The Date-Time Vocabulary

Mark H. Linehan[a], Ed Barkmeyer[b], and Stan Hendryx[c]
[a] *IBM T.J. Watson Research Center, Hawthorne, NY*
[b] *National Institute of Science and Technology, Gaithersburg, MD*
[c] *Hendryx & Associates, Sunnyvale, CA*

Abstract. The Date-Time Vocabulary is a new OMG specification that models continuous time, discrete time, the relationship of events and situations to time, language tense and aspect, time indexicals, timetables, and schedules. It offers a "business vocabulary" (i.e., a linguistically-oriented ontology), intended for application by business rules for any business domain. The vocabulary is defined in SBVR, in UML + OCL, and partially in Common Logic and in OWL.

Keywords. date-time ontology, time ontology, date-time vocabulary, SBVR, Common Logic, IKL, UML, OCL, OWL

Introduction

Much business discourse involves common, generic, cross-domain concepts such as date and time, and their relationship to situations. These concepts are characterized by frequent usage in everyday and business activities and wide application across business domains, such as finance and manufacturing. Providing a standard vocabulary for such "foundational" concepts simplifies the creation and maintenance of business rules and improves communication among and within businesses. Such a vocabulary brings together in one reference source a set of concepts that are otherwise addressed in a fragmentary, incomplete, and ambiguous manner in numerous existing standards.

This paper describes a new OMG specification that offers a foundational business vocabulary for date and time concepts that addresses these needs. This *Date-Time Vocabulary* (DTV) [1] is specified in the format, and in the "structured English" language, of the OMG *Semantics of Business Vocabulary and Business Rules* (SBVR) specification [2]. In parallel, it is specified in the *Unified Modeling Language* (UML®) [3] with constraints in the *Object Constraint Language* (OCL) [4], and in *Common Logic* (CL) [5] and the *Web Ontology Language* (OWL2) [6]. The different forms address different audiences and uses: the SBVR version targets business users who need to understand and apply the date and time concepts, the UML model provides the basis for software implementations, the CL text captures axioms about date and time in a formal reasoning language, and the OWL ontology enables adoption of these concepts by developers of OWL ontologies.

A typical scenario for applying the Date-Time Vocabulary is the development of a domain-specific business vocabulary, such as the *Financial Industry Business Objects* (FIBO) [7] effort of the Enterprise Data Management Council and OMG. The Date-Time Vocabulary could also be used by an enterprise-specific business vocabulary. Both cases need to define concepts that are based on dates and times but add domain-specific details. For example, a mortgage document applies simple date concepts such as "closing date", and much more complex date- and duration-related concepts such as "mortgage amortization schedule". The Date-Time Vocabulary provides a basis for specifying these domain concepts in a consistent manner that maximizes the semantic meaning for both human readers and semantic reasoning systems.

The paper is organized as follows: the Related Work section reviews existing standards, ontologies, and academic sources related to dates and times, and discusses how the present work builds upon them. We then discuss how we define sequences, quantities, units of measure, and mereology as infrastructure for modeling date and time concepts. Next we address our approach to continuous and discrete time. A key topic is the relationship of events and situations to time. We also discuss how this Vocabulary models language tense, language aspect, and time indexicals.

1. Related Work

At the start of this project, we surveyed numerous existing standards, ontologies, and academic papers that have addressed the topic in some manner. We summarize these sources in this section.

Date and Time concepts are extensively addressed in various established standards. Most important among these are ISO 8601, *Data Elements and Exchange Formats – Information Interchange – Representation of Dates and Times* [8], and ISO 80000-3, *Quantities and Units – Part 3: Space and Time* [9]. The former defines physical formats for exchanging dates and times among information systems – formats that are very widely reused in other standards such as XML Schema Datatypes [10]. ISO 80000-3 defines measurement units such as "second", "day", etc. Both of these standards are based on vocabularies of quantities and measurement units that are defined in *The International System of Units* (SI) [11] and the *International Vocabulary for Metrology* (VIM) [12]. None of these address time coordinate systems (time scales), which are described instead in ISO 18026, *Information Technology – Spatial Reference Model* (SRM) [13]. Developing the Date-Time Vocabulary involved reviewing and comparing these standards sources in detail, and filling in missing details for everyday but non-standardized aspects of common calendars, such as the treatment of what we call "relative time points".

