectively. Knowledge about the temporal order over time elements is represented by a table of *meets* relationships over the corresponding time elements. Hence, corresponding to the nontemporal relation scheme R(A_1, ..., A_n), we can define the temporal relation scheme as:

$$\text{T-R}(T_{\text{transaction}}, A_1, \ldots, A_n, T_{\text{valid}}),$$

together with a *meets* table:

$$\textit{meets}(T_{\text{first-argument}}, T_{\text{second-argument}}),$$

where instances of $T_{transaction}$, T_{valid}, $T_{first\text{-}argument}$ and $T_{second\text{-}argument}$ are all taken from the set of time elements, which can be intervals or points.

In what follows, we shall illustrate the ideas in terms of the simple medical record example outlined in the introduction.

For the example, if we take nontemporal attributes:

Patient, **Prescriber**, **Drug**, and **Status**,

then the database may be represented by the following schema:

medical-record($T_{transaction}$, Patient, Prescriber, Drug, Status, T_{valid}),

meets($T_{first\text{-}argument}$, $T_{second\text{-}argument}$),

Correspondingly, we may illustrate the medical-record relation representing Lee's history as the following table, MEDICAL-RECORD (Table 7.2(1)):

MEDICAL-RECORD

$T_{TRANSACTION}$	PATIENT	PRESCRIBER	DRUG	STATUS	T_{VALID}
i_{t1}	Lee	NULL	NULL	pain	i_{v1}
i_{t1}	Lee	Lee	a	worse	i_{v2}
i_{t1}	Lee	Major	b	NULL	i_{v3}
i_{t2}	Lee	NULL	NULL	pain	i_{v1}
i_{t2}	Lee	Lee	a	worse	i_{v2}
i_{t2}	Lee	Major	b	better	i_{v3}
i_{t2}	Lee	Long	b	NULL	i_{v4}
i_{t3}	Lee	NULL	NULL	pain	i_{v1}
i_{t3}	Lee	Lee	a	worse	i_{v2}
i_{t3}	Lee	Major	b&c	better	i_{v3}
i_{t3}	Lee	Long	b	worse	i_{v4}
i_{t3}	Lee	Major	b&c	NULL	i_{v5}

Table 7.2(1).

In this example there are 5 validity intervals shown:

i_{v1} -- the time when Lee first felt pain,

i_{v2} -- the time when Lee was treated by means of drug a, administered by himself,

i_{v3} -- the time when Lee was treated by drug b and drug c, prescribed by Dr. Major,

i_{v4} -- the time when Lee was treated by drug b, prescribed by Dr. Long,

i_{v5} -- the time when Lee was again treated by drug b and drug c, prescribed by Dr. Major.

There are also 3 transaction times shown:

i_{t1} -- transaction interval following data entry after first appointment with Dr. Major,

i_{t2} -- transaction interval following data entry after first appointment with Dr. Long,

i_{t3} -- transaction interval following data entry after second appointment with Dr. Major.

The graphical representation of the corresponding time network may be given as that in Figure 7.2(1).

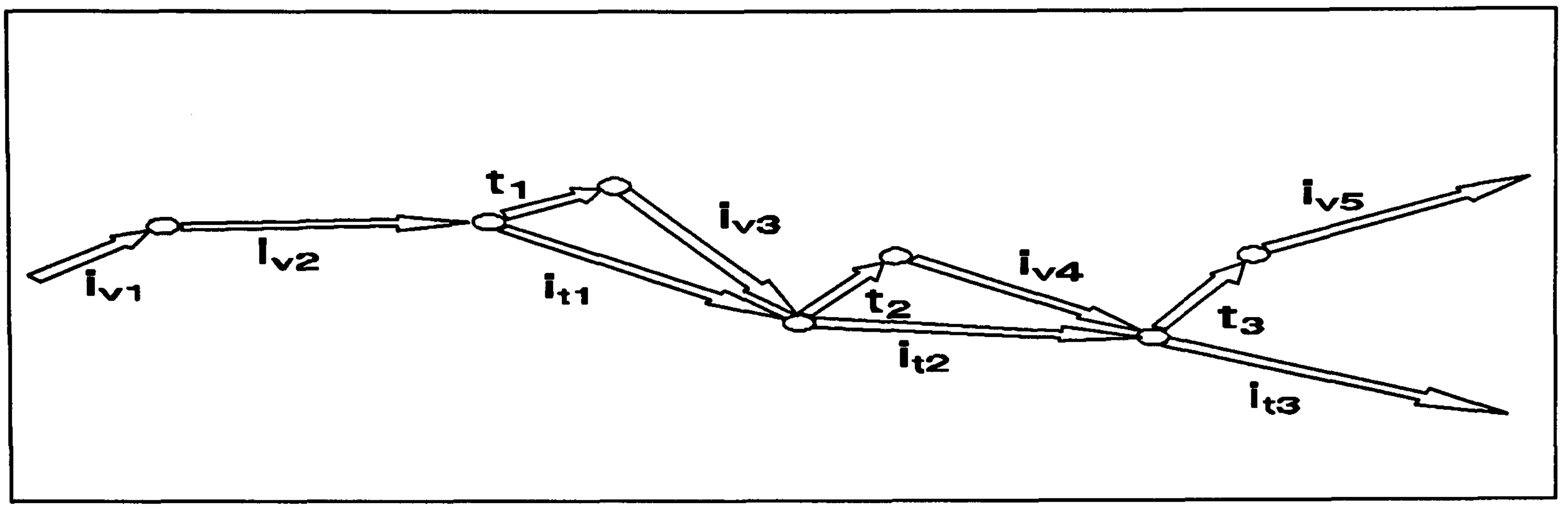

Figure 7.2(1)

The record shows three transaction intervals for the validity interval i_{v3}. The third (interval i_{t3}) corrects the second (interval i_{t2}) which incorrectly records that Lee was taking drug b alone. However, the reason for Dr. Long's prescription of drug b alone for interval i_{v4} (i.e. that he thought drug b alone had led to improvement over i_{v3}) may still be inferred from the database. This is because the transaction time allows us to retrieve the state of the database at any time in the past. Also, these transaction intervals are *consistent with the serialisation order of the transactions* (see Jensen et al.'s definition given in section 7.1). Temporal queries depending on both validity time and transaction time will be discussed in section 7.3 below.

The temporal part of the database is represented by the 8 time elements: i_{v1}, i_{v2}, i_{v3}, i_{v4}, i_{v5}, i_{t1}, i_{t2}, and i_{t3}, along with the *meets* predicate imposed over them. As a table, it may be represented as:

meets(i_{v1}, i_{v2}), *meets*(i_{v2}, t_1), *meets*(i_{v2}, i_{t1}), *meets*(t_1, i_{v3}), *meets*(i_{v3}, t_2), *meets*(i_{v3}, i_{t2}), *meets*(i_{t1}, t_2), *meets*(i_{t1}, i_{t2}), *meets*(t_2, i_{v4}), *meets*(i_{v4}, t_3), *meets*(i_{v4}, i_{t3}), *meets*(i_{t2}, t_3), *meets*(i_{t2}, i_{t3}), *meets*(t_3, i_{v5})

where t_1, t_2 and t_3 are "relative delay" times for the system. E.g., *meets*(i_{v2},t_1) and *meets*(t_1,i_{v3}) express the fact that the valid time interval i_{v3} is "*after*" i_{v2}, by means of the delay time, t_1, standing between them.

7.3 The Inference Mechanism

In this section, we consider the inference mechanism for the system. For queries to the database, we wish to support all the possible temporal relations over time elements. For example, "show me all facts *before* time t_1 or *after* time t_2", or "Is it true that fact A holds *during* time i and *before* time t?". All these temporal queries can be characterised in a general query form. The general method for query evaluation depends upon a general consistency checker for the database in terms of the necessary and sufficient conditions for the consistency of the corresponding time network (section 5.4.4 and Appendix A).

For temporal queries, temporal constraints will be added as additional querying conditions. Although the relationships between time elements are characterised in terms of the single predicate, "*meets*", we should allow a more general form of query. This should allow temporal constraints in terms of all the possible temporal predicates presented in section 4.3, over both transaction times and valid times, and any other given reference time elements.

For the general operation, **SELECT**, data and temporal selection conditions can be defined as follows:

1. Any conventional data selection condition of *SQL* is a data selection condition.

2. A temporal constraint, $\vee r_i(t, t_j)$, is a temporal selection condition, where "V" denotes the disjunction "or" of r_i's, while each r_i represents one of the binary temporal predicates governing time elements (see section 2).

3. If C_1 and C_2 are selection conditions (data or temporal) then the conjunction

   $$C_1 \wedge C_2$$

   is a selection condition. E.g.:

   $(before(T_{transaction}, t_1) \vee after(T_{transaction}, t_2)) \wedge during(t_3, T_{valid})$,

   $equal(T_{transaction}, t_2) \wedge (starts(T_{valid}, t_1) \vee during(T_{valid}, t_1) \vee finishes(T_{valid}, t_1))$

   etc.

The key problem is to test whether for a given pair of time elements, t_1 and t_2, a constraint $r(t_1,t_2)$ is satisfied. In principle we can do this by showing that, $\neg r(t_1,t_2)$, the converse constraint to $r(t_1,t_2)$, is inconsistent with the time network, by means of the consistency checker (see section 5.4.4 and Appendix A). It has been shown that the temporal predicates given in section 4.3 are mutually exclusive. So that, for example, if p is a point and i an interval, we know that precisely one of the temporal predicates in the set:

$$R = \{before,\ meets,\ starts,\ during,\ finishes,\ met\_by,\ after\}$$

must apply for p and i. Hence for $r_0 \in R$:

$$\neg r_0(p,i) \Leftrightarrow r_1(p,i) \vee r_2(p,i) \vee \dots \vee r_6(p,i)$$

where $\{r_1, r_2, \dots, r_6\} \cup \{r_0\} = R$. Hence, to show $r_0(p,i)$ we simply show that $r'(p,i)$ is inconsistent for $r' = r_1, r_2, \dots, r_6$, i.e., r' can be any one of the temporal relationships between point p and interval i, other than r_0. For instance, we may show *before*(p,i) by means of showing that *meets*(p,i), *starts*(p,i), *during*(p,i), *finishes*(p, i), *met-by*(p,i) and *after*(p,i) are all inconsistent with the database.

The general treatment of the temporal constraints can then be handled by conventional logical operations over the results of individual constraint evaluations.

It is interesting to note that improvement may be made to the evaluation of the individual converse constraint $\neg r(k_1,k_2)$. The method given above equates $\neg r(k_1,k_2)$ with the disjunction of $r'(k_1,k_2)$ for $r' \in R \setminus \{r\}$. This leads to several different consistency checks. However, for each temporal predicate the number of tests may be reduced considerably (to a maximum of 2) by expressing $\neg r$ directly in terms of *meets*. For instance:

$$\neg before(k_1,k_2) \Leftrightarrow \quad meets(k_0,k_2) \wedge meets(k_0,k) \wedge meets(k_1,k_n) \wedge meets(k,k_n)$$
$$\vee\ meets(k_1,k_2)$$

Hence, for example, the six tests for the constraint $\neg r_0(p, i)$ illuminated above may be reduced to two.

Other operations over temporal relations, such as (CARTESIAN) **PRODUCT**, **UNION**, **PROJECT** and **JOIN**, etc., may be also formally defined. For example, we can define the (CARTESIAN) **PRODUCT** operation as below:

Let

$$R_1(T_{transaction}, A_1, \ldots, A_{n1}, T_{valid})$$

and

$$R_2(T_{transaction}, B_1, \ldots, B_{n2}, T_{valid})$$

be the temporal schemes of two temporal relations, R_1 and R_2.

For convenience of expression, we define the "common part" of two time elements t_1 and t_2, denoted by, $common(t_1, t_2)$, as below:

$$equal(t_1, t_2) \vee starts(t_1, t_2) \vee during(t_1, t_2) \vee finishes(t_1, t_2)$$

$$\Rightarrow common(t_1, t_2) = t_1,$$

$$starts(t_2, t_1) \vee during(t_2, t_1) \vee finishes(t_2, t_1),$$

$$\Rightarrow common(t_1, t_2) = t_2,$$

$$overlaps(t_1, t_2) \Rightarrow common(t_1, t_2) = t, \text{ where } t_1 = t' \oplus t \wedge t_2 = t \oplus t'',$$

$$overlaps(t_2, t_1) \Rightarrow common(t_1, t_1) = t, \text{ where } t_2 = t' \oplus t \wedge t_1 = t \oplus t'',$$

Otherwise, $common(t_1, t_2)$ = NULL.

N.B. From the above definition, we know that $common(t_1, t_2) = common(t_2, t_1)$.

The (Cartesian) product of R_1 and R_2 can now be defined as a temporal relation, $R_1 \times R_2$, with the temporal scheme:

$$R_1 \times R_2(T_{transaction}, A_1, ..., A_{n1}, B_1, ..., B_{n2}, T_{valid})$$

where the tuples of $R_1 \times R_2$, together with their transaction time and valid time, are made up in the following way:

For each temporal tuple $(t_{a(i),transaction}, a_{i,1}, \ldots, a_{i,n1}, t_{a(i),valid})$ in R_1,

and each temporal tuple $(t_{b(j),transaction}, b_{j,1}, \ldots, b_{j,n1}, t_{b(j),valid})$ in R_2, if

$$common(t_{a(i),transaction}, t_{b(j),transaction}) \neq \text{NULL}$$

and

$$common(t_{a(i),valid}, t_{b(j),valid}) \neq \text{NULL},$$

then there will be a temporal tuple in the cartesian product relation $R_1 \times R_2$:

$$(t_{transaction(i,j)}, a_{i,1}, \ldots, a_{i,n1}, b_{j,1}, \ldots, b_{j,n2}, t_{valid(i,j)}),$$

where $t_{transaction(i,j)} = common(t_{a(i),transaction}, t_{b(j),transaction}),$

$$t_{valid(i,j)} = common(t_{a(i),valid}, t_{b(j),valid}).$$

N.B. In a similar fashion to the **CARTESIAN PRODUCT** operation, **JOIN** operations can be defined, by adding some corresponding *join conditions*, where a join condition can be any conventional data condition, or a temporal constraint.

7.4 Examples of Retrieval

In this section we give some examples of temporal predicates used as temporal constraints.

Query 1: Retrieve Lee's medical history, as known before the second appointment with Dr. Major.

SELECT $T_{TRANSACTION}$, PRESCRIBER, DRUG, STATUS, T_{VALID}

FROM MEDICAL-RECORD

WHERE PATIENT = "Lee" **AND**

(*before*($T_{TRANSACTION}$, i_{t3})

$\vee$ *meets*($T_{TRANSACTION}$, i_{t3})).

Result:

$T_{TRANSACTION}$	PRESCRIBER	DRUG	STATUS	T_{VALID}
i_{t1}	NULL	NULL	pain	i_{v1}
i_{t1}	Lee	a	worse	i_{v2}
i_{t1}	Major	b	NULL	i_{v3}
i_{t2}	NULL	NULL	pain	i_{v1}
i_{t2}	Lee	a	worse	i_{v2}
i_{t2}	Major	b	better	i_{v3}
i_{t2}	Long	b	NULL	i_{v4}

Query 2: Is it true that Lee took drug a after i_{t1} ?

SELECT *

FROM MEDICAL-RECORD

WHERE PATIENT = "Lee" **AND**

DRUG = a **AND**

$T_{TRANSACTION} = i_{t3}$ **AND**

after(T_{VALID}, i_{t1}).

Result : Null.

Query 3: Was it known over i_{t2} that Lee got better using drug b alone?

SELECT STATUS, T_{VALID}

FROM MEDICAL-RECORD

WHERE PATIENT = "Lee" **AND**

DRUG = b **AND**

$T_{TRANSACTION} = i_{t2}$.

Result:

STATUS	T_{VALID}
better	i_{v3}
null	i_{v4}

This query shows exactly the state of Dr. Long's knowledge according to the transaction time i_{t2}. It shows his belief concerning the past interval, i_{v3}, and his prescription for the coming interval, i_{v4}.

Of course, the tables given in the examples cannot present all of the information required by a user of the system, since the temporal relationship of the terms, i_{v1}, i_{v2}, ..., i_{t1}, i_{t2}, ..., t_1, t_2, ..., needs the graph of Figure 7.2(1) as part of the user interface. As a matter of presentation, it may be convenient to display more than just the given time elements in order to allow the user to relate them to other special reference time elements which exist in the network. Particularly, if absolute-time-stamped elements are added to the network, as in the next section, these may be displayed as reference elements.

7.5 Integrating absolute and relative temporal information

Section 7.3 outlines an inference mechanism for retrieval subject to temporal constraints, from the purely relative temporal database. This inference mechanism is based on the consistency checking algorithm given in section 5.4.4 (Appendix A): The necessary and sufficient condition for a purely relative time network without any duration constraints to be consistent are two-fold. The first is the requirement that points do not meet points. The second condition states that the modified graph of the corresponding network must be acyclic.

The inclusion of absolute temporal information into the representation may be handled by the addition of known durations to weight the arcs of the graph. In this case, the alternative consistency checker given in section 5.4.3 (see also Appendix B) may be employed. This general consistency checking algorithm, addressing mixed absolute and relative time, involves a search for cycles, and the construction of a numerical constraint for each cycle. The existence of a solution to this set of constraints implies the consistency of the system. Hence, the consistency checker for a random set of known durations is in fact a linear programming problem.

However, the retrieval situation for a absolute-time-stamped system is computationally efficient, since all start- and end-times are known, and hashing or tree-search schemes may be based on known time points. In order to reduce the computational overhead of consistency checking for the case supporting absolute duration information, in what follows, we introduce a class of databases, termed *time-limited* databases. Databases of this type allow both absolute time value and relative time information in a form suitable for many practical applications. The main idea is that, in many applications, the relative temporal knowledge may be very limited, while most data is stamped with absolute time values. The graphical structure of such a system is that of a single absolute-time-stamped chain, **C**, with occasional attached groups of purely relative time knowledge, $\mathbf{Q}_i$, i = 1, 2, ..., g, which we shall term "relative sub-networks". An assumption is made that each relative sub-network is *time-limited*: i.e. two definite times with absolute values may be determined that are earlier and later respectively than the sub-network. In this case, each relative sub-network may be spanned by an absolute-time-stamped interval which starts and finishes simultaneously with it. We shall show that retrievals may be made first from the absolute-time-stamped chain including spanning intervals, and then, if necessary, by relative inference over the union of two relative sub-networks.

In an absolute-time-stamped temporal system the time elements consist of a sequence of time points with absolute time values (reals or rationals), separated by time intervals. We shall term this set of elements an *absolute-time-stamped chain*, and any ordered union of elements in an absolute-time-stamped chain as an *absolute-time-stamped element* (In particular, any element in an absolute-time-stamped chain is an absolute-time-stamped element).

Definition: Let **C** be a special time network consisting of a sequence of time elements of which all points are stamped with absolute time values. **C** is called an *absolute-time-stamped chain* if:

$$\forall c \in \mathbf{C}(\quad c \in \mathbf{I} \Rightarrow \exists p_1, p_2 \in \mathbf{C} \cap \mathbf{P}(meets(p_1, c) \wedge meets(c, p_2))$$
$$\wedge\ c \in \mathbf{P} \Rightarrow \exists i \in \mathbf{C} \cap \mathbf{I}(meets(i, c) \vee meets(c, i))\ )$$

where **I** is the set of intervals, and **P**, the set of points.

N.B. According to the above definition, the first and last elements of an absolute-time-stamped chain must be both time points.

We now consider a time network graph **G** containing a single absolute-time-stamped chain **C**. If we let **Q** be the subgraph of **G** consisting of non-absolute-time-stamped elements in **G**, we may present **G** as $\mathbf{G} = \mathbf{C} \sqcup \mathbf{Q}$, where $\mathbf{C} \sqcup \mathbf{Q}$ represents the graph union of **C** and **Q**. Additionally, let **Q** be decomposed into subgraphs $\mathbf{Q}_1$, $\mathbf{Q}_2$,..., $\mathbf{Q}_g$, such that:

- $\mathbf{Q}_i$ is a connected sub-graph of **G**,
- $\mathbf{Q}_i$ and $\mathbf{Q}_j$ are only connected through nodes in **C**,
- For each element t in $\mathbf{Q}_i$ there exists a directed path from a node in **C** to t, and a directed path from t to another node in **C**,

where i,j = 1,2, ..., g.

The first two of these properties define what is meant by a relative sub-network, Q_i. Such a sub-network is a connected set of non-absolute-time-stamped arcs, which is isolated from other such sets in the sense that they are only connected in **G** by means of their connection to the absolute-time-stamped chain itself. Practically it is often the case that the temporal relationships between several linked events are known but that these events are unconnected with other sets of events. For example, in the illustration of section 7.4, the events are connected through a single patient. Any connection with other patient events is made through absolute time values, i.e. through the absolute-time-stamped chain)

It should be noted that all graphs $\mathbf{G} = \mathbf{C} \sqcup \mathbf{Q}$ may be decomposed into $\mathbf{Q}_i$ satisfying the first two properties, the only question being the size of the relative subgraphs. In the extreme case, there is the trivial decomposition with just **Q** itself satisfying the properties.

The third property is the reasonable practical assumption that each relative sub-network may be time-limited in some way. That is, for any set of linked temporal events, some absolute time bounds, however wide, may be established. For example, in the medical illustration it may not be known exactly when the events took place, but the month or year, will surely be known.

We shall term a relative sub-graph $\mathbf{Q}_i$ which satisfies all of the above three properties *time limited.* Figure 7.5(1) shows a time network containing an absolute-time-stamped chain **C**:

$$\mathbf{C} = p_1, i_1, p_2, i_2, p_3, i_3, p_4, i_4, p_5$$

with the following two time-limited sets:

$Q_1 = \{i_5, t_1, t_2, i_6\}$,

$Q_2 = \{i_7, t_3, t_4, t_5, t_6, t_7, t_8, t_9, t_{10}, i_8\}$.

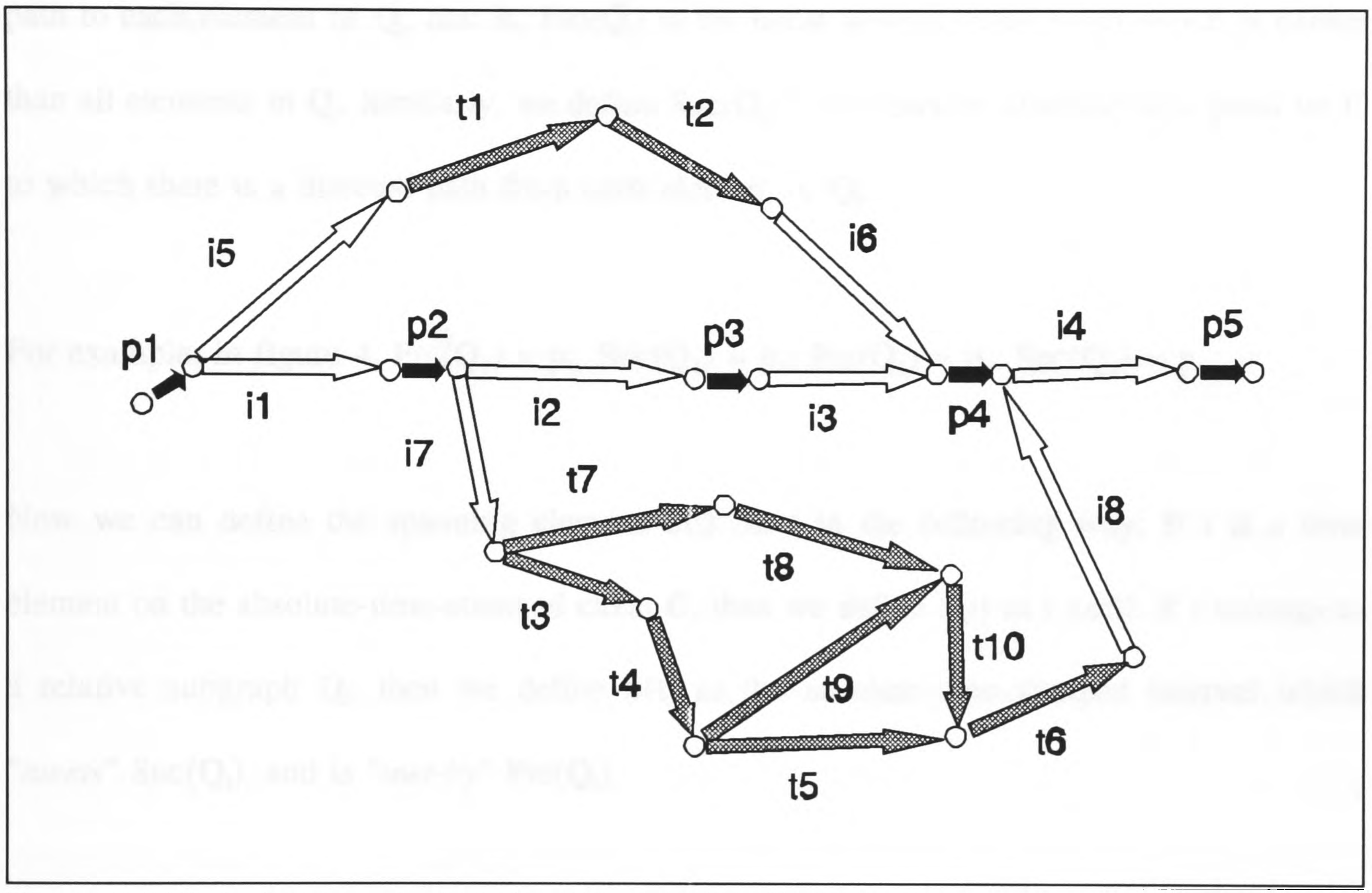

Figure 7.5(1) A time-stamped chain with two time-limited sets

N.B. In the above graphical representation, while points are still denoted as single barred arcs, and intervals as double barred arcs, time elements which are not known to be points or intervals are represented as shadowed double barred arcs.

We term the decomposition:

$$\mathbf{G} = \mathbf{C} \sqcup \mathbf{Q}_1 \sqcup \mathbf{Q}_2 \sqcup \ldots \sqcup \mathbf{Q}_g$$

a time-limited decomposition of **G**, where $\mathbf{Q}_1$, $\mathbf{Q}_2$, ..., $\mathbf{Q}_g$ are time-limited relative subgraphs.

We now need to define a spanning element S(t) for each time element t in G. First, for each relative subgraph $\mathbf{Q}_i$, we define Pre($\mathbf{Q}_i$) as the latest point on **C** from which there is a directed path to each element of $\mathbf{Q}_i$, that is, Pre($\mathbf{Q}_i$) is the latest absolute-time point which is earlier than all elements in $\mathbf{Q}_i$. Similarly, we define Suc($\mathbf{Q}_i$) as the earliest absolute-time point on **C** to which there is a directed path from each element of $\mathbf{Q}_i$.

For example, in figure 4, Pre($\mathbf{Q}_1$) = p_1, Suc($\mathbf{Q}_1$) = p_4; Pre($\mathbf{Q}_2$) = p_2, Suc($\mathbf{Q}_2$) = p_5.

Now we can define the spanning element S(t) for t in the following way: If t is a time element on the absolute-time-stamped chain **C**, then we define S(t) as t itself. If t belongs to a relative subgraph $\mathbf{Q}_i$, then we define S(t) as the absolute-time-stamped interval which "*meets*" Suc($\mathbf{Q}_i$), and is "*met-by*" Pre($\mathbf{Q}_i$).

We are interested in the evaluation of constraints such as: *before*(t_1,t_2), *during*(t_1,t_2), ..., etc. from a network *A*. For clarity, we introduce the network name as a third argument to these temporal predicates, so that *before*(t_1,t_2, *A*) means that t_1 is *before* t_2 in network *A*.

It is now straightforward to demonstrate the following three theorems, which may form the basis of retrieval mechanisms for a time-limited database. Theorems (T1) and (T2) show that we can evaluate *r*(t_1,t_2, **G**) by first testing *before*(S(t_1),S(t_2), **C**) and *after*(S(t_1),S(t_2), **C**) over the absolute-time-stamped network **C**, using the absolute-time-stamping retrieval algorithms (since $S(t_1)$ and $S(t_2)$ are both stamped with absolute time values). If these relations are not satisfied, then we need the full refutation mechanism to evaluate the other predicates. However, theorem (T3) shows that we only need to do this over the union of **C** with the two time-limited sets $\mathbf{Q}(t_1)$ and $\mathbf{Q}(t_2)$, where $\mathbf{Q}(t)$ denotes the relative sub-network containing t.

(T1). $before(S(t_1),S(t_2), \mathbf{C}) \Rightarrow before(t_1,t_2, \mathbf{G})$,

(T2). $after(S(t_1),S(t_2), \mathbf{C}) \Rightarrow after(t_1,t_2, \mathbf{G})$,

(T3). $\neg(before(S(t_1),S(t_2), \mathbf{C}) \vee after(S(t_1),S(t_2), \mathbf{C})) \wedge r(t_1,t_2, C \sqcup \mathbf{Q}(t_1) \sqcup \mathbf{Q}(t_2)) \Rightarrow r(t_1,t_2, \mathbf{G})$,

where r is any one of the possible temporal relations between time elements t_1 and t_2.

The "relative" retrieval algorithm is thus reduced to a search for cycles over the graph $\mathbf{C} \sqcup \mathbf{Q}(t_1) \sqcup \mathbf{Q}(t_2)$. The complexity of this algorithm is dependent on the size of $\mathbf{Q}(t_1)$, and $\mathbf{Q}(t_2)$ only, since there can be no cycles involving the part of **C** that is "*before*" the earlier one of $\mathrm{Pre}(\mathbf{Q}(t_1))$ and $\mathrm{Pre}(\mathbf{Q}(t_2))$, or "*after*" the later one of $\mathrm{Suc}(\mathbf{Q}(t_1))$ and $\mathrm{Suc}(\mathbf{Q}(t_2))$.

CHAPTER 8

CONCLUSIONS

Reasoning about and with time has been recognized to be of relevance to several distinct areas in computer science since the early 70s. In particular, researchers have found that the understanding and treatment of time plays an increasingly important role in the domain of artificial intelligence. This has come about because computer science as a whole is both highly formal and deeply rooted in the practice of everyday life, so that a formalism designed to handle the pervasive feature of time has an important natural role in many fields. However, the theoretical nature of time is a question with a long philosophical tradition and the literature is full of disputes and contradictory theories [Gal87]. When designing a system for temporal reasoning, we are faced with a choice of the underlying time structure - for different objectives or motivations, different people may have different approaches.

In the past two decades, many temporal systems have been proposed in order to address the problem of modelling human temporal concepts in a natural way. These various models are similar in many respects, but there are subtle differences in terminology and basic theory which derive from the differences in approach.

In this study, the examination of some existing representative temporal systems and the development of a new general temporal theory have been undertaken. This general theory may be seen as an extension of Allen and Hayes' and Vilain's theories.

One major advantage of the new theory is its approach which addresses both intervals and points as primitive time elements on an equal footing, and allows time structures such as *linear/non-linear*, and *dense/discrete*, etc. These issues may be separately specified by means of additional axioms. On the other hand, the axiomatisation provides a satisfactory characterisation for the *open* and *closed* nature of primitive intervals, which has been a problematic question of time representation in many incomplete knowledge systems.

