An Algebraic Formulation of Temporal Knowledge for Reasoning About Recurring Events

Robert A. Morris and **William D. Shoaff** and **Lina Khatib**
Florida Institute of Technology
Computer Science Program
150 W. University Blvd.
Melbourne, FL 32901

Abstract

We formulate an algebra of binary temporal relations between events the number of occurrences of which is unknown, but which are known to recur in time. Ontologically, we view the temporal domain of recurring events as a set of (convex) intervals of unknown cardinality. Thus, an explicit, declarative representation of time adequate to solve reasoning problems involving recurring events indicates a form of higher-order reasoning about relations between collections, or sequences of intervals. The relations defined here are each formed out of convex interval relations by applying an operator like "always" or "sometimes". An operator of this kind imposes a mapping between subintervals of one set of intervals and subintervals of another. Although in general higher order reasoning problems are undecidable, a restricted system for reasoning about binary relations between collections of intervals is demonstrably tractable. The means to achieving this desirable result is to impose restrictions on the kinds of mappings imposed by the temporal operators. Despite this restriction in expressive power, non-trivial reasoning problems can be solved utilizing this class of relations.

1 Setting the stage

A temporal reasoning problem can be defined abstractly as the process of inferring, from an initial *specification* of temporal information, other useful information. Useful information includes, e.g., determining whether the specification is consistent, computing the deductive closure of the set of statements making up the specification, or generating a *scenario* consistent with the specification.

Many formalisms for representing and reasoning with temporal information have been proposed in the literature. The most notable include Allen's interval algebra [1], the point algebra, [14], [13], and the Time Map Manager [4]. None of these frameworks allow for an explicit representation of temporal information about events in which the number of time periods described is unknown. Yet, there are many reasoning problems whose solutions involve a cognitive process of abstraction from the number of times an event is to occur. For example, scheduling one's office hours will typically consider things like the temporal constraints between the holding of hours and other events in a given work day. The number of times office hours will actually be held may not be known (i.e. the schedule may assign hours for an indefinite future). Indeed, such knowledge is irrelevant to the task at hand.

Recently, there has been a small, yet significant numbers of researchers who have recognized the need for formalizing knowledge of this kind (e.g. [7], [8], [12]). More recently, a framework for representing and reasoning about recurring intervals of time has been introduced by the authors [10]. This paper extends the model introduced there.

The focus of this model is on the recurrence of *binary temporal relations* between recurring events. Such recurrence can be viewed as involving mapping operations between collections of intervals implied by each relation in the class. It can be shown that distinct classes of relations for recurring events can be constructed based on different ways of characterizing these mappings. The underlying framework for solving temporal reasoning problems involving this type of information consists of a formalization based on the interval algebra [1]. To this framework, the following is introduced in what follows:

1. A model for reasoning about relations between recurring events based on the following hypotheses (section 2):
    - That statements about recurring event pick out an entity we call an *n-interval*;
    - That ascribing relations between recurring events imply a sort of *correlation* among sub-intervals of n-intervals; and
    - That a class of relations can be constructed out of temporal operators on con-

vex temporal relations;

2. A algebraic formalization of a class of relations between n-intervals, as well as a set of operations performed on them (section 3);

This paper concludes with an informal discussion of a way of strengthening the proposed model by incorporating more information from a temporal specification.

2 An ontology of recurrence

Time here is viewed as a linear order on a domain consisting of minimal time units identified with the natural numbers under the ordering $<$. The temporal domain for expressions about recurring events is viewed here as a set of collections of intervals. More precisely, each element in the domain of discourse is a set of intervals of the form $I = \{\langle I_1^-, I_1^+\rangle, \langle I_2^-, I_2^+\rangle, \ldots \langle I_n^-, I_n^+\rangle\}$, representing a (possibly) non-convex interval with n convex components, or *subintervals*, starting at I_1^- and ending at I_n^+. We call such a set an *n-interval* [1], with the ordering $I_i^- < I_i^+$ for $i = 1 \ldots n$. $I_i = \langle I_i^-, I_i^+\rangle$ represents the ith convex subinterval of I.