The Date-Time Vocabulary addresses the same general topics as the well-known *OWL-Time* [14], [15], [16]. Beyond the basic difference that Date-Time provides a business vocabulary and a UML model in addition to an OWL ontology, there are a few more subtle distinctions. In particular, the Date-Time Vocabulary avoids the concept of "time instant", instead depending upon mereological and ordering concepts to model many of the relationships among time intervals. Date-Time also provides a generic basis for defining any type of calendar (e.g., financial calendars, religious calendars) in addition to the standard calendar, whereas OWL-Time models only the standard calendar. Date-Time addresses language tense and aspect, which are not discussed in OWL Time. Two time units – "month" and "year" – require special treatment because of the varying numbers of days in each. Date-Time provides a complete formulation of these topics, as discussed below.

Upper ontologies, such as SUMO [17], and DOLCE [18], integrate temporal concepts into very complete top-down organizations of domain-independent knowledge. The temporal aspects of these ontologies generally address the same concept set as OWL-Time. Our approach differs from these primarily in being founded on relationships among "time intervals" instead of "time instants". Like SUMO and unlike DOLCE, we define special temporal relations in order to assign time to any relationship, state or event.

We reviewed numerous academic papers about ontologies of time. We incorporated the well-known Allen [19] relationships directly into our vocabulary. Hayes [20] identifies six importantly different meanings of the term "time". We address all six concepts below, but many of our terms are different. To prevent confusion, we avoid the plain term "time" altogether. One of these concepts is what we call the *Time Axis* (Hayes: "time-plenum"), defined by a dictionary as "the indefinite continued progress of existence and events in the past, present and future, regarded as a continuum" [21]. The term *time interval* refers to a segment of the Time Axis (Hayes: "time-interval"). As explained below, we avoid the "time instant" concept (Hayes: "timepoint"), but we use the term *time point* (Hayes: "time coordinate") to mean a position on a time scale, and *time coordinate* to mean a name for a time point. The concept we call *duration* is the quantity kind that is called "time" in the SI system [11] (Hayes: "time-dimension"). A property whose nature is a duration is a *particular duration* (Hayes: "duration"), a category of particular quantity.

We also drew from works in logic, linguistics and philosophy by Sowa [22], Parsons [23], Davidson [24], Menzel [25], and others.

2. Infrastructure

We developed infrastructural vocabularies for sequences, quantities and units, and a simple mereology independent of time, as a foundation for dates and times. These vocabularies are modeled in SBVR, UML, Common Logic, and (in the future) OWL, and are modularized, so that they can be used independently of the rest of the Date-Time Vocabulary, in other modeling projects.

The international standards listed above define "time" as a "kind of quantity", along with "length", "mass" and others, that are foundational elements of a "system of quantities" that enable the formulation of derived quantities (e.g., velocity, which is length / time). In such a system, a specific amount of time (e.g., 5 minutes) is a "quantity" whose kind is "time". Although a general model for quantities is outside the scope of this vocabulary, we provide a formal foundation for the basic concepts of the VIM [12] that we use to characterize quantities of time. Some future effort, such as the QUOMOS project [26], may develop a more complete ontology for quantities that is consistent with the international standards.

Our "quantities" vocabulary distinguishes several meanings of "quantity" that are identified in VIM:

- *particular quantity* – a measurable property of a particular thing, for example "the particular duration of this meeting".
- *quantity kind* or *kind of quantity* – a category of particular quantities in which all members are mutually comparable. For example, the particular duration of a

meeting may be compared with the lifetime of a battery – they are both instances of quantity kind "time" – but not with the weight of the battery, which is a different kind of quantity.

- *quantity* – the abstract amount that makes particular quantities equal, e.g., the amount of quantity kind "length" that is both the height and width of a given cube.
- *measurement unit* —a particular quantity that is chosen to be a reference amount of a quantity kind, as "second" is the international standard measurement unit for time.
- *quantity value* is the expression of a quantity in terms of a measurement unit and a number that is the ratio of the quantity in question to that unit. Different measurement units can be used for quantification, resulting in different but equivalent quantity values for one particular quantity. Quantity values can be atomic (e.g., "5 minutes") or compound (e.g., "5 hours 5 minutes").

Other ontologies that we have examined, including both QUDT [27] and QUDV [28], combine two or more of these concepts. We find that confusing.