It is shown that this general theory is powerful enough to subsume the examined representative theories, such as Allen and Hayes's interval-based theory, Bruce's and McDermott's point-based theories, and the interval and point-based theory of Vilain. It retains the appealing characteristics of these existing temporal theories, but without bearing their corresponding deficiencies discussed in the literature survey.

As applied to computer based systems, a finite time network based on the theory is specially addressed, where a consistency checker in two different forms is provided for cases with, and without, duration reasoning, respectively. Based on the new time axiomatisation, a temporal logic for temporal reasoning about propositions whose truth values are associated with particular intervals/points is presented with an explicit definition of its syntax and semantics. It is shown that the logic is more expressive than that of existing systems, such as Allen's interval-based logic, Galton's revised temporal theory, Shoham's point-based interval logic, and Haugh's MTA based logic; and that problems with these systems are satisfactorily solved.

As an application of the temporal theory, a new architecture for a temporal database system is proposed. The system allows the expression of relative temporal knowledge of data transaction and data validity times. A general retrieval mechanism is presented for a database with a purely qualitative temporal component which allows queries with temporal constraints in terms of any logical combination of Allen's temporal relations. To reduce the computational complexity of the consistency checking algorithm when quantitative time duration knowledge is added, a class of databases, termed time-limited databases, is introduced. This class allows absolute-time-stamped and relative time information in a form which is suitable for many practical applications, where qualitative temporal information is only occasionally needed, and the efficient retrieval mechanisms for absolute-time-stamped databases may be adapted. The implementation of such a time-limited database remains as future work in the application area of the study.

REFERENCES

[Ahn86] Ahn Ilsoo: "Towards An Implementation Of Database Management Systems With Temporal Support", *CH2261-6/86/0000/0374$01.00 IEEE*, pp.374-381, 1986.

[AlH85] Allen J. F. and Hayes P. J.: "A Common-Sense Theory of Time", *Proc. 9th IJCAI*, Los Angeles, california, pp.528-531, 1985.

[AlH89] Allen J. F. and Hayes P. J.: "Moments and Points in an Interval-based Temporal-based Logic", *Comput. Intell. (Canada)*, Vol.5, no.4, pp.225-238, Nov. 1989.

[All81] Allen J. F.: "An interval-based representation of temporal knowledge", *Proc. 7th IJCAI*, pp.221-226, 1981.

[All83] Allen J. F.: "Maintaining Knowledge about Temporal Intervals", *communication of ACM.*, Vol.26, pp.832-843, 1983.

[All84] Allen J. F.: "Towards a General Theory of Action and Time", *Artificial Intelligence*, 23(2), pp.123-154. 1984.

[Ari86] Ariav G.: "A Temporally Oriented Data Model", *ACM Transactions on Database Systems*, Vol.11, No.4, pp.499-527, 1986.

[Ari91] Ariav G.: "Temporally Oriented Data Definitions: Managing Schema Evolution In Temporally Oriented Databases", *Data & Knowledge Engineering*, 6, pp.451-467, 1991.

[BBG91] Bernard, D., Borillo, M. and Gaume, B.: "From event calculus to the scheduling problem. Semantics of action and temporal reasoning in aircraft maintenance", *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 1(1991), pp.195-221.

[Bee89] Beek P. V.: "Approximation Algorithms For Temporal Reasoning", *Proc. 11th IJCAI*, pp.1291-1296, 1989.

[Bee90] Beek P. V.: "Reasoning About Qualitative Temporal Information", *Proc. 8th AAAI*, pp.728-734, 1990.

[Bee92] Beek P. V.: "Reasoning About Qualitative Temporal Information", *Artificial Intelligence*, 58, pp.297-326, 1992.

[Ben827] Ben-Zvi J.: "The time relational model", Ph.D. disseration, Computer Science Department, Univ. California, Los Angeles (1982).

[BoG90] Borillo, M.; Gaume, B.: "Spatiotemporal reasoning based on an extension of event calculus", in *Proceedings of the Third COGNITIVE Symposium*, Madrid, Spain (1990), Kohonen, T. and Fogelman-Soulie F. (eds.), pp.337-344.

[Boy59] Boyer C. B.: *The History of the Calculus and Its Conceptual Development*, Dover, New York. (Reprint of 1949).

[Bru72] Bruce B. C.: "A Model for Temporal References and Application in a Question Answering Program", *Artificial Intelligence*, 3, pp.1-25, 1972.

[BTK91] Bacchus F., Tenenberg J. and Koomen J. A.: "A non-reified temporal logic", *Artificial Intelligence*, 52(1991), pp.87-108.

[Car79] Carre B.: *Graphs and Networks*, rendon Press, Oxford, 1979.

[Cha88] Chaudhuri S.: "Temporal Relationships in Databases", *Proc. 14th VLDB Conference*, Los Angeles, California, pp.160-170, 1988.

[Cli85] Clifford J.: "Toward an algebra of historical relational databases", *Proc. ACM-SIGMOD Int. Conf. on Management of Data*, Austin, USA, pp.247-265, 1985.

[Cli87] Clifford J.: "Historical Databases--It's About Time", Graduate School of Business Administration, New York Univ., NY (1987).

[ClI93] J. Clifford and T. Isakowitz, On the semantics of Transaction Time and Valid Time in Bitemporal Databases, *Proceedings of the international Workshop on an Infrastructure for temporal database*, Ed. R. T. Snodgrass, Arlington, TX. (1993).

[CLR89] Cormen T. H, Leiserson C. E. and Rivest R. L.: *Introduction to Algorithms*, The MIT Electrical Engineering and Computer Science Series, The MIT Press, Cambridge, Massachusetts, London, England, 1989.

[ClW83] Clifford J. and Warren D.: "Formal Semantics For Time In Databases", *ACM Transactions on Database Systems*, Vol.8, No.2, pp.214-254, 1983.

[DeM87] Dean T. L. and Mcdermott D. V.: "Temporal Data Base Management", *Artificial Intelligence*, 32, pp.1-55, 1987.

[Dil90] Diller A.: *Z--AN Introduction to Formal Methods*, John Wiley & Sons Ltd., Baffins Lane, 1990.

[DMP91] Dechter R., Meiri I. and Pearl J.: "Temporal Constraint Networks", *Artificial Intelligence*, 49, pp.61-95, 199[illegible].

[Fun83] FunK K. H.: "Theories, Models, and Human-Machine Systems", *Mathematical Modelling*, Vol.4, pp.567-587, 1983.

[GaM91] Gabbay D. and McBrien P.: "Temporal Logic & Historical Databases", *Proc. 17th VLDB Conference*, Barcelona, pp.423-430, 1991.

[Gad86] Gadia S. K.: "Toward a Multihomogeneous Model for a Temporal Database", *IEEE CH2261-6/86/0390$01.00*, pp.390-397, 1986.

[Gad88] Gadia S. K.: "A Homogeneous Relational Model and Query Languages for Temporal Databases", *ACM Transactions on Database Systems*, Vol.13, No.4, pp.418-448, 1988.

[Gal84] Galton A.: *The Logic of Aspect*, Clarendon Press, Oxford, 1984.

[Gal87] Galton A.: "Temporal logic and computer science: an overview", *Temporal Logic and their Applications*, ed. A. Galton, Academic Press Limited, pp.1-52, 1987.

[Gal90a] A. Galton, *Logic for Information Technology*. JOHN WILEY & SONS Ltd., Baffins Lane, 1990.

[Gal90b] Galton A.: "A Critical Examination of Allen's Theory of Action and Time", *Artificial Intelligence*, 42, pp.159-188, 1990.

[HaS91] Halpern J. Y. and Shoham Y.: "A Propositional Model Logic of Time Intervals", *Journal of the Association for Computing Machinery*, Vol.38. No.4. October, 1991, pp.935-962.

[Hau87] Haugh B. A.: "Non-Standard Semantics for the Method of Temporal Arguments", *Pro. 10th IJCAI*, Vol.1, Milan, August 23-28, 1987, pp.449-455.

[Hay78] Hayes P.: "The Naive Physics manifesto", *Expert systems in the microelectronic age*, Michie D. ed., Edinburgh, 1978.

[JCG92] C. S. Jensen, J. Clifford, S. K. Gadia, A. Segev and R. T. Snodgrass, A Glossary of Temporal Database Concepts, *SIGMOD RECORD*, Vol.21, No.3, 1992, pp.35-43.

[JoM80] Jones S. and Mason P. J.: "Handling the time dimension in a data base", *Proc. Int. Conf. on Data Bases*, Aberdeen, British Computer Society, 1980, pp.65-83.

[KaG77] Kahn K. M. and Gorry A. G.: "Mechanizing Temporal Knowledge", *Artificial Intelligence*, 9, 1977, pp.87-108.

[Kni92] Knight B.: "A Deductive Approach To Temporal Databases", *The Computer Journal*, Vol.35, No.4, Aug. 1992, pp.A395-A402.

[KnM92] Knight B. and Ma J.: "A General Temporal Model Supporting Duration Reasoning", *AI Communication Journal*, Vol.5, No.2, June, 1992, pp.75-84.

[KnM93] Knight B. and Ma J.: "An Extended Temporal System Based on Points and Intervals", *Information System*, Vol.18, No.2, 1993, pp.111-120.

[KoS86] Kowalski R. and Sergot M.: "A Logical-based Calculus of Events", *New Generation Computing*, 1986, pp.67-95.

[Kou93] Koubarakis M.: "Representation and Querying in Temporal Databases: the Power of Temporal Constraints", *Proceedings of the IEEE Conference on Data Engineering*, pp.327-334 (1993).

[Kow92] Kowalski R.: "Database Updates In The Event Calculus", *The Journal of Logic Programming*, (1992), pp.121-146.

[Lad86] Ladkin P.: "Time Representation: A Taxonomy Of Interval Relations", *Proc. 5th AAAI*, 1986, pp.360-366.

[Lad87] Ladkin P.: "Models of Axioms for Time Intervals", *Proc. 6th AAAI,* Vol.1, 1987, pp.234-239.

[Lad92] Ladkin P.: "Effective Solution Of Qualitative Interval Constraint Problems", Artificial Intelligence 57 (1992), pp.105-124.

[LaM88] Ladkin P and Maddux R.: " The Algebra of Constraint Satisfaction Problems and Temporal Reasoning", Technical Report (1988), Kestrel Institute, Palo Alto, Calif.

[LiB92] Ling D. H. O. and Bell D. A.: "Modelling and Managing Time in Database Systems", The Computer Journal, Vol.35, No.4, 1992, pp.332-341.

[Lip64] Lipschutz S.: Theory and Problems of Set Theory and Related Topics, Schaum Publishing Co., 257 Park Avenue South, New York, 1964.

[Lon89] Long D.: "A review of temporal logics", The Knowledge Engineering Review, vol.4:2, 1989, pp.141-162.

[Mai92] Maiocchi R.: "Automatic Deduction of Temporal Information", *ACM Transactions on Database Systems*, Vol.17, No.4, pp.647-688, 1992.

[MaK94] Ma J. and Knight B.: "A General Temporal Theory", *The Computer Journal,* 37(2), 1994, pp.114-123.

[Mcc63] McCarthy J.: "Situation, Actions, and Causal Laws, Memo 2, Stanford Artificial Intelligence Project, 1963.

[Mcd82] McDermott D. V.: "A Temporal Logic for Reasoning about Processes and Plans", Cog. Sci., 6(1982), pp.101-155.

[McH69] McCarthy J. and Hayes P. J.: "Some Philosophical Problems from the Standpoint of Artificial Intelligence", *Machine Intelligence*, B. Meltzer and D. Michie, (eds.), 4(1969), Edinburgh U.p., pp.463-502.

[Mck88] Mckenzie L. E.: "An algebraic language for query and update of temporal databases", Ph.D. dissertation, Computer Science Department, Univ. of North Carolina at Chapel Hill (1988).

[McS91] Mckenzie L. E. and Snodgrass R. T.: "Supporting Valid Time in an Historical Relational Algebra", Tech. Rep. TR91-15. Department of Computer Science, Univ. of Arizona (1991).

[Rei87] Reichgelt H.: "Semantics For Reified Temporal Logic", Advanced in Artificial Intelligence, Proceedings of the 1987 AISB Conference, University of Edinburgh, 6-10 April 1987.pp.49-61.

[Sad87] Sadri F.: "Three Recent Approaches to Temporal Reasoning", *Temporal Logic and their Applications*, ed. Galton A., Academic Press, 1987, pp.121-168.

[Sho87a] Shoham Y.: "Reified Temporal Logics: Semantical and Ontological Considerations", *Advances in Artificial Intelligence - II*, B. Du Boulay, D. Hogg and L. Steels (Editors), Elsevier Science Publishers B. V. (North-Holland), (1987) pp.183-190.

[Sho87b] Shoham Y.: "Temporal Logics in AI: Semantical and Ontological Considerations", *Artificial Intelligence*, 33 (1987) pp.89-104.

[SnA86] Snodgrass R. T. and Ahn I.: "Temporal Databases", *Computer (IEEE)*, 1986, pp.35-42.

[Sno87] Snodgrass R. T.: "The Temporal Query Language Tquel", ACM Transactions on Database Systems, Vol.12, No.2, 1987, pp.247-298.

[Sri88] Sripada S. M.: "A Logical Framework for Temporal Deductive Database", In *Proceedings of the VLDB* (Los Angles), 1988, pp.171-182.

[ThL91] Theodoulidis C. and Loucopoulos P.: "The time dimension in conceptual modelling", Information Systems, Vol.16, No.3, 1991, pp.273-300.

[Tsa87] Tsang E. P. K.: "Time Structure for AI", *The 10th IJCAI*, Milan, August 23-28, 1987, pp.456-461.

[ViK86] Vilain M. B. and Kautz H.: "Constraint Propagation Algorithms for Temporal Reasoning", *Proc. 5th AAAI,* 1986, pp.377-382.

[Vil82] Vilain M. B.: "A System for Reasoning about Time", *Proc. 1st AAAI,* 1982, Pittsburgh, PA. pp.197-201.

APPENDIX

PUBLISHED RESULTS

(with Brian Knight)

Appendix A

An Extended Temporal System Based on Points and Intervals

(*Information System*, Vol. 18, No. 2, pp.111-120, 1993)

Abstract

This paper proposes axioms for temporal systems based on a discrete set of intervals and points which are treated equally as primitive elements. Temporal ordering is specified by means of the primitive relation 'meets'. The axioms, and a graphical representation of temporal knowledge, are specified formally by using the **Z** language. A consistency condition for a temporal database is given in terms of the cyclic properties of the graphical representation, and an algorithm for consistency checking is provided. Formal proofs of Allen's transitivity table for interval relations are given.

The paper addresses some hitherto unresolved issues in the use of interval based systems for temporal databases and proposes a solution. These issues are the problems involved in modelling 'open' and 'closed' intervals.

Key words: **Z**, intervals, points, temporal database, incomplete knowledge, consistency checking.

1.Introduction

Our intention in this paper is to propose a model for a discrete system of primitive elements which may be points or intervals, by providing axioms for a single relation, 'meets', over the elements. This gives a basis for discrete temporal reasoning. The axioms allow modelling of both open and closed intervals, and also allow intervals to be unspecified as to inclusion or exclusion of end points. A necessary and sufficient condition for the consistency of an incomplete temporal system is formulated in terms of graph cycles. This model may be viewed as an extension of Allen's interval based logic (see [All83,84] and [AlH89]). The extension allows the inclusion of points as primitives, and is provided with a consistency checking algorithm which may form the basis of an inference system.

Allen introduced his temporal logic in order to provide a framework for the treatment of two major subareas of artificial intelligence: natural language processing and problem solving. Instead of adopting time points (or states which are associated with time points), he took intervals as the primitive temporal quantity and introduced thirteen (mutually exclusive) relations between any two intervals. The axioms of Allen's system are formulated in terms of a 144 element transitivity table which is derived from a study of intervals on the real line. However, no formalism is adopted for the concept of open and closed intervals - a weakness which causes some problems for modelling. Inference in this system is performed by calculation of the transitive closure under this table.

Allen specifically excludes time points in claiming that any quantity of time must be *subdivisible.* This ruling eliminates the possibility of instantaneous events from Allen's treatment. The contention is that instantaneous moments are not necessary for the qualitative modelling of temporal data, and that instantaneous events may be considered as small intervals at some appropriate grain-size for the interval. Allen gives the example of a light which is switched on. To model such a system, we may need two intervals: one where the light is off 'meeting' one where it is switched on. This may be enough for our modelling purposes and, if so, we do not need to state whether the intervals are open or closed. By this means, we avoid awkward questions about the end-points: if the intervals are closed then there is a point at which the light is on and off, similarly if they are both open, then at this point the light is neither on nor off. Allen's argument is that if we really want to model the switching process itself, then we need to examine in more detail the physics, and smaller intervals at this event will again be sufficient.

However, there are some difficulties with this approach in the qualitative modelling of everyday occurrences. In qualitative modelling of physical processes, we often wish to impose 'landmark' points which by definition separate two intervals. For instance, consider the action of throwing a ball up into the air. The motion of the ball can be modelled by a quantity space of three elements: going-up, stationary, and going-down. Intuitively, there are intervals for going up, and going down. However, there is no interval, however small over which the ball is neither going up nor going down. Hence the third element in the quantity space is best modelled as a point - a 'landmark' point which separates two other intervals. We may deduce from this example that qualitative modelling appears naturally to require a discrete quantity space of both points and intervals.

Vilain ([Vil82]) has proposed an extension of Allen's system which includes points. In this, Allen's 13 interval relations are extended to 26 relations between points and intervals. However, this leads to an increase in computation for closure, and a corresponding overhead for consistency maintenance. Vilain and Kautz ([ViK86]) have demonstrated that such a calculation is NP-hard, as opposed to an alternative point-based system which is $O(n^3)$. Unfortunately, the problems involved in modelling 'open' and 'closed' intervals still exist in Vilain's system.

In Allen's system, consistency checking is performed by formation of the transitive closure according to a transitivity table which describes the composition of the thirteen (mutually exclusive) relations. If no conflict is found according to the exclusivity, then the system is consistent. However, he has not provided a formal consistency checker. We give here a formal proof that the transitivity table follows from our axioms, by using the necessary and sufficient condition of consistency in terms of the acyclicity of the "meets" relation.

Intuitively, the starting point for the axiomatisation proposed in this paper is a view of time as a totally ordered discrete set, E, of fundamental elements, which may either be points or intervals. We postulate that this fundamental set exists, sufficient for our modelling purposes, although we do not necessarily have full knowledge of E. That is, although we know that there exists a total ordering of E, we do not necessarily know what it is in any situation. Modelling our knowledge of a situation means expressing a (partial) knowledge of the temporal ordering of E. For example, such knowledge may be expressed in terms of Allen's interval relationships, but as we show later, these may all be expressed in terms of a single relation: "meets". The relation "meets" is defined as the successor relation under the total ordering, so that each element "meets" a next element (except possibly for the last element). We may also define compound elements from two elements which meet. When elements e1 and e2 meet in E, we may construct a compound element k = e1 + e2, and this in turn may be used to construct other compound elements.

This view of time is one which to some extent is forced upon us by the practicality of the computer based modelling approach. We have to store elements as a discrete finite set, and the semantics of any database of time elements will naturally assume a total ordering at some fundamental level. However, the database represents our state of knowledge about temporal events, and this may well be incomplete. The knowledge incorporated in the database will in general not be the fundamental elements in E, but rather of the compound elements k. The existence of E is simply a belief which may be used to test the consistency of a database. This is the approach we take to consistency here : if we can show that E exists, then the database is consistent, if we can show it does not exist, then it is not.

So far, no distinction has been made between points and intervals made in the above description. Intuitively, the distinction usually is connected with the decomposability: intervals are assumed to be decomposable if required, and points are not. However, in this treatment, we do not make further commitment as to decomposability of time elements. Here, we are more interested in the order relation between elements. The differentiating property which is proposed here is that although intervals may succeed intervals, points are not successors to points. This characteristic, which is later built into the axioms, is in line both with modelling requirements where points are defined as separators or end-points of intervals, and with the denseness of points on the real line. But this is the only extra requirement which is made of elements if they are to be points. The system overcomes many of the problems involved with open and closed intervals. We show by example in section 4 how open and closed intervals may be modelled as and when required using this system.

We give below a formal description of the axioms for the system outlined here, and present a representation of a database of elements in terms of a graph. We also show how a simple necessary and sufficient condition for consistency in terms of cycles of the graph may be derived from the supposition that an underlying totally ordered fundamental set E must exist for any database. This condition may be used to prove many results. In particular, the system of Allen and that of Vilain may be derived from the axioms. The consistency condition is shown to be of order O(n) for a graph with n arcs.

In section 2, we give a formal definition of the system using **Z**. A **Z** schema consists of variable type definitions and first order logic conditions on variables (see, e.g. [Dil90]). Schemas are named, and the inclusion of a schema name in another schema implies that it is to be inherited. All of the type definitions in the schemas given below are in terms of a fundamental primitive type: **TimeElement**. In section 3.1 we give a schema for incomplete temporal knowledge and define what is meant by consistency for incomplete temporal knowledge. In order to give an algorithm for consistency checking, it is necessary to provide a graphical representation for temporal systems. Schemas for such a representation are given in section 3.2. In section 3.3, we provide a necessary and sufficient condition for consistency and it is illustrated by means of a simple example. Section 4 formally defines Allen's 13 primitive relations in terms of schemas in **Z** and addresses the problem of modelling 'open' and 'closed' intervals. Allen's transitivity closure is formally and intuitively proved in section 5. Finally, section 6 provides a summary and concluding remarks.

2. Definition of Temporal System

The definition of a temporal system consists first of a definition of an underlying well-ordered discrete set E. The elements of E may be points or intervals. The temporal system is then defined as the closure of E under the binary operations of combining adjacent elements.

Let **Point** be the type of all time points which are not decomposable and **Interval**, the type of all time intervals which are decomposable if required (that is, Allen's primitive intervals). We use **TimeElement** to denote the type of all time elements each of which is either a point or an interval.

We define the elementary temporal set E as:

Elementary_E

E : sq**TimeElement**

$\neg\exists n \in \mathbf{N} \bullet (En, En+1 \in \text{Point})$

Nb. The above definition implies that E is a set *similar to* an *initial segment* ([Lip64]) of the natural numbers, and no two points are next to each other. We use a ordered pair (t_1, t_2) to represent that t_2 is the successor of t_1 under the immediate successor relation over E.

The definition of M_E, the set of successor relations over E, is captured in the following schema:

M_E

Elementary_E;

M_E : **TimeElement** $\leftrightarrow$ **TimeElement**

$(e_1, e_2) \in M_E \Leftrightarrow \exists n \in N \bullet (e_1 = En \wedge e_2 = En+1))$

The following schema lies at the heart of the proposed time representation. In effect the set T is defined as the closure of the elementary set E under the composition operator "+", which allows composition of successive TimeElements. Thus we may regard any element of T as a unique combination of fundamental elements of E :

$$e_1 + e_2 + \ldots + e_r$$

The sets M is similarly defined as extensions of the fundamental sets M_E

(T, M)

Elementary_E; M_E;

+ : **TimeElement** $\times$ **TimeElement** $\rightarrow$ **TimeElement**

$\doteq$: **TimeElement** $\leftrightarrow$ **TimeElement**

T : P**TimeElement**

M : P(**TimeElement** $\times$ **TimeElement**)

r_1, r_2, r : **N**

$s_1 : 1..r_1$

$s_2 : 1..r_2$

$T \supseteq E$

$M \supseteq M_E$

$(t_1,t_2) \in M \Leftrightarrow t_1 + t_2 \in T \bullet (\forall t_a, t_b \in T \bullet ((t_a, t_1) \in M \Leftrightarrow (t_a, t) \in M$

$(t_2, t_b) \in M \Leftrightarrow (t, t_b) \in M))$

$t_{1,s1}, t_{2,s2} \in E$

$t_{1,1} + \ldots + t_{1,r1} \doteq t_{2,1} + \ldots + t_{2,r2} \Leftrightarrow r_1 = r_2 \wedge \exists n \in N \bullet (E_n = t_{1,1} \wedge E_n = t_{2,1})$

$t \in T \Rightarrow \exists e_1, \ldots, e_r \in E \bullet (t \doteq e_1 + \ldots + e_r)$

Nb. 1.The operator "+" is associative but not commutative and (T, M) is the closure of E under +

2.In the case that $t \doteq t_1 + t_2$, we say that t is decomposable into t_1 and t_2; Without confusion, $t_1 \doteq t_2$ is simply written as $t_1 = t_2$.

3. Incomplete Temporal System

The set T includes E and all the intervals and points which can be formed from it by means of +. However, in an application neither the fundamental set E nor the complete set T may be known. A database of "facts" about T will express knowledge that is incomplete in several ways. For example, the system may have incomplete knowledge about the successor relation (in order to describe the context conveniently, we name the successor relation as "meets" and use $m(t_1, t_2)$ to denote "t_1 meets t_2"), and the database will often contain redundancy, as when facts are known about two elements without the knowledge that they are actually identical. To allow for possible duplicate elements, the basic structure of the database is that of a bag, rather than a set.

3.1 Definition of (K, M_K)

Accordingly, we use (K, M_K) to express incomplete (and possibly inconsistent) temporal knowledge, where:

```
┌──── (K, M_K) ────────────────────────────
│ K : bagTimeElement
│ M_K : bag(TimeElement × TimeElement)
├──────────────────────────
│ (k1, k2)∈ M_K ⇒ k1, k2∈ K
└──────────────────────────────────────────
```

Nb. i) K expresses our knowledge of what intervals and points are there;

ii) M_K expresses our knowledge as to how the intervals or points in K meets each other;

There is no restriction in the schema (K, M_K) to ensure that the knowledge is consistent. For example, the sets:

$K = \{k_1, k_2\}, M_K = \{ (k_1, k_2), (k_2, k_1) \}$

conform to the schema although (K, M_K) represents knowledge of two TimeElements which are successors to each other. Such knowledge should be regarded as inconsistent according to the schema which follows. In this, consistent partial temporal knowledge is defined by insisting that the elements of the pair (K, M_K) are elements of (T, M) which is derived from an underlying (E, M_E).

```
┌──── Consistent(K, M_K) ──────────────────
│ (K, M_K); (T, M)
├──────────────────────
│ k∈K ⇒ k∈ T
│ (k1, k2)∈ M_K ⇒ (k1, k2)∈ M
└──────────────────────────────────────────
```

3.2 A Graphical Representation

The schema of section 3.1 specify what is meant by consistent partial temporal knowledge, but do not provide an algorithm for checking the consistency of any (K, M_K). In order to prove consistency, we need to find an elementary pair (E, M_E), and from it a closure *(T, M)* so that the *Consistent(K, M_K)* schema is satisfied. In 3.3, we will give such an algorithm in terms of a graphical representation of (K, M_K). However, we must first define the graphical representation itself.

The graph is one in which TimeElements are represented by directed arcs. The meets relation M_K is represented by the nodes of the graph: if $(k_1, k_2) \in M_K$ then k_1 is the in-arc to a node, and k_2 is the out-arc from the node. All TimeElements which are known to meet k_2 will be in-arcs to the node, and all TimeElements which k_1 meets will be out-arcs from the node. Although this representation is intuitively straightforward, the definition of nodes in **Z** is more involved.

The following schemas define nodes of the graphical representation in terms of two equivalence classes of TimeElements: Eq_in and Eq_out. Effectively, Eq_in is a class of TimeElements known to meet a common element, and Eq_out is a class of TimeElements known to be met by a common element. Nodes are then defined as pairs of equivalence classes.

Some difficulty is encountered for nodes not meeting any other node, but this is resolved by extending the equivalence relation to include these (by means of the final clause in *Eq_in* and *Eq_out*). In order to deal with this problem, we give the definitions of two equivalence relations over K, Eq_in_K and Eq_out_K, in terms of the following schemas:

Eq_in_K

(K, M_K);
Eq_in : **TimeElement** $\leftrightarrow$ **TimeElement**

$\forall t_1, t_2 \in K \bullet (t_1 \text{ Eq\_in } t_2 \Leftrightarrow (\exists t \in \textbf{TimeElement} \bullet ((t_1, t), (t_2, t) \in M_K)) \vee t_1 = t_2)$

and

Eq_out_K

(K, M_K);
Eq_out: **TimeElement** $\leftrightarrow$ **TimeElement**

$\forall t_1, t_2 \in K \bullet (t_1 \text{ Eq\_out } t_2 \Leftrightarrow (\exists t \in \textbf{TimeElement} \bullet ((t, t_1), (t, t_2) \in M_K)) \vee t_1 = t_2)$

According to these two equivalence relations, we get two sets of equivalent classes of K, K_Eq_in and K_Eq_out, respectively:

K_Eq_in

$$
\begin{array}{|l}
\textit{Eq\_in\_K};\\
r_1 : \mathbf{N}\\
i : 1..r_1\\
K_{Eq\_in,i} : \text{bag}\,\mathbf{TimeElement}\\
\hline
K = K_{Eq\_in,1} \uplus \ldots \uplus K_{Eq\_in,r1}\\
\forall i, j : 1..r_1 \bullet (i \neq j \Rightarrow K_{Eq\_in,i} \cap K_{Eq\_in,j} = \varnothing)\\
\forall k_1, k_2 \in K \bullet (k_1 \text{ Eq\_in } k_2 \Leftrightarrow \exists l : 1..r_1 \bullet (k_1, k_2 \in K_{Eq\_in,\,l})
\end{array}
$$

K_Eq_out

$$
\begin{array}{|l}
\textit{Eq\_out\_K};\\
r_2 : \mathbf{N}\\
j : 1..r_2\\
K_{Eq\_out,j} : \text{bag}\,\mathbf{TimeElement}\\
\hline
K = K_{Eq\_out,1} \uplus \ldots \uplus K_{Eq\_out,r2}\\
\forall i, j : 1..r_2 \bullet (i \neq j \Rightarrow K_{Eq\_out,i} \cap K_{Eq\_out,j} = \varnothing)\\
\forall k_1, k_2 \in K \bullet (k_1 \text{ Eq\_out } k_2 \Leftrightarrow \exists l : 1..r_2 \bullet (k_1, k_2 \in K_{Eq\_out,\,l})
\end{array}
$$

Now, we can characterise the nodes of graph-(K, M_K) in terms of the following schemas:

mid_Node_of_K

$$
\begin{array}{|l}
\textit{K\_Eq\_in};\ \textit{K\_Eq\_out};\\
n(K_{Eq\_in,i}, K_{Eq\_out,j}) : \mathbf{bag}\text{TimeElement} \leftrightarrow \mathbf{bag}\text{TimeElement}\\
\hline
\exists k_i \in K_{Eq\_in,i}\ \exists k_j \in K_{Eq\_out,j} \bullet ((k_i, k_j) \in M_K)
\end{array}
$$

in_Node_of_K

$$
\begin{array}{|l}
\textit{K\_Eq\_in};\ \textit{K\_Eq\_out};\\
n(\varnothing, K_{Eq\_out,j}) : \text{bag}\,\mathbf{TimeElement}\\
\hline
k \in K_{Eq\_out,j} \Rightarrow \neg\exists k' \in \mathbf{TimeElement} \bullet (k', k) \in M_K
\end{array}
$$

and

out_Node_of_K

| *K_Eq_in*; *K_Eq_out*;
| $n(K_{Eq\_in,i}, \varnothing)$: bag**TimeElement**

| $k \in K_{Eq\_in,i} \Rightarrow \neg\exists k' \in$ **TimeElement** $\bullet$ $(k, k') \in M_K$

3.3 Consistency Condition

To draw inferences from (K, M_K), we must rely on the assumed properties of T and M. A consistency checker is needed which will establish whether a pair (K, M_K) is consistent with our basic assumptions about T and M.