The number of possible binary relations between n-intervals is exponential in the maximum number of the subintervals of each n-interval. Specifically, Ligozat [8] and, independently, one of the authors of this paper [6], have proven the following:

Remark 1 *Let $R^{nc}(n,m)$ represent the number of possible binary relations between two n-intervals of size n and m. Then*

$$\begin{aligned} R^{nc}(n,m) &= \sum_i \binom{2m}{i}\binom{2n+2n-i}{2m} \\ &= \sum_i 2^i \binom{2m}{i}\binom{2n+2m-i}{i} \end{aligned}$$

This result implies that formalizing the reasoning process for manipulating elements of this kind will be difficult. Fortunately, it is possible to trim this domain of discourse significantly by incorporating into the formal representation presuppositions which occur in natural language discourse about recurring events. More specifically, many relations between recurring events picked out in natural language discourse involve an implicit partitioning of time into meaningful segments. One way in which time is partitioned emerges in the act of predicating a temporal relation between a pair of recurring events. For example, consider the meaning of the sentence "Meetings *always* precede lunch". This meaning involves the predication of a relation *always precedes* between two recurring events of unknown number, viz., *meetings* and *lunches*. Suppose there are in fact at least three meetings and lunches. This sentence might be true despite the fact that, for example, the third meeting comes after, not before, the first lunch. Informally, the truth condition of this sentence implies a *pairing* of subintervals imposed by the relation *precedes*. In this paper, we will say that the paired subintervals are *correlated*, and that the relation *induces* the correlation.

[1] This term is used differently by Ligozat [8], viz., to denote an interval consisting of n points.

In the proposed model, correlation is given an implicit, axiomatic definition. Formally, correlation establishes a mapping between sets of intervals making up n-intervals. There are many ways of defining this mapping. The approach taken here defines correlation in two stages.

Definition 1 (Local Correlation) *Let $COR_{x,y}$, be a family of functions defined between convex parts of n-intervals. Where I and J denote such a pair, $COR_{I,J}$ designates a set of ordered pairs $\langle i,j\rangle, i \in I, j \in J$. $COR_{I,J}$ is restricted to be one-one; thus, each subinterval I_x of I can be correlated with at most one subinterval J_y of J, and visa versa.*

Local correlation allows for the characterization of mappings between subintervals of pairs of n-intervals. To generalize the reasoning process to arbitrary finite collections of intervals, we define *global correlation*, COR:

Definition 2 (Global Correlation) *Let $COR = \bigcup_{I,J} COR_{I,J}$. Global correlation is an equivalence relation.*

It should be stressed that the restrictions imposed in the definitions of local and global correlation (specifically, that local correlation is a 1-1 function and that global correlation is an equivalence relation) are the result of the need for a useful, yet computationally manageable approach to reasoning about recurrence. Although these restrictions are reflected in presuppositions in natural language discourse, there is no claim here to have accurately formalized all of these presuppositions.

With an account of correlation in hand, it is possible to identify a class of relations that describe patterns of binary relationships between correlated convex parts of two recurring events. Attention is restricted to relations expressed in natural language by applying temporal adverbs to qualitative binary temporal relations, as introduced by Allen [1]. Reviewing the account introduced in [10], we distinguish among 3 *(basic) relation forming operators* for constructing n-interval relations: *sometimes*, *always*, and *only*. Their meaning is illustrated in Figure 2(a)–(c). These operators are used to construct what will be termed *basic relations* between two n-intervals.

Notice that Figure 2(a) also illustrates a scenario in which "Meetings only precede or overlap cocktails" is true. In English, a more complete description of the temporal relations between meetings and cocktails illustrated by this figure would be the sentence

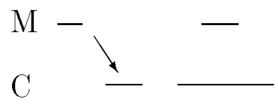

Figure 2a. Cocktails *sometimes* after meetings (meetings sometimes before cocktails).

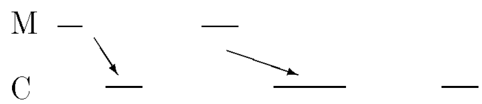

Figure 2b. Cocktails *always* after meetings (meetings only before cocktails).

Figure 2d. Cocktails *always or only* after meetings (meetings always or only before cocktails).