We developed a robust vocabulary of sequences as a basis for modeling time scales, such as calendars and the 24-hour time-of-day scale. Our sequence model covers finite and indefinite sequences, sequences with repeated or unique or missing members, and regular sequences (even though some of these features are not used in our Date-Time ontology). The model includes relationships among sequence members, and the important idea of an *index* – a number used to identify a position in a sequence. This gives rise to the idea of an *origin* – a reference position in the sequence that establishes the numbering. The model supports zero-origin sequences (as in seconds, minutes, and hours), one-origin sequences (as with days and months), and arbitrary-origin sequences (as with years).

Mereology, the philosophical and mathematical theory of part-whole relationships, is a foundation of our continuous time model. We define a vocabulary and a rigorous set of axioms for the *part of* relationship, corresponding to the "Minimal Mereology" of Casati and Varzi [29]. *Part of* is reflexive (every thing is part of itself), antisymmetric (if two things are part of each other, then they are the same thing), and transitive (if a first thing is part of a second thing which itself is part of a third thing, then the first thing is part of the third thing), thus defining a partial order among parts and wholes. We define *overlaps* (two things overlap if there is a third thing that is part of both) and *proper part* relationships (a thing is a part of a whole if it is not the whole). We include a weak supplementation axiom (if a first thing is a proper part of a second thing then there is a third thing that is a proper part of the second thing and that does not overlap the first thing). We axiomatize these relationships in both the SBVR Structured English syntax and in Common Logic Interchange Format (CLIF) to provide definitional rigor. We rely upon this general mereology plus the "is before" relationship among time intervals, as described below, to define the complete properties of time intervals.

3. Duration

Duration is a kind of quantity. Its instances are amounts of time.

The formal model of duration is a vector space over the mathematical real numbers, because that model (with addition and subtraction of durations, and multiplication of durations by scalar real numbers) supports the "ratio" model for quantity values that represent durations. We define the vector space relationships (e.g., "duration2 equals number times duration1") in both SBVR and CLIF.

We use *time unit* to refer to measurement units for duration quantities, and we use *duration value* to refer to the quantity values. The standard measurement unit for expressing durations is the second. Other commonly used time units (minute, hour, day, week) are defined in terms of second. We call these *precise time units* because they are well-defined and directly comparable.

In addition, we define two *nominal time units – month* and *year* – that are not true measurement units because they vary in the number of days and thus are not directly comparable to other units. For example, comparing the duration value "1 month" to "30 days" is not meaningful because it depends upon the month (e.g., January or February).

We adopted what we believe is a novel approach to comparability for quantity values involving nominal time units. We define these units in terms of sets of days: a month is {28 days, 29 days, 30 days, 31 days}, and a year is {365 days, 366 days}. We note that multiples of months or years are not simple multiples of the days in the sets because of the cyclic nature of the calendar. For example, no two successive years both contain leap days, so "2 years" is {730 days, 731 days}, not {730 days, 732 days}. We develop an algorithm that converts any multiple of months or years to a corresponding set of days based on this observation. The algorithm is particularly useful in converting compound durations that involve months or years to sets of durations using precise time units.

Converting duration values to a common measurement unit is necessary for their arithmetic or comparison. Handling the nominal time units as sets permits results that otherwise would be meaningless. For example, "3 years 1 day" is {1096 days, 1097 days} and thus is necessarily greater than 1095 days. This enables logical reasoning systems to reach conclusions that might be unreachable in other designs.

4. Time intervals

Much of the existing ontological work on time, e.g., [16], is based on the accepted mathematical model of time as a linear continuum, in which "time instants" -- points in time -- can be labeled with scalar multiples of some time unit, e.g., the second. In that model, "time intervals" are segments of the continuum that correspond to mathematical intervals of the real numbers. This analytical model is useful for physics and engineering, because it supports mathematical models of physical phenomena.

We adopt a discrete model of time in which there are only time intervals. A *time interval* is a continuous finite segment of the Time Axis. Every time interval has an arbitrary positive particular duration. No time interval has zero duration. Thus, there are no "points in time" that correspond to mathematical real numbers. (We use the term "time point", but with a different meaning, described in 5.1) This discrete time model supports time concepts that are used in everyday discourse. In particular, for every application, there is some duration for which smaller intervals are not distinguished, and in many cases not distinguishable.. Events occur within some small time interval rather than at some "point in time". This avoids the mathematical conundrums that arise from including idealized points in time in the ontology, notably the problem of time intervals overlapping at a single point. This simplifies the axiomatization of time.