In general, a pair (K, M_K) is consistent if we can add to K and make any necessary equality assignments, and add to M_K, so that the resulting pair (T, M) is a closure for some (E, M_E) satisfying the conditions in section 2. A necessary and sufficient condition for consistency may be given in terms of the graphical representation introduced in 3.2.

Let G be the graph of (K, M_K), then (K, M_K) is consistent if and only if:

(I.1) G^r is acyclic, where G^r is the associated <u>reduced graph</u> formed from G by merging two nodes connected by a point in G and removing the corresponding arc.

(I.2) there are no nodes that are both in-node and out-node to two point-arcs in G.

In fact, by a standard result in graph theory (e.g. see Carre[79]), we can show that the nodes in G, the (K, M_K)-graph, can be ordered in such a way that in- and out- nodes of any intervals in E are successors, as follows:

i) Set variable n = 1

ii) Select any node in the reduced graph G^r without in-arc. Such a node exists since G^r is acyclic (See Carre[79], or any standard graph theory book).

iii) Number this node n.

iv) Remove this node and associated arcs from G^r to form graph $G^{r'}$. $G^{r'}$ is also acyclic. Set G^r to $G^{r'}$, increment n by <u>**2**</u> if the deleted node is formed from a pair of notes in G, otherwise, increment n by <u>**1**</u>.

v) Repeat from ii) until G^r is empty.

vi) Form arcs between consecutive integer nodes. In the case that integer n+1 is missed between n and n+2 in the reduced graph G^r, then the consecutive integers n and n+1 are associated with the corresponding pair of notes in G, the (K, M_K)-graph.

Then the arcs between consecutive integer nodes form the set E, and M_K is formed by the natural order over these integers. Finally, the closure (T, M) can be formed under $\oplus$.

Hence the proof of consistency is a test of the graph for the existence of a cycle. Using a breadth first search scheme, the complexity of this test is O(N), where N is the number of arcs in G.

As an example of the consistency condition, we take a case where a database is consistent if an element t_a is not known to be a time point, but inconsistent if it is.

E.G :

meets(t_0, t_a),
meets(t_a, t_n),
meets(t_0, t_b),
meets(t_b, t_c),
meets(t_c, t_n).

If t_a is not known to be a point then the (K, M_K)-graph shown in Fig3.3(a) is non-cyclic, and the system is consistent.

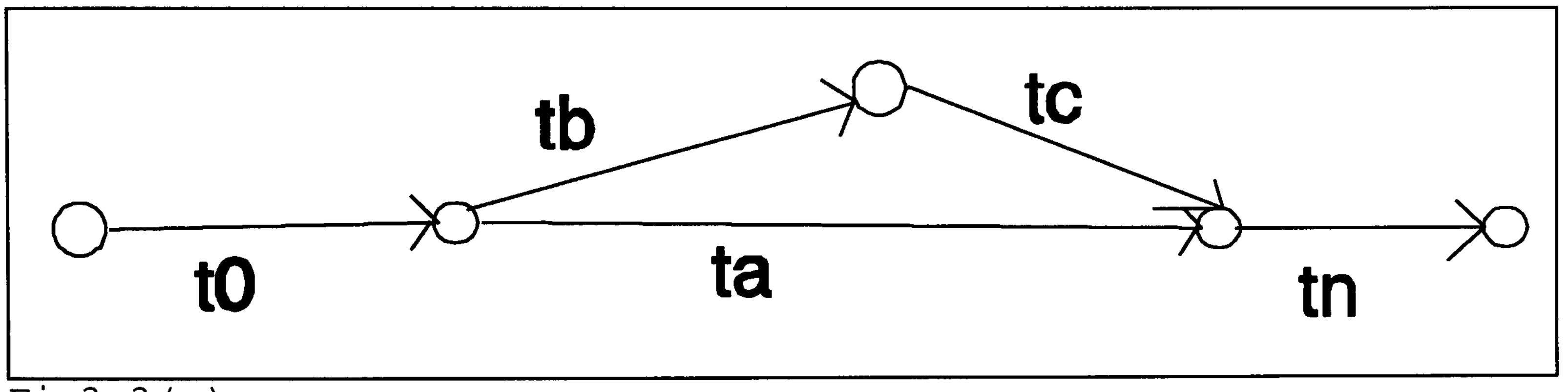

Fig3.3(a)

However, if t_a is a point, then we have the reduced graph in Fig3.3(b), which shows cyclicity, and we deduce that the system is inconsistent.

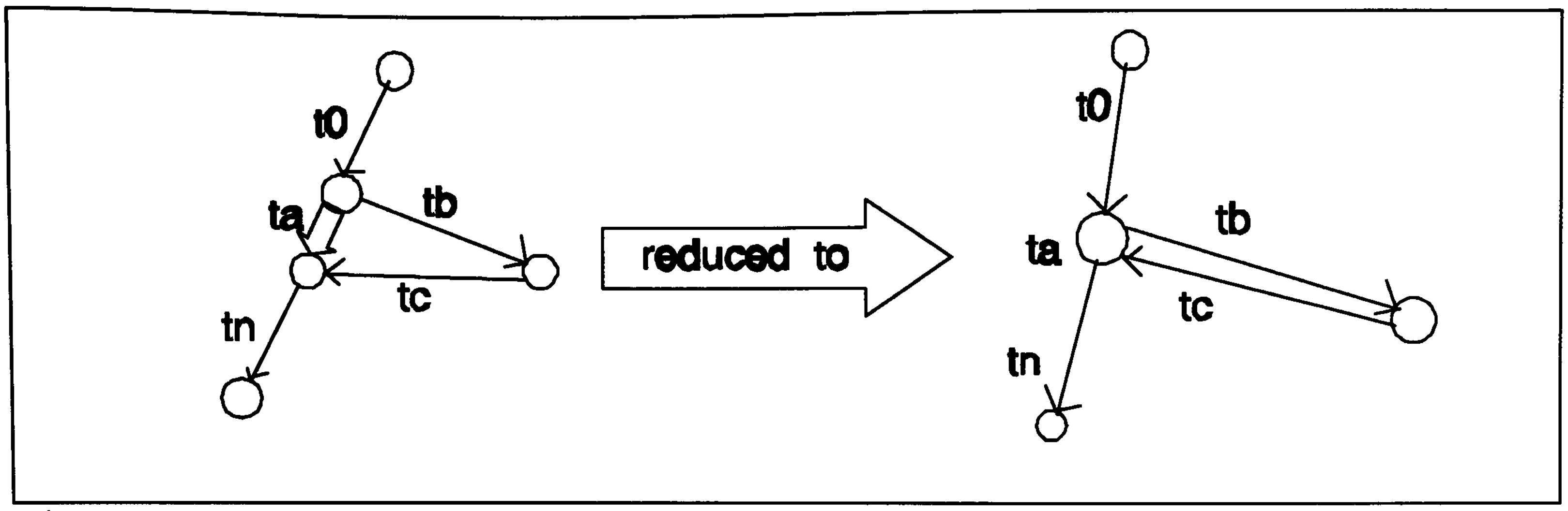

Fig3.3(b)

We can see why this is so intuitively by noticing that in Fig3.3(a):

$$t_a = t_b + t_c \tag{3.3.1}$$

This is consistent until we add the fact that t_a is non-subdivisible. Since equation 3.3.1 states that t_a is subdivisible, we reach an inconsistency when t_a is a point.

4. Allen's Primitive Relations - 'Open' and 'Closed' Intervals

In [All83] and [All84], Allen introduced thirteen relations between his primitive intervals that are formulated later in terms of the single relation "meets" in [AlH89]. Here, we have extended the primitive element to include points. We can also characterise these thirteen primitive relations between our TimeElements in terms of the following schemas:

Equal

Equal(t_a, t_b), Equal(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$t_a \doteq t_b$

Before, After

(T, M);

Before(t_a, t_b), After(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$\exists t \in$ **TimeElement** $\bullet$ ((t_a, t), (t, t_b) $\in$ M)

Meets, Met-by

(T, M);
Meets(t_a, t_b), Met-by(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$(t_a, t_b) \in \mathbf{M}$

Overlaps, Overlapped-by

(T, M);
Overlaps(t_a, t_b), Overlapped-by(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$\exists t_1, t_2, t_3 \in$ **TimeElement** $\bullet$ $(t_a \doteq t_1 + t_2 \wedge t_b \doteq t_2 + t_3)$

Starts, Started-by

(T, M);
Starts(t_a, t_b), Started-by(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$\exists t \in$ **TimeElement** $\bullet$ $t_b \doteq t_a + t$

During, Contains

(T, M);
During(t_a, t_b), Contains(t_b, t_a) : **TimeElement** $\leftrightarrow$ **TimeElement**

$\exists t_1, t_2 \in$ **TimeElement** $\bullet$ $t_b \doteq t_1 + t_a + t_2$

Finishes, Finished-by

(T, M);
Finishes(t_a, t_b), Finished-by(t_a, t_b) : **TimeElement** $\leftrightarrow$ **TimeElement**

$\exists t \in$ **TimeElement** $\bullet$ $t_b \doteq t + t_a$

In [AlH89], Allen and Hayes formulated the thirteen primitive relations between intervals in terms of the single relation "meets". However, here we have extended the primitive element to include points. In this case, some of Allen's relationships are consistent and some are inconsistent when they involve points.

For example, let $p \in$ Point, then:

Before(p, t_b) (see Fig4.1) is consistent.

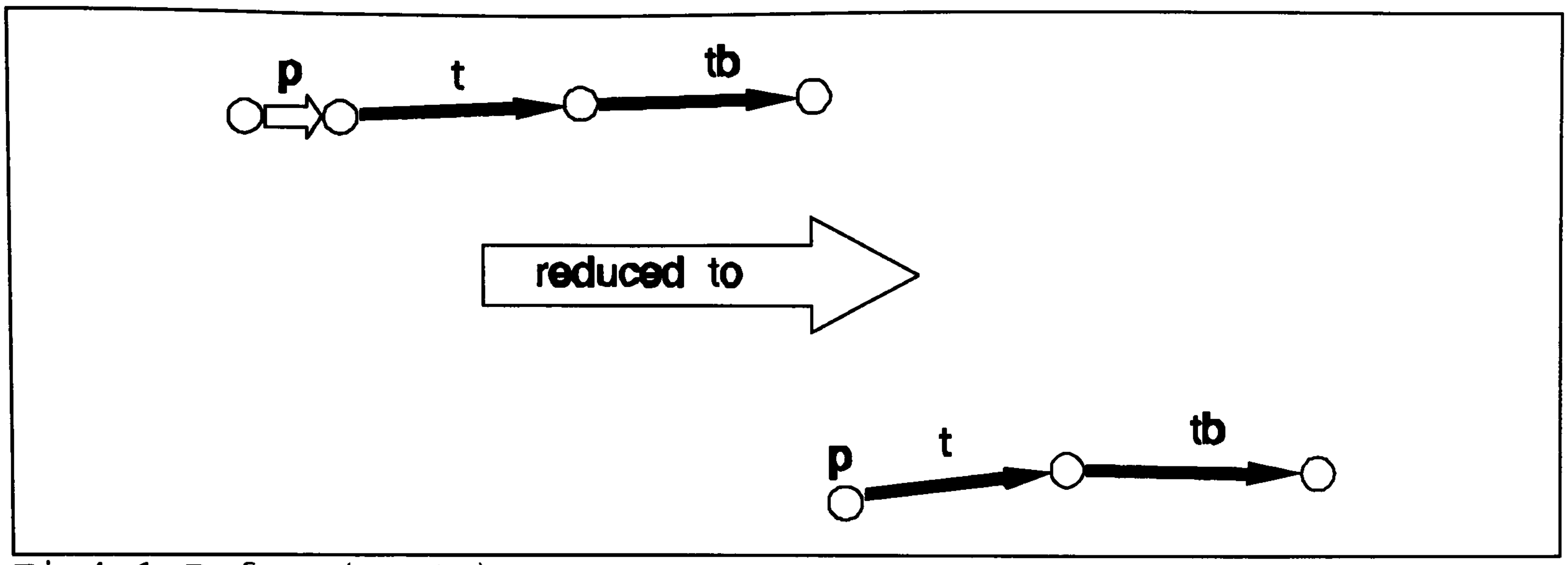

Fig4.1 Before(p, t_b)

Meets(p,t_b) (see Fig4.2) is also consistent and implies that t_b is *open* at p.

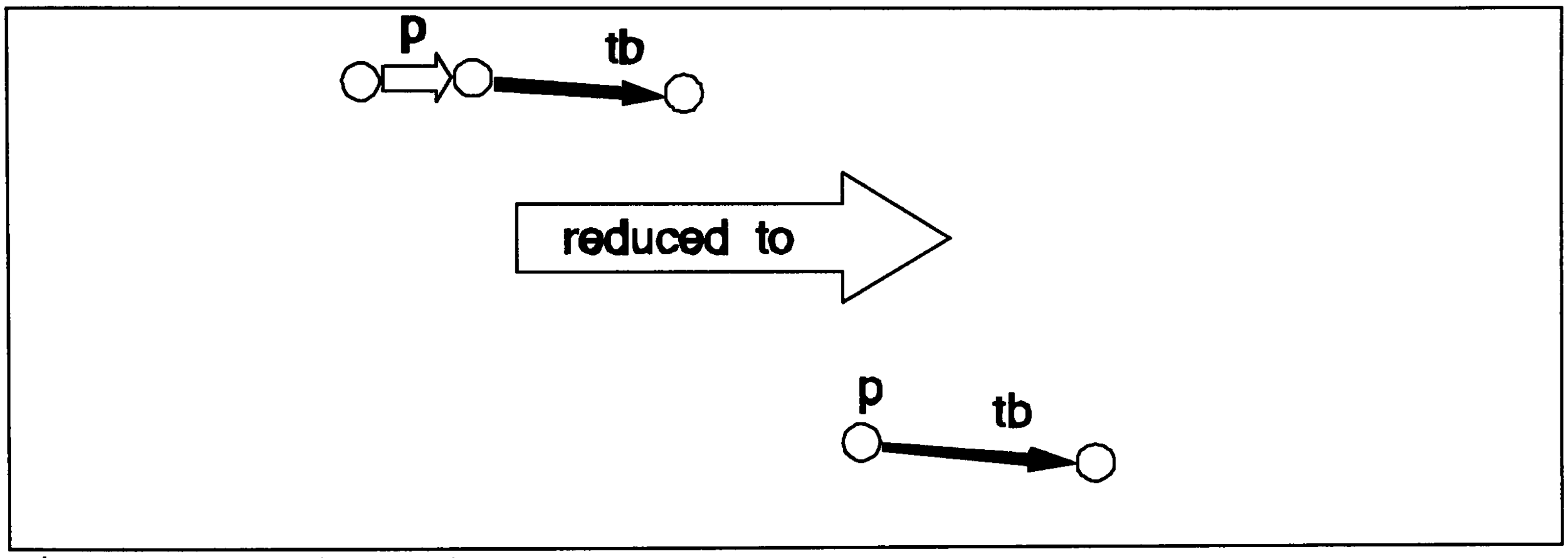

Fig4.2 Meets(p, t_b)

However, consider the following case:

Overlaps(p, t_b) (see Fig4.3).

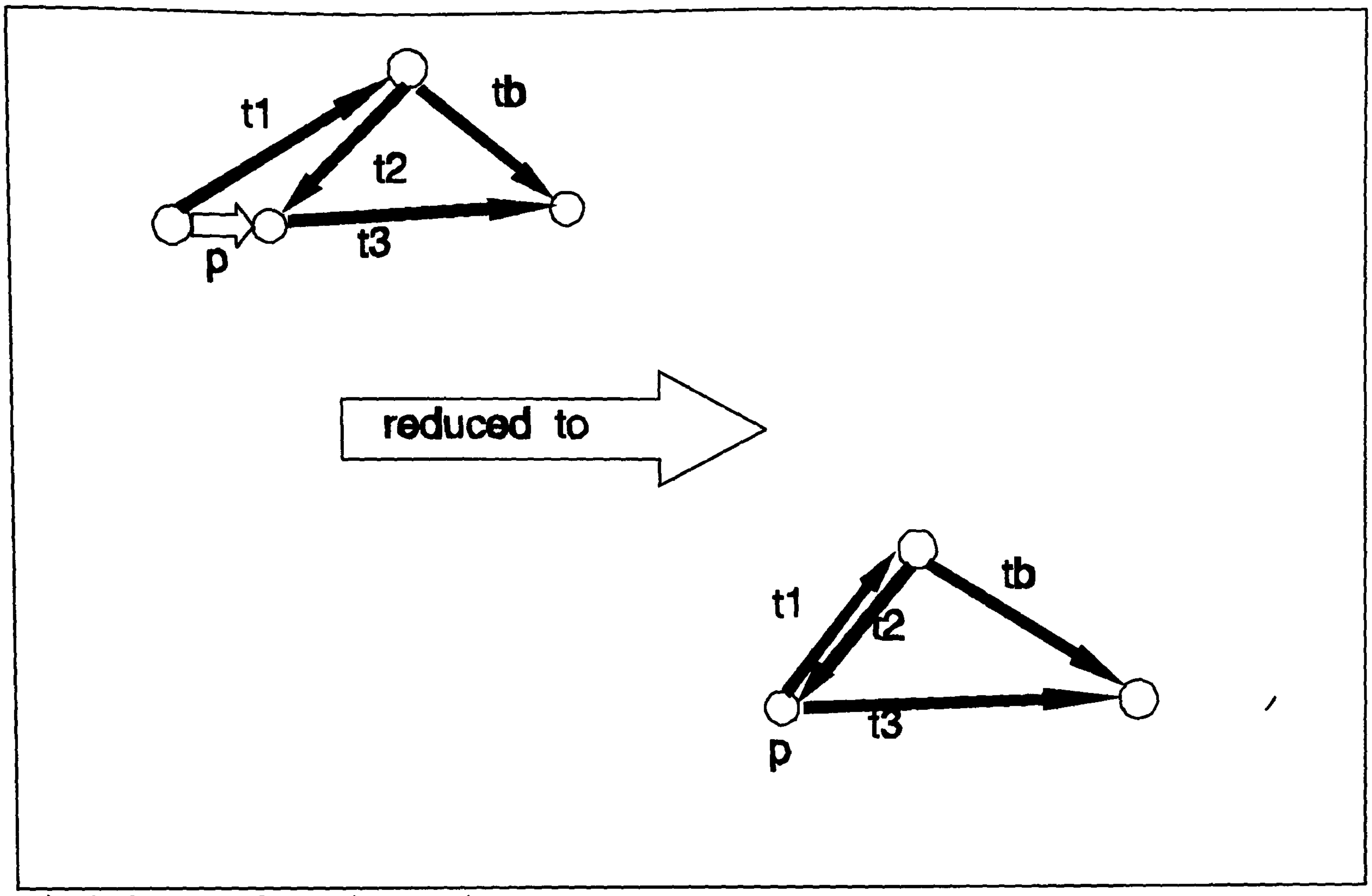

Fig4.3 Overlaps(p, t_b)

In this case the reduced graph is cyclic, so that the system is inconsistent. This is what we would expect intuitively: t_1 and t_2 are closed at p and t_3 is open at p; m(t_1, t_2) and m(t_2, t_3) assert t_1 = [p, X}, t_2 = {X, p], t3 = (p, Y} are consecutive intervals, where "{" represents either open or closed. This is obviously impossible.

We may also show in a similar fashion that other relations such as Starts(p,t_b), During(p, t_b), Finishes(p, t_b) are consistent, but overlapped_by(p,t_b) is inconsistent.

5. Composition of Allen's Primitive Relations

Allen and Hayes show that the transitivity table in [All83,84] is a result of the axiomatization in [HaH89], following the intuitive reasoning by possible cases which was used to construct the table originally. However, a full proof is not possible without a formal consistency checker since they are not able to show why other cases are not possible. We give here two formal proofs that the transitivity table follows from the axioms, by using the necessary and sufficient condition of consistency in terms of acyclicity of "meets", and the similar function between E and an initial segment of **N**, respectively.

For example, consider the transitivity:

Before(t_a, t_b), During(t_b, t_c).

Using the necessary and sufficient condition of consistency in terms of acyclicity of "meets", we can prove that the possible relation between t_a and t_c is Before(t_a, t_c), or Overlaps(t_a, t_c), or Meets(t_a, t_c), or During(t_a, t_c), or Starts(t_a, t_c), as follows:

$$\text{Before}(t_a, t_b) \wedge \text{During}(t_b, t_c) \Leftrightarrow \exists t, t_1, t_2 \in T \bullet (m(t_a, t) \wedge m(t, t_b) \wedge t_c = t_1 + t_b + t_2)$$

It is possible for us to take:

(1) $t = t' + t_1$ which gives $m(t_a, t')$ and $m(t', t_c)$, i.e. Before(t_a, t_c). From Fig5.1, we know this case is consistent. Hence we have shown that Before(t_a, t_c) is one possible case under the conditions Before(t_a, t_b) and During(t_b, t_c).

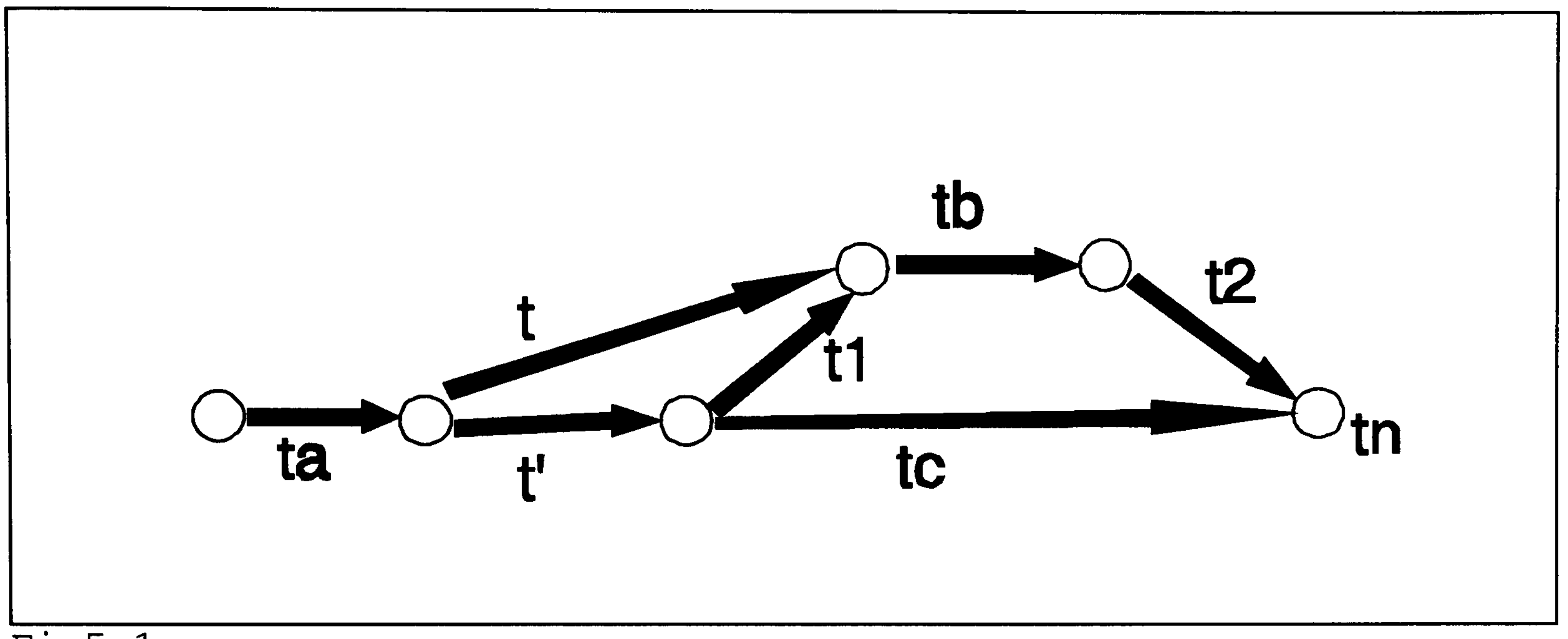

Fig5.1

In the similar way, we can show that Overlaps(t_a, t_c), M(t_a, t_c), During(t_a, t_c) and Starts(t_a, t_c) are also possible cases, as follows:

(2) $t_a = t_3 + t_4 \wedge t_c = t_4 + t_5 \Rightarrow$ Overlaps(t_a, t_c).

(3) $t_1 = t \Rightarrow m(t_a, t_c)$, i.e. M($t_a$, t_c).

(4) $t_1 = t_3 + t_a + t \Rightarrow t_c = t_3 + t_a + (t + t_b + t_2)$, i.e. During($t_a$, t_c).

(5) $t_1 = t_a + t \Rightarrow t_c = t_a + (t + t_b + t_2)$, i.e. Starts($t_a$, t_c).

Additionally, we can prove that there is no other possible relation between t_a and t_b as follows:

(6) After (t_a, t_c) $\Rightarrow \exists t' \in \text{TimeElement}(m(t_c, t') \wedge m(t', t_a))$.

However, $m(t_a, t)$, $m(t, t_b)$, $m(t_b, t_2)$, $m(t_2, t')$, $m(t', t_a)$ form a cycle: $t_a, t, t_b, t_2, t', t_a$, which shows inconsistency (see Fig5.2).

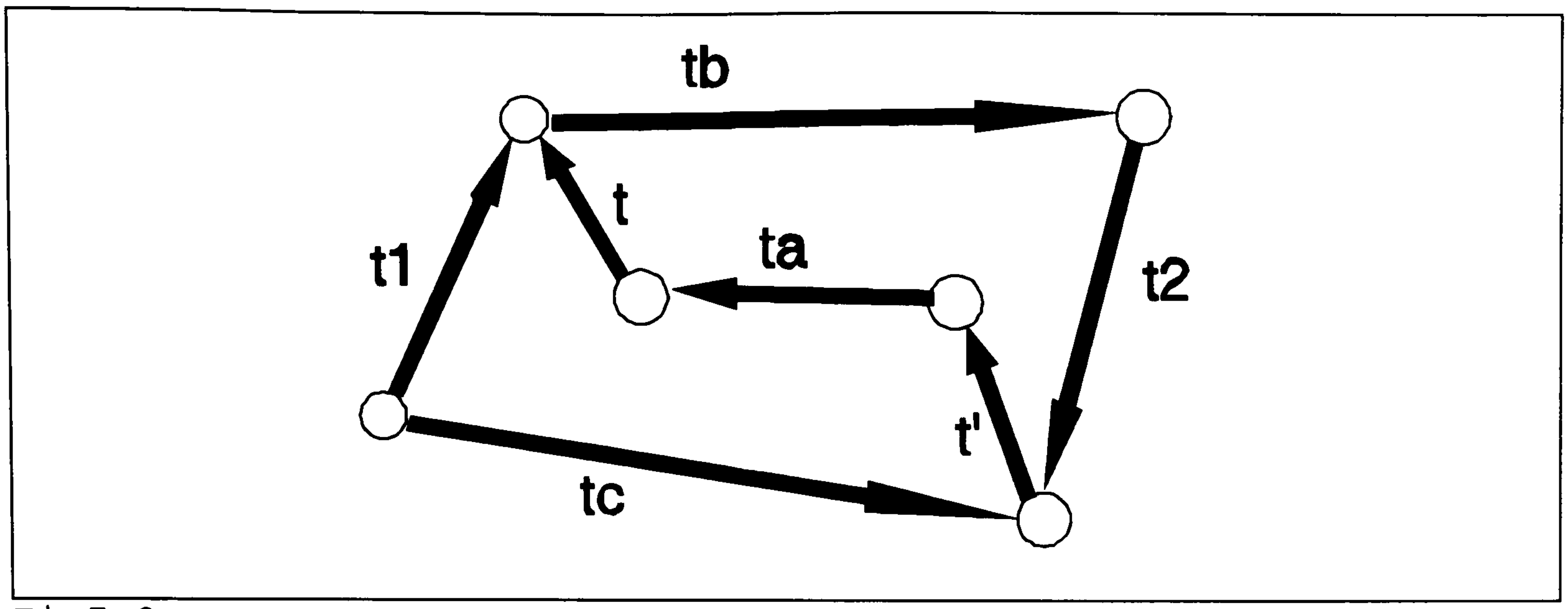

Fig5.2

Similarly, for other cases:

(7) Met-by(t_a, t_c) $\Rightarrow$ $m(t_c, t_a)$, so that there is a cycle: t_a, t, t_b, t_2, t_a, which shows inconsistency.

(8) Overlapped-by(t_a, t_c) $\Rightarrow$ $\exists t_3, t_4, t_5 \in$ TimeElement($t_c = t_3 + t_4$, $t_a = t_4 + t_5$), which forms a cycle: t_5, t, t_b, t_2, t_5, and shows inconsistency.

(9) Started-by(t_a, t_c) $\Rightarrow$ $\exists t_3 \in$ TimeElement($t_a = t_c + t_3$), which forms a cycle: t_3, t, t_b, t_2, t_3, and shows inconsistency.

(10) Contains(t_a, t_c) $\Rightarrow$ $\exists t_3, t_4 \in$ TimeElement($t_a = t_3 + t_c + t_4$), which forms a cycle: t_4, t, t_b, t_2, t_4, and shows inconsistency.

(11) Finishes(t_a, t_c) $\Rightarrow$ $\exists t_3 \in$ TimeElement($t_c = t_3 + t_a$), which forms a cycle: t, t_b, t_2, t, and shows inconsistency.

(12) Finished-by(t_a, t_c) $\Rightarrow$ $\exists t_3 \in$ TimeElement($t_a = t_3 + t_c$, which forms a cycle: t, t_b, t_2, t, and shows inconsistency.

All the entries of Allen's transitivity table have been checked in the above way. From our assumption we know that a point does not meet or be met-by another point, and from our axioms we have proved that a point will not overlap or be overlapped-by other interval. Hence, we can extend Allen's system to include time points (this will overcomes the problems involved in the need to model 'open' and 'closed' intervals, by allowing knowledge of interval end-points to be expressed explicitly) and prove (in terms of the condition of acyclicity of "meets") that the corresponding transitivity of the extended model is just as same as that one of Allen.

6. Conclusions

In this paper, a formal specification in **Z** for a temporal system based on both intervals and points has been provided. A significant feature of this approach is that, in common with Allen's system, intervals do not need to be defined as point pairs. It has been shown here that specialisation of the system to intervals leads to Allen's system. The model allows reasoning on the primitive elements, and provides an extension of Allen's system that includes both intervals and points as primitive elements. The formulation of axioms by means of a single relation allows a graphical representation of the temporal database entities, and this in turn allows an efficient consistency checker in terms of a search for graphical cycles. It has been proved that Allen's and Vilain's truth-propagation inference may be derived.

