Deriving Explanations from Partial Temporal Information

Jixin Ma, Brian Knight, Miltos Petridis

The University of Greenwich
London SE10 9SL, United Kindom
{j.ma, b.knight, m.peridis}@gre.ac.uk

Abstract. The representation and manipulation of natural human understanding of temporal phenomena is a fundamental field of study in Computer Science, which aims both to emulate human thinking, and to use the methods of human intelligence to underpin engineering solutions. In particular, in the domain of Artificial Intelligence, temporal knowledge may be uncertain and incomplete due to the unavailability of complete and absolute temporal information. This paper introduces an inferential framework for deriving logical explanations from partial temporal information. Based on a graphical representation which allows expression of both absolute and relative temporal knowledge in incomplete forms, the system can deliver a verdict to the question if a given set of statements is temporally consistent or not, and provide understandable logical explanation of analysis by simplified contradiction and rule based reasoning.

Keywords: Derivation, Explanation, Temporal Reasoning.

1 Introduction

The notion of time plays a vital and ubiquitous role of a common universal reference. In particular, many Artificial Intelligence systems need to deal with the representation and reasoning about time in modeling natural phenomena and intelligent human activities. Other disciplines such as information systems, cognitive science, linguistics, philosophy and history have the idea of time deeply integrated in as well. For instance, in database systems, time-dependent data may be stamped with time elements, denoting the valid time and/or transaction time of that data [6].

It has been noted that absolute-time-stamping of temporal data provides an efficient indexing method for temporal systems, but suffers from the requirement that precise time values for all temporal data need to be available. Generally speaking, in many AI systems, temporal knowledge can be uncertain and incomplete. For instance, we may only know that event *A* happened before event *B*, without knowing their precise starting and finishing time, or what happened between them. Incomplete relative temporal knowledge such as this is typically derived from humans, where complete and absolute temporal information is rarely available and remembered for knowledge representation and reasoning. Allen's interval-based time theory [1] is a representative example of temporal systems addressing relative temporal relations including "Meets", "Met_by", "Equal", "Before", "After", "Overlaps", "Overlapped_by", "Starts", "Starts_by", "During", "Contains", "Finishes" and "Finished_by". It has been claimed in the literature that time intervals are more suited for expression of common sense temporal knowledge, especially in the domain of linguistics and artificial intelligence. In addition, approaches like that of Allen [1,2] that treat intervals as primitive temporal elements can successfully overcome/bypass puzzles like the *Dividing Instant Problem* [1,4,5,10,11], which is in fact an ancient historical puzzle encountered when attempting to represent what happens at the boundary point that divides two successive intervals. However, as Galton shows in his critical examination of Allen's interval logic [5], a theory of time based only on intervals is not adequate for reasoning correctly about continuous change. In fact, many common sense situations suggest the need for including time points in the temporal ontology as an entity different from intervals. For instance, it is intuitive and convenient to say that instantaneous events such as "The database was updated at 00:00am", "The light was automatically switched on at 8:00pm", and so on, occur at time points rather than intervals (no matter how small they are). Therefore, for general treatments, it is appropriate to include both points and intervals as primitives in the underlying time model, for making temporal reference to instantaneous phenomena with zero duration, and periodic phenomena which last for some positive duration, respectively.

The objective of this paper is to promote a framework to assist referring and deriving logical explanations with respect to partial temporal information. In section 2, a time theory based on both points and intervals as the temporal primitive is introduced. Section 3 presents a graphical representation of partial temporal information. The necessary and sufficient condition for the consistency of a temporal reference is discussed in section 4.1, where section 4.2 and section 4.3 provides explanations to the consistent case and inconsistent case, respectively. An illustrating example as the application of the framework in the legal field of witness evidence is presented in section 4.3. Finally, section 5 concludes the paper.