The mathematical algebra of time intervals is based on two fundamental ontological predicates: "time interval *is before* time interval" and "time interval *is part of* time interval". Each of these is a partial ordering of time intervals.

Before: The "time interval is before time interval" relation reflects relative placement of the time intervals on the Time Axis. It is a partial ordering. *is before* is axiomatized as anti-reflexive (no time interval is before itself), strongly antisymmetric (if time interval t1 is before time interval t2, t2 is not before t1), and transitive (if time interval t1 is before time interval t2 and t2 is before time interval t3, then t1 is before t3).

The partial ordering under *is before* is only partial because of overlaps. Non-overlapping time intervals are totally ordered by *is before*.

Part of: Applied to time intervals, the *part of* relationship described above represents the containment of one time interval within another, such as a day within a week. In addition to general mereological axioms, the following apply to time intervals:

There are no atoms – no smallest time intervals: Every time interval has a proper part. Thus, there are no time intervals whose particular duration is zero.

The intersection of two time intervals exists if and only if they overlap. If they overlap, there is a maximal time interval that is the intersection.

In general, there is no complement -- the difference between a time interval and one of its proper parts might be two disjoint time intervals, one on each side.

For any two time intervals, there is a time interval that contains both, but that is trivially satisfied by the indefinite interval *forever* that covers the entire Time Axis. A useful union is defined below.

Allen Relations: From these two basic relationships, we axiomatize a complete characterization of the possible relationships among time intervals, following the approach of James Allen [18]. Allen defines 13 possible relationships between an ordered pair of time intervals, six of which are the inverses of six others. We renamed several from Allen to avoid terminology conflicts.

A time interval t1 is *properly before* a time interval t2 if t1 is before t2 and there is an intervening time interval. The inverse relationship to *properly before* is called *properly after*, but it is not necessary to introduce the relation.

A time interval t1 *meets* a time interval t2 if t1 is before t2 and there is no intervening time interval. The inverse relationship is called *is met by*.

A time interval t1 *starts* a time interval t2, if t1 is a proper part of t2 and they begin simultaneously, that is, no time interval that is part of t2 is before t1.

Similarly a time interval t1 *finishes* a time interval t2, if t1 is a proper part of t2 and they end simultaneously, no time interval that is part of t2 is after t1.

Time intervals that start and finish a longer time interval have complements within the longer time interval. That is, if time interval t1 starts time interval t2, there is a time interval, called the *start complement* of t1 in t2, that is met by t1 and finishes t2. A similar axiom asserts the existence of a *finish complement* of a time interval that finishes t2. Previous axioms imply that some proper part of t2 exists within the start complement interval, but not that it complements the t1 interval that starts t2.

A time interval t1 *properly overlaps* a time interval t2, if they overlap and some part of t1 is before t2 and some part of t2 is after t1. Unlike *overlaps*, this is not a symmetric relationship – the ordering is significant. The inverse is called *is properly overlapped by*.

One can infer the existence of the complementary part of t1 that meets t2 and the complementary part of t2 that is met by t1 from axioms given above.

A time interval t1 *is properly during* t2 if it is a proper part and there are parts of t2 on both sides. The inverse relationship is called *properly includes*.

The remaining relationship is *equals*. The final Allen axiom is: Any two time intervals are related in exactly one of the above ways, including *equals*.

Convex hull: The Allen relations allow us to axiomatize the existence of a useful form of union of two time intervals, called the *sum* or *convex hull* – the smallest time interval that contains both.

For any time intervals t1 and t2, there is a time interval t1+t2 with the following properties:

- if t1 is before t2 or t1 properly overlaps t2, t1 starts t1+t2 and t2 finishes t1+t2
- if t1 is after t2 or t1 is properly overlapped by t2, t2 starts t1+t2 and t1 finishes t1+t2
- if t1 is part of t2, t1+t2 = t2
- if t1 includes t2, t1+t2 = t1

Note that t1+t2 is well-defined. There are no other possible choices for the relationship between t1 and t2, and in the case t1 = t2 (which satisfies both of the last two conditions), t1 = t1+t2 = t2.