6. References.

[Lip64] Lipschutz S.
Set Theory and Related Topics, Schaum publishing Co. New York, 1964.

[Car79] Carre B.
Graphs and Networks, Clarendon Press, Oxford, 1979.

[Vil82] Vilain M. V.
"A System for Reasoning about Time", Proc. AAAI-82, Pittsburgh, PA. pp.197-201.

[All83] Allen J. F.
"Maintaining Knowledge about Temporal Intervals", communication of ACM. Nov. 1983, Vol.26, pp.123-154.

[All84] Allen J. F.
"Towards a General Theory of Action and Time", Artificial Intelligence 23 (2), July 1984, pp.123-154.

[ViK86] Vilain M. V. and Kautz H.
"Constraint Propagation Algorithms for Temporal Reasoning", Pro. AAAI-86 (1986).

[AlH89] Allen J. F. and Hayes P. J.
"Moments and Points in an Interval-based Temporal-based Logic", Comput. Intell. (Canada), Vol.5, no.4, pp.225-238, Nov. 1989.

[Dil90] Diller A.
Z--AN Introduction to Formal Methods, John Wiley & Sons Ltd., Baffins Lane, 1990, London: George Allen and Unwin, 1990.

Appendix B

A General Temporal Model Supporting Duration Reasoning

(*AI Communications*, Vol.5 No.2, pp.75-84, 1992)

Abstract

This paper proposes axioms for a temporal system based on a discrete set of primitive elements, which may be intervals or points, supporting duration reasoning. It is shown that this system can be interpreted in various possible models. The proposed system overcomes problems involved in the need to model 'open' and 'closed' intervals, by allowing knowledge of interval end-points to be expressed explicitly.

The axioms are formulated in terms of a single relation 'meets', and formalise an intuitive consistency condition for a temporal database: that a well-ordered sequence of fundamental elements exists which underlies the database. A graphical representation of the database is given, in terms of which a necessary and sufficient consistency condition for the existence of a well-order is proved.

Key words: time, temporal logic, incomplete knowledge, intervals, points, duration reasoning, consistency.

1.Introduction

There are many approaches to temporal systems, led initially by Russell [14] who took the first order logic approach. Of these, some describe computer systems whose primitive representation of time is in terms of points and inherit axioms accordingly. This is the case with the "time specialist" of Kahn and Gorry [11], the naive physics of Hayes [10], and the time map of McDermott and Dean [6,13]. However, there are problems with such systems in attempting to model many of the qualitative concepts needed in a temporal database where attributes are assigned to time *intervals*. For these systems, interval-based logic has been introduced according to two different views of intervals. One view takes intervals as primitive objects, such as the models of Allen [1-3] and of Vilain [15,16]. Another view, such as the CHRONOS system of Bruce [4], the propositional modal logic of Halpern and Shoham [9], and the TCSP of Dechter, Meiri and Pearl [7], takes the problems involving constraints on pairs of time points, which can be considered as weaker.

Allen introduced his temporal logic in order to provide a framework for the treatment of two major subareas of artificial intelligence: natural language processing and problem solving. Instead of adopting time points (or states which are associated with time points), he took intervals as the primitive temporal quantity and introduced (Allen[1,2]) thirteen (mutually exclusive) relations between any two intervals, which are formally defined in terms of the single relation 'meets' (Allen[3]). In the former papers (e.g., Allen[1,2]), Allen specifically excludes time points in claiming that any quantity of time must be *subdivisible.* However, in the recent paper (Allen[3]), he defines point as the "meeting place" of intervals and uses the concept of a "moment" , i.e. a very short interval which is indivisible, in order to model instantaneous events. The contention is that nothing can be true at a point, for a point is not an entity at which things happen or are true. The most obvious structural difference between points and moments is that moments can meet other intervals (especially, other moments), and hence stand between them, while points are not treated as primitive objects and cannot meet anything. Allen gives the example of a light which is switched on. To model such a system, we may need two intervals: one where the light is off 'meeting' one where it is switched on. This may be enough for our modelling purposes and, if so, we do not need to state whether the intervals are open or closed. By this means, we avoid awkward questions about the end-points: if the intervals are closed then there is a point at which the light is on and off, similarly if they are both open, then at this point the light is neither on nor off. Allen's argument is that if we really want to model the switching process itself, then we need to examine in more detail the physics, and smaller intervals at this event will again be sufficient.

However, the problem to model open or closed intervals still exists, and there are some difficulties with this approach in the qualitative modelling of everyday occurrences. In qualitative modelling of physical processes, we often wish to impose 'landmark' points which by definition separate two intervals. For instance, consider the example mentioned by Allen: throwing a ball up into the air. The motion of the ball can be modelled by a quantity space of three elements: going-up, stationary, and going-down. Intuitively, there are intervals for going up, and going down. However, there is no interval, however small, over which the ball is neither going up nor going down. The property of being stationary is naturally associated with a point, rather than a moment, a 'landmark' point which separates two other intervals. We may deduce from this example that qualitative modelling appears naturally to require a discrete quantity space of both points and intervals.
Vilain [15] has described an extension of Allen's system which includes points. In this, Allen's 13 interval relations are extended to 26 relations between points and intervals. However, this leads to an increase in computation for closure, and a corresponding overhead for consistency maintenance. Vilain and Kautz [16] have demonstrated that such a calculation is NP-hard, as opposed to an alternative point-based system which is $O(n^3)$. Unfortunately, the problems involved in modelling 'open' and 'closed' intervals still exist.

Our intention is to propose a model by providing axioms for a discrete system of primitive elements which may be points or intervals, and a single relation over the elements. This gives a basis for discrete temporal reasoning. These axioms allow modelling of both open and closed intervals, or allow intervals to be unspecified. Further, they support duration reasoning, which has been a problematic aspect in many temporal systems. A necessary and sufficient condition for the consistency of an incomplete temporal system can be formulated in terms of LP problem. In the case that there is no duration constraints, the corresponding condition is formulated in terms of graph cycle, and the consistency checking is $O(n)$, where n is the number of graphical arcs. In Allen's system, consistency checking is performed by formation of the transitive closure according to a transitivity table with 144 entries which describes the composition of the thirteen (mutually exclusive) relations. If no conflict is found according to the exclusivity, then the system is consistent. However, they have not provided a formal consistency checker. We give here a formal proof that the transitivity table follows from the axioms, by using the necessary and sufficient condition of consistency in terms of acyclicity of 'meets'.

Intuitively, the starting point for this axiomatisation is a view of time as a well-ordered discrete set, E, of fundamental elements which may either be points or intervals with a duration assignment, and E has no *limit* element (see e.g. Lipschutz[12]). We postulate that this fundamental set exists, sufficient for our modelling purposes, although we do not necessarily have full knowledge of E. That is, although we know that there exists a well-ordered E, we do not necessarily know what it is in any situation. Modelling our knowledge of a situation means expressing a (partial) knowledge of the temporal ordering of E. For example, such knowledge may be expressed in terms of Allen's interval relationships which may all be expressed in terms of a single relation: 'meets'. The relation 'meets' is defined as the immediate successor relation under the well-order, so that each element 'meets' a next element (except possibly for the last element). We may also define compound elements from two elements which meet. When elements e1 and e2 meet in E, we may construct a compound element k = e1 ⊕ e2, and this in turn may be used to construct other compound elements.

This view of time is one which to some extent is forced upon us by the practicality of the computer based modelling approach. We have to store elements as a discrete finite set, and the semantics of any database of time elements will naturally assume a well-order at some fundamental level. However, the database represents our state of knowledge about temporal events, and this may well be incomplete. The knowledge incorporated in the database will in general not be the fundamental elements in E, but rather of the compound elements k. The existence of E is simply a belief which may be used to test the consistency of a database. This is the approach we take to consistency here : if we can show that E exists, then the database is consistent, if we can show it does not exist, then it is inconsistent.

Excepting the axiom that the duration of a interval is positive while the duration of a point is zero, the differentiating property between interval and point which is proposed here is that although intervals may succeed points or intervals, points are not successors to points, although they can meet (or met-by) other intervals. This characteristic, which is later built into the axioms, is in line both with modelling requirements where points are defined as separators or end-points of intervals, and with the denseness of points on the real line. But this is the only extra requirement which is made of elements if they are to be points. According to their definitions, points, as our primitive elements, are different from either Allen's points or moments. It seems that Allen's moments may be taken as the elementary intervals in E. We show by example in section 3.1 and section 4.2 how open and closed intervals may be modelled as and when required using this system.

In section 2, we give a formal definition of the system. In section 3 it is shown how incomplete knowledge may be represented, and how inference may be performed on it. A necessary and sufficient condition for consistency is provided, and is illustrated by means of an example. In section 4 a limited case without duration is examined; a formal discussion of the treatment of open and closed intervals is given in section 4.2. Further, section 4.3 proves Allen's transitivity closure formally by using the consistency checker. Section 5 provides a summary and concluding remarks.

2. Definition of Temporal System

The definition of a temporal system supporting duration reasoning consists first of a definition of an underlying well-ordered discrete set E without *limit* elements. The elements of E may be points or intervals with a duration assignment. The temporal system is then defined as the closure of E under the binary operations of combining adjacent elements and corresponding duration addition.

Let I be a set whose elements are intervals with a positive duration and P be a set whose elements are points with a zero-length duration.

Let $E = I \cup P$, which is *similar* to an *initial segment* of the set of natural numbers with the natural order (see Lipschutz[12]), with an *immediate successor* relation 'meets', such that:

$$\forall i \in I \; (\exists e \in E \; (\text{meets}(i, e) \vee \text{meets}(e, i) \;) \wedge (\; 0 < \text{duration}(i) \in \mathbf{R} \;) \;) \qquad (2.1.1)$$
$$\wedge \; \forall p \in P \; (\exists i \in I \; (\text{meets}(p, i) \vee \text{meets}(i, p) \;) \wedge (\; 0 = \text{duration}(p) \in \mathbf{R} \;) \;) \qquad (2.1.2)$$

We shall notate the "duration" assignment to the elements of E as D_E and the unique successor relation 'meets' over E as M_E.

The above definitions imply the existence of a well-order over a discrete system E without limit elements, and no two points are next to each other.

A temporal system supporting duration reasoning is a triad (T, M, D), where $T \supset E$, M is a successor relation 'meets' over T and D is a duration assignment to the elements of T with the following properties:

1) M coincides with M_E over E, that is, $M|_E = M_E$;
2) D coincides with D_E over E, that is, $D|_E = D_E$;
3) $t_1,t_2 \in T$, and $\text{meets}(t_1,t_2) \Rightarrow t \in T$, where

   for all $t_a,t_b \in T$:
   $\text{meets}(t_a,t_1) \Leftrightarrow \text{meets}(t_a,t)$,
   $\text{meets}(t_2,t_b) \Leftrightarrow \text{meets}(t,t_b)$.

Define:

$$t = t_1 \oplus t_2,$$
$$\text{duration}(t) = \text{duration}(t_1) + \text{duration}(t_2).$$

Nb. The operator "$\oplus$" is associative but not commutative and "+" is just the normal addition operator of real numbers.

4) (T, M, D) is the closure of (E, M_E, D_E) under $\oplus$ and +.

In the case that $t = t_1 \oplus t_2$, we say that t is decomposable into t_1 and t_2, and t_1 is equal to t_2 means they are the same elements, denoted by $t_1 = t_2$.

Nb. If we take $D = \emptyset$, we get the limited system defined in section 4.

3. Incomplete Temporal System

The set T includes E and all the intervals and points which can be formed from it by means of $\oplus$ and +. However, in an application neither the fundamental set E nor the complete set T may be known. A database of "facts" about T will express knowledge that is incomplete in several ways. For example, the database may contain knowledge of duration for only some of its elements, and may have incomplete knowledge about the 'meets' relation. In addition, the database will often contain redundancy, as when facts are known about two elements without the knowledge that they are actually identical. For example, we may know that meets(a,b) and duration(c) =1, without knowing that a and c are the same element. To allow for possible duplicate elements, the basic structure of the database is that of a bag, rather than a set.

Accordingly, we take the following representation to express incomplete temporal knowledge :

i) $K = K_1 \uplus K_2 \uplus ... \uplus K_p$, where $K_i \subseteq K_{i+1} \subseteq T$, $i = 1, ..., p-1$; and "$\uplus$" represents the *bag* union. K is called a bag (see Diller[8]). We use K^e to denote the ordinary set union of K_i, that is, $K^e = \cup_i K_i$, hence $K^e \subseteq T$;
ii) $M_K = M|_{K1} \uplus M|_{K2} \uplus ... \uplus M|_{Kp}$;
iii) $D_{K0} \sqsubseteq D_K = D|_{K1} \uplus D|_{K2} \uplus ... \uplus D|_{Kp}$; here, "$\sqsubseteq$" represents the *sub-bag* relation.

Nb. i) expresses our knowledge of what intervals and points are there;
ii) expresses our knowledge as to how the intervals or points in K meet each other;
iii) expresses our knowledge of duration over a sub-bag K_0 of K.

3.1 An Example

Before giving a formal development of the axioms, we illustrate the main ideas by means of the simple example, which we shall later relate to the modelling of open and closed intervals. Consider knowledge represented by (K, M_K, D_{K0}), where:

$K = [\![k_0, k_1, k_2, k_3, k_4, k_5, k_6, k_7]\!]$,
$M_K = [\![$ meets(k_0, k_1), meets(k_0, k_2), meets(k_1, k_3),
meets(k_1, k_4), meets(k_2, k_5), meets(k_3, k_5),
meets(k_4, k_6), meets(k_5, k_6), meets(k_5, k_7) $]\!]$
$D_{K0} = [\![$ duration$(k_4)=1$, duration$(k_5)=1$, duration$(k_7)=0$ $]\!]$.

The system may be represented by a graph as in Fig3.1.

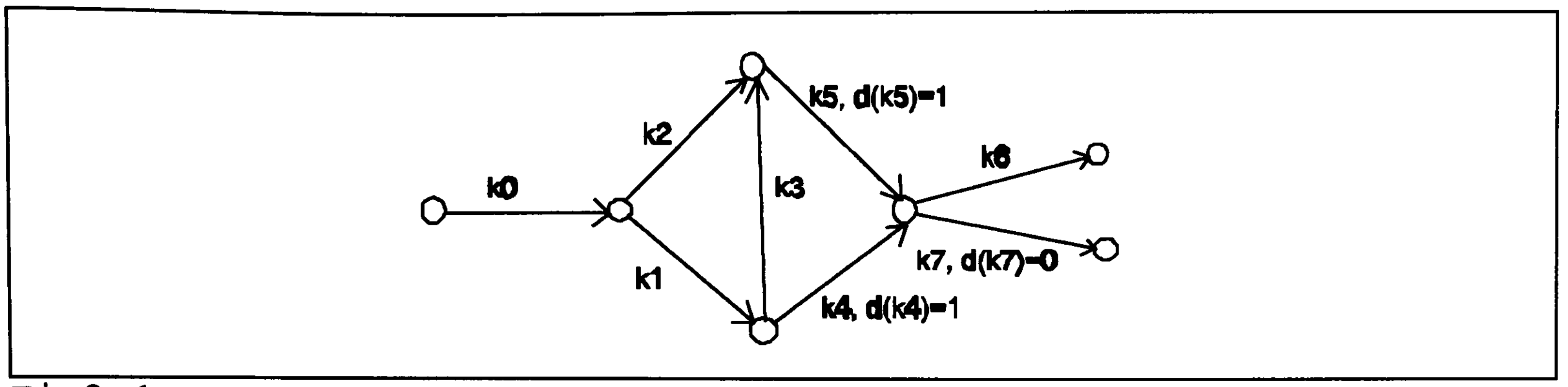

Fig3.1

Here, the arcs of the graph represent the elements, and meets(k_i,k_j) is represented by k_i being in-arc and k_j being out-arc to a common node n. For elements in K where there is a duration assignment, the arcs are weighted by the duration.

In order to demonstrate the consistency of (K, M_K, D_{K0}), consider the triad (E,M_E,D_E):

E = {k_0, k_1, k_3, k_5, k_7},
M_E = {meets(k_0, k_1), meets(k_1, k_3),
meets(k_3, k_5), meets(k_5, k_7)};
D_E = {duration(k_0)=0, duration(k_1)=1, duration(k_3)=0, duration(k_5)=1, duration(k_7)=0}.

If we let I = {k_1, k_5} be the set of intervals, and P = { k_0, k_3, k_7} be the set of points, then E = I $\cup$ P, and E satisfies conditions (2.1.1) and (2.1.2). There is a unique successor 'meets' over E, and no two points are next to each other.

Finally, we can show that k_2, k_4 and k_6 are in the closure T of E:

$k_1 \oplus k_3 = k_2$, $k_3 \oplus k_5 = k_4$, $k_6 = k_7$; and:
meets(k_0, k_2), meets(k_2, k_5), meets(k_1, k_4), meets(k_4, k_6),
meets(k_4, k_7), meets(k_5, k_7);
duration(k_2) = 1 = duration(k_1) + duration(k_3);
duration(k_4) = 1 = duration(k_3) + duration(k_5).
duration(k_6) = 0 = duration(k_7)

The example illustrates how "open and closed" intervals may be expressed. For example, consider again the quantity space for the motion of the ball described in section 1, and let k_1, k_3, k_5 be: going-up, stationary, and going-down respectively. In this case, k_4 represents "stationary $\oplus$ going-down". In terms of a point based system, the element k_4 represents the interval [1, 2), whereas k_5 represents (1, 2). The other elements are given in the table:

Element	point based element	
k_0	0	start
k_1	(0,1)	going-up
k_2	(0,1]	going-up $\oplus$ stationary
k_3	1	stationary
k_4	[1,2)	stationary $\oplus$ going-down
k_5	(1,2)	going-down
$k_6(=k_7)$	2	end

The effect of restriction 2.1.2 is to prohibit two points with no interval in between ; i.e an equivalent of a denseness postulate for point based systems. This would rule out as infeasible an assignment of 0 to either of the arcs k_1 or k_5; so that qualitative knowledge that an interval with positive duration is expressed in (K,M_K,D_{K0}). In a point-based system, it is implicit that (x, x), (x, x] and [x, x) are impossible, while [x, x] may be used to represent time point x.

3.2 A Necessary and Sufficient Condition for Consistency

To draw inferences from (K, M_K, D_{K0}), we must rely on the assumed properties of T, M and D. A consistency checker is needed which will establish whether a triad (K, M_K, D_{K0}) is consistent with our basic assumptions about T, M and D.

In general, a triad (K, M_K, D_{K0}) is consistent if we can add to K and make any necessary equality assignments, and add to M_K and to D_{K0}, so that the resulting triad (T, M, D) is a closure for some (E, M_E, D_E) satisfying the conditions in section 2. A necessary and sufficient condition for consistency may be given in terms of the graphical representation introduced in 3.1. For convenience, we adopt the notation that k_{ij} represents an arc from node n_i to node n_j, and d_{ij} represents the duration of this arc. We let G be the graph of (K, M_K, D_{K0}).

Let $N = \{ n_1, n_2, \ldots\ldots , n_S \}$ be the nodes in G. The system (K, M_K, D_{K0}) is consistent if and only if:

(I) There is a solution $(x_{i1,j1}, \ldots , x_{iq,jq})$ for unknown durations $(X_{i1,j1,}, \ldots , X_{iq,jq})$ which forms a $D_K \sqsupseteq D_{K0}$, where $x_{i,j} \geq 0$, such that:

(I.1) for each simple circuit in G_K, the directed sum of weights is zero.
(I.2) $\text{duration}(k_{ij}) + \text{duration}(k_{jh}) > 0$.

Otherwise, the system is inconsistent.

<u>Proof of sufficiency:</u>

i) We first show that if (I) holds, then a function f of N into **R** exists:

$$N \ni n \text{-------}> f(n) \in \mathbf{R}, \text{ such that:}$$

(II.1) If $k_{ij} \in K$, then:

$f(n_i) \leq f(n_j)$,

$f(n_j) - f(n_i) = \text{duration}(k_{ij}) \in D_K$;

(II.2) If k_{ij}, $k_{jh} \in K$, then:

$f(n_h) - f(n_i) > 0$.

Nb. condition (II.2) preserves that no two points "meet" each other.

To show this, we assume G to be connected by means of 'meets' (the extension to a graph with several connected components is straightforward).

Let y_{ij} denote the duration assignment for $k_{ij} \in K$, where

$y_{ij} = d_{ij}$, if $d_{ij} \in D_{K0}$;
$y_{ij} = x_{ij}$, otherwise.

Now take a spanning tree of G (i.e. a tree joining all the nodes of G, formed by removing some arcs from G). Selecting any node n_0 as origin, a unique path is determined by the spanning tree between n_0 and any other node n (Fig3.2). We may take f(n) as the directed sum of the weighted arcs from n_0 to n along this path.

With this assignment, condition (II.1) follows immediately for all arcs on the spanning tree. For any arc k_{ij} not on the spanning tree, we consider the circuit formed by k_{ij} together with the spanning tree. Applying condition (I.1), we have:

$y_{ij} - f(n_j) + f(n_i) = 0$,

i.e. (II.1) again holds.

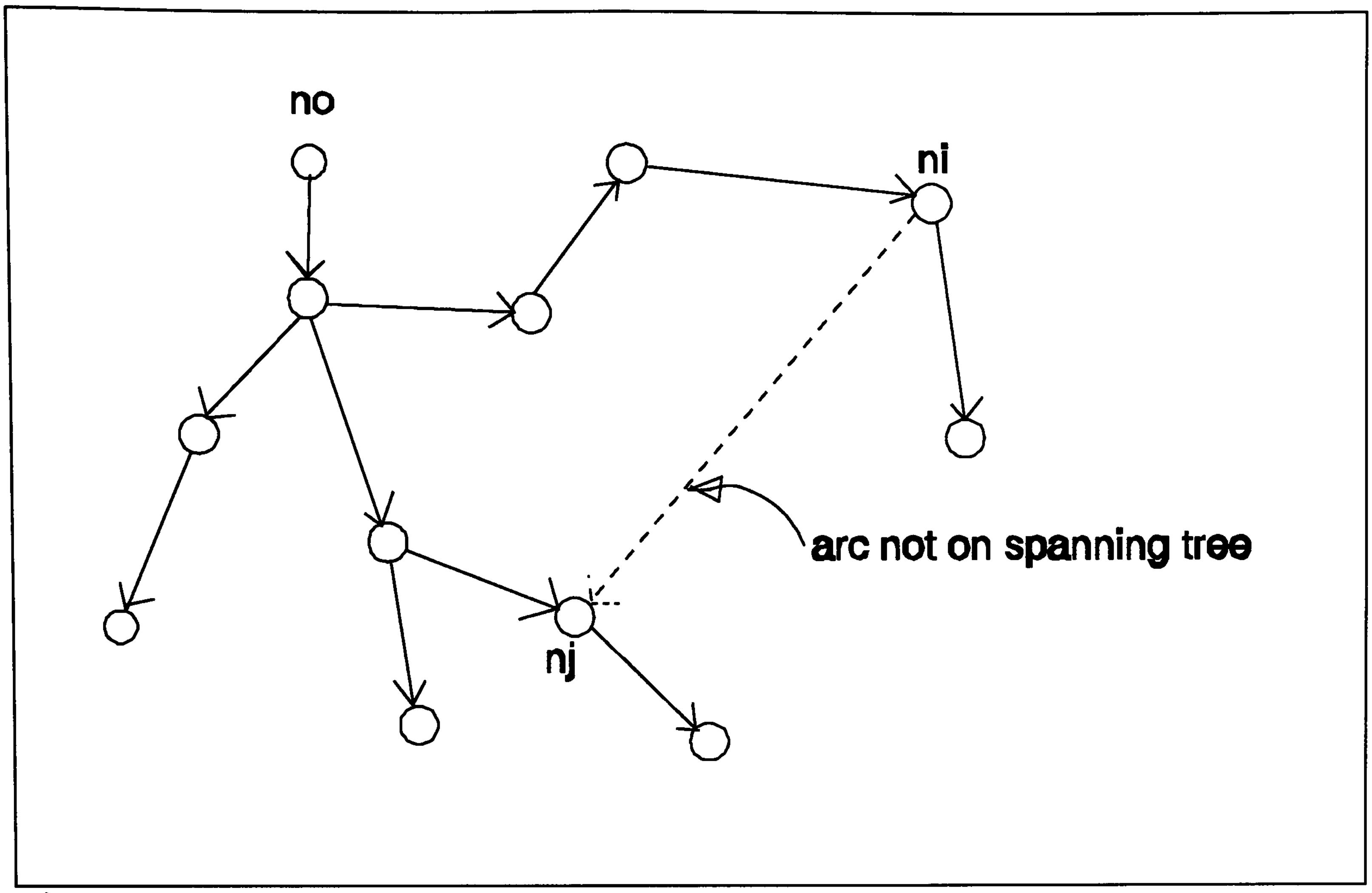

Fig3.2

Additionally, it is clear that condition (I.2) <==> (II.2).

ii) We now show that f(n) may be used to construct (E, M_E, D_E). In effect, the function assigns a time measure to the nodes. However, care must be taken to deal with points: if a number of nodes are assigned the same f(n), then we must be sure that we can construct an E without two consecutive points. In the procedure that follows, we show how this may be done.

(1) Define equivalent classes N_1, N_2, ... , N_{s1}, $s_1 \leq s$ among N as:
n_i, n_j belong to the same class N_r <==> $f(n_i) = f(n_j)$;

(2) The nodes within any class N_i are of three types:
(i) those that are in-nodes to zero duration arcs in K,
(ii) those that are out-nodes to zero duration arcs in K, and
(iii) those that are not in- or out- nodes to zero duration arcs in K

Condition I.2 ensures that there are no nodes that are both in-node and out-node to two zero duration arcs. The in-node and out-node to a zero duration arc will be in the same equivalence class, and the in-node must be ordered before the out-node. Accordingly we subdivide each class Ni into two subsets: N_i^1 containing nodes of type (i) and N_i^2 containing nodes of type (ii) and (iii).

(3) The graph of E is now formed over the set of subclasses as nodes. The successor

relation is defined by the natural ordering of equivalence classes according to f, and by the rule that N_i^2 is the 'successor' to N_i^1. Duration assignment to E is defined by duration $(N_j, N_{j+1}) = f(N_{j+1}) - f(N_j)$, where N_{j+1} is the successor to N_j in G.

iii) Finally we show that (K, M_K, D_{K0}) is in the closure of (E, M_E, D_E). We let e_{lm}^{rs} be the arc in the closure of E between node N_l^r and N_m^s. We make the following equality assignments over K:

$k_{ij} = e_{lm}^{rs}$ if $n_i \in N_l^r$ and $n_j \in N_m^s$,

With this assignment, k_{ij} is in the closure of E, and

$$\text{duration}(e_{lm}^{rs}) = f(N_m^s) - f(N_l^r) = d_{ij}.$$

Proof of necessity:

If (E, M_E, D_E) exist, satisfying condition 1) - 4) in section 2, then condition 3) shows directly that condition (I.1) holds, and since no two points meet each other in E, (I.2) holds.

3.3 An Illustration

Now we use the example given in section 3.1 again to illustrate the procedure of establishing the elementary triad (E, M_E, D_E):

There are two elementary circuits in G to consider. Setting the directed sum of weights in each of these equal to zero, we get 2 independent constraints:

$$\text{duration}(k_3) + \text{duration}(k_5) - \text{duration}(k_4) = 0,$$
$$\text{duration}(k_1) + \text{duration}(k_3) - \text{duration}(k_2) = 0.$$

By inspection, one consistent solution is:

$$\text{duration}(k_3) = 0, \text{duration}(k_1) = 1, \text{duration}(k_2) = 1, \text{duration}(k_0) = 0, \text{duration}(k_6) = 0.$$

(Nb. There may be other consistent solutions, for instance:

$$\text{duration}(k_3) = 0, \text{duration}(k_1) = 1.8, \text{duration}(k_2) = 1.8,$$
$$\text{duration}(k_0) = 3, \text{duration}(k_6) = 10.)$$

Correspondingly, let N = {n_1, n_2, n_3, n_4, n_5, n_6, n_7} be the nodes in graph (K, M_K, D_K), see Fig3.3:

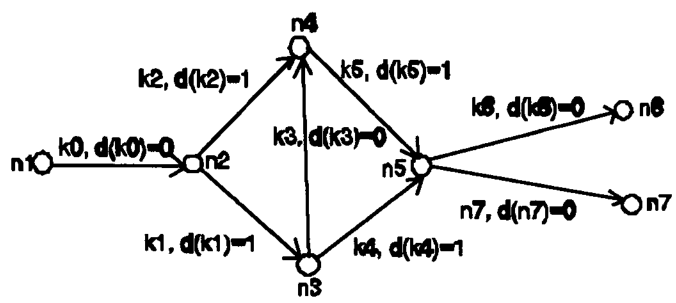

Fig3.3

the function f of N into **R** may be defined as:

$f(n_1) = 0$, $f(n_2) = 0$, $f(n_3) = 1$, $f(n_4) = 1$,
$f(n_5) = 2$, $f(n_6) = 2$, $f(n_7) = 2$.

which satisfies conditions I.1 and I.2.

(1) equivalent classes:
$N_1 = [\![n_1, n_2]\!]$,
$N_2 = [\![n_3, n_4]\!]$,
$N_3 = [\![n_5, n_6, n_7]\!]$;

(2) $N_1^1 = [\![n_1]\!]$, $N_1^2 = [\![n_2]\!]$, $N_2^1 = [\![n_3]\!]$,
$N_2^2 = [\![n_4]\!]$, $N_3^1 = [\![n_5]\!]$, $N_3^2 = [\![n_6, n_7]\!]$

(3) E = {k_0, k_1, k_3, k_5, k_7},
M_E = {meets(k_0, k_1), meets(k_1, k_3), meets(k_3, k_5), meets(k_5, k_7)};
D_E = {duration(k_0)=0, duration(k_1)=1, duration(k_3)=0, duration(k_5)=1, duration(k_7)=0}.

It is easy to see (E, M_E, D_E) satisfies the conditions given in the schemas of section 2.

4. A Limited Case

Now, we limit the temporal system (T, M, D) in a simple case: D = ∅, that is, there is no constraints about the length of the intervals or points. We denote this limited model to be a pair (T, M), where T and M are defined as in section 2 but excludes anything related to "duration". The distinction between intervals and points is definitely that: intervals are decomposable while points are not. Our intention here is to show that the pair (T, M) may be seen as an extended model to those of Bruce and Allen.

4.1 Limited Temporal Knowledge.

The set T includes E and all the intervals and points which can be formed from it by means of ⊕. As we mentioned in section 3, in an application neither the fundamental set E nor the complete set T may be known. Our basic assumption is that only a finite set of time intervals and points will be needed for modelling purposes, so that we may assume that E exists. The state of temporal knowledge as represented in a database will be represented as (K, M_K), K and M_K are defined as in section 3.

A consistency condition may be derived from the basic assumptions. An incomplete limited-system (K, M_K) will be consistent if we can add to K and make any necessary equality assignments, and add to M_K , so that the resulting pair (T, M) is a closure for some (E, M_E) satisfying:

1') M coincides with M_E over E, that is, $M|_E = M_E$;
2') (T, M) is the closure of (E, M_E) under $\oplus$, where $\oplus$ is defined as in section 2.