2 The Time Basis

In this paper, we shall simply adopt the general time theory proposed in [8], which takes a nonempty set, *T*, of primitive time elements, with an *immediate predecessor* relation, Meets, over time elements, and a *duration assignment* function, Dur, from time elements to non-negative real numbers. If Dur(t) = 0, then t is called a point; otherwise, that is Dur(t) >0, t is called an interval. The basic set of axioms concerning the triad (*T*, Meets, Dur) is given as below [8]:

T1. $\forall t_1,t_2,t_3,t_4(\text{Meets}(t_1, t_2) \wedge \text{Meets}(t_1, t_3) \wedge \text{Meets}(t_4, t_2) \Rightarrow \text{Meets}(t_4, t_3))$

That is, if a time element meets two other time elements, then any time element that meets one of these two must also meets the other. This axiom is actually based on the intuition that the "place" where two time elements meet is unique and closely associated with the time elements [3].

T2. $\forall t \exists t_1,t_2(\text{Meets}(t_1, t) \wedge \text{Meets}(t, t_2))$

That is, each time element has at least one immediate predecessor, as well as at least one immediate successor.

T3. $\forall t_1,t_2,t_3,t_4(\text{Meets}(t_1, t_2) \wedge \text{Meets}(t_3, t_4) \Rightarrow$
$\text{Meets}(t_1, t_4) \nabla \exists t'(\text{Meets}(t_1, t') \wedge \text{Meets}(t', t_4)) \nabla \exists t''(\text{Meets}(t_3, t'') \wedge \text{Meets}(t'', t_2)))$

where ∇ stands for "exclusive or". That is, any two meeting places are either identical or there is at least a time element standing between the two meeting places if they are not identical.

T4. $\forall t_1,t_2,t_3,t_4(\mathrm{Meets}(t_3, t_1) \wedge \mathrm{Meets}(t_1, t_4) \wedge \mathrm{Meets}(t_3, t_2) \wedge \mathrm{Meets}(t_2, t_4)) \Rightarrow t_1 = t_2)$

That is, the time element between any two meeting places is unique.

N.B. For any two adjacent time elements, that is time elements t_1 and t_2 such that Meets(t_1, t_2), we shall use $t_1 \oplus t_2$ to denote their ordered union. The existence of such an ordered union of any two adjacent time elements is guaranteed by axioms T2 and T3, while its uniqueness is guaranteed by axiom T4.

T5. $\forall t_1,t_2(\mathrm{Meets}(t_1, t_2) \Rightarrow \mathrm{Dur}(t_1) > 0 \vee \mathrm{Dur}(t_2) > 0)$

That is, time elements with zero duration cannot meet each other.

T6. $\forall t_1,t_2(\mathrm{Meets}(t_1, t_2) \Rightarrow \mathrm{Dur}(t_1 \oplus t_2) = \mathrm{Dur}(t_1) + \mathrm{Dur}(t_2))$

That is, the "ordered union" operation over time elements is consistent with the conventional "addition" operation over the duration assignment function, i.e., Dur.

Analogous to the 13 relations introduced by Allen for intervals [1,2], there are 30 exclusive temporal order relations over time elements including both time points and time intervals, which can be derived from the single Meets relation and classified into the following 4 groups:

- Order relations relating a point to a point:
  {Equal, Before, After}
- Order relations relating an interval to an interval:
  {Equal, Before, After, Meets, Met_by, Overlaps, Overlapped_by, Starts, Started_by, During, Contains, Finishes, Finished_by}
- Order relations relating a point to an interval:
  {Before, After, Meets, Met_by, Starts, During, Finishes}
- Order relations relating an interval to a point:
  {Before, After, Meets, Met_by, Started_by, Contains, Finished_by}

3 A Graphical Representation

In general, the temporal order relation between two time elements can be any one of those 30 as classified in section 2. However, as shown in [8], analogous to Allen and Hayes's approach [3], all the temporal can be defined as derived relations in terms of the single "Meets" relation. In fact, such definitions are straightforward. For example, "Before" can be defined as:

$$\mathrm{Before}(t_1, t_2) \Leftrightarrow \exists t(\mathrm{Meets}(t_1, t) \wedge \mathrm{Meets}(t, t_2))$$

Therefore, in general, a set of temporal statements, which we shall call a scenario, can be expressed as a collection of propositions, together with the corresponding temporal reference denoted as a triad (T, M, D), where T is a collection of time elements, expressing the knowledge of what time elements are involved with respect to the given scenario (possibly incomplete); M is a collection of Meets relations over T, expressing the knowledge (possibly incomplete) as to how the time elements in T are related by means of the primitive immediate predecessor relationship; and D is a collection of duration assignments (possibly incomplete) to time elements in T.