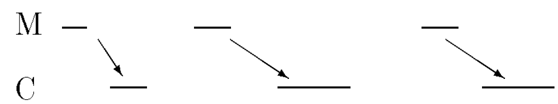

Figure 2e. Cocktails *always and only* after meetings (meetings always and only before cocktails).

"Meetings only precede or overlap cocktails *and furthermore* they sometimes overlap cocktails and (furthermore) sometimes precede cocktails". Intuitively, the role of this operator is to impose a sort of commitment to there being at least one pair of subintervals of pairs of n-intervals involved in a temporal relation. For example, the relation *always before or during and furthermore sometimes during* is a more restricted relation than *always before or during*, since the former, but not the latter, implies at least one pair of intervals in the relation *during*. It turns out that inferences involving recurring events can be formalized in terms of this operator; therefore, we include it as part of the representation of n-interval relations.

Additional relation forming operators are defined from the basic operators. *Always and only* is a specialization of the basic operators *always* and *only*. *Always or only* is a generalization of these operators. They are illustrated in Figure 2(d)-(e). Always and only establishes a 1-1 relationship between correlated intervals. I Always or only R J implies that any correlated pairs $\langle i, j\rangle \in I \times J$ must be induced by R.

Of course, other relation-forming operators could be identified (e.g *mostly*). To keep the discussion brief, we limit our attention to formulating reasoning systems using the 5 operators defined in this section.

3 Formalizing a class of n-interval relations

In this section, we define a class **NR** of n-interval relations involving the relation-forming operators defined in the previous section. In addition, a set of *relation-inferring operators* will be defined. These operators will be employed by a reasoner in order to solve reasoning problems involving recurring events and their relations. The results of performing these operations to members of **NR** will implicitly incorporate assumptions about correlation relations between subintervals of n-intervals.

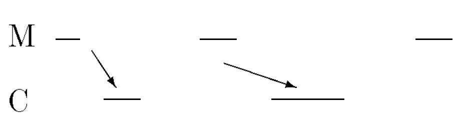

Figure 2c. Cocktails *only* after meetings (meetings always before cocktails).

Definition 3 (Basic N-interval Relations) *Let* **CR** *be the set consisting of the* 2^{13} *convex interval relations (finite disjunctions of the 13 primitives* [2] *expressed as sets). The set of basic N-interval relations consists of relations that result from applying one of five* **relation-forming** *operators Σ (sometimes),* Π *(always), Ω (only), Υ (always or only) or Θ (always and only) to an element of* **CR**.

Generalizing, each n-interval relation in **NR** can be expressed as a restricted form of finite conjunction (using the *and furthermore* operator) of basic relations. Formally:

Definition 4
(The set NR) *If* $OP \in \{\Pi, \Omega, \Theta\, \Upsilon\}, R, S_1, \ldots S_k \in$ **CR**, $S_i \subset R, \forall i = 1 \ldots k$, *and, for all* $i, j = 1 \ldots n, i \neq j, S_i \neq S_j$, *then* **NR** *contains the relation* $OP(R) \oplus \Sigma(S_1) \oplus \ldots \oplus \Sigma(S_k)$ *(where $\oplus$ is the symbol for "and furthermore"). $OP(R)$ is called the* leading relation, *and the others the* Σ-relations.

Notice that restrictions are placed on the Allen (convex) relations S_i, S_j appearing in pairs of Σ-relations $\Sigma(S_i), \Sigma(S_j)$ within an n-interval relation. We call constraints of this kind Σ-*restrictions*. These restrictions are adopted for various reasons. One reason is to maintain consistency: for example, the relation *always before and furthermore sometimes after* defines the empty relation (hence the restriction that $\forall i = 1 \ldots n, S_i \subset R$). A second reason for adding Σ-restrictions is to avoid redundancy: for example, the relation *always before or during and furthermore sometimes before and furthermore sometimes before* is reducible to *always before or during*

[2] The 13 primitives will have the following abbreviations: b (before), d (during), m (meets), o (overlaps), s (starts), = (equals), f (finishes), bi (before inverse), etc.

and furthermore sometimes before; hence the restriction that $\forall i, j = 1 \ldots n, i \neq j, S_i \neq S_j$. Finally, Σ-restrictions can be applied to further limit the size of the set of recurrence relations being represented, e.g., in order to simplify the reasoning process. For example, the set of legal temporal specifications for a given application may not require a relation like $\Pi\{b, m, o\} \oplus \Sigma\{o, m\} \oplus \Sigma\{b, m\}$. Such a relation can be approximated by the simpler relation $\Pi\{b, m, o\} \oplus \Sigma\{m\}$. In general, it may be sufficient to restrict the Σ-relations to those applied to atomic Allen relations. These issues will be explored in future work.