The necessary and sufficient condition for consistency which has been derived in section 3.2 can be reformulated in a more convenient form for this limited case in terms of graphs. Let $\underline{G}$ be the graph of (K, M_K), then (K, M_K) is consistent if and only if:

> (I.1)' $\underline{G}^r$ is acyclic, where $\underline{G}^r$ is the associated reduced graph formed from $\underline{G}$ by merging two nodes connected by a point in $\underline{G}$ and removing the corresponding arc.
> (I.2)' there are no nodes that are both in-node and out-node to two point-arcs in $\underline{G}$.

The above conclusion can be directly derived from the result given in section 3.2. In fact, (I.1)'and (1.2)' ensure that we can always make the duration assignment satisfy condition (I.1) of section3.2; inversely, if there is a cycle in $\underline{G}^r$, then the directed sum of weights for the simple circuit derived from this cycle must be bigger than zero, which is contrary to (I.1). On the other hand, it is clear that (I.2)' <==> (I.2).

As an example of the consistency condition, we take a case where a database is consistent if an element t_a is not known to be a time point, but inconsistent if it is.

E.G:
meets(t_0, t_a),
meets(t_a, t_n),
meets(t_0, t_b),
meets(t_b, t_c),
meets(t_c, t_n).

If t_a is not known to be a point then the corresponding graph shown in Fig4.1(a) is non-cyclic, and the system is consistent.

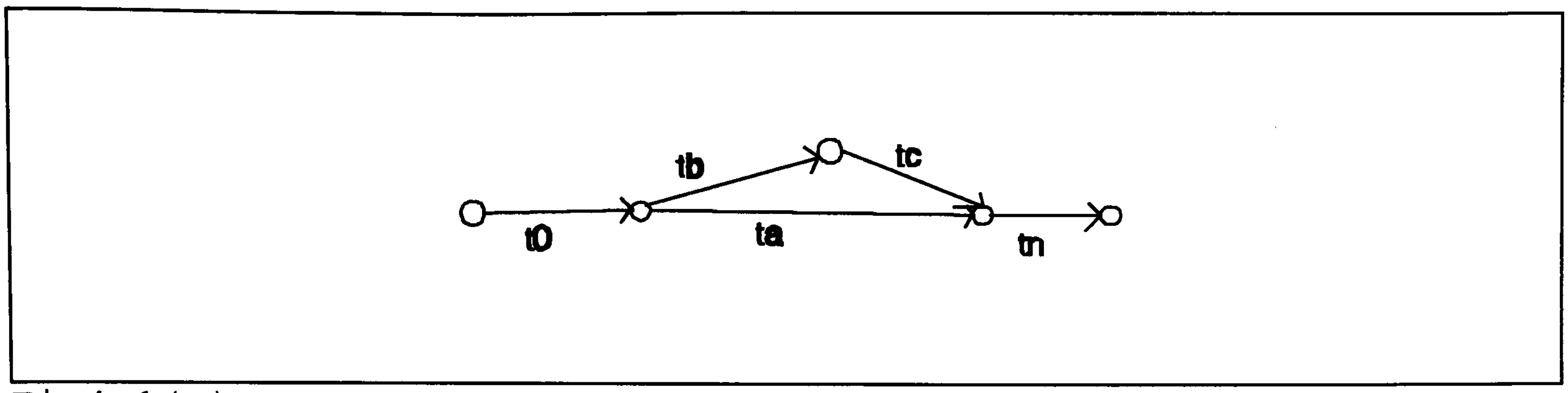

Fig4.1(a)

However, if $t_a \in P$, then we have the reduced graph in Fig4.1(b), which shows cyclicity, and we deduce that the system is inconsistent.

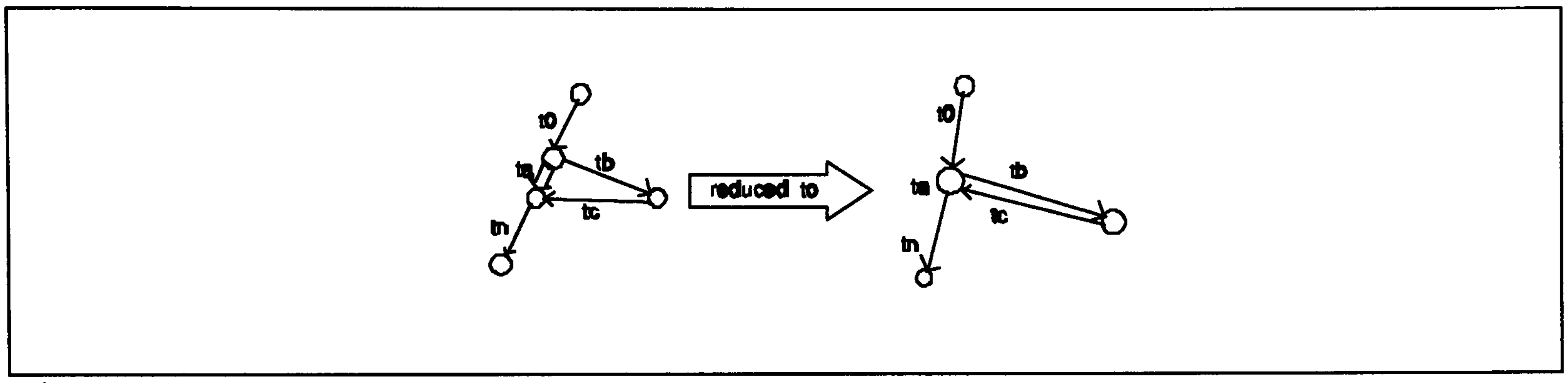

Fig4.1(b)

We can see why this is so intuitively by noticing that in Fig4.1(a):

$$t_a = t_b \oplus t_c \qquad (4.1)$$

This is consistent until we add the fact that t_a is non-decomposable. Since equation 4.1 states that t_a is decomposable, we reach an inconsistency when $t_a = p \in P$.

4.2 'Open' and 'Closed' Intervals

In [3], Allen and Hayes formulated the thirteen primitive relations between intervals in terms of the single relation 'meet'. However, here we have extended the primitive element to include points. In this case, some of Allen's relationships are consistent and some are inconsistent when they involve points.

For example, let $p \in P$:

Before(p, t_b) :- $\exists t \in T$:
meets(p,t),
meets(t,t_b).

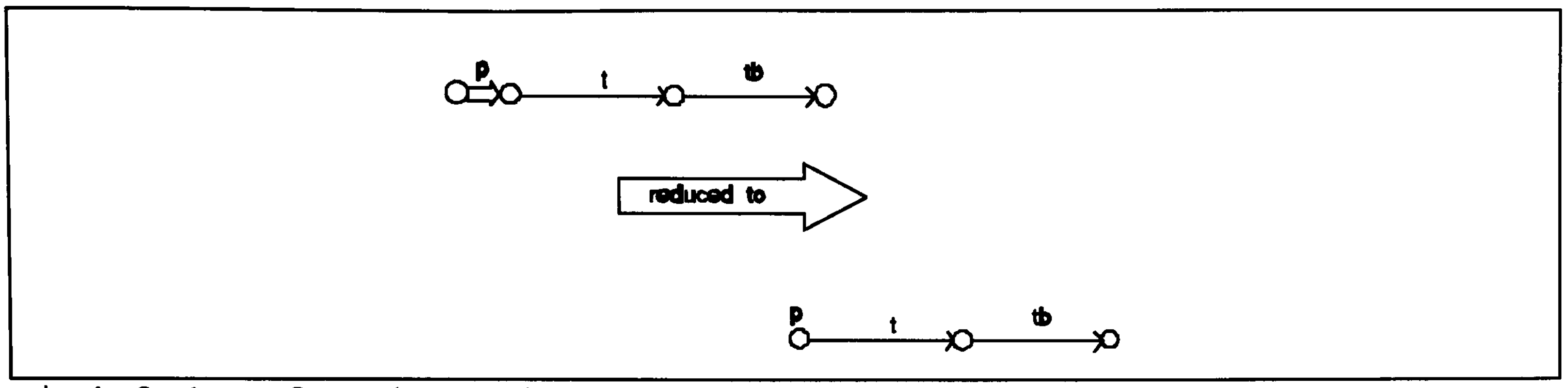

Fig4.2.1 Before(p, t_b)

i.e. This is consistent.

Meets(p,t_b) :-
meets(p,t_b).

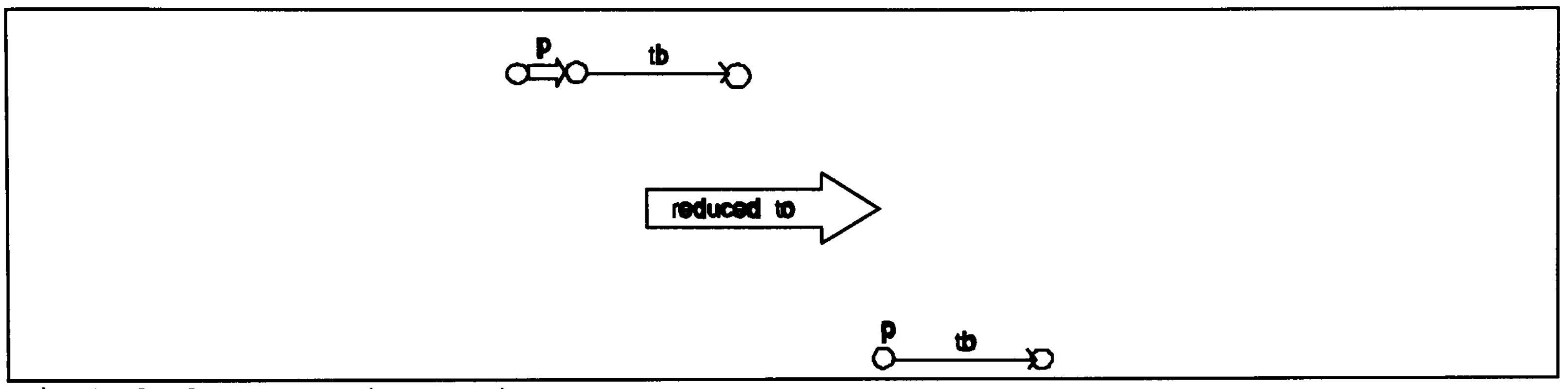

Fig4.2.2 Meets(p, t_b)

i.e. This is also consistent, and implies that t_b is open at p.

However, consider the following case:

$Overlaps(p, t_b)$:- $\exists t_0, t_1, t_2, t_3, t_n \in T$:
meets(t_0, p),
meets(p, t_3),
meets(t_3, t_n),
meets(t_0, t_1),
meets(t_1, t_b),
meets(t_b, t_n),
meets(t_1, t_2),
meets(t_2, t_3).

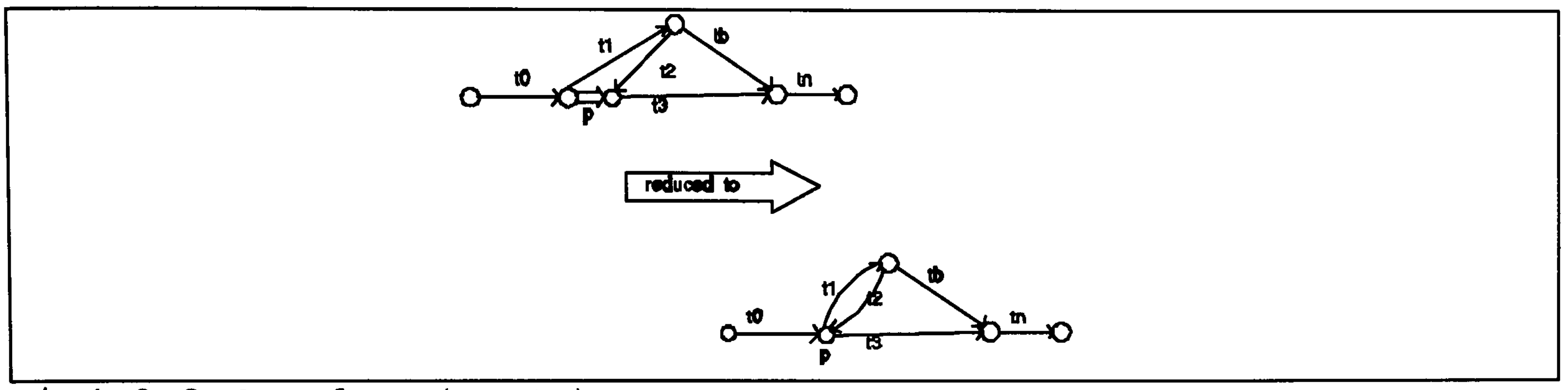

Fig4.2.3 Overlaps(p, t_b)

In this case the reduced graph is cyclic, so that the system is inconsistent. This is what we would expect intuitively: t_0 is open at p, therefore t_1 and t_2 are closed at p and t_3 is open at p. Therefore the last two predicates, meets(t_1, t_2) and meets(t_2, t_3), assert t_1 = [p, X}, t_2 = {X, p], t3 = (p, Y} are consecutive intervals, where '{' represents either open or closed. This is obviously impossible.

We may also show in a similar fashion that other relations such as Starts(p,t_b), During(p, t_b), Finishes(p, t_b) are consistent, but overlapped_by(p,t_b) is inconsistent.

4.3 Composition of Allen's Primitive Relations

Allen and Hayes show that the transitivity table in [1,2] is a result of the axiomatization in [3], following the intuitive reasoning by possible cases which was used to construct the table originally. However, a full proof is not possible without a formal consistency checker since they are not able to show why other cases are not possible. We give here a formal proof that the transitivity table follows from the axioms, by using the necessary and sufficient condition of consistency in terms of acyclicity of 'meets'.

For example, consider the transitivity:

Before(t_a, t_b), During(t_b, t_c).

We can prove that the possible relation between t_a and t_c is Before(t_a, t_c), or Overlaps(t_a, t_c), or Meets(t_a, t_c), or During(t_a, t_c), or Starts(t_a, t_c), as follows:

Before(t_a, t_b) and During(t_b, t_c) :- $\exists$t', t_0, t_1, t_2, t_n $\in$T:
meets(t_a, t'), meets(t', t_b),
meets(t_0, t_1), meets(t_1, t_b), meets(t_b, t_2), meets(t_2, t_n),

meets(t_o, t_c), meets(t_c, t_n).

(1) when t' = $t_0 \oplus t_1$, then meets(t_a, t_0), meets(t_0, t_c),
i.e: Before(t_a, t_c) (see Fig4.3.1).

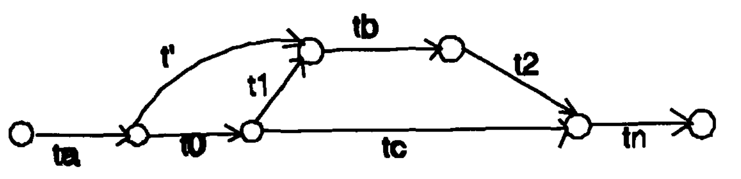

Fig4.3.1

(2) when $t_a = t_0 \oplus t_3$, $t_c = t_3 \oplus t_4$, $t_1 = t_3 \oplus t'$, where $t_3 \in T$, then:
$t_4 = t' \oplus t_b \oplus t_2 \in T$,
meets(t_0', t_a), meets(t_a, t_4), meets(t_4, t_n),
meets(t_0', t_0), meets(t_0, t_c), meets(t_c, t_n),
meets(t_0, t_3), meets(t_3, t_4),
i.e: Overlaps(t_a, t_c).

(3) when $t_a = t_0$, $t_1 = t'$, then meets(t_a, t_c),
i.e: Meets(t_a, t_c).

(4) when $t_1 = t_3 \oplus t_a \oplus t'$, where $t_3 \in T$, $t_4 = t' \oplus t_b \oplus t_2 \in T$, then:
meets(t_0, t_3), meets(t_3, t_a), meets(t_a, t_4), meets(t_4, t_n), meets(t_0, t_c), meets(t_c, t_n),
i.e: During(t_a, t_c).

(5) when $t_c = t_3 \oplus t_4$, $t_a = t_3$, $t_1 = t_3 \oplus t'$, where $T_3 \in T$, $t_4 = t' \oplus t_b \oplus t_2 \in T$, then:
meets(t_0, t_a), meets(t_a, t_4), meets(t_4, t_n), meets(t_0, t_c), meets(t_c, t_n),
i.e: Starts(t_a, t_c).

On the other hand, we can prove that there is no other possible relation between t_a and t_b as follows:

(6) If After (t_a, t_c), then:
$\exists t'' \in T$: meets(t_c, t''), meets(t'', t_a).
However,
meets(t_a, t'), meets(t', t_b), meets(t_b, t_2), meets(t_2, t''), meets(t'', t_a)

form a cycle: t_a, t', t_b, t_2, t'', t_a, which shows inconsistency (see Fig4.3.2).

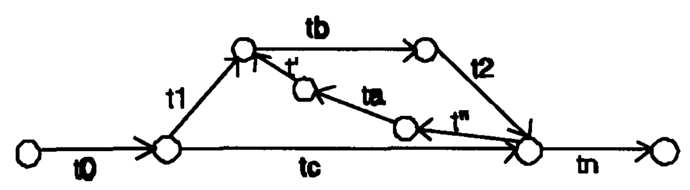

Fig4.3.2

Similarly, for other cases:

(7) If Met-by(t_a, t_c), then meets(t_c, t_a), so that there is a cycle: t_a, t', t_b, t_2, t_a, which shows inconsistency.

(8) If Overlapped-by(t_a, t_c), then $\exists t_0'$, t_3, t_4, t_5, $t_n' \in T$ such that:
meets(t_0', t_c), meets(t_c, t_5), meets(t_5, t_n), meets(t_0', t_3), meets(t_3, t_a), meets(t_a, t_n), meets(t_3, t_4), meets(t_4, t_5),
which forms a cycle: t_5, t', t_b, t_2, t_5, and shows inconsistency.

(9) If Started-by(t_a, t_c), then $\exists t_0'$, t_3, $t_n' \in T$ such that:
meets(t_0', t_c), meets(t_c, t_3), meets(t_3, t_n'),
meets(t_0', t_a), meets(t_a, t_n'),
which forms a cycle: t_3, t', t_b, t_2, t_3, and shows inconsistency.

(10) If Contains(t_a, t_c), then $\exists t_0'$, t_3, t_4, $t_n' \in T$ such that:
meets(t_0', t_3), meets(t_3, t_c), meets(t_c, t_4), meets(t_4, t_n),
meets(t_0', t_a), meets(t_a, t_n'),
which forms a cycle: t_4, t', t_b, t_2, t_4, and shows inconsistency.

(11) If Finishes(t_a, t_c), then $\exists t_0'$, t_3, $t_n' \in T$ such that:
meets(t_0', t_3), meets(t_3, t_a), meets(t_a, t_n'),
meets(t_0', t_c), meets(t_c, t_n'),
which forms a cycle: t', t_b, t_2, t', and shows inconsistency.

(12) If Finished-by(t_a, t_c), then $\exists t_0'$, t_3, $t_n' \in T$ such that:
meets(t_0', t_3), meets(t_3, t_c), meets(t_c, t_n'),
meets(t_0', t_a), meets(t_a, t_n'),
which forms a cycle: t', t_b, t_2, t', and shows inconsistency.

All the entries of Allen's transitivity table have been checked in the above way. From our assumption we know that a point does not meet or be met-by another point, and from our axioms we have proved that a point will not overlap or be overlapped-by other interval or point. Hence, we can extend Allen's system to include time points (this will overcomes the problems involved in the need to model 'open' and 'closed' intervals, by allowing knowledge of interval end-points to be expressed explicitly) and prove (in terms of the condition of acyclicity of ‘meets’) that the corresponding transitivity of the extended model is just as same as that one of Allen.

5. Conclusions

In this paper, a temporal system has been outlined which uses intervals and points as primitives. A significant feature of this approach is that, in common with Allen’s system, intervals do not need to be defined as point pairs. It has been shown here that specialisation of the system to intervals leads to Allen’s system. The model allows reasoning on the primitive elements, and provides an extension of Allen’s system that includes duration. If the system is limited to points, then Bruce’s system, and Dechter, Meiri and Pearl’s STP, are also shown to be special cases.

The axioms given in section 2 imply a linear time ordering, but it is intended to extend them to include branching time, and to include TCSP of Dechter, Meiri and Pearl. The formulation of axioms by means of a single relation allows a graphical representation of the temporal database entities, and this in turn allows an efficient consistency checker in terms of a solver for an LP corresponding to the duration assignment. As a limited case, when no duration is assigned to time elements, the consistency checker is just a search for graphical cycles. It has been proved that Allen’s and Vilain’s truth-propagation inference may be derived.

6. References.

[1] Allen J. F.
"Maintaining Knowledge about Temporal Intervals", communication of ACM., Nov. 1983 Vol.26, 123-154

[2] Allen J. F.
"Towards a General Theory of Action and Time", Artificial Intelligence, 23 (2), July 1984, 123-154

[3] Allen J. F. and Hayes P. J.
"Moments and Points in an Interval-based Temporal-based Logic", Comput. Intell. (Canada), Vol.5, no.4, 225-238 (Nov. 1989)

[4] Bruce B. C.
"A Model for Temporal References and Application in a Question Answering Program", Artificial Intelligence, 3(1972), 1-25

[5] Carre B.
Graphs and Networks, rendon Press, Oxford

[6] Dean T. L. and Mcdermott D. V.
"Temporal Data Base Management", Artificial Intelligence, 32(1987), 1-55

[7] Dechter R. Meiri I. and Pearl J.

"Temporal Constraint Networks", Artificial Intelligence, 49(1991), 61-95

[8] Diller A.
Z--AN Introduction to Formal Methods, John Wiley & Sons Ltd., Baffins Lane, 1990

[9] Halpern J. Y. and Shoham Y.
"A Propositional Model Logic of Time Intervals", Journal of the Association for Computing Machinery, Vol. 38. No. 4. Octber 1991. pp. 935-962

[10] Hayes P.
"The Naive Physics manifesto", Expert systems in the microelectronic age, Michie D. ed., Edinburgh, 1978

[11] Kahn K. M. and Gorry A. G.
"Mechanizing Temporal Knowledge", Artificial Intelligence, 9(1977), 87-108

[12] Lipschutz S.
Theory and Problems of Set Theory and Related Topics, Schaum Publishing Co., 257 Park Avenue South, New York.

[13] McDermott D. V.
"A Temporal Logic for Reasoning about Processes and Plans", Cog. Sci., 2(3), 1982

[14] Russell B.
Principles of Mathematics, London: George Allen and Unwin, 1903.

[15] Vilain M. V.
"A System for Reasoning about Time", Proc. AAAI-82, Pittsburgh, PA. 1982, pp.197-201

[16] Vilain M. B. and Kautz H.
"Constraint Propagation Algorithms for Temporal Reasoning", Pro. AAAI-86, (1986).

Appendix C

A General Temporal Theory

(*the Coputer Journal*, 37(2), pp.114-123, 1994)

Abstract

In this paper, a first-order theory of time is proposed as an underlying framework for most of the representative temporal models in artificial intelligence. The theory treats both points and intervals as primitive on an equal footing, and is shown to be powerful enough to subsume the interval based theories of Allen and Hayes, the point based theories of Bruce, of McDermott, and the interval & point based theories of Vilain and of Knight and Ma. The approach is different from that of Ladkin, of Van Beek, of Dechter, Meiri and Pearl, and of Maiocchi, which is either to construct intervals out of points, or to treat points and intervals separately. Formal definitions are presented to characterise the *open* and *closed* nature of primitive intervals. The axiomatisation allows non-linear time structures such as *branching time* and *parallel time*. Additional axioms specifying the *linearity* and *density* of time are separately presented.

Key words: time, theory, temporal systems, intervals, points.

1. Introduction

The essential role of time in the modelling of natural processes has given rise in recent years to a body of artificial intelligence research into temporal theory. This research has led to a variety of temporal systems, attempting to capture the primary elements of time, for application to the modelling of human activities, such as problem solving, natural language understanding, planning and the qualitative modelling of physical processes. Amongst the temporal systems which have been proposed, many have been based on axioms stated in first order predicate logic. Although these systems show considerable commonality in structure, they also show considerable differences in formalisation.

Probably the most important structural variation between first order theories is in their treatment of time intervals. Many theories, such as those of Bruce [3], of Ladkin [13,14], of Dechter et al. [17], and of Maiocchi [18], are based on points as the basic primitive element. In these theories, intervals are defined in terms of points, usually by means of beginning and ending points. However, as Allen has commented [7], modelling intervals by taking their bounding-points can lead to problems: the annoying question of whether bounding-points are in the interval or not must be addressed, seemingly without any satisfactory solution. If intervals are all closed then adjacent intervals have bounding-points in common, which when adjacent intervals correspond to states of truth and falsehood of some property, can lead to situations in which a property is both true and false at an instant. Similarly, if intervals are all open, there will be points at which the truth or falsity of a property will be undefined. The solution in which intervals are all taken as semi-open (e.g., see the definition of intervals in Maiocchi's *TSOS* [18]), so that they sit conveniently next to one another, seems arbitrary and unsatisfactory. Other theories, predominantly that of James Allen [7,8], treat intervals as primitive, and points are relegated to a subsidiary status as "meeting places" of intervals. Other theories again, e.g. that of Vilain [11], and that of Knight and Ma [4], treat both intervals and points as primitive on an equal footing.

In van Beek's temporal framework [15,16], both time intervals and time points are addressed by means of the interval algebra, *IA*, and point algebra, *PA*, respectively. However, it is interesting to note that, in this framework, *IA* and *PA*, deal with temporal relations between intervals, and relations between points separately, that is, the interval-based framework *IA* deals with the thirteen temporal relations (defined by Allen [7]) between intervals only, while the point-based framework *PA* deals with temporal relations between points only, which are addressed in Vilain and Kautz's point algebra [12]. Relations between intervals and points, such as that proposed by Vilain in [11], are not addressed at all. Again, as in Dechter et al.'s framework, time intervals are not defined as primitive. Indeed, time intervals, and temporal relations between intervals are defined in terms of points (rationals), and the corresponding order relations between points, respectively.

Theories also differ in their treatment of the two important issues of linearity and density of time elements. Most systems assume that time is linear, i.e., that all time elements are ordered along a single time line. However, non-linear temporal structures are also proposed in some systems. For instance, McDermott's temporal logic allows time to branch into the future. In Bruce's system, branching or linear time may also be specified by the user, although the axioms required for such a specification are not given.

The question of the density of time elements depends on the type of primitives assumed for the system. For interval based systems, a dense system is taken to be one where every interval is (infinitely) decomposable. For point based systems, a dense system is one in which between any two points on the same time line, there is a third. Bruce's proposed system leaves the density question open, whereas McDermott's system is decidedly dense. Knight and Ma's system is decidedly not dense, being a discrete system: i.e., one where every time element has a unique predecessor and successor fundamental time element. Other systems, such as that of Allen and Hayes, permit a mixture of dense and discrete time elements.

Finally, there is a difference between systems in their ability to model the "open" and "closed" nature of intervals. Allen's system allows only intervals of indeterminate type: since points are not allowed, there is no definition of open\closed intervals. In Vilain's system, although both points and intervals are taken as primitive, it is still not possible to characterise the open and closed nature of intervals. However, Knight and Ma's temporal model allows modelling of open and closed intervals, and it can be shown that the characterisation is in agreement with the conventional concepts of open, semi-open, and closed intervals which are constructed out of points.

The importance of treating points and intervals as primitives on an equal footing lies in the need for the temporal theory to model the way things happen in time. Both Allen [7,8] and McDermott [5] give examples of properties defined over time, and many AI applications involve continuous change of variables in time. Galton [1] has shown that time-points are needed in order to accommodate the representation of facts concerned with continuous change, and has proposed a revision of Allen's system to this effect.

It is the objective of this paper to provide a general axiomatic framework to serve as a unifying basis for these temporal systems. The axiomatisation may be seen as an extension of Allen and Hayes' theory [10], to include points as primitive objects. A discussion of the implications of including points as primitive, and of distinguishing points from moments is given in section 2. Here, a problem with Allen's interval based logic concerning reasoning about continuous change is examined. The discussion indicates that points are necessary as primitive objects for the correct modelling of continuous change. There follows a discussion of some limitations of Allen and Hayes' axiom, <M6>, which states that moments never meet moments. It is shown that this axiom leads to the conclusion that we can have neither a completely discrete nor a completely dense system which contains moments. However, if we revise Allen's and Hayes' system to include points and limit <M6> to points, rather than moments, this objection does not apply.

We present the main body of the general axiomatisation for a temporal frame based on both interval and points in section 3. These axioms are independent of the specification of density and linearity. Additional axioms are provided in section 4 to specify the linearity and density of time. Definitions are also given for the open and closed nature of an interval. A classification of all possible temporal relations over intervals and points is presented in section 5. In section 6 we give various models to illustrate the theory. We present a completely dense model and a completely discrete model of the theory. We further show how other temporal systems may be subsumed by the theory, with the appropriate denseness and linearity axioms. It is also shown that, assertions about the instantiation in time of properties and occurrences may be naturally hanged on the temporal frame.

2. Allen and Hayes' Axiomatisation of Time based on Intervals

Allen and Hayes' theory of time is based on a nonempty class, **I**, of *time intervals*, and is axiomatised in terms of the single temporal relation "*meets*" between intervals. The set of axioms is proposed first in [9], and then revised in [10], as follows:

<M1> $\forall i,j,k,l \in \mathbf{I}(meets(i,j) \wedge meets(i,k) \wedge meets(l,j) \Rightarrow meets(l,k))$

<M2> $\forall i,j,k,l \in \mathbf{I}(meets(i,j) \wedge meets(k,l) \Rightarrow$
$meets(i,l)$
$\nabla\ \exists m \in I(meets(i,m) \wedge meets(m,l))$
$\nabla\ \exists n \in I(meets(k,n) \wedge meets(n,j)))$

N.B. In this paper, "∇" means exclusive disjunction.

<M3> $\forall i \in \mathbf{I} \exists j,k \in \mathbf{I}(meets(j,i) \wedge meets(i,k))$

<M4> $\forall j,k \in \mathbf{I}(\exists i,l \in \mathbf{I}(meets(i,j) \wedge meets(j,l) \wedge meets(i,k) \wedge meets(k,l)) \Rightarrow j = k)$

N.B. In this paper, we follow Allen and Hayes' notation that "j = k" means j and k represent the same time element.

<M5> $\forall i,j \in \mathbf{I}(meets(i,j) \Rightarrow$
$\exists k \in \mathbf{I} \forall m,n \in \mathbf{I}(meets(m,i) \wedge meets(j,n) \Rightarrow meets(m,k) \wedge meets(k,n))$

Axiom <M1> states that the "place" where two intervals meet is unique and closely associated with the intervals. The role of <M2> is to ensure that meeting places are totally ordered. <M3> makes every interval have at least one neighbouring interval preceding it, and another succeeding. <M4> simply says that there is only one time interval between any two meeting places. Finally, <M5> states that if two meeting places are separated by a sequence of intervals, then there is an interval which connects these two meeting places. Hence, with axiom <M4> and the definition of equality, for any two adjacent intervals, i and j, the ordered union of i and j may be written as i + j.

