Temporal Pattern Mining and Reasoning using Reference Event based Temporal Relations (RETR)

V. Uma[1] and Dr. G. Aghila[2]
[1]Assistant Professor, [2]Professor,
Department of Computer Science,
Pondicherry University, India

ABSTRACT
Reasoning and mining over temporal patterns has to be more specific and accurate for efficient knowledge discovery. The patterns expressed with the temporal interval relationships of Allen have the problem of aggregating many events under one relation "before" and hence leads to ambiguity. In this work an attempt has been made to overcome this by augmenting the "before" relation with four contextual Reference Event based Temporal Relations (RETR) so that temporal ordering of events can be identified more efficiently and hence resulting in effective knowledge discovery. Moreover, Roddick refined temporal relations by considering the midpoints for equal length intervals wherein the temporal relations proposed in this paper extending the "before" relation can be applied both to equal, unequal length intervals with midpoints. The superiority of this novel form of representing temporal knowledge can be proved by incorporating these topological temporal relations in a time ontology which eventually would result in efficient reasoning. This has been demonstrated by presenting a real life data set from medical domain.

Keywords
Allen Interval Algebra, Knowledge representation, Temporal reasoning and mining, Knowledge discovery.

1. INTRODUCTION
The representation of temporal knowledge and its reasoning is very much required in linguistics, artificial intelligence and philosophy. Allen [1] proposed Interval Algebra (IA) network with 13 basic relations for representing and reasoning temporal information between two intervals. The temporal relations specified in Allen's logic are before(<), meets(m), starts(s), finishes(f), overlaps(o), during(d), after(>), met-by(mi), started-by(si), finished-by(fi), overlapped-by(oi), contains(di) and equal(=). The vital areas of Artificial Intelligence such as scheduling, planning and NLP requires temporal concepts to be represented and reasoned out more efficiently [2], and this is particularly evident in the field of medicine where the temporal representation and reasoning is crucial for decision support-related tasks [3, 12]. Reasoning with incomplete temporal knowledge [4], refinement of "equal" [5] "Overlap" [6] relations, introduction of "coincide" operator [7], proposal of "containment" relationship [8] were the extensions made to represent temporal knowledge.

However, some extension to Allen interval algebra is required to overcome the vagueness, ambiguity present in mining the relation between two events [9] particularly when they have the temporal relation "before". The meaning of relationship "before" is context dependent and hence it is difficult to quantify and resolve [10]. Therefore the ambiguous and context dependent relationship has to be represented efficiently to deduce the facts and temporal relationship among the events successfully. The reference event based temporal relations (RETR) has been proposed in this work as an extension to "before" relation by linking the events to a reference event. The relationship is determined based on how the events under considerations are related to the reference event. These relationships will not propagate as they are contextual. These RETR would be crucial in temporal pattern mining and reasoning as it helps in efficient representation of temporal relations.

The Allen's temporal relations along with the RETR can be used to represent the relations between events in time ontology. Time ontology is a explicit specification of a conceptualization of temporal knowledge. In this work, the benefit that would result in medical diagnosis with the implementation of additional contextual relations has been demonstrated as a case study. The events in the case study can be considered as concepts and the temporal relations between them can be represented in time ontology. In the field of medicine using the time ontology, decision support can be efficiently provided by means of temporal reasoning.

The remainder of this paper is organised as follows. The related works concerning tine interval representation are reviewed in section 2. Section 3 discusses the potential problem of Allen IA and introduces the concept of Reference event based refinement of temporal relations. Section 4 evaluates the new approach with an illustration of its execution over a medical dataset. Section 5 explains the significance of the new relations when incorporated in time ontology. Section 6 provides concluding remarks and future work.

2. RELATED WORKS
Allen[1] proposed nonoverlapping set of 13 interval-interval relationships to describe relative positioning of two intervals. Allen's temporal relations was modified to reason with incomplete knowledge, specifically with coarse knowledge about temporal relationships by Freksa[3]. He represented knowledge about time in terms of relationships between beginnings and endings of events called "semi intervals" to accommodate vagueness in one of the end points. Compared to Allen only 2 time points were required to represent two intervals rather than four. Older, younger, head to head, tail to tail, survives, survived by, born before death, died afterbirth, precedes, succeeds were the relations specified by Freksa.

