Concise Papers

Linear Temporal Sequences and Their Interpretation Using Midpoint Relationships

John F. Roddick and Carl H. Mooney

Abstract—The temporal interval relationships formalized by Allen, and later extended to accommodate semi-intervals by Freksa, have been widely utilized in both data modeling and artificial intelligence research to facilitate reasoning between the relative temporal ordering of events. In practice, however, some modifications to the relationships are necessary when linear temporal sequences are provided, when event times are aggregated, or when data is supplied to a granularity which is larger than required. This paper discusses these modifications and outlines a solution to this problem which accommodates any available knowledge of interval midpoints.

Index Terms—Temporal reasoning, temporal uncertainty, Allen temporal relationships, Freksa semi-intervals.

1 Introduction

In 1983, Allen outlined a closed, nonoverlapping set of 13 interval-interval relationships; a set which can completely characterize the relative relationship between two temporal (or directional 1D) intervals [1]. This work has subsequently been extended in a number ways, including spatial and uncertain data; see [2], [3], [4] for surveys on the area.

The reasoning between such relationships is facilitated by a transitivity matrix, which, given the relationship between intervals A and B and B and C, provides the subset within which any relationship between A and C must fall. This work was later extended by Freksa to accommodate vagueness in one or more of the end points (termed semi-intervals) [5]. The Allen relationships, and to a lesser extent those of Freksa also, have been widely used in research in both data modeling, particularly temporal data modeling and temporal databases, and artificial intelligence, particularly temporal reasoning and decision support, and to a lesser extent in other fields such as data mining.

In practice, however, some modifications to the relationships are necessary when the input provided is a linear temporal sequence, when it has been aggregated into unordered blocks, or when the granularity adopted is larger than is required for reasoning. For example, with events time-stamped to the day, two events recorded on consecutive days can be as much as 48 hours or as little as a minute apart. Moreover, no ordering can be implied for events with the same date.

As can be seen in Fig. 1, four (meets, equals, starts, and finishes) of the 13 interval-interval relationships (plus their inverses) require at least one set of the endpoints of the intervals to be simultaneous, which when presented with a single data stream (as is common with, for example, real-time sensor data) cannot occur—tokens typically occur in an ordered sequence, even when, at a fine level of detail, that order is arbitrary. The problem becomes evident when this capricious ordering hampers reasoning and/or mining over the data. In practice, situations such as these occur frequently.

- *The authors are with the School of Informatics and Engineering, Flinders University, PO Box 2100, Adelaide, SA 5001, South Australia. E-mail: {roddick, carl.mooney}@infoeng.flinders.edu.au.*

Manuscript received 31 July 2003; revised 2 Feb. 2004; accepted 31 Mar. 2004; published online 18 Nov. 2004.
For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org, and reference IEEECS Log Number TKDE-0136-0703.

This paper discusses this issue, suggests a suitable interpretation, and provides a general solution involving the utilization of midpoint data.

2 Disposing of Implied Order and Implied Simultaneity

The two related problems can be briefly summarized as follows:

1. In cases where tokens are provided as a linear (nontime-stamped, unidentified) stream originating from n independent sensors, we must assume that any sequence of n tokens between two tokens in the stream originating from the same source took place simultaneously and that any apparent order is as a result of polling delays.
2. Alternatively, when data is supplied with a larger than required granularity, we cannot assume that those tokens time-stamped at t_i are necessarily all closer to each other than those time-stamped at t_{i-1} or t_{i+1}. This applies also to events placed unordered in ordered blocks of input.

The simple solution to the first problem (and one used in some research) has been to fragment the stream into sections of at least $2n+1$. However, this does not eliminate the problem as this is a moving window, as shown in Fig. 2, which can span any selected fragment size. Moreover, a large fragment (as with a large granularity) may also serve to create the second problem and act to imply simultaneity where none exists.

In some cases, data is provided in blocks in which events are unordered, even if the blocks themselves are ordered. In this case, there is again a window, as shown in Fig. 3. Assuming n sensors, we would ideally want to use the moving window IW_k, but in this case, we must use the larger window W_k, which encompasses any block covered by IW_k.