A limitation of Allen and Hayes' theory, expressed by Tsang [6], is that the axioms are not primitive enough for extensions. For example, linearity might be hoped to be removed from the axiomatisation in order to address the issues such as **branching time** and **parallel time**. However, Tsang points out that it is difficult to see which axiom in Allen and Hayes' axiom set entails linearity. Allen and Hayes conclude that the linearity assumption is characterised by means of axiom <M4> in the revised version of the set of their axioms [10]. However, it is indeed axiom <M2>, rather than <M4>, that entails the linearity of time. In fact, if we remove <M2> from the set of axioms, then the time may be **circular**, **parallel**, **branching**, as shown in Figure 1. In this graphical representation, the arcs of the graph represent time intervals, and the relation *meets*(i,j) is represented by i being in-arc and j being out-arc to a common node:

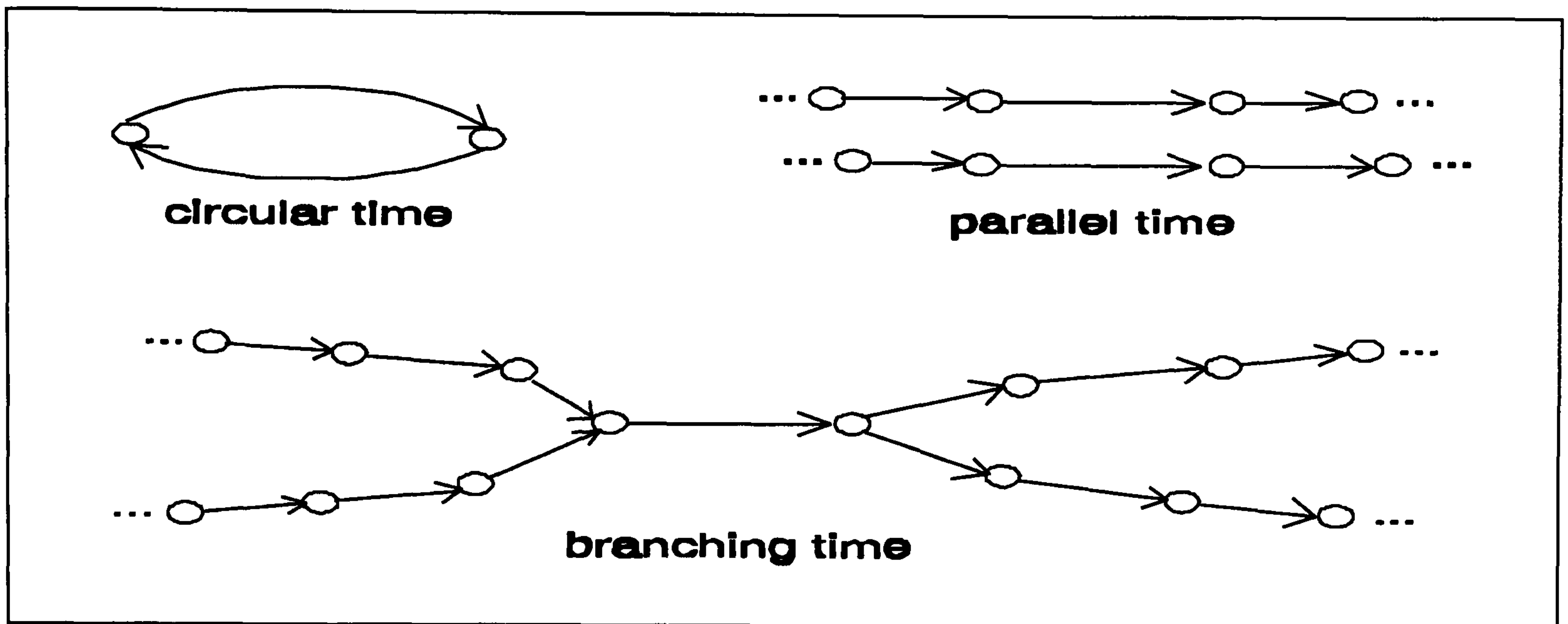

Figure 1

Another limitation of Allen and Hayes' time theory is that it takes only intervals, rather than points, as primitive time elements, although points are later introduced as the "meeting places" of intervals at a subsidiary status within the theory. Their contention is that nothing can be true at a point, for a point is not an entity at which things happen or are true. However, as Galton shows in his critical examination of Allen's interval logic [1], the theory of time based on intervals is not adequate, as it stands, for reasoning correctly about continuous change. We may illuminate the problem involved with reference to time points by means of the following example of a ball thrown vertically into the air: The motion may be described qualitatively by the use of two intervals, interval i where the ball is going up, and interval j where the ball is coming down. According to classical physics, there is a point p at which the ball is stationary. As Allen suggested, in the interval calculus, we have two alternatives: we may assume that there is a very small interval where the ball is stationary, or we may assume that interval i "*meets*" interval j. The first alternative does not seem tenable, being inconsistent with the laws of physics, no matter how small the interval. The second alternative also gives problems, since the interval calculus allows us to combine two intervals which meet, that is, i + j = k (see [8] and [10]): in Allen's logic, the formula HOLDS(*pro*, I) is used to say that the property *pro* holds <u>during</u> the interval I. More precisely, what it says is that *pro* holds <u>throughout</u> that interval [1]. However, although the property "*ball_in_motion*" holds throughout both of intervals i and j, that is:

HOLDS(*ball_in_motion*,i),
HOLDS(*ball_in_motion*,j)

we cannot assert that

HOLDS(*ball_in_motion*,i+j),

since the property "*ball_in_motion*" does not hold throughout the whole combined interval k, within which there is a point p at which the ball is stationary.

To characterise the times that some "instant-like" events occupy, Allen and Hayes introduce the idea of *very short intervals*, called *moments*. A moment is simply a non-decomposable time interval. The important distinction between moments and points is: although being non-decomposable, moments are defined by having extent and by means of having distinct beginning and end points (just as for other intervals [10]), while points are defined by having no extent.

Relating to the meets relation, another obvious difference between points and moments is that moments can meet other intervals, and hence stand between them, while points are not treated as primitive objects and cannot meet anything. However, as Allen and Hayes themselves point out, a theory incorporating granularity involves introducing a "*tolerance relation*" that defines when two times are indistinguishable. For example, two intervals, i and j, might be indistinguishable if their beginning points are at most a moment apart, and likewise for their end points. To ensure that the tolerance relation is an equivalence relation, Allen and Hayes propose axiom <M6>, which insists that moments never meet:

<M6> $\forall m,n \in \mathbf{I}(moment(m) \wedge moment(n) \Rightarrow \neg meets(m,n))$

where *moment*(m) is defined by:

$$\forall m \in \mathbf{I}(moment(m) \Leftrightarrow \neg\exists i,j \in \mathbf{I}(m = i + j))$$

Allen and Hayes declare that their formulation permits either discrete or continuous time models, as well as more exotic models that may alternate between continuous and discrete stretches of time. Unfortunately, axiom <M6> leads to another limitation to the primitive time elements: for any interval, either it is non-decomposable, that is, a moment, or it must be infinitely decomposable. For, if it is only finitely decomposable, then it must be the sum of a finite number of moments which would meet one another, contrary to <M6>. This precludes discrete models from the theory containing axiom <M6>. In addition, dense models of the theory, i.e where all intervals are infinitely decomposable, permit no moments at all, so that <M6> is only vacuously true. Hence models of the theory including <M6> which contain moments can be neither dense nor discrete.

However, although <M6> appears to bring little benefit in the form that is presented here, dealing with moments, it is shown in the next section to play a critical role in a general theory if it is applied to "*time points*". In this case the axiom does not limit the interval structure at all.

3. An Axiomatisation of Time based on Intervals and Points

As discussed in the above section, Allen and Hayes' time theory is not primitive enough for extensions, and is not adequate for reasoning correctly about continuous change. Our objective is to develop and explore a first-order theory of time which should be more general as an underlying framework for most of representative temporal models in artificial intelligence. The new time theory may be seen as an extension of Allen and Hayes' axiomatisation by means of some additional axioms relating to the inclusion of time points as primitive elements, and generalisation of Allen and Hayes' axiomatisation by removing the linearity of time in order to allow non-linear time structures such as branching time, parallel time, etc.

We start the formal theory by posing a nonempty set, **T**, of objects that we shall call **time-elements**, and a function d from **T** to $\mathbf{R}_0^+$, the set of non-negative real numbers. A time-element, t, is called a (time) interval if $d(t) > 0$, otherwise, t is called a (time) point. According to this classification, the set of time-elements, **T**, may be expressed as $\mathbf{T} = \mathbf{I} \cup \mathbf{P}$, where **I** is the set of intervals, and **P** is the set of points. As in Allen and Hayes' approach, at this early stage we do not make any commitment as to whether all time intervals are decomposable or not. The density question will be addressed by some further axioms.

In order to define the primitive order over time elements, we adopt Allen and Hayes' axiomatisation for the single relation "*meets*" between intervals while axiom <M2> will not be included in the first place. Since the time elements may now be not only intervals but also points, some critical axioms are necessary relating to the treatment of points. The whole set of axioms for the "*meets*" relation over **T** are listed below, where axioms <A1>, <A2>, <A3> and <A4> correspond to Allen and Hayes' <M1>, <M3>, <M4> and <M5> in the above section, respectively:

<A1> $\forall t_1,t_2,t_3,t_4 \in \mathbf{T}(meets(t_1,t_2) \wedge meets(t_1,t_3) \wedge meets(t_4,t_2) \Rightarrow meets(t_4,t_3))$

<A2> $\forall t \in \mathbf{T} \exists t',t'' \in \mathbf{T}(meets(t',t) \wedge meets(t,t''))$

<A3> $\forall t_1,t_2 \in \mathbf{T}(\exists t',t'' \in \#\#\mathbf{T}\#\#(\quad meets(t',t_1) \wedge meets(t_1,t'')$
$\wedge\, meets(t',t_2) \wedge meets(t_2,t'')) \Rightarrow t_1 = t_2)$

<A4> $\forall t_1,t_2 \in \mathbf{T}(meets(t_1,t_2) \Rightarrow$
$\exists t \in \mathbf{T} \forall t',t'' \in \mathbf{T}(\quad meets(t',t_1) \wedge meets(t_2,t'')$
$\Rightarrow meets(t',t) \wedge meets(t,t''))$

N.B. For any two time elements, t_1 and t_2, such that $meets(t_1,t_2)$, axioms <A4> and <A3> ensure that there is a unique time element corresponding to the ordered union of t_1 and t_2. Following Allen and Hayes' notation, we shall still indicate it as i + j, which will always imply that $meets(i,j)$.

<A5> $\forall t_1,t_2 \in \mathbf{T}(\mathit{meets}(t_1,t_2) \Rightarrow t_1 \in \mathbf{I} \vee t_2 \in \mathbf{I})$

<A6> $\forall t_1,t_2 \in \mathbf{T}(\mathit{meets}(t_1,t_2) \Rightarrow d(t_1 + t_2) = d(t_1) + d(t_2))$

Axiom <A5> is based on the intuition that points will not meet other points, that is, between any two time points, there is a time interval. This is indeed very similar to Allen and Hayes' <M6> which states that moments never meet other moments. However, unlike <M6>, <A5> does not imply the limitation that any decomposable interval must be **infinitely** decomposable. Additionally, axiom <A5> does not affect whether the set of points is dense or not. This issue will be depend on a further assumption ensuring that "within" any time interval, there is a time point (see section 6). Axiom <A6> ensures that the addition operation, "+", over time elements is consistent with the function d, which we shall call the **duration assignment** over **T**.

This is the complete fundamental set of axioms concerning the *meet* relation. We denote this set as **A**, and use a pair, (**T**,*meet*), to represent the temporal frame defined by the axiomatisation.

4. Some Further Issues

The axiomatisation proposed in the above section defines a general temporal frame based on both intervals and points as primitive objects. In this section, we address some further issues relating to the structure of the frame.

Open and Closed Nature of Intervals: Although intervals are taken in the theory as primitive, that is there are no definitions about the ending-points for intervals, the axiomatisation allows the expression of the "open" and "closed" nature of intervals. For example, to represent the quantity space for the motion of the ball described in section 2, we may relate *ball_going_up*, *ball_stationary*, and *ball_coming_down* to interval i_1, point p, and interval i_2, respectively, where $\mathit{meets}(i_1,p)$, $\mathit{meets}(p,i_2)$. Intuitively, $t = p + i_2$ relates to *ball_stationary-ball_coming_down*. In Figure 2 (for clarity, we denote points with bold arcs), since i_1 has point p as its immediate successor, we may view i_1 as "right-open" at p, and similarly, i_2 as "left-open" at p. Since interval t ($= p + i_2$) and point p have the same immediate predecessor, i_1, we may view t as "left-closed" at p.

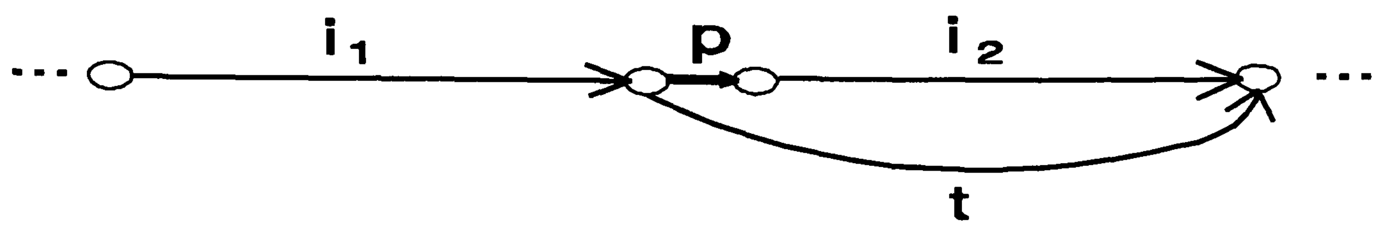

Figure 2

Formally, the open and closed nature of primitive intervals may be defined as follows:

interval i is **left-open** at point p iff
$meets(p, i)$;

interval i is **right-open** at point p iff
$meets(i,p)$;

interval i is **left-closed** at point p iff
$\exists i' \in \mathbf{I}(meets(i',i) \wedge meets(i',p))$;

interval i is **right-closed** at point p iff
$\exists i' \in \mathbf{I}(meets(i,i') \wedge meets(p,i'))$.

It is easy to see that "left-open" and "left-closed" (symmetrically, "right-open" and "right-closed") are exclusive to each other under the axiomatisation. In fact, if interval i is left-open at point p_1, and left-closed at point p_2, then by the above definition, we get:

$meets(p_1,i) \wedge meets(i',i) \wedge meets(i',p_2)$, where $i' \in \mathbf{I}$

Hence, by axiom <A1> we can infer that $meets(p_1,p_2)$, which is contradictory to axiom <A5>.

The above interpretation of the "open" and "closed" nature of primitive intervals is in fact in line with the conventional meaning of the open and closed nature for point-based intervals. For instance, point-based interval $(p_1, p_2]$ is "left-open" at point p_1 , since intuitively p_1 is an immediate predecessor of interval $(p_1, p_2]$; similarly, $(p_1, p_2]$ is "right-closed" at p_2, since both point p_2 and interval $(p_1, p_2]$ have the same immediate successor, $(p_2, \_\}$.

<u>Linearity of Time</u>: Time is usually considered as having a *linear* structure. This corresponds to the classical physical model of time, where the structure is that of the real line, extending infinitely in both directions.

The **(full) linearity** of a temporal frame (**T**,*meets*) can be characterised by adding an axiom, <A_{Linear}>, to **A**, the set of axioms proposed in section 3:

<A_{Linear}>

$\forall t_1,t_2,t_3,t_4 \in \mathbf{T}(\mathit{meets}(t_1,t_2) \wedge \mathit{meets}(t_3,t_4) \Rightarrow$
$\quad \mathit{meets}(t_1,t_4)$
$\quad \triangledown\ \exists t' \in \mathbf{T}(\mathit{meets}(t_1,t') \wedge \mathit{meets}(t',t_4))$
$\quad \triangledown\ \exists t'' \in \mathbf{T}(\mathit{meets}(t_3,t'') \wedge \mathit{meets}(t'',t_2)))$

N.B. The axiom <A_{Linear}> is in fact the axiom <M2> (see section 2) for Allen and Hayes' interval-based theory. The "exclusive ors" in this axiom have some quite powerful consequences. In particular, they ensure that there can be no **circular**, **parallel**, and **branching** times. The following lemma is straightforward (see [10]):

<Lemma[1]> $\quad \forall t \in \mathbf{T}(\neg \mathit{meets}(t,t))$

This lemma ensures that there is no possibility of circular time.

However, without <A_{Linear}>, a temporal frame usually allows branching into both the past and the future. Branching temporal frames offer an attractive way to handle possible worlds, uncertainty about the past or the future and the effects of alternative actions when planning. A temporal frame which allows branching into the future but not into the past is called **left-linear** (see Figure 3). This may be characterised by adding to **A**, the axiom <$A_{\mathrm{L\text{-}Linear}}$>, rather than the stronger axiom <A_{Linear}>:

<$A_{\mathrm{L\text{-}linear}}$>

$\forall t_1,t_2,t_3,t_4,t \in \mathbf{T}(\mathit{meets}(t_1,t_2) \wedge \mathit{meets}(t_2,t) \wedge \mathit{meets}(t_3,t_4) \wedge \mathit{meets}(t_4,t) \Rightarrow$
$\quad \mathit{meets}(t_1,t_4)$
$\quad \triangledown\ \exists t' \in \mathbf{T}(\mathit{meets}(t_1,t') \wedge \mathit{meets}(t',t_4))$
$\quad \triangledown\ \exists t'' \in \mathbf{T}(\mathit{meets}(t_3,t'') \wedge \mathit{meets}(t'',t_2)))$

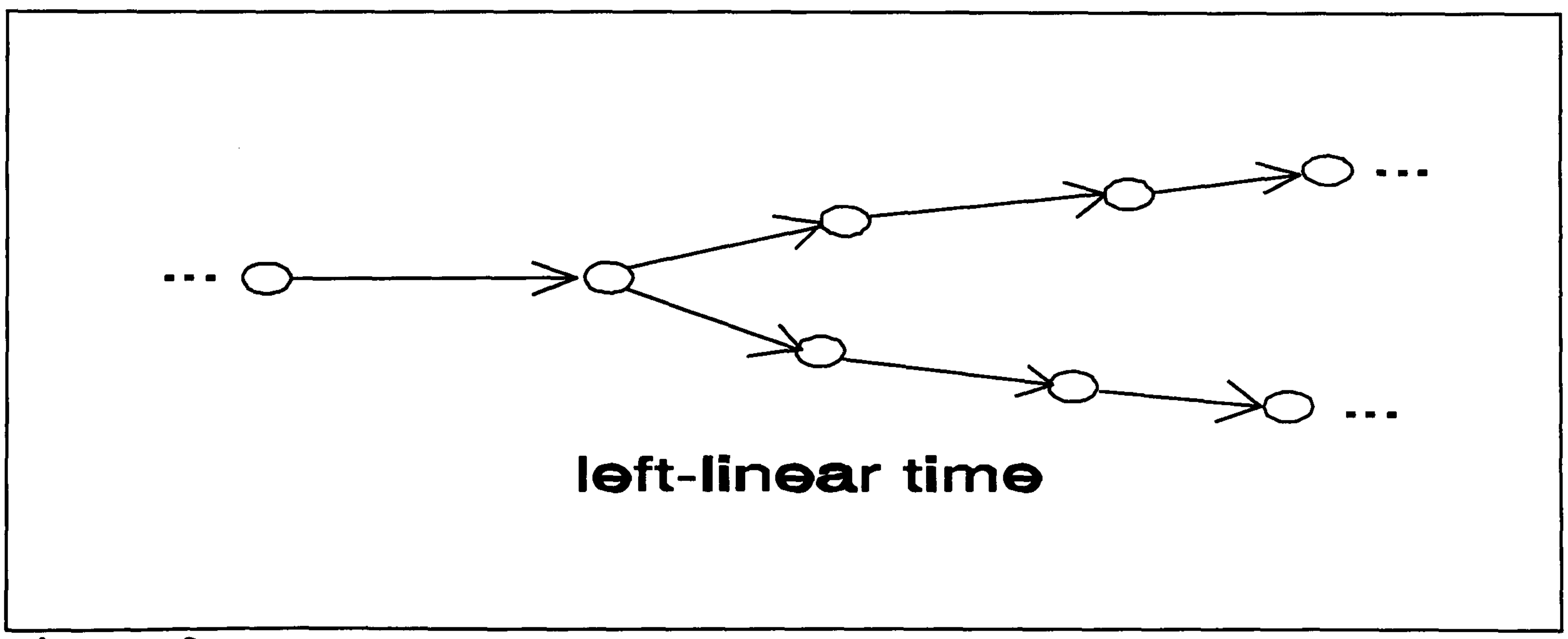

Figure 3

Analogously, **right-linearity** is defined by means axiom $<A_{R\text{-}Linear}>$:

$<A_{R\text{-}Linear}>$

$$\begin{aligned}
&\forall t,t_1,t_2,t_3,t_4 \in \mathbf{T}(\mathit{meets}(t,t_1) \wedge \mathit{meets}(t_1,t_2) \wedge \mathit{meets}(t,t_3) \wedge \mathit{meets}(t_3,t_4) \Rightarrow \\
&\qquad \mathit{meets}(t_1,t_4) \\
&\qquad \vee \exists t' \in \mathbf{T}(\mathit{meets}(t_1,t') \wedge \mathit{meets}(t',t_4)) \\
&\qquad \vee \exists t'' \in \mathbf{T}(\mathit{meets}(t_3,t'') \wedge \mathit{meets}(t'',t_2)))
\end{aligned}$$

As Galton puts it in [2], it is interesting to note that **left-linearity** and **right-linearity** together just fail to imply (**full**) **linearity**, the exception being the case of parallel time lines as shown in Figure 4.

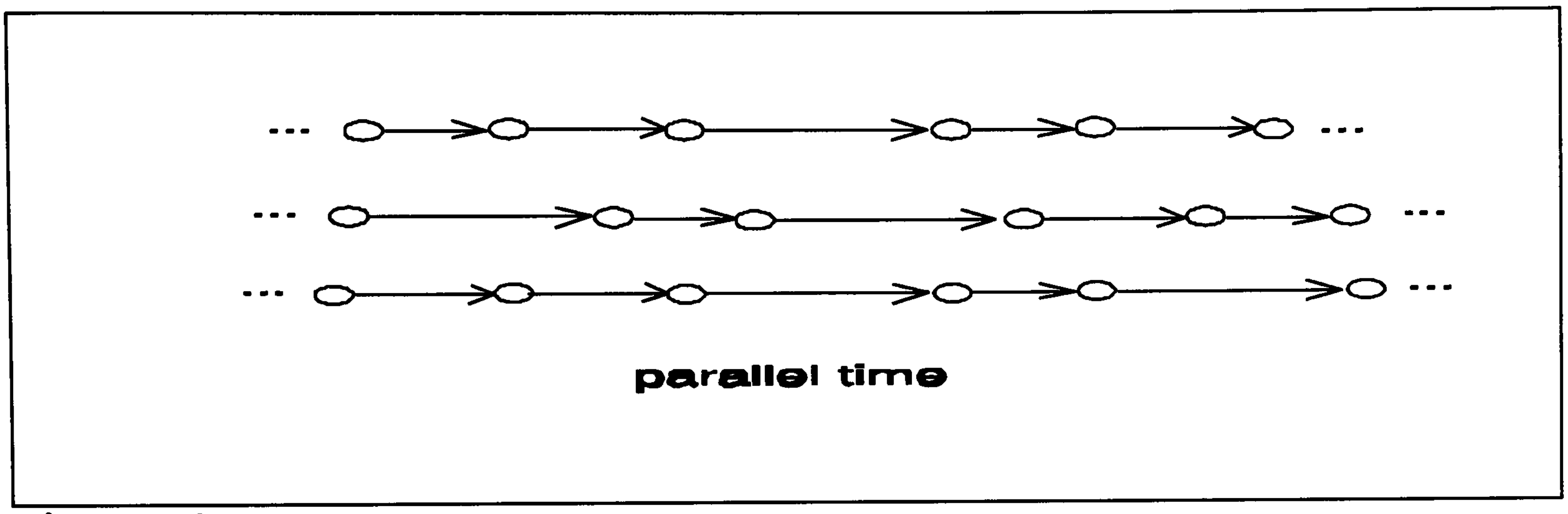

Figure 4

Parallel temporal frames provide a way of modelling separate and asynchronous processes, and might prove useful in developing logics for reasoning about parallel computation and concurrent processes.

Dense and Discrete Time: According to Axiom <A2>, for each time-element t, there is a time-element which "*meets*" it, and another one which it "*meets*". Therefore, in particular, axiom <A4> and <A5> additionally ensure that, between any two distinct time points on the same time line, there is always a time interval. However, for time intervals, can we always assume that any interval can be decomposed into two distinct contiguous intervals? If so, we say that the set of time elements forms a dense system.

We may use the following axiom to characterise the density of a temporal frame (**T**,*meets*):

$<A_{Dense}>$

$\forall i \in \mathbf{I} \exists t_1,t_2 \in \mathbf{T}(i = t_1 + t_2)$

We can show that axiom $<A_{Dense}>$ implies that each time interval can be decomposed into two distinct contiguous intervals. In fact, assume interval $i = t_1 + t_2$; if t_1 is a point, then by axiom <A5>, t_2 must be an interval; hence, by $<A_{Dense}>$, $t_2 = t' + t''$, where $t', t'' \in \mathbf{T}$. By <A4> and <A3>, we get $i = t_1 + t' + t''$. Since t_1 is a point, axiom <A5> implies that t' must be an interval; hence $i_1 = t_1 + t'$ is an interval, and $i = i_1 + t_2$. Similar discussion applies to the case that t_2 is a point which implies that t_1 must be an interval.

The **discreteness** of a temporal frame (**T**,*meets*) can be characterised by means of adding two axioms, <$A_{L\text{-Discrete}}$> and <$A_{R\text{-Discrete}}$> to **A**:

<$A_{L\text{-Discrete}}$>
$\forall t \in \mathbf{T} \exists t_1 \in \mathbf{T}(meets(t_1,t) \wedge \neg\exists t_2,t_3 \in \mathbf{T}(t_1 = t_2 + t_3))$

<$A_{R\text{-Discrete}}$>
$\forall t \in \mathbf{T} \exists t_1 \in \mathbf{T}(meets(t,t_1) \wedge \neg\exists t_2,t_3 \in \mathbf{T}(t_1 = t_2 + t_3))$

Axiom <$A_{L\text{-Discrete}}$> entails the **left-discreteness**, and Axiom <$A_{R\text{-Discrete}}$> entails the **right-discreteness** of a temporal frame. By taking t to be a non-decomposable interval (or moment, termed by Allen and Hayes) in the above axioms, since t_1 is by definition a moment, we see that <$A_{L\text{-Discrete}}$> or <$A_{R\text{-Discrete}}$> implies that each moment has a predecessor moment or successor moment respectively. Hence, Allen and Hayes' <M6> is inconsistent with the discreteness axioms.

It is interesting to note that there may exist temporal frames which are neither dense, nor discrete. In such a frame, there may be some intervals which are finite sums of moments. However, this case is axiomatically consistent with our axiom <A5>, but not consistent with Allen and Hayes' <M6>, which implies that each decomposable interval must be infinitely decomposable.

5. Derived Temporal Relations over Time Elements

In terms of the primitive relation "*meets*", we may induce the complete set of possible relationships over time elements by means of the following definitions:

$EQUAL(t_1,t_2) \Leftrightarrow t_1 = t_2$,

$BEFORE(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(meets(t_1,t) \wedge meets(t,t_2))$,

$OVERLAPS(t_1,t_2) \Leftrightarrow \exists t,t',t'' \in \mathbf{T}(t_1 = t' + t \wedge t_2 = t + t'')$,

$START(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(t_2 = t_1 + t)$,

$DURING(t_1,t_2) \Leftrightarrow \exists t',t'' \in \mathbf{T}(t_2 = t' + t_1 + t'')$,

$FINISHES(t_1,t_2) \Leftrightarrow \exists t \in \mathbf{T}(t_2 = t + t_1)$,

$MEETS(t_1,t_2) \Leftrightarrow meets(t_1,t_2)$,

$AFTER(t_1,t_2) \Leftrightarrow BEFORE(t_2,t_1)$,

$OVERLAPPED\text{-}BY(t_1,t_2) \Leftrightarrow OVERLAPS(t_2,t_1)$,

$STARTED\text{-}BY(t_1,t_2) \Leftrightarrow STARTS(t_2,t_1)$,

$CONTAINS(t_1,t_2) \Leftrightarrow DURING(t_2,t_1)$,

$FINISHED\text{-}BY(t_1,t_2) \Leftrightarrow FINISHES(t_2,t_1)$,

$MET\text{-}BY(t_1,t_2) \Leftrightarrow MEETS(t_2,t_1)$,

> N.B. Since points are now allowed, the above 13 relations have somewhat different "feel" to Allen's 13 temporal relations between intervals. For instance, if i_1 and i_2 are open intervals separated by a point p, then we have $BEFORE(i_1,i_2)$, although this situation looks very like i_1 "*meets*" i_2 in Allen's system. Again, if i_1 is right-closed, and i_2 is left-closed at point p, respectively, according to the above definitions, we have $OVERLAPS(i_1,i_2)$, but again it "looks" like the two intervals meeting. Additionally, from the above definitions, any open interval is "*DURING*" its closure. What all this means is that, taking both intervals and points as primitive time-elements, we have more than 13 significantly different relationships to considered, because, for example, from almost any point of view, the first case mentioned above (i.e., $MEETS(i_1,p) \wedge MEETS(p,i_2)$) is no more similar to the case of two intervals separated by a third <u>interval</u> (a necessary condition of *BEFORE* in Allen's system) than it is to the case of two intervals strictly meeting.

As Allen and Hayes show in [10], all the thirteen relations may hold in the case that only intervals are taken as time elements. However, when we examine the general case where elements may also be points, some of these relationships hold and some do not hold.

For example, let $p \in \mathbf{P}$:

$MEETS(p,t_2)$ may hold for time elements $t_2 \in \mathbf{T}$ according to the axiomatisation.

However, consider the following case:

$OVERLAPS(p,t_2) \Leftrightarrow \exists t,t',t'' \in \mathbf{T}(p = t' + t \wedge t_2 = t + t'')$,

On the one hand, by axiom <A6>, $d(p) = d(t') + d(t)$; and the assumption that p is a point gives:

$$d(t') + d(t) = d(p) = 0 \qquad (1)$$

On the other hand, axiom <A5> ensures that at least one of t' and t is an interval, hence:

$$d(t') + d(t) > 0 \qquad (2)$$

(1) and (2) show that $OVERLAPS(p,t_2)$ can not hold.

It is straightforward to prove in a similar fashion that all the possible relations over intervals and points may be classified into the following four groups:

Point - Point:

{*EQUAL, BEFORE, AFTER*}
which relate points to other points;

Interval - Interval:

{*EQUAL, BEFORE, MEETS, OVERLAPS, STARTS, DURING, FINISHES, FINISHED-BY, CONTAINS, STARTED-BY, OVERLAPPED-BY, MET-BY, AFTER*}
which relate intervals to intervals;

Point - Interval:

{*BEFORE, MEETS, STARTS, DURING, FINISHES, MET-BY, AFTER*}
which relate points to intervals;

Interval - Point:

{*BEFORE, MEETS, FINISHED-BY, CONTAINS, STARTED-BY, MET-BY, AFTER*}
which relate intervals to points.