Ultsch[4] proposed Unification based temporal Grammar(UTG) a hierarchical pattern language. The UTG event patterns are more robust than Allen's relations, by matching interval boundaries with a threshold. This method combined more or less simultaneous intervals into Allen's "equal" relation by using chord length which guided the mining process in filtering out irrelevant information. Moerchen[5] modified UTG and introduced an operator coincide to describe coincidence of intervals. He proposed the Time series knowledge representation (TSKR) and proved that the patterns created, better explained the semantics of the temporal concepts and conceptually helped in pattern mining with more efficiency and accuracy.

Villafane[8] proposed containment relationship for a series of interval events and the algorithm to mine the relationship. Arun K.Pujari proposed an Interval and duration network[9] with 25 basic relations extending Allen's IA to handle qualitative information of time interval and duration. {b,bi,m,mi,o,oi} relations was extended with {<,=,>} to relate the durations of intervals.

Wu et al[6] has pointed out two ambiguous problems of Allen IA. First is representation of same relationship among events by different temporal pattern and the second disadvantage is same temporal pattern can represent different relationship among events.

To overcome the ambiguity problem Roddick[7] extended Allen interval algebra's 13 relations to 49 relations by considering the midpoints of relations along with their endpoints. The expanded set divides overlaps relationship into nine different types, during relationship into seven and starts and finishes relationship into three each. Considering their inverses 49 interval-interval relationships with midpoints were proposed for intervals of varying lengths. For intervals of identical lengths 11 relationships were proposed by extending the overlaps and overlapped by relationship.

The extended overlap relations has been considered for equal, non equal intervals in this work as the start time of the intervals only play a major part in extending the "before" relation.

Ontology of time is gaining importance as it is necessary to realize the Semantic Web. They play an important role in Natural Language Processing applications, data mining, decision support systems etc [11]. In the existing time ontologies in DAML[13], KSL [14], OWL [15], KIF [16], Cyc [17] the topological temporal relations between events are represented efficiently.

In this paper an attempt has been made to analyse the "before" relation in-depth to prove that the extensions made improves the efficiency of temporal pattern mining and reasoning when incorporated in time ontology.

The next section describes the proposed idea of extending Allen's interval algebra by augmenting the "before" relationship.

3. REFERENCE EVENT BASED TEMPORAL RELATIONS (RETR) - REFINEMENT OF ALLEN'S INTERVAL ALGEBRA (IA)

Allen's IA describes the "before" relation as an interval relation between two events (A,B) if $A_{endtime} < B_{starttime}$. The figure 1 shows the state interval sequence diagrammatically.

Event Name	Start Time	End Time
A	4	8
B	10	14

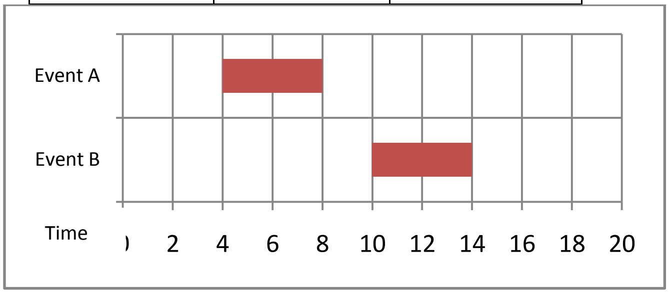

Figure 1: Diagrammatic representation of Allen's "before" relation

The relationship "before" persists between A and B for any difference between the end time of A and start time of B as it does not capture the implicit meaning of the "before" relation. Figure 2 shows the presence of ambiguity in Allen's IA.

Event Name	Start Time	End Time
A	4	8
B	12	16
C	14	18

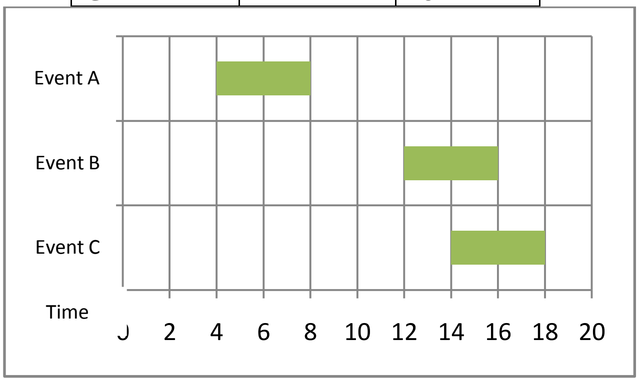

Figure 2: Diagrammatic representation of "before" relation highlighting the vagueness in Allen's IA

Figure 2 clearly shows that $A_{endtime}<B_{starttime}$. So again the temporal relation between A and B is "before". It doesn't reflect the increase in time duration between the events A and B. The relation between A and C is also "before" as $A_{endtime}<C_{starttime}$. The same "before" relation has been represented by different temporal patterns and the presence of ambiguity in Allen IA can be established.