A temporal reference (T, M, D) can be intuitively expressed in terms of a directed and partially weighted graph [7], in which time elements are denoted as arcs, and the immediate predecessor relationship over times is denoted by the node structure where Meets(t_i, t_j) is represented by t_i being an in-arc and an t_j being out-arc to a common node, respectively. For time elements with known duration, the corresponding arcs are weighted by their durations respectively. For example, consider temporal reference (T, M, D), where

T = {t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9};
M = {Meets(t_1, t_2), Meets(t_1, t_3), Meets(t_2, t_5), Meets(t_2, t_6), Meets(t_3, t_4), Meets(t_4, t_7), Meets(t_5, t_8), Meets(t_6, t_7), Meets(t_7, t_8);
D = {Dur(t_2) = 1, Dur(t_4) = 0.5, Dur(t_6) = 0, Dur(t_8) = 0.3}

The graphical representation of temporal reference (T, M, D) is shown in Fig. 1:

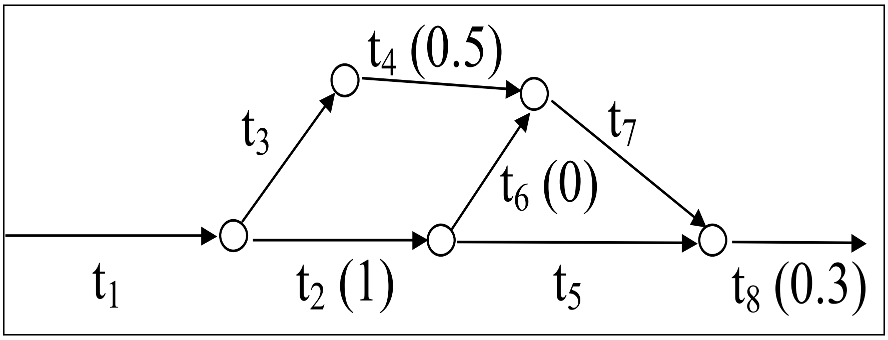

Fig. 1. Graph representation of (T, M, D)

4 Deriving Verdicts and Explanations

As mentioned in the above, temporal information in general can be uncertain and incomplete in various ways. Also, for a given scenario, the corresponding temporal reference itself may be temporal inconsistent. Therefore, first of all, a mechanism which can provide the verdict as for the question if a temporal reference is temporal consistent or not is expected.

4.1 Checking the temporal consistency

The necessary and sufficient condition for the consistency of a general temporal reference, (T, M, D), can be given as below (see details in [7]):

1) For each simple circuit in the graph of (T, M, D), the directed sum of weights is zero;
2) For any two adjacent time elements, the directed sum of weights is bigger than zero.

Here, condition 1) guarantees that there exists a valid duration assignment function Dur to the time elements in T agreeing upon D; and condition 2) ensures that no two time points meet each other, that is between any two time points, there is an interval standing between them [8].

The consistency checking for a temporal reference with duration constraints involves searching for simple circuits, and constructing a numerical constraint for each circuit. The existence of a solution(s) to this set of constraints implies the consistency of the system, and each solution gives a possible case for the corresponding temporal scenario that can subsume the addressed temporal reference. Hence, the consistency checker for a random temporal reference is in fact a linear programming problem.

In fact, in the graph presented in Fig. 1 of section 3, there are two simple circuits as shown in Fig. 2.

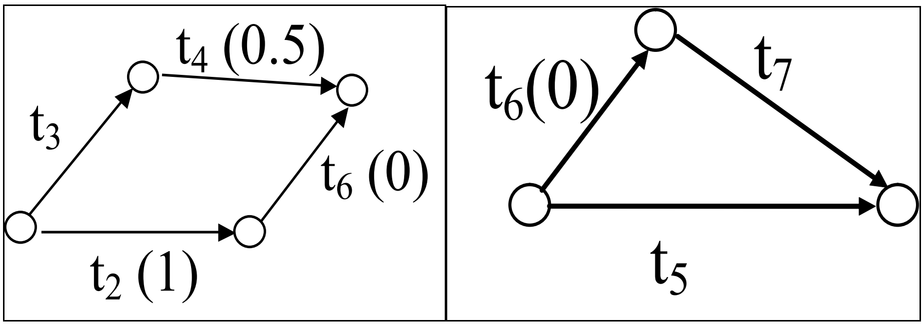

Fig. 2. The two simple circuits

Setting the directed sum of weights in each of these two circuits as 0, we get 2 independent constraints:

$$Dur(t_2) + Dur(t_6) = Dur(t_3) + Dur(t_4)$$
$$Dur(t_5) = Dur(t_6) + Dur(t_7)$$

We can easily find a solution, for instance: $Dur(t_3) = 0.5$, $Dur(t_5) = Dur(t_7) = 1$. Actually, the duration assignment to t_5 and t_7 can be any positive real number, provided that $Dur(t_5) = Dur(t_7)$.

In some special cases where only relative temporal knowledge are addressed, that is there is no duration constraint involved, temporal reference (T, M, D) is reduced to a pair (T, M) and the consistency checking can be reformulated in a simpler form. In fact, (T, M) is consistent if and only if:

1)' There are no nodes with at least one point in-arc and at least one point out-arc;
2)' The associated reduced graph is acyclic, where the associated reduced graph is formed by means of removing every point arc in the graph of (T, D), and merging any two nodes connected by the point arc.

Here, again, condition 1)' preserves that no two time points meet each other, while condition 2)' preserves that time points are not decomposable, and excludes any circular time structure.

For example, consider a relative temporal reference (T, M), where

$T = \{t_1, t_2, t_3, t_4, t_5, t_6\}$
$M = \{Meets(t_1, t_2), Meets(t_1, t_3), Meets(t_2, t_6),$
$Meets(t_3, t_4), Meets(t_4, t_5), Meets(t_5, t_6)\}$

The graphical representation of temporal reference (T, M) is shown in Fig. 3:

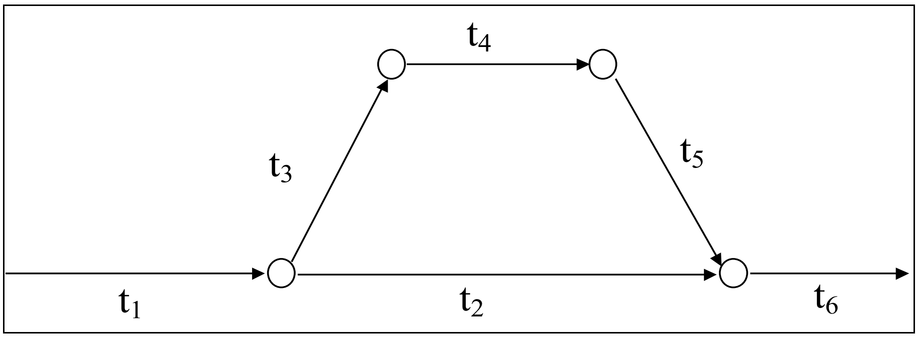

Fig. 3. Graph representation of (T, M)

If t_2 is not known to be a time point, then the corresponding graph shown in Fig. 3 is acyclic, and hence the temporal reference is consistent.

However, if t_2 is stated to be a point, then the graph in Fig. 3 is reduced to the graph as shown in Fig. 4.