According to the above classification, there are in total 30 possible temporal relations over time-elements which may be both intervals and points. It is interesting to note that, however, in Vilain's interval & point based system [11], only 26 of these 30 temporal relations are addressed. There is a critical omission from the primitive relations between points and intervals in Vilain's system, for the "*MEETS*" relation is defined only between intervals and is not allowed between points and intervals. This omission leads to some difficulties in modelling the "open" and "closed" nature of intervals (see section 4).

6. Models of the Theory

Since the time theory itself characterises a very general temporal structure, we may interpret the axiomatisation in various temporal models: dense or discrete, linear or branching, etc.

As an example of dense and linear models of the axiomatisation, consider the interpretation in which the set of time points, **P**, is the set of all real numbers; and the set of time intervals, **I**, is the set of periods which are constructions over all possible point-pairs, $p_1,p_2 \in \mathbf{P}$ such that $p_1 < p_2$, with the following structures:

$$(p_1,p_2,\text{open},\text{open}) =_{def} \{ r \in \mathbf{R} \mid p_1 < r < p_2 \},$$

$$(p_1,p_2,\text{open},\text{closed}) =_{def} \{ r \in \mathbf{R} \mid p_1 < r \leq p_2 \},$$

$$(p_1,p_2,\text{closed},\text{open}) =_{def} \{ r \in \mathbf{R} \mid p_1 \leq r < p_2 \},$$

$$(p_1,p_2,\text{closed},\text{closed}) =_{def} \{ r \in \mathbf{R} \mid p_1 \leq r \leq p_2 \},$$

where "$<$" and "$\leq$" are the ordinary ordering relations over the set, **R**, of real numbers.

N.B. Here, we represent the interval structure by means of the extra primitives: left-type, l, and right-type, r, which take values from a set **Type** $=_{def}$ {open, closed}. There are thus four types of intervals based on points. For convenience of expression, we may denote a point, p, as (p,p,closed,closed), that is, a special period whose left ending point and right ending point are identical, with "closed" type for both left-type and right-type.

The duration assignment function d is simply defined by:

$$d((p_1,p_2,\_,\_)) = p_2 - p_1.$$

We may define the *meets* relation over $\mathbf{T} = \mathbf{P} \cup \mathbf{I}$ as following:

$$\begin{aligned} &meets((p_{11},p_{12},l_1,r_1), (p_{21},p_{22},l_2,r_2)) \Leftrightarrow \\ &\qquad p_{12} = p_{21} \wedge r_1 = \text{open} \wedge l_2 = \text{closed} \\ &\qquad \vee\ p_{12} = p_{21} \wedge r_1 = \text{closed} \wedge l_2 = \text{open} \end{aligned}$$

It is easy to see that this model satisfies axioms <A1> - <A6>. Additionally, the (full) linearity axiom, <A_{Linear}>, and the dense axiom, <A_{Dense}>, are also satisfied. Hence, the above structure forms a dense and linear temporal model of the theory.

A discrete model satisfy axioms <A1> - <A6>, <A_{Linear}>, <$A_{L\text{-}Discrete}$> and <$A_{R\text{-}Discrete}$> can be constructed by simply limiting all elements of **P** to be integers in the above model, although the internal points of intervals are still reals. It is interesting to note that in such a discrete model, although points never meet each other, intervals are not necessarily infinitely decomposable. For instance, according to our axiomatisation, interval (6,8,open,closed) can be only decomposed into at most four (non-decomposable) time elements:

$$\begin{aligned} &(6,8,\text{open},\text{closed}) = \\ &\qquad (6,7,\text{open},\text{open}) \\ &\qquad + (7,7,\text{closed},\text{closed}) \\ &\qquad + (7,8,\text{open},\text{open}) \\ &\qquad + (8,8,\text{closed},\text{closed}) \end{aligned}$$

However, this model will not be valid for Allen and Hayes' axiomatisation including <M6> (see section 2), which implies that if an interval is decomposable then it must be infinitely decomposable. (Otherwise, if it is only finitely decomposable, then it must be the sum of a finite number of moments which would meet one another, contrary to <M6>.)

N.B. As mentioned in section 2, in order to interpret Allen and Hayes' axioms in discrete models, their axiom <M6> must be excluded. In another word, axiom <M6> is inconsistent with discrete times. However, the above example shows that the axiom <A5> in our axiomatisation can be satisfied by discrete models.

In what follows, we shall show that our axiomatisation is powerful enough to subsume many representative temporal systems, such as: the point based systems of Bruce, of McDermott, Allen's logic of intervals and Galton's revised theory, and the point & interval based theories of Vilain, of Knight and Ma.

Bruce's point based system:

Bruce's *time-system* is simply a set of time points with a partial order (see [3]). In our theory, we may define a partial order, "$\leq$", over the set of points, **P**, as:

$$p_1 \leq p_2 \Leftrightarrow EQUAL(p_1,p_2) \vee BEFORE(p_1,p_2),$$

where *EQUAL* and *BEFORE* are introduced as in section 5. Hence, the sub-frame, $(\mathbf{P},\leq)$, of the temporal frame (T,*meets*) defined by the axiomatisation, forms a temporal system of Bruce.

In a similar way, we may define Bruce's 7 binary relations over *time-segments* (see [3]), in terms of the temporal relations over intervals introduced in section 5.

> N.B. As discussed in the introduction, the temporal theories of Ladkin [13,14], of Dechter et al. [17], and of Maiocchi [18] are similar to that of Bruce in the sense that intervals are defined to be constructed out of points. Hence, in a similar way, we may induce the corresponding time model for each of these temporal frameworks.

McDermott's temporal logic:

McDermott develops a first-order temporal logic to provide a versatile "common-sense" model for temporal reasoning. The theory assumes a "*no later than*" ordering relation over a dense collection of states (points), which is axiomatised to give rise to a left linear (branching into future) time structure. That is, there are many possible futures branching forward in time from the present. Each single branch, called a "*Chronicle*", consists of a dense set of states and is isomorphic to the real line (see [5]). Consider the temporal frame axiomatised by axioms <A1>-<A6>, $<A_{L\text{-linear}}>$, and the following additional axioms $<A_{P\text{-Dense}}>$ which states that there is always a time point during any time interval.

$<A_{P\text{-Dense}}>$
$$\forall i\in \mathbf{I}\exists p\in \mathbf{P}\exists i_1,i_2\in \mathbf{I}(i = i_1 + p + i_2)$$

By consideration of axioms <A2> and <A5>, we can infer that axiom $<A_{P\text{-Dense}}>$ ensures that between any two distinct time points on the same time line, there is a third. In fact, axiom $<A_{P\text{-Dense}}>$ is stronger than axiom $<A_{Dense}>$ (see section 4), since it is clear that $<A_{P\text{-Dense}}>$ implies $<A_{Dense}>$.

In the same way as for Bruce's partial order, we may also define the "no later than" relation over time points in terms of relations *EQUAL* and *BEFORE*. In this way, we may take McDermott's time structure as a model of the above theory by addressing only time points and the "no later than" relation.

Allen and Hayes' interval based theory:

Since the axiomatisation proposed in this paper may be seen as an extension of Allen and Hayes' interval based temporal theory [10], it is straightforward to subsume Allen and Hayes' theory by taking the set of time points to be empty, and including the linearity axiom $<A_{Linear}>$

in the fundamental axiomatisation. Of course, in this case, axiom <A5> will become vacuous.

N.B. Allen and Hayes' temporal theory presented in [10] only handles time as a pure abstraction, although Allen's interval based temporal logic is originally supposed to be set up as a framework on which to hang assertions about the instantiation in time of properties and occurrence [8]. In Allen's interval based logic, there are a small number of predicates among which HOLDS is one of the most important. To secure the interpretation of HOLDS (see section 2), Allen introduces the following axiom:

$$\text{HOLDS}(pro, \text{i}) \Leftrightarrow \forall \text{i'} \in \mathbf{I}(IN(\text{i'}, \text{i}) \Rightarrow \text{HOLDS}(pro, \text{i'})$$

where *IN* is defined in terms of the temporal relations over intervals, as below:

$$IN(\text{i'}, \text{i}) \Leftrightarrow DURING(\text{i'}, \text{i}) \wedge STARTS(\text{i'}, \text{i}) \vee FINISHES(\text{i'}, \text{i})$$

The negation of a property is then characterised by the axiom

$$\text{HOLDS}(\neg pro, \text{i}) \Leftrightarrow \forall \text{i'} \in \mathbf{I}(IN(\text{i'}, \text{i}) \Rightarrow \neg\text{HOLDS}(pro, \text{i'}))$$

However, in [1], Galton has shown that there are some problems with reasoning correctly about continuous change in Allen's logic, (in particular, with Allen's property-negation), and suggested the way out: instantaneous property-ascriptions.

As Galton puts it, the problems with Allen's system can be traced to the assumption that all properties should receive a uniform treatment with respect to the logic of their temporal incidence. Galton's starting point is then to distinguish sharply between two kinds of properties, i.e., *states of position* and *states of motion*, which have different temporal logics: States of position can hold at isolated points; and if a state of position holds throughout an interval, then it must hold at the limits of that interval. States of motion cannot hold at isolated points, that is, if a state of motion holds at a point then it must hold throughout some interval within which that point falls. Additionally, Galton defines three types of statement by the forms

$$\text{HOLDS-ON}(pro,\text{i}),\ \text{HOLDS-IN}(pro,\text{i}),\ \text{and HOLDS-AT}(pro,\text{p}),$$

which assert that a property, *pro*, holds <u>throughout</u> an interval, <u>during</u> an interval (i.e. at some time *DURING* an interval, not necessarily through all of it), and <u>at</u> a point, respectively, while in Allen's logic, there is only one way, HOLDS, of ascribing properties to times, that is, HOLDS-ON.

Since our general temporal theory allows both intervals and points, it is straightforward to form Galton's revised temporal theory. For example, we may formally characterise a state of position s_p by:

$$\begin{aligned}\forall i \in \mathbf{I} \forall p \in \mathbf{P}(\quad & \text{HOLDS-ON}(s_p, i) \\ & \wedge (\mathit{MET\text{-}BY}(i,p) \vee \\ & \quad \mathit{MEETS}(i,p) \vee \\ & \quad \mathit{STARTED\text{-}BY}(i,p) \vee \\ & \quad \mathit{FINISHED\text{-}BY}(i,p)) \\ & \Rightarrow \text{HOLDS-AT}(s_p, p) \;)\end{aligned}$$

and a state of motion s_m by:

$$\forall p \in \mathbf{P}(\text{HOLDS-AT}(s_m, p) \Rightarrow \exists i \in \mathbf{I}(\mathit{DURING}(p,I) \wedge \text{HOLDS-ON}(s_m, i)))$$

It is interesting to note that, the definitions relating to the open and closed nature of intervals given in section 4 provide another formal and intuitive characterisation for the distinction between *states of position* and *states of motion*: States of position can hold at isolated points; and if a states of position holds on an interval, then it must hold on the closure of that interval. States of motion hold only on open intervals. For instance, in the example of a ball thrown vertically into the air described in section 2, the property *ball_stationary* is a state of position, while *ball_going_up* and *ball_coming_down* are states of motion.

In a similar way, we may axiomatise other results for general properties (see [8] and [1]), as well as other issues such as processes and events, etc. However, since the main objective of this paper is to present a general time theory at some abstract level, here, we will not go further on addressing these broader issues.

Vilain's interval & point based system:

Noting that intervals are not the only mechanism by which human beings understand time, another common construct being that of time points, Vilain proposes a system which handles time points in much the same way that it handles intervals [11,12]. This system is arrived at by expanding Allen's 13 temporal relations over intervals to 26, which are primitively defined to relate points to points, intervals to intervals, intervals to points, and points to intervals. It is interesting to note that all Vilain's 26 temporal relations form a subset of the set of those 30 relations we introduced in section 5. The excluded four relations in Vilain's system are: *MEETS*, *MET_BY* that relate points to intervals, and *MEETS*, *MET_BY* that relate intervals to points (see section 5). Hence, if we employ the following more strict axiom instead of <A5>:

$$\forall t_1, t_2 \in \mathbf{T}(\mathit{meets}(t_1, t_2) \Rightarrow t_1 \in \mathbf{I} \wedge t_2 \in \mathbf{I})$$

then we get Vilain's temporal system. The above axiom ensures that if two time elements meet each other, then both of them must be intervals.

Knight and Ma's temporal model:

Knight and Ma [4] have proposed a temporal model akin to that presented here, taking both points and intervals with duration assignments as primitive time elements. However, this model addresses only finite linear sets of time elements. Hence, it is possible to consider it as a specialisation of the time theory to a finite set of time elements. In fact:

Assume (**T**,*meets*) is the temporal frame defined by axioms <A1> - <A6>, A_{Linear}, <$A_{L\text{-}Discrete}$> and <$A_{R\text{-}Discrete}$>. The discreteness property of the temporal frame allows us to form a nonempty finite set $\mathbf{T_f} \subset \mathbf{T} = \mathbf{I} \cup \mathbf{P}$, such that:

$\mathbf{T_f} = \{t_1, t_2, ..., t_n\}$;
$meets(t_i,t_{i+1})$, $i = 1, 2, ..., n\text{-}1$;
$meets(t_i,t_{i+1}) \Rightarrow t_i \in \mathbf{I} \vee t_{i+1} \in \mathbf{I}$.

These theorems are precisely the axioms for Knight & Ma's set **E**, of "*fundamental time elements*" (which may be thought as Allen and Hayes's "*moments*", see [10]). Additionally, it is easy to see that the limitation of axioms <A4>, <A5> and <A6> onto $\mathbf{T_f}$ precisely gives the definition of the *closure* of E, under the binary operations of combining adjacent time elements and corresponding addition of duration, that is, the so-called temporal system.

It is interesting to note that, in computer-based modelling approach, a database consists of only a finite (discrete) set of elements, that is, the database models only a finite subset of the fundamental (dense or discrete) set of primitive elements. The existence of complete set of primitive elements is a belief which may be used to test the consistency of the database. Hence, with this meaning, the consistency checker provided in [4] may be used for any finite temporal sub-frame defined by the axiomatisation.

7. Conclusions

In this paper, we have proposed a general time theory which may be seen as an extension of Allen and Hayes' axiomatisation by the addition of axioms relating to the inclusion of time points as primitive elements. This theory allows other first order temporal systems as models. It unifies a variety of temporal concepts into a single framework. We have attempted to define key concepts and terms with respect to the axiomatic system. And in addition, we have separated axioms for linearity and for density from the main body of axioms, since these appear to be most "user dependent". The resulting theory presents a unified view of what is currently a disparate field.

Acknowledgements

We would like to express our thanks to the referees for their helpful comments and suggestions during the preparation of this paper.

References

[1] A. Galton, A Critical Examination of Allen's Theory of Action and Time. *Artificial*

Intelligence, **42**, pp.159-188 (1990).

[2] A. Galton, *Logic for Information Technology*. JOHN WILEY & SONS Ltd., Baffins Lane, 1990.

[3] B. C. Bruce, A Model for Temporal References and Application in a Question Answering Program. *Artificial Intelligence*, **3**, pp.1-25 (1972).

[4] B. Knight and J. Ma, An Extended Temporal System Based on Points and Intervals. *Information System*, **18**, pp.111-120 (1993).

[5] D. V. McDermott, A Temporal Logic for Reasoning about Processes and Plans. *Cognitive Science*, **6**, pp.101-155 (1982).

[6] E. P. K. Tsang, Time Structure for AI. *Proceedings of the IJCAI*, **10**, pp.456-461 (1987).

[7] J. F. Allen, Maintaining Knowledge about Temporal Intervals. *Communication of ACM*, **26**, pp.123-154 (1983).

[8] J. F. Allen, Towards a General Theory of Action and Time. *Artificial Intelligence*, **23**, pp.123-154 (1984).

[9] . J. F. Allen, A Common-Sense Theory of Time. *Proceedings of the IJCAI*, **9**, pp.528-531 (1985).

[10] J. F. Allen and P. J. Hayes, Moments and Points in an Interval-based Temporal-based Logic. *Computational Intelligence (Canada)*, **5**, pp.225-238 (1989).

[11] M. B. Vilain, A System for Reasoning about Time. *Proceedings of AAAI*, **1**, pp.197-201 (1982).

[12] M. B. Vilain and H. Kautz, Constraint Propagation Algorithms for Temporal Reasoning. *Proceedings of AAAI*, **5**, pp.377-382 (1986).

[13] P. Ladkin, Models of Axioms for Time Intervals. *Proceedings of AAAI*, **6**, pp.234-239 (1987).

[14] P. Ladkin, Effective Solution of Qualitative Interval Constraint Problems. *Artificial Intelligence*, **52**, pp.105-124 (1992).

[15] P. V. Beek, Approximation Algorithms For Temporal Reasoning. *Proceedings of the IJCAI*, **11**, pp.1291-1296 (1989).

[16] P. V. Beek, Reasoning About Qualitative Temporal Information. *Artificial Intelligence*, **58**, pp.297-326 (1992).

[17] R. Dechter, I. Meiri and J. Pearl, Temporal Constraint Networks. *Artificial Intelligence*, **49**, pp.61-95 (1991).

[18] R. Maiocchi, Automatic Deduction of Temporal Information. ACM Transactions on Database Systems, **4**, pp.647-688(1992).

Appendix D

Time Representation: A Taxonomy Of Temporal Models

(*AI Review*, Vol.7, pp.401-419, 1994)

Abstract

The objective of the paper is to provide a taxonomy of temporal systems according to three fundamental considerations: the assumed axiomatic theory, the expressiveness, and the mechanisms for inference which are provided. There is an discussion of the significance of the key features of the taxonomy for computer modelling of temporal events. A review considers the most significant representative systems with respect to these issues, including those due to Bruce, Allen and Hayes, Vilain, McDermott, Dechter et. al., Kahn and Gorry, Kowalski and Sergot, Bacchus et. al., and Knight and Ma. A tabular comparison of systems is given according to their main structural features. In conclusion, the characteristics of a general axiomatic system capable of representing all the features of these models is discussed.

Key Words: time representation, temporal system, axioms, semantic analysis.

1. Introduction

In this article we consider the characteristics of modelling systems which have been proposed for capturing the temporal properties of events and processes in computer based systems. The objective is to give a taxonomy of systems, according to some fundamental features. We start with a brief discussion of what the fundamental features are, and why they are important for use in computer modelling.

Basic to all computer systems dealing with temporal events is an assumed theory of time. We require that this theory satisfies our intuitive notions of time, so that we can say that the real world is a model of the theory. By this, following the ideas of Suppes [26], and of Funk [10], we mean that the statements of the theory may be interpreted as true in the real world. The most common theoretical basis is the standard time-point system assumed by classical physics. In this theory, the time domain consists of a continuum of time points, isomorphic to the real line. Time intervals are taken as intervals on the real line, and duration of intervals is the real number difference of their start and end points. However, for many applications, particularly those in artificial intelligence and natural language understanding, the time-point system is not ideal for either the expression of temporal facts, or for the storage and organisation of incomplete temporal knowledge. For these applications, other theories have been proposed, for example, based on time intervals as primitive rather than time points.

The importance of the theory to a database system is as a basis for reasoning over the database. Inference may be performed over the stored data, by logical deduction from the axioms of the assumed theory. In some systems, no formal mechanism for inference is proposed, it being left to the user to draw inference from the database. In other systems a deduction system is proposed, in which rules are provided that allow deduction of true facts by forward chaining from the database. It is a characteristic of these systems that they are undirected, and do not allow the specific determination of a given query. Finally, some systems provide a consistency checker, and allow deduction by refutation. In these systems the user may enter a specific query, and the system checks whether it, or its negation, is inconsistent with the database. In this way the system may deduce whether a fact is: known true, known false, or unknown.

This view of temporal systems leads us to attempt a characterisation of temporal systems according to three basic elements, as follows:

- *The assumed axiomatic theory*: For all of the systems which we shall consider, there exists an underlying theoretical basis. For some systems this basis is formally described, and for others it remains assumed as intuitively agreed.

- *The expressiveness of the modelling language*: A computer based system may be viewed as a model of the fundamental theory, in the form of a finite data base of temporal facts. Given that the model is incomplete by reason of storage limitations, there is a drive for efficient storage and ##retrieval of incomplete temporal knowledge. Expressive modelling languages allow the storage of temporal information which is incomplete in various fashions.

- *The reasoning mechanisms which are provided*: Deductive inference may be performed on the stored data, with reference to the underlying theory, so that any fact which can be proved from the axioms of the theory and the stored temporal database may be assumed true by inference. In this way, the axioms plus database may be viewed as a deductive system from which facts may be retrieved by inference.

In summary, we can capture some fundamental characteristics of existing temporal systems with respect to the following set of questions:

- What are the assumed primitives?
- Is there a formal theory?
- What are its good/bad features of expression?
- What is its application domain?
- What reasoning mechanisms are there?
- Is there a consistency checker?

In section 2, we address the major issues of theory which characterise systems at a fundamental level. The questions of expressiveness and inference are particular to proposed systems, and these are discussed in a review of some representative temporal models in section 3. Section 3.10 provides a summary table characterising these models. In section 4, the characteristics of a general axiomatic system capable of representing the features of these models is discussed.

2. Major Theoretical Issues

The theoretical nature of time is a question with a long philosophical tradition and the literature seems full of disputes and contradictory theories. This contrasts sharply with the commonly held view of time, which allows people to cope easily with time in their everyday life. However, there are several major issues which should be addressed in terms of the theoretical basis of proposed systems. These issues are as follows:

2.1 The Primitive Nature Of Time

This is the issue of what should be taken as the primitive elements of time. There are three known choices: points, intervals, or both ##. Additionally, there are two fundamentally different treatments of interval based systems. In the first, intervals are assumed to consist of points, and hence, the corresponding systems may be considered as models of point-based time theories. An example of this kind of interval is the *time-segment* of Bruce's model for temporal references [8]. However, as Allen has commented [1,2], modelling intervals by taking their ending-points can lead to problems: the annoying question of whether ending-points are in the interval or not must be addressed, seemingly without any satisfactory solution. The second treatment takes intervals as primitive objects without any definitions of the "ending-point" and "internal-point" structures. Allen's interval logic [1,2,3], Vilain's temporal system [27,28], Knight and Ma's extended temporal model [15,16], are examples that treat intervals as primitive.

2.2 Ordering Relations

Whatever primitive time elements are taken, all time systems must adopt axioms defining some sort of ordering relations. Two fundamental issues are associated with time ordering: the density of time elements, and the linearity of the time axis. We discuss these issues in the following sections.

2.2.1 Density of time

The density question is associated with the choice of whether the set of time-elements should be modeled as a continuum (such as rationals or reals) or as a discrete set (such as integers). For time-points, can we always assume that between any two distinct time-points there is at least another time-point? For time-intervals, can we always assume that any interval can be decomposed into two distinct contiguous intervals? If so, then the primitive elements form a dense system. The alternative assumption is that of discrete time, "whereby each time (except the first and last if there is a beginning or end to time) is sandwiched between unique **previous** and **next** times" [12].

The fact that the database must consist of a finite set of time-elements has no bearing on the density question at all, which is a question of the assumed theory only. This theoretical issue impinges upon the inferencing mechanisms which may be used to derive facts from the database, insofar that the denseness assumption is needed to prove the consistency algorithms.

2.2.2 Linearity of Time

This issue refers to whether the time axis can be always considered as *linear* or *non-linear*. Linear time corresponds to the classical physical model of time, where the structure is that of the real line, extending indefinitely in both directions. The majority of time modelling approaches consider the time axis as being linear, that is, there is a total order over the whole set of time elements. However, non-linear time structures have been proposed, where the fundamental order relation allows topologies such as *branching time*, *parallel time* and *circular time*, etc.

It is questionable whether computer based systems really require non-linearity to be built into the temporal axioms, since its raison d'etre appears to be involved with a lack of knowledge of temporal events, rather than with our intuition about time itself. For example, parallel time lines have been proposed as a way of modelling separate parallel processes. However, it is a limitation in our knowledge which gives rise to the parallelism. We believe that the two processes are actually operating in the same linear time - it is just that we have no knowledge of synchronisation. We do not need a theory of parallel time lines for this application; what we need is a model which allows us incomplete knowledge of synchronisation over a single linear time line. Similarly, branching time is proposed as a useful model to handle possible worlds, uncertainty about the past or the future and the effects of alternative actions when planning. However, it is arguable whether we need a theory which assumes that time itself branches in order to model possible worlds, rather than a model which expresses our limited knowledge of causality in possible worlds over a single linear time. For most applications, linearity is sufficient at the theoretical level. This corresponds with the usual assumption of classical physics where all events may be universally synchronised with a single time measure. Only if we wish to model relativity would we be unable to assume synchronisation of distant events at a theoretical level.

2.3 Duration Assignment

In most applications, it is expected that a temporal system can support duration reasoning. For example, if it is known that interval I_a and interval I_b start together and that the duration of I_a is greater than duration of I_b, we may infer that I_b finishes before I_a. This inference can be made by use of duration knowledge.

The duration assignment to time elements may be characterised by a function from the set of time elements to $\mathbf{R}_0^+$, the set of non-negative real numbers. Intuitively, of course, the duration of the points should be zero, while the durations of intervals are positive. For point-based intervals, their durations may be derived from the distance between their greatest lower bound and least upper bound. However, for systems which treat intervals as primitive, their durations may be directly defined by an abstract function from intervals to positive reals. Given a duration assignment over time elements, some corresponding operators, such as *addition*, may be required to be defined, providing consistency of the whole system.

3. Some Representative Models

In this section, we review some representative temporal models, with respect to the fundamental issues addressed in the introduction.

3.1 Bruce's *Temporal Model*

An early attempt at mechanizing part of the understanding of time within an artificial intelligence is Bruce's model for temporal references [8]. In this system a formal framework, based upon first-order logic, is established for the analysis of tenses, time relations, and other references to time in natural language. The axioms of the framework are based on the following definitions: A *time-system* is a pair, (*time*, $\leq$), where *time* is a set whose elements are called *time-points*, and $\leq$ is a partial order over *time*. Because there is nothing that has been defined about *time* other than that it is partially ordered by $\leq$, the theory allows linear time or branching time, discrete time or dense time. The theory is thus more general than that for the standard point-based system, and inferencing mechanisms must be built on weaker axioms.

Bruce then defines point-based intervals, termed *time-segments,* as chains which are convex in the sense that there are no points missing within the chains, where a chain is a totally ordered subset of *time-points*. The related issues about time-segments, such as: density and linearity, may hence be derived from the corresponding issues of the time-points which make up the time-segments. The ordering relations between segments are also inherited from the partial order over the time points. Bruce gives seven binary relations between *time-segments,* which can be derived from the ordering relations over their greatest lower bounds and the least upper bounds: *Before, During, Same-time, Overlaps, After, Contains* and *Overlapped.* In terms of these binary relations, a *tense* is defined as a special n-ary relation on time-segments with the following form:

$$R_1\_R_2\_...\_R_{n-1}(S_1,S_2,...,S_n) \equiv R_1(S_1,S_2) \wedge R_2(S_2,S_3) \wedge ... \wedge R_{n-1}(S_{n-1},S_n)$$

where each S_i is a time-segment and R_i is a binary relation between S_i and S_{i+1}. S_1 is called the *time of speech*, S_2, ..., S_{n-1} are called the *times of reference*, and S_n is called the *time of event*. For example, the following sentence

- He will have been going to be going to go

has the tense

$$\mathit{Before\_After\_Before\_Before}(S_1,S_2,S_3,S_4,S_5) \equiv$$
$$\mathit{Before}(S_1,S_2) \wedge \mathit{After}(S_2,S_3) \wedge \mathit{Before}(S_3,S_4) \wedge \mathit{Before}(S_4,S_5)$$

where S_1 is the time of speech, S_2, S_3, S_4 are reference times, and S_5 is the time of event.

Bruce provides a natural language system, termed *CHRONOS*, which consists of a simple English sentence parser, a theorem prover, and a database of facts and events. The system accepts facts about events from the user and the information which is given by tense and time relations can be combined with other facts to allow inferences about the temporal ordering of events. However, a consistency checker for the database has not been provided explicitly. No heuristics are used in searching the network of temporal ordering links. Additionally, as argued by Allen (see next section), there are some problems in dealing with the treatment of open or closed intervals. Mechanisms for duration reasoning are not specified, although these may be defined by introducing a mapping from the time-points to the reals.

3.2 The *Interval Logic* Of Allen

Allen introduces his temporal logic in order to provide a framework for the naive treatment of two major subareas of artificial intelligence: natural language processing and problem solving. Instead of adopting time points (or states which are associated with time points), he takes intervals as the primitive temporal quantity, as being the natural means of human reference to time. As an example, in [2], Allen gives the following story:

Ernie entered the room and picked up a cup in each hand from the table. He drank from the one in the right hand, put the cups back on the table, and left the room.

In this account we can identify several time intervals, e.g.: the time Ernie was in the room, the time between entering the room and picking up each cup, the time between putting down the cups and leaving the room, and many others. However, the claim is that intervals are sufficient for modelling all the temporal references in human accounts such as this. Even references to apparent point events, such as the time Ernie entered the room, or the time that he put down a cup, are best modelled as small time intervals. The argument is put forward that all apparently instantaneous events can be decomposed further if we examine them more closely. For example, "entering the room" may be decomposed into: opening the door, moving through the doorway, and closing the door. And again, "opening the door" can be decomposed into turning the handle and pushing the door open. As Allen puts it [2]:

> *There seems to be a strong intuition that, given an event, we can always "turn up the magnification" and look at its structure.*

In order to express temporal relationships over time intervals, Allen took originally as primitive a set of nine (mutually exclusive) basic binary relations between any two intervals [1], extended later to 13 [2]: *Equal, Before, Meets, Overlaps, Starts, Started-by, During, Contains, Finishes, Finished-by, Overlapped-by, Met-by, After*. These are based on Bruce's seven relationships, but whereas Bruce's relations are derived from the partial order within a point-based theory, Allen's are taken as primitive.

These relationships are later formally defined in terms of the single primitive relation "*Meets*" by Allen and Hayes [3]. This is done by positing the existence of related intervals for some relations. For example:

$$Before(i_1,i_2) \Leftrightarrow \exists i(Meets(i_1,i) \wedge Meets(i,i_2))$$

In Allen's system, consistency checking is performed by formation of the transitive closure, according to a transitivity table with 144 entries which describes the composition of the thirteen (mutually exclusive) relations. If no conflict is found according to the exclusivity, then the system is consistent. For example, for the system:

$$Before(a,b),\ Before(b,c)$$

we may use the transitivity entry:

$$Before(i_1,i_2) \wedge Before(i_2,i_3) \Rightarrow Before(i_1,i_3)$$

to deduce that *Before*(a,c). Hence facts may be derived by forward chaining from the database, using the transitivity rules (termed truth propagation by Allen). Possible inconsistencies in a database can also be established by truth propagation. For example, from:

$$Before(a,b),\ Before(b,c),\ Before(c,a)$$

we can deduce *Before*(a,c) from the first two predicates, and *After*(a,c) from the third. Hence we have two distinct relations between a and c, which are not allowed due to the exclusivity of temporal relations.