Thus "before" relation doesn't differentiate events occurring after a short time or long time. This can be resolved if events are considered with respect to the context or otherwise reference events. Figure 3 shows the relationship between A and B by considering reference event C.

Event Name	Start Time	End Time
A	4	8
B	11	15
C	10	14
D	20	24
E	12	16

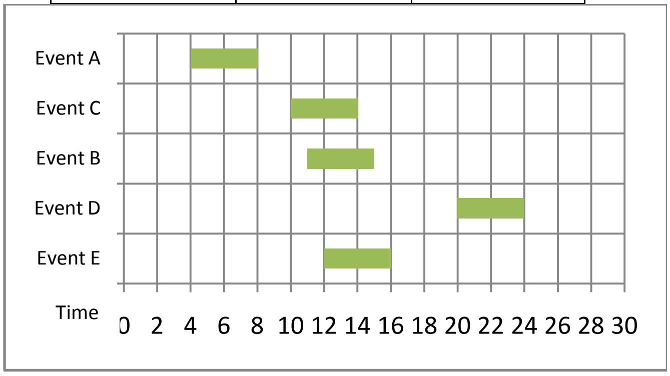

Figure 3: State interval sequence representation of the events using reference event

The Figure 3 clearly shows that the event A has occurred "shortly before B", "extremely before" D and "quite before" E. This distinction can be made only with a reference event C. In Roddick's interval relation [6] midpoints of intervals were used to extend "overlap" relation. By considering the midpoint of the reference event, "before" relation can be extended to "shortly before(sb)", "quite before(qb)", "long before(lb)" and "extremely before(eb)".

In Figure 3 considering C as the reference event, A can be related to B as "A shortly before B" since event B starts before the C_{mid}, whereas E starts at C_{mid} and so "A quite before E" and D starts after end of C and so A is "extremely before D".

BrC / ArB	<	m	so	mo	lo
<	<	<	<	<	<
m	<	<	<	<	<
so	<	<	<	<	<,m,so
mo	<	<	<	m	so
lo	<	<	<,m,so	so	so,mo,lo

Figure 4: Trasitivity Matrix for Allen IA

The above transitivity matrix in figure 4 shows the constraint propagation matrix for some of the Allen's temporal relations for equal length intervals [6]. The matrix elements clearly show that most of the relations between A and C ends up in "before" relation. The figure 5 shows the transitivity matrix when a reference event and RETR are considered.

RefrE2 / E1r Ref	<	m	so	mo	Lo
<	<(eb)	<(eb)	<(lb)	<(qb)	<(sb)
M	<(eb)	<(eb)	<(lb)	<(qb)	<(sb)
So	<(eb)	<(eb)	<(lb)	<(qb)	<(sb),m,so
Mo	<(eb)	<(eb)	<(lb)	m	So
Lo	<(eb)	<(eb)	<(lb),m ,so	so	so,mo,lo

Figure 5: Transitivity matrix with augmented "before" relation of Allen

Figure 5 shows that relations between events E1, E2 has been specified with the additional temporal relations and hence temporal sequence can be mined more efficiently. If the reference event based extension of "before" relation is not considered most of the relations will be identified as "before" which would not be efficient when considered with respect to temporal pattern mining. Only 5 relations of Allen(<,m,so,mo,lo) are considered in figure 5 as the new temporal concepts does not contribute to other relations. Propagation of relations doesn't hold for new temporal concepts as they are context dependent.

The temporal relationships identified by considering C as reference event for the state interval sequence specified in figure 3 is shown in figure 6. Events B,D,E are with "after" relation with event A . The new extensions to "before" relation can provide more information about the relationship between temporal events.

CrE2 / E1rC	A	B	C	D	E
A	=	<(jb)	<	<(eb)	<(qb)
B	>(ja)	E	Loi	<	So
C	>	Lo	E	<	So
D	>(ea)	>	>	E	>
E	>(qa)	Soi	soi	<	E

Figure 6: Temporal Relationship between the events with reference event C

The distinction between the relationship of event A with events B,D,E is possible only by using RETR added to "before" relation.