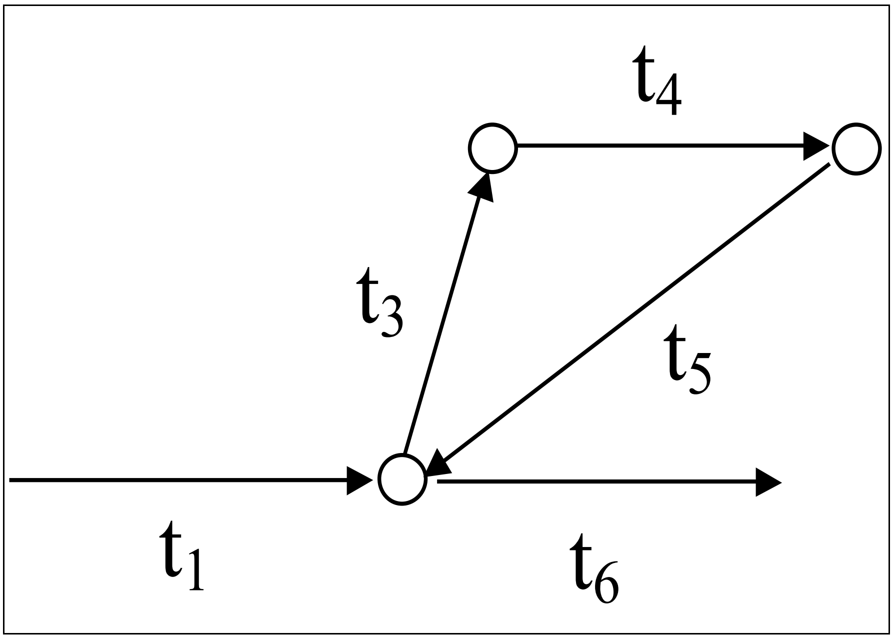

Fig. 4. The reduced graph

In the reduced graph in Fig. 4, there is a cycle, i.e., t_3 -> t_4 -> t_5 -> t_3. Therefore the temporal reference is inconsistent.

Now, further investigations are needed to explain what does the verdict that a give scenario is temporal consistent or inconsistent mean in the context.

4.2 Explanation to the consistent case

As mentioned in section 4.1, the consistency checking for a general temporal reference is in fact a linear programming problem, where each solution to the linear programming problem gives a possible case for the corresponding temporal scenario that can subsume the addressed temporal reference. In the case where the temporal reference is consistent, there exists at least one solution to the linear programming problem. Of course, if the solution(s) is unique, we can use this solution construct the corresponding complete temporal reference which is also unique.

However, in general cases where a verdict that the temporal reference is consistent has been reached, there may be more than one, or even an infinite number of solutions to the corresponding linear programming. This may be due to various forms of incompleteness of the corresponding temporal reference, e.g., some referencing time elements may be missing, the duration of some time elements may be unknown, and so on. Therefore, we can only construct the possible complete scenarios which can subsume the addressed temporal reference.

In this case, we can at least find the minimal model(s) among these complete scenarios by means of defining and calculating the similarity degree between the complete temporal references and the original partial temporal reference.

Since each temporal reference can be expressed as a directed and partially weighted graph, the problem of matching temporal references can be transformed into conventional graph matching.

4.3 Explanation to the inconsistent case

In the case where a verdict that the temporal reference is inconsistent has been reached, we can simply analyse and identify the linear equations which make the corresponding linear programming unsolvable, which, in turn, will identify and explain which part(s) of the temporal reference actually leads to the inconsistency.

4.4 An illustrating example

In As an example, consider the situation where two persons, Peter and Jack, are suspected of committing a murder during the daytime. In court, Jack and Peter gave the following statements, respectively:

- Peter's statements:

  I got home with Jack before 1pm. We had our lunch, and when Jack left I put on a video. The video lasts 2 hours. Before it finished, Robert arrived. When the video finished we went to the train station and waited until Jack came at 4 pm.

- Jack's statements:

  Peter and me went to his home and arrived there before 1pm. When we finished our lunch there, Peter put on a video, and I left and went to the supermarket. I stayed there for between 1 and 2 hours. Then I drove to my home to collect some mail. It takes between 1.5 to 2 hours to reach my home, and about the same to the train station. I arrived at the train station at 4 pm.

- In addition, being a witness, Robert made these statements:

  I left home at 2 pm and went to Peter's house. He was playing a video, and we waited till it finished. Then we went together to the train station and waited for Jack. Jack got to the train station at 4pm.