For data with too large a granularity, there is again a window, this time of those events that have an implied simultaneity where none may exist. For example, in Fig. 3, event i is closer to k than event j despite being given a different timestamp (i.e., in a different block). We can consider this case as being analogous to the blocked events problem, albeit that the blocks may be sparse.

One possible solution is to utilize modified point-interval transitivity graphs to accommodate the window. In 1982, Vilain [6] discussed five point-interval temporal relationships (shown also in Fig. 1). Thus, one solution is to consider two events as being simultaneous if any of the tokens j starts, finishes, or is during the interval represented by the moving window W_k of the other. This must be performed twice, once for each end of W_k.

Alternatively, as we propose in this paper, we can extend Allen's interval-interval algebra to consider not only the endpoints of intervals but also their midpoints, effectively extending his 13 relationships to 49, as shown in Fig. 4. The expanded set divides the overlap relationship in nine different types, the during relationship into seven, and the starts and finishes relationships into three each. Together with their inverses and the five unchanged relationships, there is a closed set of 49 variable length, interval-interval relationships with midpoints. These can be viewed as a further restriction of Allen's relationships in the same way that Allen's are more restricted than those of Freksa.

However, because W_j and W_k are identical in length, of the 13 Allen relationships, six of them (during, contains, starts, started by, finishes, and finished by) are not required. Instead, we can create a closed set of 11 relationships by extending the overlaps/overlapped by relationship, as shown in Fig. 5.

Relationship	Inverse Relationship	Schematically	Conditions	Endpoint Constraints
a Before b	b After a	a b	a.end < b.start	< < < <
		a b		
a Meets b	b Met by a	a b	a.end = b.start and a.start < b.start	< < = <
a Overlaps b	b Overlapped by a	a b	a.start < b.start < a.end < b.end	< < > <
a Equals b		a b	a.start = b.start and a.end = b.end	= < > =
		a b	a.start = b.start = a.end = b.end	= = = =
a Starts b	b Started by a	a b	a.start = b.start and a.end < b.end	= < > <
a Finished by b	b Finishes a	a b	a.start < b.start and a.end = b.end	< < > =
a Contains b	b During a	a b	a.start < b.start and a.end < b.end	< < > >
a Begins b	b Begun by a	a b	a.start = a.end = b.start	= < = <
a Ended by b	b Ends a	a b	a.start = a.end = b.end	< < = =

Fig. 1. Temporal interval-interval, point-interval, and point-point.

More formally, two nonzero-length intervals X and Y can be considered to have a temporal relationship based on the relative positions of their endpoints and midpoints. Since $X_{start} < X_{mid} < X_{end}$ and $Y_{start} < Y_{mid} < Y_{end}$, the 3^9 combinations are reduced to 49. Requiring the intervals to be fixed length introduces further constraints, such as $X_{start} = Y_{start} \rightarrow X_{mid} = Y_{mid} \wedge X_{end} = Y_{end}$, which reduces the combinations to 11. As for the Allen constraints, transitive relationships can be calculated through a transitivity table, as outlined in Fig. 6.

In Fig. 5, we can see that in the context of a data stream, only MediumOverlap, LargeOverlap, and Equals results in the midpoint of one interval being within the other. Thus, using this new algebra, we note that if the relationship between W_i and W_j, for two tokens i and j, is SmallOverlap for any given number of sensors n, then the implied point-point relationship is Before, as the end-point of one does not overlap with the midpoint of the other. For Medium-Overlap and LargeOverlap, for any given number of sensors n, the implied point-point relationship is *potentially* Equals, implying possible simultaneity.

Finally, while it is beyond the scope of this short paper to discuss a full extension to Allen's algebra with midpoints, it should be noted that some increase in specificity can be gained by expanding the model to all 49 relationships. For example, it can be shown from the augmented transitivity table in Fig. 6, for the relationships given in Fig. 5, that while in Allen's algebra

$$A\ o\ B \wedge B\ o\ C \rightarrow A\ o, m, <\ C, \quad (1)$$

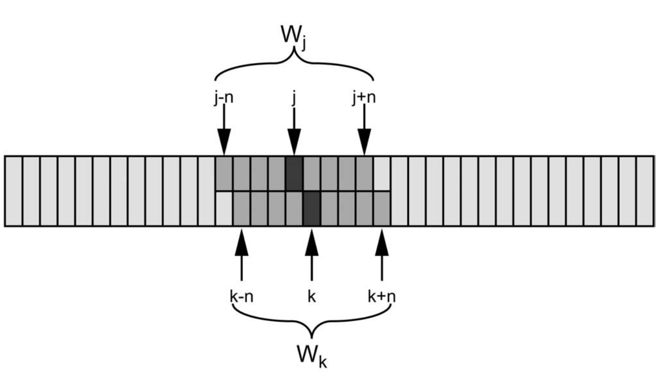

Fig. 2. Moving window of potentially simultaneous tokens where $n = 4$.