Figure 5 above showed temporal relations for equal length intervals. When unequal length intervals are considered the matrix formed is shown in figure 7. Figure 7 shows that when reference event has a relation before or meets with E1, and when reference event has any of the relation before, meets, small overlap, medium overlap, large overlap, during (with start time beforemid, atmid, aftermid) with E2 then the additional temporal patterns can contribute valuable information for unequal intervals.

<table>
<tr><td>Ref r E2 / E1 r Ref</td><td><</td><td>M</td><td>so,during
(ref.mid<E2.start)
&&((ref.end<E2.end)||
(ref.end>E2.end)</td><td>mo,during
(ref.mid=E2.start)
&&((ref.end<E2.end)||
(ref.end>E2.end)</td><td>lo,during
(ref.mid>E2.start)
&&((ref.end<E2.end)||
(ref.end>E2.end)</td></tr>
<tr><td><</td><td><(eb)</td><td><(eb)</td><td><(lb)</td><td><(qb)</td><td><(sb)</td></tr>
<tr><td>M</td><td><(eb)</td><td><(eb)</td><td><(lb)</td><td><(qb)</td><td><(sb)</td></tr>
</table>

Figure 7: Transitivity matrix for unequal length intervals

4. CASE STUDY

A medical dataset shown in Table 1 has been taken for proving the significance of the extended temporal relations. Temporal mining and reasoning in medicine is a challenging area. Information extraction in medical datasets requires robust mining and reasoning technique.

Table 1. Patient details

Patient Id	Disease(Event) name	Event code	Start Time (days)	End Time (days)
1	Attack of Jaundice	A	10	25
1	Suffering from Abdominal Pain	B	10	25
1	Blood test	C	26	27
1	Treatment for Jaundice	D	28	35
1	Cancerous growth	E	70	200
1	Pain in upper abdomen	F	95	210
1	Passing of high colored urine	G	100	210
1	Sense of Indigestion	H	135	220
1	Loss of weight	I	170	260
1	Deepening of Jaundice	J	180	202
1	Onset of Pruritis	K	180	205
1	Investigation for Jaundice	L	190	192
1	Palliative Surgery&Biopsy	M	195	198
1	Oncology treatment	N	200	290

The pictorial representation of the history of the patient is given in Figure 8 which helps in visualising the relative position of the intervals. The above dataset be the input to the decision support system considered here.

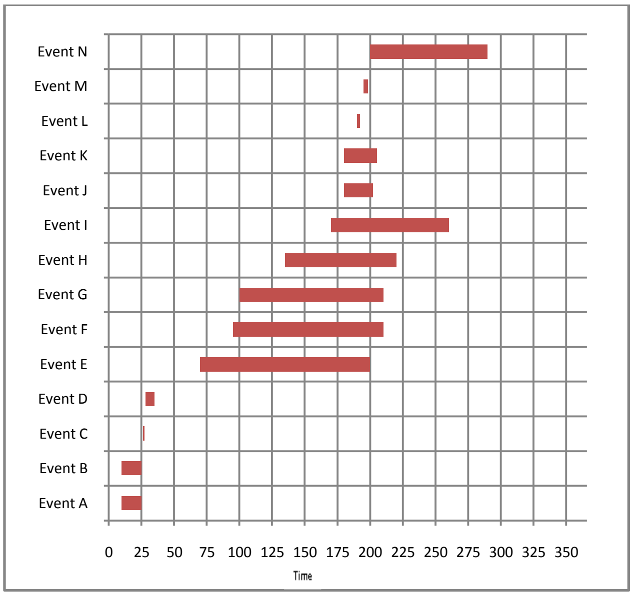

Figure 8: Relative Position of events in medical dataset

5. RETR IN TIME ONTOLOGY

The above dataset can be represented using time ontology. The events are the concepts and the relations between the concepts are temporal. The events are associated with properties like start time and end time. This can be implemented in any of the time ontologies in DAML, OWL, Cyc, KSL, KIF. The Table 2 shows the list of queries and the answers provided by different time ontologies and the answer that would be provided if RETR relations are incorporated in any of the above time ontologies.

So it is possible to identify the relative occurrence of the events within a reference event with the extended “before” relation which would lead to more clear temporal semantics enhancing temporal reasoning.

6. CONCLUSION

New temporal relations have been proposed for the formulation of temporal knowledge. It is better interpretable than Allen’s “before” relation. This novel concept of reference event based relations identifies the temporal relationships more efficiently in multivariate data and contributes to identify linear ordering of events. The efficacy of mining and reasoning the proposed relations has been demonstrated to be more meaningful than Allen’s “before” relations with a real life data.