The temporal reference of the above temporal information involves the following time elements:

- i_1: the time (interval) over which Peter and Jack went to Peter's home;
- 1pm: the time (point) before which they arrived at Peter's home;
- i_2: the time (interval) over which Peter and Jack had lunch;
- i_3: the time (interval) over which Peter played the video ($Dur(i_3) = 2$);
- i_4: the time (interval) over which Jack went to the supermarket;
- p_1: the time (point) when Robert arrived at Peter's house;
- i_5: the time (interval) over which Peter and Robert went to the train station;
- i_6: the time (interval) over which Peter and Robert waited for Jack at the train station;
- 4pm: the time (point) when Jack arrived at the train station;
- i_7: the time (interval) over which Jack stayed in the supermarket ($1<Dur(i_7)< 2$);
- i_8: the time (interval) over which Jack drove to his home ($1.5<Dur(i_8)< 2$);
- i_9: the time (interval) over which Jack collected some post from his home;
- i_{10}: the time (interval) over which Jack drove to the train station ($1.5<Dur(i_{10})<2$);
- 2pm:the time (point) when Robert left home;
- i_{11}: the time (interval) over which Robert went to Peter's house;
- i_{12}: the time (interval) over which Peter and Robert watched the video together;
- i_{13}, i_{14}, ..., i_{27}: some extra relative time elements which are used for expressing the correspondingly relative duration knowledge, e.g., with i_{19}, i_{20}, i_{21}, i_{22}, and $Dur(i_{19}) = 1.5$ and $Dur(i_{21}) = 2$, we can express that $1.5 < Dur(i_8) < 2$ (Picture 3)

The graphical representation of the corresponding temporal reference for the above legal statements can be shown as Fig. 5 as below:

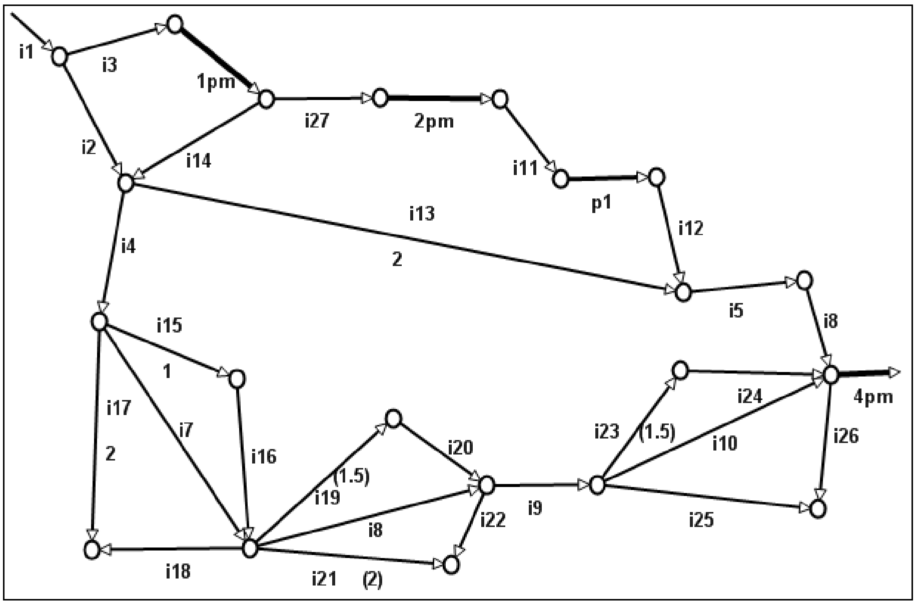

Fig. 5. (T, M, D) of the legal statements

From Fig. 5, we see that there are three time elements (i.e., two intervals, i_{11} and i_{12}, and one points, p_1) standing between 2pm and 4pm. Since each interval has a positive duration and each point has a non-negative duration, we can infer that:

$$Dur(i_5) + Dur(i_6) < 2$$

in addition, since $Dur(i_3) = 2$, hence

$$Dur(i_3) + Dur(i_5) + Dur(i_6) < 2 + 2 = 4$$

However,

$$Dur(i_4) + Dur(i_7) + Dur(i_8) + Dur(i_9) + Dur(i_{10}) > 0 + 1 + 1.5 + 0 + 1.5 = 4$$

Therefore, for the simple circuit, i.e., i_3, i_5, i_6, i_{10}, i_9, i_8, i_7, i_4, as shown below in Fig. 6, there does not exist any duration assignment over T such that

$$Dur(i_3) + Dur(i_5) + Dur(i_6) = Dur(i_4) + Dur(i_7) + Dur(i_8) + Dur(i_9) + Dur(i_{10})$$

In other words, there is no solution to the following linear equation:

$$Dur(i_3) + Dur(i_5) + Dur(i_6) - Dur(i_4) - Dur(i_7) - Dur(i_8) - Dur(i_9) - Dur(i_{10}) = 0$$