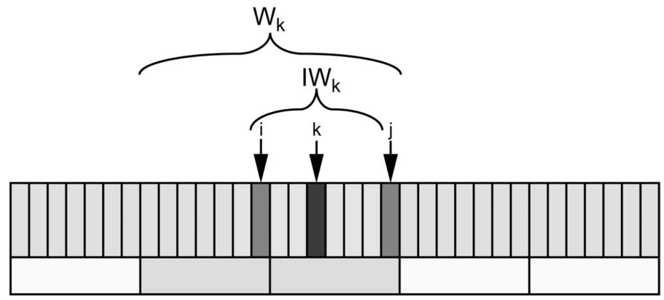

Fig. 3. Moving window over tokens with larger than required granularity (or tokens in blocks).

using midpoints can offer the ability to refine this to

$$A\ so\ B \wedge B\ so\ C \rightarrow A\ <\ C. \quad (2)$$

That is, information about midpoints can refine the set of relationships returned. It should be also noted, however, that in cases when midpoint data is unavailable, the proposal degrades gracefully to the algebra given by Allen.

3 Practical Application

Allen and Freksa's model has been applied widely [2], [3], [4]. In principle, there is no reason why, given the appropriate data, most of these applications should not use and benefit from this extended model. Midpoints are known, for example, whenever the end-points are given in absolute time.

In Section 1, we discussed some general scenarios in which such a model may be applied. One specific application investigated has been in sequence mining—specifically that of detecting and characterizing interacting sequences in very long strings [7]. For some data sets, we cannot assume that just because a token occurs after another token in the string, the events they represent occurred in that order. For example, tokens collected though polling N event monitors could, in the worst case, occur $N-1$ tokens out of position. We have thus used this technique to accommodate this uncertainty.

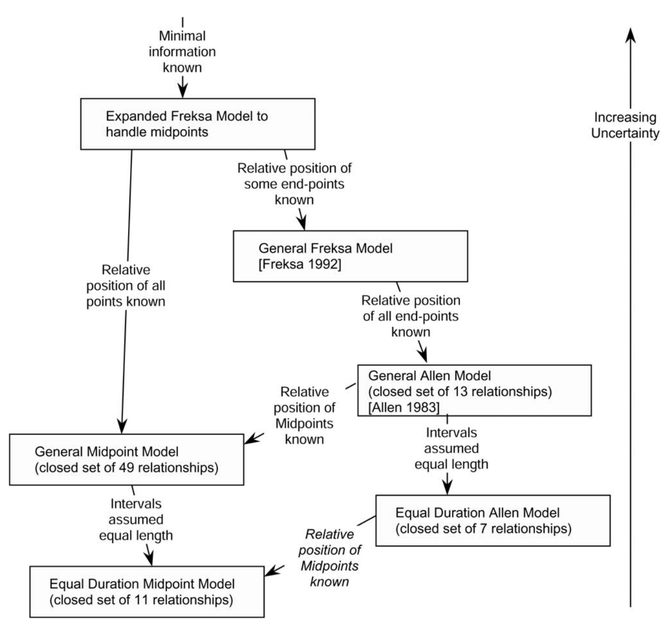

Fig. 4. Models of temporal interval relationships.

Relationship	Sym	Inverse Relationship	Sym	Schematically	Conditions
a Before b	<	b After a	>	a b	a.end < b.start
a Meets b	m	b Met by a	mi	a b	a.end = b.start
a SmallOverlap b	so	b SmallOverlapInverse a	soi	a b	a.end > b.start, a.mid < b.start
a MediumOverlap b	mo	b MediumOverlapInverse a	moi	a b	a.mid = b.start
a LargeOverlap b	lo	b LargeOverlapInverse a	loi	a b	a.mid > b.start, a.start < b.start
a Equals b	=			a b	a.start = b.start

Fig. 5. Equal-length interval-interval relationships with midpoints.