Allen and Hayes show that the transitivity table in [2] is a result of the their axioms in [3], following the intuitive reasoning by possible cases which has been used to construct the table originally. Additionally, in [2], Allen has suggested that duration reasoning may also be incorporated into the interval-based system by giving examples of rules for duration reasoning. For example:

$$During(a,b) \vee Starts(a,b) \vee Finishes(a,b) \Rightarrow duration(a) < duration(b)$$

However no comprehensive mechanism has been proposed, and hence the duration reasoning is rather weak.

The most disputed aspect of Allen's system is its exclusion of time points as primitive, although in the later paper [3], Allen and Hayes define a point as the "*meeting place*" of intervals or as a maximal set, termed "*nest*", of intervals that share a common intersection, at a subsidiary status within the theory; and use the concept of a "moment", i.e., a very short interval which is non-decomposable, to model some instantaneous events. Their contention is that nothing can be true at a point, for a point is not an entity at which things happen or are true [2]. Except for the assumption that moments have positive length, while points have zero length, another obvious structural difference between points and moments is that moments are treated as primitive objects, and hence can meet other intervals (although they are not allowed to meet other moments), while points are not treated as primitive objects and cannot meet anything [3].

However, as Galton shows in his critical examination of Allen's interval logic, Allen's theory of time is not adequate, as it stands, for reasoning correctly about continuous change [11]. This problem stems from Allen's determination to base his theory on time intervals rather than on time points, either banishing points entirely, or, latterly, relegating them to a subsidiary status within the theory. The following example of a ball thrown vertically into the air intuitively shows the problem involved with references to time points: The motion may be described qualitatively by the use of two intervals, interval i_1 where the ball is going up, and interval i_2 where the ball is coming down. According to classical physics, there is a point where the ball is stationary. In the interval calculus, we have two alternatives: we may assume that there is a small interval where the ball is stationary, or we may assume that interval i_1 "*Meets*" interval i_2. The first alternative does not seem tenable, not being consistent with the laws of physics. On the other hand, the second alternative also gives problems, since the interval calculus allows us to combine two intervals which meet; that is, $i_1 \oplus i_2 = i_3$. However, although both of the intervals i_1 and i_2 have the property "ball-in-motion", the combined interval i_3 doesn't have this property.

3.3 Vilain's *Temporal System*

Noting that intervals are not the only mechanism by which human beings understand time, another common construct being that of time points, Vilain [27,28] proposes a system which handles time points in much the same way that it handles intervals. The logic of points is arrived at by expanding Allen's logic of intervals: adding new primitive relations and composition rules over them to Allen's interval logic. The new primitive relations may be classified into three groups:

(Point-Point)	*Equal, Before, After,* which relate points to other points;
(Interval-Point)	*Before, Started-by, Contains, Finished-by, After,* which relate intervals to points;
(Point-Interval)	*Before, Starts, During, Finishes, After,* which relate points to intervals.

The mechanism by which Vilain's system makes deductions about points is an extension of that which it uses to make deductions about intervals. In an approach similar to that of Allen, the system maintains a "complete picture" of all relations over intervals and points by means of a transitive closure operation. The operation is performed over the expanded set of composition rules in the newer logic.

However there is a critical omission from the primitive relations between points and intervals in Vilain's system; for the "*Meets*" relation is defined only between intervals and is not allowed between points and intervals. Hence, the problems in modelling continuous change by Allen's system mentioned by Galton in [11] still exist in Vilain's system. For example, the system is still not capable of modelling the processes of a ball thrown vertically into the air: Let interval i_1 refer to *ball-going-up*, point p refer to *ball-stationary*, and interval i_2 refer to *ball-coming-down*. On the one hand, it is easy to see that p is neither in i_1 nor i_2. On the other hand, according to Vilain's classifications of relations over points and interval, point p is not allowed to meet or be met-by any interval. Hence, we deduce that p is after i_1 and before i_2, that is, there is another time element between i_1 and p, and another time element between p and i_2. This is obviously contrary to our intuition of the processes.

N.B. In [5], Beek has proposed an interval-based framework, *IA*, and point-based framework, *PA*, for representation of and reasoning about incomplete and indefinite qualitative temporal information. However, it is interesting to note that the frameworks, *IA* and *PA*, deal with temporal relations between intervals, and relations between points separately, that is, the interval-based framework *IA* deals with the thirteen temporal relations (defined by Allen [2]) between intervals only, while the point-based framework *PA* deals with temporal relations between points only, which are addressed in Vilain and Kautz's point algebra [28]. Relations between intervals and points, such as that proposed in [27], are not addressed at all. Additionally, like Dechter et. al.'s framework (see next section), time intervals are not defined as primitive. Indeed, time intervals, and temporal relations between intervals are defined in terms of points (rationals) and the corresponding order relations between points, respectively.

3.4 Dechter , Meiri and Pearl's *TCSP*

Dechter, Meiri and Pearl [9] have presented a unified approach to temporal reasoning based on constraint-network formalism. In this framework of temporal constraint satisfaction problems (*TCSP*), variables represent time points, and temporal information is represented by a set of unary and binary constraints, each specifying a set of permitted intervals. The unique feature of this framework lies in the inclusion of duration information, namely, time differences between events. Algorithms are presented for performing some reasoning tasks, such as finding all feasible times at which a given event can occur, finding all possible relationships between two given events, and generating one or more scenarios consistent with the information provided. A *TCSP* involves a set of variables, $X_1, ..., X_n$, having continuous domains; each variable represents a time point. Each constraint is represented by a set of intervals: $\{I_1, ..., I_n\}$, where these intervals are similar to Bruce's time-segments, that is, they are point-based, may be closed, open, or semi-open. A simple temporal problem (*STP*) is a *TCSP* in which all constraints specify a single interval. The duration of an interval may be defined by the distance between its greatest lower bound and least upper bound. Relations between intervals, such as the thirteen relations defined by Allen, may be derived from the known total order relation among their greatest lower bound and least upper bound. Consistency checking for a *TCSP* is transformed to a corresponding examination of its graphic representation.

The theory is formally stated, with points and real numbers as primitives, and intervals being constructed out of points. It assumes a dense set of time-elements, but time may be branching. Duration reasoning is encompassed by the system by means of a consistency checking algorithm. The limitation of the *TCSP* model is it's assumption that point based intervals have the same open\closed nature, that is, either intervals are all assumed to be closed, or they are all assumed to be open (semi-open). This assumption can lead to problems: if intervals are all closed then adjacent intervals have ending-points in common, which when adjacent intervals correspond to states of truth and falsehood of some property, can lead to situations in which a property is both true and false at an instant. Similarly, if intervals are all open, there will be points at which the truth or falsity of a property will be undefined (The solution in which intervals are all taken as semi-open, so that they sit conveniently next to one another, seems arbitrary and unsatisfactory).

3.5 The *Time Specialist* Of Kahn And Gorry

In order to store, retrieve, and reason about temporal information, Kahn and Gorry [14] have designed and implemented a module, called the *time specialist,* to maintain separate mechanisms for dealing with dated and undated information. The time specialist is endowed with the capacity to order temporal facts in three major ways:

(1) relating events to dates,
(2) relating events to special "reference events",
(3) relating events together into before-after chains.

The time specialist can answer different types of questions such as:

- Did event X happen at time expression T?
- When did event X happen?
- What happened at time expression T?##

The time specialist is able to make deductions and check whether## they are consistent with the facts known in the database. However, it is weak if the time indications are not definite. Also, each of the three methods to organize temporal statements has its own special data structures and routines to work with those structures. For a given set of temporal facts, it is up to the user, not the time specialist, to choose the most appropriate methods.

The time specialist can check the consistency of the latest fact with facts previously accepted, and try to resolve inconsistencies through interaction with the user. In such an interaction, the user may withdraw either the new fact, or some old facts whose removal would lead to consistency. However, removing old facts may involve undoing some prior deductions. In order to be able to do this, a deduced fact is marked by those facts used to deduce it.

No formal theory is stated as a basis for the time specialist. The basis for temporal reasoning is contained in the algorithms which make up the system.

3.6 The *Temporal Logic* of McDermott

McDermott [22] has developed a first-order temporal logic to provide a versatile "common-sense" model for temporal reasoning. In accordance with the "naive physics" advocated by Hayes [13], McDermott adopts an infinite collection of states (points) as the primitive temporal elements and adds several crucial axioms. Every state has a time of occurrence, $d(s)$, a real number called its *date*. Time is assumed to be a continuum, with an infinite numbers of states between any two distinct states, where states are partially ordered by the "*no later than*" order relation "$\leq$". The future (not the past) is branching, that is, there are many possible futures branching forward in time from the present. Each single branch, called a

"*Chronicle*", consists of a connected series of states and is isomorphic to the real line. Developing his theory, McDermott examines three major problems that a temporal reasoning system must face: reasoning about causality, reasoning about continuous change, and planning actions.

McDermott's system has formal axioms with time-points (states) and reals as primitives. The theory assumes a partial ordering relation, which gives rise to branching time. Reasoning is via the assumed theory of the real numbers, and no special mechanisms are needed. We can represent a time state, s, as the pair (C_s, t), where $t = d(s)$ and C_s is the set of chronicles that s belongs to. Possible events may be associated with time states.

For illustration, we shall consider the example of a man, called John, planning a trip to the theatre. We assume that a decision will be made to go by train or bus. If the decision is made to go by train at time s_{train1}, where $d(s_{train1}) = t_1$, then John will arrive at the theatre at time s_{train2}, and the play will start at time s_{train3}, where $d(s_{train3}) = t_3$. All of these time states lie on a chronicle c_{train}. Alternatively, if the decision is made to go by bus at time s_{bus1}, where $d(s_{bus1}) = t_1$, then he will arrive at the theatre at time s_{bus2}, and the play will start at time s_{bus3}, where $d(s_{bus3}) = t_3$. All of these time states lie on chronicle c_{bus}. These events and states may be represented by the following data:

(decides-to-take-train,	c_{train},	t_1)
(arrives-at-theatre-by-train,	c_{train},	t_2)
(play-starts,	c_{train},	t_3)
(decides-to-take-bus,	c_{bus},	t_1)
(arrives-at-theatre-by-bus,	c_{bus},	t_2')
(play-starts,	c_{bus},	t_3)

Here, s_{train1} has been represented by the pair (c_{train}, t_1), s_{train2} by (c_{train}, t_{train2}) etc.

In this example, illustrated in Figure 1, we see that time states divide into two separate chronicles c_{train} and c_{bus}, from the state s_0 as a result of the John's decision. Although it is obviously possible for us to compare times on different chronicles by means of the t component, McDermott uses the "*no later than*" relation over time states which is restricted to states on the same chronicle. This is to prevent us from making "no later than" comparisons for events which cannot both occur in reality. For example, we are not allowed to ask whether he arrives at the theatre by bus before he arrives by train, since he cannot do both. These two events are said to be in different possible worlds (i.e. chronicles).

McDermott also provides axioms which ensure that chronicles branch only into the future, and this limits the expressiveness of the logic. For, in the example, we have the event "play starts" on two different chronicles which cannot be compared. Using McDermott's logic we must view these as two separate events: "play starts after John's arrival by train", and "play starts after john's arrival by bus". Since we may judge that the play is independent of John, we may wish to join the two chronicles at the state that play starts.

It is in fact arguable whether we need to consider time as branching in order to model possible worlds. In fact, it is possible to conceptualise the world number, or chronicle, as

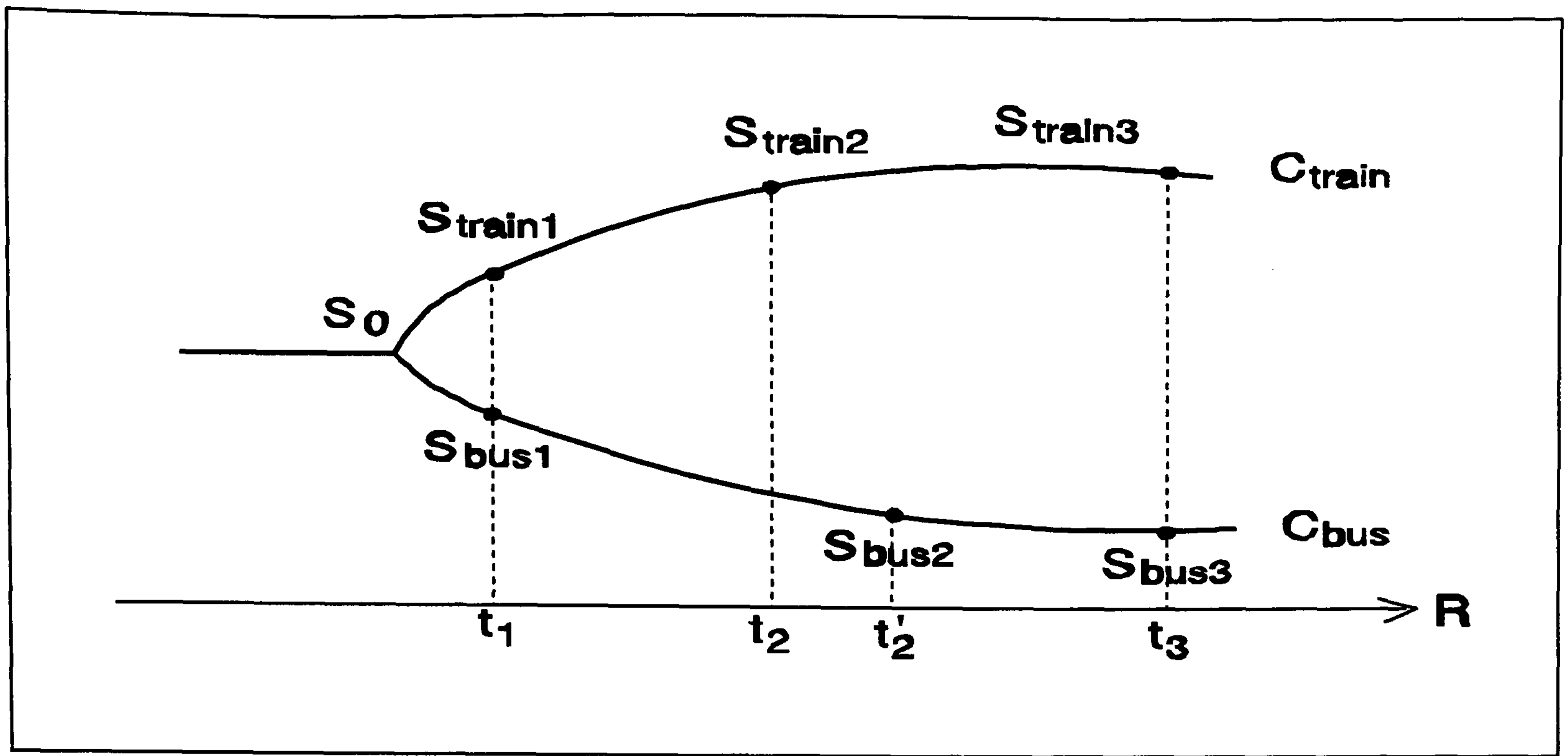

Figure 5

related to the event data, and not to the time. For example, we can regard the predicate:

(decides-to-take-train, c_{train}, t_1)

as relating:

(event, possible_world, time)

rather than:

(event, time_state)

In this case, time elements are standard linear dense time points, and the axioms for chronicles can be specified independently of those for time.

3.7 Kowalski And Sergot's *Event Calculus*

The *event calculus* of Kowalski and Sergot [17] is an approach for representing and reasoning about time and events within a logic programming framework. It is based in part on the situation calculus [20,21], but focuses on the concept of events as highlighted in semantic network representations of case semantics. Its main intended application is the representation of events in updating databases [18] and discourse representation.

Primitives of the theory are events, which are considered to be structureless "points" in time, where "point" is used here only to convey the lack of internal structure. Events start and finish periods of time, during which states are maintained. Events are considered to be after the time periods that they finish and before the time periods that they start, not fully contained within either of these periods.

Sadri [23] has illustrated a number of the general characteristics of the event calculus:

(1) Event descriptions can be assimilated in any order, independent of the order in which events actually take place.
(2) Events can be used for temporal references and need not be associated with absolute times.
(3) Events can be simultaneous.

(4) Events can be partially ordered.
(5) All updates are additive. The effect of deletion is obtained by adding information about the end of periods.
(6) The event calculus rules are in Horn clause logic augmented with negation by failure.
(7) The event calculus allows events to be input with incomplete descriptions.

In [18], Kowalski specially investigates the case of the event calculus connected with database updates. The way in which relational databases, historical databases, modal logic, and the situation calculus deal with database updates is discussed in detail. It is claimed that the event calculus may overcome the computational aspects of the frame problem in the situation calculus, and it can be implemented with an efficiency approaching that of destructive assignment in relational databases. Bernard et. al. [6] have recently presented an adaptation of the event calculus to the problem of determining the temporal structure of operations that must be performed during the realization of some complex objectives. In [7], an extension to Kowalski's event calculus model is proposed by Borillo and Gaume, by means of the additional spatial component, and the introduction of uncertainty and a general abstract relation among propositions.

The formal theory of Kowalski and Sergot's *event calculus* may be taken as the Horn clause system plus negation by failure. However, the use of negation by failure introduces a procedural element into the axioms. In this respect, the system is thus akin to the time specialist, in that the theory is presented in terms of algorithms.

3.8. Bacchus, Tenenberg and Koomen's *BTK*

Bacchus, Tenenberg and Koomen present a many-sorted temporal logic, termed *BTK* [4], for reasoning about propositions whose truth values might change as a function of time. In order to provide a clear semantics and a well-studied proof theory, Bacchus et. al. partition both the universe of discourse and the symbols of their language into two sorts, temporal and non-temporal, by which time is given a special syntactic and semantic status without having to resort to reification. In *BTK*, propositions are associated with time objects by including temporal arguments to the functions and predicates, where terms and wffs are defined in the standard fashion, with the only restriction being that arguments of the correct sort must be given for each function and predicate.

Actually, *BTK* is sorted in much the same way as Shoham's *reified logic* [24,25]. Unlike Shoham's first-order logic in which propositions are expressed just with respect to a pair of time points (denoting a time interval), propositions in *BTK* can be expressed and interpreted with respect to any number of temporal arguments: there is neither a syntactic commitment to the number of temporal objects that any function or predicate may depend upon, nor is there any commitment to interpreting the temporal objects as either intervals or points.

It is interesting to noted that, in their paper [4], Bacchus et. al. have shown that Shoham's logic can in fact be subsumed by *BTK* by defining two transformations, a syntactic transformation, π_{syn}, and a semantic transformation, π_{sem}. π_{syn} maps sentences of Shoham's logic to sentences of *BTK*, while π_{sem} maps models of Shoham's logic to models of *BTK*. Additionally, Bacchus et. al. argue that Shoham's categorization of propositions over point-based time intervals may also be translated to *BTK*, and the ontology of *BTK* is richer since it allows time intervals to be the primitive temporal objects rather than being defined as pairs of time points.

The major difficulty involved in reasoning in a *BTK* system lies in reasoning with the temporal terms, while the complexity of reasoning is highly dependent on the nature of the temporal domain. However, in *BTK*, there is no axiomatisation characterising the time structure. This question is left open, so that the temporal domain of *BTK* may be defined to be any temporal structure which can be characterised by a set of axioms, for example that of Bruce [8], of Allen and Hayes [3], or of McDermott [22]. A complete proof theory may then be generated by adding the axioms for the temporal domain to the fundamental axiomatisation of the logic.

3.9. Knight and Ma's *ETM*

As mentioned in section 3.2 and 3.3, there are some difficulties with Allen's and Vilain's approaches in the qualitative modelling of everyday occurrences. The authors have proposed an extended temporal model, *ETM* [15], which treats both intervals and points as primitive time elements on the same footing, and supports duration reasoning and consistency checking.

The definition of a temporal system supporting duration reasoning consists first of a definition of an underlying well-ordered set E. The elements of the elementary set E may be both points and intervals with a duration assignment which is defined by a mapping from the primitives to the non-negative reals. The temporal system is then defined as the closure, T, of E under the binary operations $\oplus$, representing the combination of adjacent elements, and the conventional addition of the corresponding durations. This model provides axioms for a single successor relation, termed "*Meets*", over time intervals and points, and supports duration reasoning, which has been a problematic aspect in many temporal systems. Excepting the axiom that the duration of an interval is positive while the duration of a point is zero, the differentiating property between intervals and points which is proposed is that while intervals may meet points or intervals, points are not allowed to meet points, although they can meet (or be met-by) intervals. This characteristic is in line both with modelling requirements where points are defined as separators or end-points of intervals, and with the denseness of points on the real line. But this is the only extra requirement which is made of elements if they are to be points. According to their definitions, points, as primitive elements of *ETM*, are different from either Allen's points or moments. It seems that Allen's moments may be taken as the elementary intervals in E.

An intuitive graphical representation of an incomplete temporal system, (K, M_K, D_K), is introduced in terms of a directed, partially weighted graph, where K is a set of time elements, and M_K, D_K are the "*Meets*" knowledge and duration knowledge over K, respectively. And necessary and sufficient condition for the consistency of an incomplete system (K,M_K,D_K) [15], and the corresponding limited system (K,M_K) [16], is formally presented.

If we let intervals i_1 and i_2 refer to *ball-going-up*, *ball-coming-down* respectively, and point p refer to *ball-stationary*, we can now satisfactorily model the processes of a ball thrown into the air (see section 3.2 and 3.3) as: *Meets*(i_1,p) and *Meets*(p,i_2).

Additionally, although intervals are taken as primitive, as in Allen's system, the *ETM* allows formal expression of open and closed nature of intervals with the following meaning:

interval i is left-open at point p iff *Meets*(p, i);

interval i is right-open at point p iff *Meets*(i, p);

interval i is left-closed at point p iff $\exists$i'(*Meets*(i', i) $\wedge$ *Meets*(i', p));

interval i is right-closed at point p iff $\exists$i'(*Meets*(i, i') $\wedge$ *Meets*(p, i')).

which is in fact consistent with the conventional meaning of the "open" and "closed" nature for point-based intervals.

In terms of "*Meet*", 30 relations over intervals and points may be formally defined. This is indeed an extension of Vilain's primitive relations (see section 3.3), by means of adding four critical relations: *Meets*, *Met-by* that relate intervals to points, and symmetrically, *Meet*, *Met-by* that relate points to intervals.

The consistency condition given in *ETM* implies an inferencing mechanism including duration reasoning. It is straightforward to prove that all Allen's duration reasoning rules are explicit results of the inferencing mechanism, by using the consistency condition.

A limitation of this system is the assumption of discreteness. The theory on which the system is based assumes a discrete set as time domain. However, since any computer based system must be in a finite form, this requirement does not in fact place any restriction on the application field. In section 4, it is proposed that the system may also be based on a dense time domain.

3.10 Overview of Models

Based on the above discussions, we present an overview of these representative temporal models in terms of Table 1:

Issue → / **Model ↓**	**Primitive**	**Ordering Relation**	**Theory**	**Duration Reasoning**	**Inference Mechanism**
Bruce's ***CHRONOS***	point-based intervals	7 binary relations	formal	no	refutation (no consistency checker provided)
Allen & Hayes' ***interval logic***	primitive intervals	13 binary relations formed by "Meets"	formal	weak	deductive rules (transitivity table)
Vilain's ***temporal system***	primitive intervals and points	26 intuitive binary relations	no formal axioms	no	deductive rules (transitivity table)
Dechter, Meiri & Pearl's ***TCSP***	point-based intervals	total order	formal	yes	refutation (consistency checker provided)
Kahn & Gorry's ***time specialist***	points (event dates)	"before-after" chains	no formal axioms	no	not formal
Mcdermott's ***temporal logic***	points (states)	"no later than" binary order	formal	no	assumed theory of the real numbers
Kowalski & Sergot's ***event calculus***	points (events)	partial order	no formal axioms	no	resolution (negation by failure)
Bacchus, Tenenberg & Koomen's ***BTK***	not specified	not specified	not presented	no	none
Knight & Ma's ***ETM***	primitive intervals and points	successor relation	formal	yes	refutation (consistency checker provided)

Table 1

4. Conclusion

In this paper, we have examined the bases of various temporal systems, concentrating on the differences of approaches taken. However, apart from differences of terminology, the models show a commonality of structure at a fundamental level. All the systems rely on theories based on a primitive set of time elements, which may be points, intervals, or both of them. The systems are axiomatised by primitive order relations over the time elements. This suggests the question as to whether a general axiomatic system is possible, which will express this common structure at a theoretical level. We first discuss the properties that we might wish for a general axiomatic system, and then how it might be possible to achieve it.

To start we ask the question as to what we might ideally require of a general axiomatic system. Firstly, we might require that it should take both intervals and points as primitive time elements, and thus allow point-based, interval-based, or point- and interval- based models. Secondly, primitive order relations should be defined over the primitive time elements, from which the order relations for the main temporal systems, outlined in section 2, may be derived. For point based systems such as those of Bruce, of Dechter, Pearl and Meier, and of McDermott, the primitive order is "no later than". For interval based systems, it is "*Meets*", in terms of which thirteen possible temporal relations can be defined. Thirdly, a primitive duration function is needed, assigning a real number to each time element.

To ensure the generality of the axiomatisation, it should allow discreteness or denseness of time, which could be specified by additional axioms if required. It should also provide a special axiom for linearity of time, without which the time structure is branching. Finally, we would like a consistency checking algorithm for any finite database of temporal facts to be proved from the axioms, so that inference by refutation is possible.

In [3], Allen and Hayes have provided formal axioms for their interval based system, including a special axiom, <M2>, for the linearity of time. The primitive order assumed in Allen's theory is the "*Meets*" relation between time intervals, which may be used to define all the thirteen possible temporal relations between intervals. In *ETM*, it is also used to define the three relations between two points: *Before*, *Equal* and *After*. Hence, the fundamental order relation for time point systems, "*no later than*", may be defined in terms of "*Meets*".

The problem is that neither Allen's system, nor *ETM* are a sufficient basis for this purpose: Allen's system does not include points, while *ETM* deals with finite sets of time elements only.

However, Allen and Hayes' interval based axiomatisation of time [3] is in fact very appropriate for extending to a general time theory. What is needed is an extension of Allen's axiomatisation to include points. In *ETM* there is a critical axiom that a point cannot meet another point, and it seems that this is likely also to be necessary for the general axiomatisation. In fact, in [19], the authors have proposed a general temporal theory which addresses both intervals and points as primitive time elements of equal footing. The fundamental axiomatisation is independent of the specification of density and linearity, while additional axioms specifying the *linearity* and *density* of time are separately presented. It is shown that Allen and Hayes' interval based theory may be subsumed, and the authors's *ETM* may be taken as a special finite model of the general theory.

5. References

[1] Allen J. F.: "An interval-based representation of temporal knowledge", *Proc. 7th Int. Joint Conf. on AI*, 1981, pp.221-226.

[2] Allen J. F.: "Maintaining Knowledge about Temporal Intervals", *communication of ACM.*, Nov. 1983, Vol.26, pp.123-154.

[3] Allen J. F. and Hayes P. J.: "Moments and Points in an Interval-based Temporal-based Logic", *Comput. Intell.* (Canada), (Nov. 1989), Vol.5, no.4, pp.225-238.

[4] Bacchus F., Tenenberg J. and Koomen J. A.: "A non-reified temporal logic", *Artificial Intelligence*, 52(1991), pp.87-108.

[5] Beek P. V.: "Reasoning About Qualitative Temporal Information", *Artificial Intelligence*, 58, pp.297-326, 1992.

[6] Bernard, D., Borillo, M. and Gaume, B.: "From event calculus to the scheduling problem. Semantics of action and temporal reasoning in aircraft maintenance", *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 1(1991), pp.195-221.

[7] Borillo, M.; Gaume, B.: "Spatiotemporal reasoning based on an extension of event calculus", in *Proceedings of the Third COGNITIVE Symposium*, Madrid, Spain (1990), Kohonen, T. and Fogelman-Soulie F. (eds.), pp.337-344.

[8] Bruce B. C.: "A Model for Temporal References and Application in a Question Answering Program", *Artificial Intelligence*, 3(1972), pp.1-25.

[9] Dechter R., Meiri I. and Pearl J.: "Temporal constrain network", *Artificial Intelligence*, 49(1991), pp.61-95.

[10] Funk K. H.: "Theories, Models, And Human_Machine Systems", *Mathematical Modelling*, Vol.4, pp.567-587, 1983.

[11] Galton A.: "A Critical Examination of Allen's Theory of Action and Time", *Artificial Intelligence*, 42(1990), pp.159-188.

[12] Galton A.: *Logic for Information Technology*, JOHN WILEY & SONS Ltd., 1990.

[13] Hayes P.: "The Naive Physics manifesto", *Expert systems in the microelectronic age*, Michie D. ed., Edinburgh, 1978.

[14] Kahn K. M. and Gorry A. G.: "Mechanizing Temporal Knowledge", *Artificial Intelligence*, 9(1977), pp.87-108.

[15] Knight. B. and Ma. J.: "A General Temporal Model Supporting Duration Reasoning", *AI Communication Journal*, Vol.5, 2(1992), pp.75-84.

[16] Knight B. and Ma J.: "An Extended Temporal System Based on Points and Intervals", *Information System*, Vol.18, 2(1993), pp.111-120.

[17] Kowalski R. A. and Sergot M. J.: "A logic-based calculus of events", *New Generation Computing*, 4(1986), pp.67-95.

[18] Kowalski R.: "Database Updates In The Event Calculus", *The Journal of Logic Programming*, (1992), pp.121-146.

[19] Ma J. and Knight B.: "A General Temporal Theory", *The Computer Journal* (In Press).

[20] McCarthy J.: "Situation, Actions, and Causal Laws, Memo 2, Stanford Artificial Intelligence Project, 1963.

[21] McCarthy J. and Hayes P. J.: "Some Philosophical Problems from the Standpoint of Artificial Intelligence", *Machine Intelligence*, B. Meltzer and D. Michie, (eds.), 4(1969), Edinburgh U.p., pp.463-502.

[22] McDermott D. V.: "A Temporal Logic for Reasoning about Processes and Plans", *Cog. Sci.*, 6(1982), pp.101-155.

[23] Sadri F.: "Three Recent Approaches to Temporal Reasoning", *Temporal Logic and their Applications*, ed. Galton A., Academic Press, 1987, pp.121-168.

[24] Shoham Y.: "Reified Temporal Logics: Semantical and Ontological Considerations", *Advances in Artificial Intelligence - II*, B. Du Boulay, D. Hogg and L. Steels (Editors), Elsevier Science Publishers B. V. (North-Holland), (1987) pp.183-190.

[25] Shoham Y.: "Temporal Logics in AI: Semantical and Ontological Considerations", *Artificial Intelligence*, 33(1987), pp.89-104.

[26] Suppes P.: "A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences", in *The Concept and Role of the Model in Mathematics and Natural and Social Sciences*, D. Reidel Publishing Company (1961).

[27] Vilain M. V.: "A System for Reasoning about Time", *Proc. AAAI-82*, Pittsburgh, PA. 1982, pp.197-201.

[28] Vilain M. B. and Kautz H.: "Constraint Propagation Algorithms for Temporal Reasoning", *Pro. AAAI-86*, 1986, pp.377-382.