To infer new information about temporal relations from old, Allen defined a set of operations on elements of **CR**. A similar set of operations can be defined for elements of **NR**: inverse (or converse), intersection and composition (or relative product). We call this class of operators *relation-inferring*. Applying these operators to basic relations involves two independent sub-operations: applying the corresponding operation to the convex relations, as defined by Allen [1], and applying the operator to a pair of the relation-forming operators. Furthermore, since each of the relation-inferring operators distribute over $\oplus$, defining them for the general case (i.e. between arbitrary members of **NR**) will be straight-forward.

Inverse allows a reasoner to infer, for example, from the fact that cocktails always follow or meet faculty meetings, that faculty meetings only precede or are met by cocktails. Generalizing:

Definition 5 (Inverse operator for members of NR)

$$[OP(R) \oplus \Sigma(S_1) \oplus \ldots \oplus \Sigma(S_k)]^{-1} =$$
$$OP(R)^{-1} \oplus \Sigma(S_1)^{-1} \oplus \ldots \oplus \Sigma(S_k)^{-1}$$

Defining inverse for the basic relations is straight-forward ([10]).

In [10], intersection ($\sqcap$) between basic relations was defined. For example, the equation $\Pi(R) \sqcap \Theta(S) = \Theta(R \cap S)$ allows a reasoner to infer from the fact that meetings always follow cocktails, and that they always and only either follow or begin cocktails, that they always and only follow cocktails. It is the intersection operator that introduces relations involving the *and furthermore* operator. For example, if meetings are always before or after cocktails, and later it is learned that they are, in fact, sometimes after cocktails, then the relation inferred is depicted formally as $\Pi\{b, bi\} \oplus \Sigma\{bi\}$. Generalizing:

Definition 6 (Intersection between members of NR)

$$[OP(R) \oplus \Sigma(S_1) \oplus \ldots \oplus \Sigma(S_k)]$$
$$\sqcap$$
$$[OP(R') \oplus \Sigma(S'_1) \oplus \ldots \oplus \Sigma(S'_m)] =$$
$$[OP(R) \sqcap OP(R')] \oplus$$
$$CONV_\Sigma[\Sigma(S_1) \oplus \ldots \oplus \Sigma(S_k) \oplus \Sigma(S'_1) \oplus \ldots \oplus \Sigma(S'_m)]$$

where $CONV_\Sigma$ is a procedure to apply the Σ-restrictions to a conjunction of Σ-relations.

Example 1 (Intersection of n-interval relations)

$$I\ \Pi\{b, o, m\} \oplus \Sigma\{b\} \oplus \Sigma\{m\}\ J$$
$$\sqcap$$
$$I\ \Theta\{b, s, m\} \oplus \Sigma\{m\}\ J$$
$$=$$
$$I\ \Theta\{b, m\} \oplus \Sigma\{b\} \oplus \Sigma\{m\}\ J$$

Composition of basic n-interval relations has been defined in [10], and will not be reprinted here. In the general case, composition between arbitrary relations in **NR** involves three-steps, each consisting of basic compositions. First, composition of the leading relations is performed. Second, the first leading relation is composed with each of the Σ-relations of the second. Finally, the second leading relation is composed with each of the Σ-relations of the first. The results of each step are conjoined using *and furthermore*. Formally:

Definition 7 (Composition of members of NR)

$$OP(R_0) \oplus \Sigma(R_1) \ldots \oplus \Sigma(R_n)$$
$$\otimes$$
$$OP'(R'_0) \oplus \Sigma(R'_1) \ldots \oplus \Sigma(R'_m)$$
$$=$$
$$OP(R_0) \otimes OP'(R'_0) \oplus$$
$$CONV_\Sigma[(OP(R_0) \otimes \Sigma(R'_1)) \oplus \ldots \oplus (OP(R_0) \otimes \Sigma(R'_m)) \oplus$$
$$(\Sigma(R_1) \otimes OP'(R'_0)) \oplus \ldots \oplus (\Sigma(R_n) \otimes OP'(R'_0))]$$