	<	m	so	mo	lo	soi	moi	loi	=	mi	>
<	<	<	<	<	<	<, m, so, mo, lo	<, m, so	<, m, so	<	<, m, so, mo, lo	Any
m	<	<	<	<	<	lo	mo	so	m	=	soi, moi, loi, mi, >
so	<	<	<	<	<, m, so	lo, loi, =	lo	so, mo, lo	so	loi	soi, moi, loi, mi, >
mo	<	<	<	m	so	loi	=	lo	mo	moi	soi, mi, >
lo	<	<	<, m, so	so	so, mo, lo	soi, moi, loi	loi	lo, loi, =	lo	soi	soi, mi, >
soi	<, m, so, mo, lo	lo	lo, loi, =	loi	soi, moi, loi	>	>	soi, mi, >	soi	>	>
moi	<, m, so	mo	lo	=	loi	>	mi	soi	moi	>	>
loi	<, m, so	so	so, mo, lo	lo	lo, loi, =	soi, mi, >	soi	soi, moi, loi	loi	>	>
=	<	m	so	mo	lo	soi	moi	loi	=	mi	>
mi	<, m, so, mo, lo	=	loi	moi	soi	>	>	>	mi	>	>
>	Any	soi, moi, loi, mi, >	soi, moi, loi, mi, >	soi, mi, >	soi, mi, >	>	>	>	>	>	>

Fig. 6. Constraint propagation matrix for equal length interval-interval relationships with midpoints.

In practice, the time complexity of using an 11×11 or a 49×49 element table is the same as employing Allen's 13×13 element table. The space complexity varies marginally with the obvious change in the size of the static lookup table.

4 Conclusion

This paper has demonstrated that a simple extension to Allen's interval-interval relationship algebra, to refine the overlaps relationship to consider midpoints for equal length intervals, can be used to accommodate linear (nontime-stamped) sequences of tokens and allow more appropriate reasoning over stream data, such as real-time sensor data. Although sufficient for the problem at hand, this work could be generalized to a full set of variable-length, interval-interval relationships with midpoints.

Acknowledgments

The authors would like to thank the referees for their helpful and constructive comments.

References

[1] J.F. Allen, "Maintaining Knowledge About Temporal Intervals," *Comm. ACM,* vol. 26, no. 11, pp. 832-843, 1983.
[2] P.B. Ladkin, "Models of Axioms for Time Intervals," *Proc. Sixth Nat'l Conf. Artificial Intelligence (AAAI '87),* pp. 234-239, 1987.
[3] R. Gennari, "Temporal Reasoning and Constraint Programming–A Survey," *CWI Quarterly,* vol. 11, nos. 2/3, pp. 163-214, 1998.
[4] L. Vila, "A Survey on Temporal Reasoning in Artificial Intelligence," *Artificial Intelligence Comm.,* vol. 7, no. 1, pp. 4-28, 1994.
[5] C. Freksa, "Temporal Reasoning Based on Semi-Intervals," *Artificial Intelligence,* vol. 54, pp. 199-227, 1992.
[6] M.B. Vilain, "A System for Reasoning About Time," *Proc. Nat'l Conf. Artificial Intelligence (AAAI '82),* pp. 197-201, 1982.
[7] C. Mooney and J.F. Roddick, "Mining Relationships Between Interacting Episodes," *Proc. 2004 SIAM Int'l Conf. Data Mining,* U. Dayal and M.W. Berry, eds., 2004.
[8] J.F. Allen, "An Interval-Based Representation of Temporal Knowledge," *Proc. Seventh Int'l Joint Conf. Artificial Intelligence,* A. Drinan, ed., pp. 221-226, 1981.
[9] J.F. Allen, "Towards a General Theory of Action and Time," *Artificial Intelligence,* vol. 23, no. 2, pp. 123-154, 1984.
[10] J.F. Allen and P.J. Hayes, "Moments and Points in an Interval-Based Temporal Logic," *Computational Intelligence,* vol. 5, pp. 225-238, 1989.