Future enhancement is to incorporate the temporal operators in temporal logic to identify the truth of facts by temporal reasoning and also these relations are to be used as a base for constructing time ontology.

7. REFERENCES

[1] J.F. Allen, "Maintaining knowledge about temporal intervals", *Comm. ACM,* vol. 26, no.11, pp. 832–843, 1983.

[2] Vineet Padmanabhan, Abdul Sattar, ArunK. Pujari and Chinmoy Goswamy, "Temporal Reasoning: A ThreeWay Analysis", *Proceedings of the International workshop on Temporal representation and reasoning*, pp. 183-189, 2000.

[3] M. Campos, J.M. Juarez, J. Palma, R. Marin and F. Palacios, "Avian influenza: Temporal modeling of a human to human transmission case", *International Journal of Expert Systems with Applications,*vol.38, no. 7, pp. 8865-8885, July 2011.

[4] C. Freksa, "Temporal reasoning based on semi-intervals", *Artificial Intelligence*, vol. 54, pp. 199–227, 1992.

[5] A.Ultsch, "Unification-based temporal grammar", *Technical Report 37*, CS Dept., Philipps-University Marburg, Germany, 2004.

[6] J.F. Roddick and C. Mooney, "Linear temporal sequences and their interpretation using midpoint relationships", *IEEE Transactions on Knowledge and Data Eng*, vol. 17, no.1,pp. 133–135, 2005.

[7] F. Mörchen, "A better tool than Allen's relations for expressing temporal knowledge in interval data", *12th ACM SIGKDD international conference on knowledge discovery and data mining*. pp. 25–34, 2006.

[8] R. Villafane, K. Hua, D. Tran and B. Maulik, "Knowledge discovery from series of interval events", *Journal of Intelligent Information Systems*, vol. 15, no.1, pp.71–89, 2000.

[9] S. Wu and Y. Chen, "Mining nonambiguous temporal patterns for interval-based events", *IEEE Trans. On Knowledge and Data Engineering*, vol. 19, no.6, pp. 742-758, June2007.

[10] K.M. Kahn and G.A. Gorry, "Mechanizing Temporal Knowledge," *Artificial Intelligence* vol. 9, no. 1, pp. 87-108, Aug 1977.

[11] C.Zhang, C.Cao, Y.Sui, X.Wu, "A Chinese Time Ontology for Semantic Web", *Knowledge-Based Systems* vol. 24, no. 7, pp. 1057-1074, October 2011.

[12] J.C.Augusto, "Temporal Reasoning for decision support in Medicine", *Artificial Intelligence in Medicine,* vol. 33,no. 1, January 2005.

[13] DAML-TimeHomepage, 2003. http://www.cs.rochester.edu/ ~ferguson/daml/.

[14] KSL-Time,2008. http://www.ksl.stanford.edu/ontologies/time.

[15] Time Ontology in OWL, 2006. http://www.w3.org/TR/owl-time/.

[16] Time Ontology in KIF, 1994. <http://www-cs-students.stanford.edu/~omart/timeont.html>.

[17] Time and Dates, 2002. http://www.cyc.com/cycdoc/vocab/time-vocab.html.

Table 2. Queries and the comparative results

S.No	**Query**	**Answer from DAML, OWL,Cyc, KIF, KSL**	**Answer if RETR incorporated**	**Remarks**
1	List all events that occurred "after" event D	E,F,G,H,I,J,K,L,M,N	E,F,G,H,I,J,K,L,M,N	As "after" is an Allen interval relation both methods give the same answer. Gives no idea about the ordering of events.
2	List all events that occurred "shortly after" event D with event E as reference event.	No result	F,G	Time ontology when incorporated with RETR Relations provides more specific results.
3	List all events that occurred "quite after" event D with event E as reference event.	No result	H	The fact that events F,G has started occurring before event H and so the relations <s,=,si,oi,di,mi,si, fi> does not hold between them can be reasoned out.
4	List all events that occurred "long after" event D with event E as reference event.	No result	I,J,K,L,M	The fact that events F,G,H has started occurring before I,J,K,L,M and so the relations <s,=,si,oi,di,mi,si, fi> does not hold between them can be reasoned out.
5	List all events that occurred "extremely after" event D with event E as reference event.	No result	N	The fact that events F,G,H,I,J,K,L,M has started occurring before N and so the relations <s,=,si,oi,di,mi,si, fi> does not hold between them can be reasoned out.