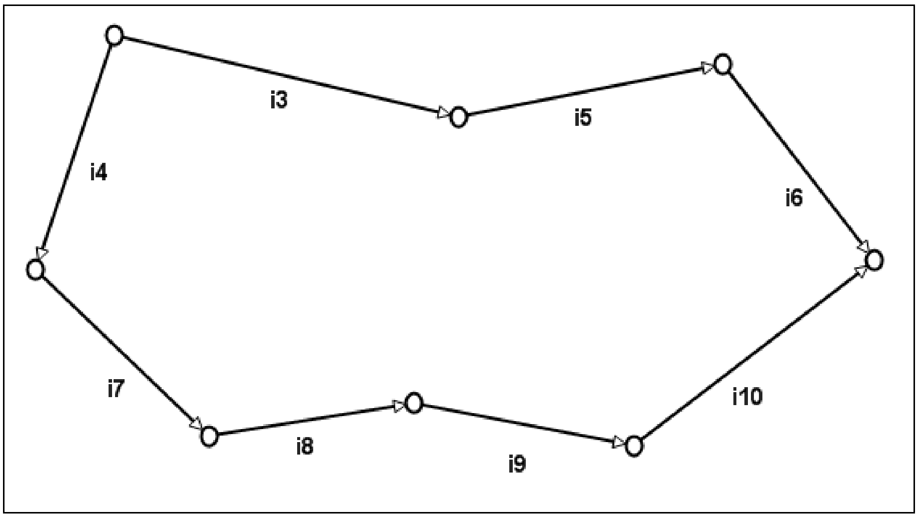

Fig. 6. A simple circuit in the legal statements

Hence, the temporal reference shown in Fig. 6 is inconsistent, and therefore we can directly confirm that some statements are untrue.

Suppose the video can be checked that it did actually last for two hours, we can confirm that there must be some falsity in either Robert's or Jack's statements. If it can be proved that Robert did leave home at 2 pm, then Jack must have lied in making his statements. Otherwise, to convince the jury that his statements are true, Jack must prove that Robert left home at some time before 2 o'clock in the afternoon.

5 Conclusion

In this paper, we introduced an inferential framework for temporal representation and temporal reasoning. It allows expression of both absolute and relative temporal knowledge, and provides graphical representation of temporal references in terms of directed and partially weighted graphs. Based on the temporal reference of a given scenario with partial temporal information, the framework can check if it is temporally consistent or inconsistent, and derive the corresponding explanations. The benefit of this approach is that the inferential framework has powerful analytic abilities, and its analysis is amenable to human scrutiny.

References

1. Allen, J.: Maintaining knowledge about temporal intervals, Communications of the ACM, 26 (11), 832-843, (1983).
2. Allen, J.: Towards a General Theory of Action and Time, Artificial Intelligence, 23, 123-154 (1984).

3. Allen, J., Hayes, P.: Moments and Points in an Interval-based Temporal-based Logic, Computational Intelligence, 5, 225-238 (1989).
4. van Benthem, J.: The Logic of Time, Kluwer Academic, Dordrech (1983).
5. Galton, A.: Critical Examination of Allen's Theory of Action and Time, Artificial Intelligence, 42, 159-188 (1990).
6. Jensen, J., Clifford, J., Gadia, S., Segev, A., Snodgrass, R.: A Glossary of Temporal Database Concepts, SIGMOD RECORD, 21(3), 35-43 (1992).
7. Knight, B., Ma, J.: A General Temporal Model Supporting Duration Reasoning, Artificial Intelligence Communication, 5(2), 75-84 (1992).
8. Ma, J., Knight, B.: A General Temporal Theory, the Computer Journal, 37(2), 114-123 (1994).
9. Ma, J., Knight, B.: A Reified Temporal Logic, the Computer Journal, 39(9), 800-807 (1996).
10. Ma, J., Knight, B.: Representing The Dividing Instant, the Computer Journal, 46(2), 213-222 (2003).
11. Vila, L.: A Survey on Temporal Reasoning in Artificial Intelligence, AI Communication**,** 7(1), 4-28 (1994).