Example 2 (Relation composition in general case) *I is always before or meets J, and furthermore it is sometimes before J. J is always and only before, finished by, or starts K and furthermore it sometimes starts K and furthermore it sometimes begins K.*

$$I \Pi\{b, m\} \oplus \Sigma\{b\}\ J$$
$$\odot$$
$$J\ \Theta\{b, fi, s\} \oplus \Sigma\{s\} \oplus \Sigma\{b\}\ K$$
$$=$$
$$I\ \Pi\{b, m\} \oplus \Sigma\{b\}\ K$$

Proof:

$$I \Pi\{b, m\} \oplus \Sigma\{b\}\ J$$
$$\odot$$
$$J\ \Theta\{b, fi, s\} \oplus \Sigma\{s\} \oplus \Sigma\{b\}\ K$$
$$=$$
$$(\Pi\{b, m\} \odot \Theta\{b, fi, s\}) \oplus$$
$$(\Pi\{b, m\} \odot \Sigma\{s\}) \oplus (\Pi\{b, m\} \odot \Sigma\{b\}) \oplus$$
$$(\Sigma\{b\} \odot \Theta\{b, fi, s\})$$
$$= \Pi\{b, m\} \oplus \Sigma\{b, m\} \oplus \Sigma\{b\}$$

By the Σ-restrictions for redundancy, this relation reduces to

$$\Pi\{b, m\} \oplus \Sigma\{b\}$$

Thus, I is always before or meets K, and furthermore it sometimes precedes K.

The set **NR** and the relation-forming operations collectively form an *n-interval algebra*. The next definition and remark summarize the results of this section.

Definition 8 (N-Interval Algebra) *The set* **NR** *(N-interval Relations) is the foundation of a pure N-interval algebra.* **NR** *consists of $5 \times (2^{13})$ basic relations that result from applying one of five* **relation-forming** *operators to an element in* **CR**. *The empty relation can be expressed in this model as $OP(\emptyset)$, where OP is any basic relation-forming operator, and $\emptyset$ is the empty Allen relation. Dually, the relation for no information can be expressed as $\Sigma(??)$, where ?? is the Allen relation for no information (i.e., the disjunction of all the atomic Allen relations. There are three* **relation-inferring** *operators defined on* **NR**: *inverse, intersection and composition.*

Remark 2 *The set* **NR** *is closed under the relation-inferring operations converse, intersection, and composition.*

4 Discussion: The temporal dimension of specifications

In reasoning about specifications of recurrence relations, a reasoner is manipulating partial descriptions of distributions and participations of relations among pairs of collections of time units. The notion of correlation was introduced to identify the restricted domain of discourse for a specification.

Based on previous efforts by the researchers [10], it has become evident that the account presented here would benefit from an investigation of the temporal information found in the structure of the recurrence specifications themselves. This echoes the sentiments of Kamp [5], whose concept of *discourse representation* has been utilized to uncover the temporal information implicit in the *order* of the statements in text. Our work has revealed that taking advantage of information supplied by *anaphoric relationships* between noun phrases used to express recurring events allows the reasoner to infer tighter constraints among recurrences.

To illustrate the sorts of contexts of interest, compare the following:

1. John sometimes calls his dad before going to work. *Then*, he misses the start of (i.e. the call overlaps) the meeting.
2. Joan sometimes goes to work before calling Jim. *Otherwise*, she calls Jim first.
3. Faculty meetings always precede seminars. *Those* meetings only overlap with lunch.

In the first passage, the adverb "then" serves to establish an anaphoric reference to the previously introduced occurrence of going to work. In the second, "otherwise", serves to prohibit this connection with the previously introduced interval, and implies at least two correlated events of callings and goings. The third sentence introduces a reference to a *subset* of n-intervals, viz., corresponding to those meetings which precede seminars.

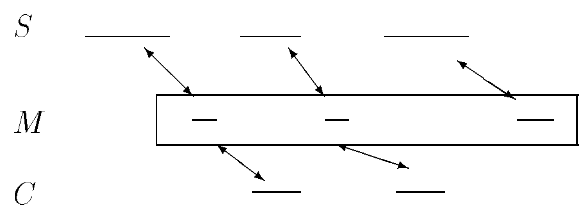

Figure 3. Illustration of the contribution of anaphora in composing relations.

Consider the following specification:

Example 3 Seminars (S) only before faculty meetings (M). Those meetings always followed by cocktails (C).
An annotated scenario appears in Figure 3. The subset of the n-interval corresponding to **meetings** indicated by the occurrence of **those** is depicted by the rectangle. With the added information supplied by the anaphoric relation between "those" and the previously introduced reference, it is possible to infer by composition that chantings are always before cocktails. The inferred relation is stronger (more restricted) than the relation inferable as a result of the version of composition introduce in section four.

In [10], an simple indexing mechanism on recurrence relations was introduced in order to represent certain forms of anaphora. Although adding this information results in a more effective reasoner, there is significant overhead incurred in manipulating contexts involving anaphora. Managing this overhead is the subject of continuing research.

5 Concluding remarks

The research presented here extends current efforts in developing interval-based systems for reasoning about recurring periods of time by allowing for an explicit representation of collections of intervals and their relations. The purpose of this extension is to expand the set of qualitative relations that can be attributed to periods of time for applications such as static scheduling, natural language processing, and temporal database querying.

References

[1] Allen, J. (1983) Maintaining Knowledge About Temporal Intervals. In Brachman, R., and Levesque, H., (eds.) *Readings in Knowledge Representation*, (San Mateo:Morgan Kaufman), 510-521.

[2] Allen, J. (1987) *Natural Language Understanding.* Benjamin/Cummings, Menlo Park.

[3] Chou, In-Cheng (1993) "Implementing Path Consistency for n-interval Relation Networks", M.S. Thesis, The Florida Institute of Technology.

[4] Dean, T. M., and McDermott, D. V. (1987) Temporal Data Base Management. *Artificial Intelligence* Vol 32, 1-55.

[5] Kamp, H. Tense in Texts. In Baeuerle, R., Schwarze, C., and Von Stechow, A., *Meaning, Use and the Interpretation of Language*, New York:de Gryter (1983), 250-269.

[6] Khatib, L. (1994) Reasoning With Non-Convex Time Intervals. Doctoral Dissertation, Florida Institute of Technology.

[7] Ladkin, P.B. (1987) Time Representation: A Taxonomy of Interval Relations. In *Proceedings of AAAI-86*, (San Matyeo:Morgan Kaufman), 360- 366.

[8] Ligozat, G., (1990) Weak Representation of Interval Algebras. In *Proceedings of the Eighth National Conference on Artificial Intelligence* (AAAI-90), Vol. 2, 715-720.

[9] Mackworth, A. (1977). Consistency in Networks of Relations. *Artificial Intelligence*, 8, 99-118. 25-32.

[10] Morris, R., Shoaff, W., Khatib, L., Path Consistency in a Network of Non-Convex Intervals.*Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, Chambery, France (1993), pp. 650-655.

[11] Morris, R., and Al-Khatib, L., (1991) An Interval-Based Temporal Relational Calculus For Events With Gaps. *The Journal of Experimental and Theoretical Artificial Intelligence*, 3, (1991), 87-107.

[12] Niezette, M., Stevenne, J-M., (1992) An Efficient Symbolic Representation of Periodic Time. *Implementing Temporal Reasoning: Workshop Notes, AAAI-1992*, 130-140.

[13] Van Beek, P. (1990). *Exact and Approximate Reasoning about Qualitative Temporal Relations.* Technical Report TR 90-29, University of Alberta, Edmonton, Alberta, Canada.

[14] Villain, M., and Kautz, H. (1986) Constraint Propagation Algorithms for Temporal Information. In *Proceedings of AAAI-86*, 377-382.

[15] Vilain, M. Kautz, H. and Van Beek, P. (1989). Constraint Propagation for Temporal Reasoning: A Revised Report. In *Readings in Qualitative Reasoning about Physical Systems*, Morgan-Kaufmann, 351-357.