It follows that for all time intervals t1 and t2, the sum t1+t2 includes t1, t2, and any time interval between them. It also follows that t1+t2 is minimal, that is, any time interval that includes both t1 and t2 must include t1+t2.

Of course, if t1 is *properly before* t2, there are intervening time intervals that are part of the sum and do not overlap either t1 or t2. In fact, one can show that there is a maximal such interval. This property is inconsistent with typical mereological "union" axioms.

Finally, the convex hull of two finite time intervals that *meet* is a useful sum – its particular duration is the sum of the particular durations. From this, we derive duration difference relationships for the start complement and finish complement.

This set of predicates and axioms is the foundation for the ontology.

5. Discrete Time

The objective of the Date Time Vocabulary is to deal with the more common commercial uses of time, which deal primarily with discrete, well-defined time intervals. Our discrete time vocabulary covers concepts that are used in every day discourse of non-zero duration.

5.1. Time Points, Time Scales, Time Periods, and Calendars

We now turn to the division of the Time Axis into discrete time intervals using various *time scales*. Each time scale is a regular sequence of *time points*, each of which is conceptually a classifier of time intervals. All of the time intervals described by a time scale have a common duration that is a particular time unit – the *granularity* of the time scale. A time point corresponds to all of the time intervals that are in the position that it represents relative to some reference time interval. For example, *Tuesday* refers to each time interval that has a particular duration of 1 day and is met by a *Monday*.

"The third Tuesday of each calendar month" describes the time intervals with index 3 in the sequence of Tuesdays contained in each calendar month.

A calendar is a schema that defines one or more related time scales of different granularities. For example, Coordinated Universal Time (UTC) is a calendar with time scales of year, month, day, hour, minute, and second granularities. Civil calendars are traditionally associated with time offsets from UTC. Calendars are related to one another by identifying common events on each, e.g., midnight in Greenwich.

We distinguish between *indefinite time scales* and *finite time scales*. Indefinite time scales extend without limit into the past and the future, representing the idea that time points such as calendar years and calendar days extend indefinitely in both directions. These time scales are defined to align the time point for 20 May 1875 to the signing of the Convention du Mètre, which is the standard reference point for the Gregorian calendar. On an indefinite time scale, every time point corresponds to exactly one time interval – the reference point is a specific interval on the Time Axis. For example, there is one and only one UTC time interval that is the "calendar year 2000".

Finite time scales have fixed size, and they have relative reference points -- they *subdivide* longer time points into sequences of shorter time points. For example, the finite time scale *hour of minutes* contains 60 time points and subdivides each time point that is an *hour of day*. Time points that are on finite time scales identify *sets* of time intervals because such time scales "repeat" – the reference point for each subdivision is the start of the time point being subdivided. We call these *relative time points*. For example, there is a "15th day of the month" in every calendar month, and a March in every calendar year. This structure captures the complex containment relationships among time scales.

We define all the time units and corresponding time scales specified in ISO 8601 – seconds, minutes, hours, days, weeks, months, and years. This goes beyond XML Schema datatypes [10], which does not address weeks. We also specify the Internet Time Scale, which formalizes the scheme used by the Network Time Protocol [30].

The choice of time scale establishes the precision of time references. For example, if a railway schedule is given to the minute, a train arriving within that minute is on time, and a train arriving in the next minute is 1 minute late, even if the actual elapsed time between the two is 10 seconds. The schedule specifies the named minute in which the event occurs. This is the nature of the discrete view of time.

5.2. Time Coordinates

We distinguish between time points and names for time points (*time coordinates*) – we say a time coordinate *indicates* a time point – because some time points have multiple names. This is particularly true of time points that have names as well as indices, and those that are indicated by *compound time coordinates*, which have multiple name components. For example, "February 1" "indicates" the same time point as "the 32^{nd} day of the year", which corresponds to one time interval in each calendar year. The relationship between time coordinates and time points uses an SBVR feature called "reference schemes", similar to OWL reference keys: a property or combination of properties whose values uniquely identify individuals. We define time coordinates as keys for time points. Thus we have a three-level structure with time coordinates that "indicate" time points whose extensions are the actual time intervals.

We also distinguish between *absolute time coordinates* and *relative time coordinates*: the former are on indefinite time scales and always mention a calendar year, whereas the latter are on finite time scales. Each absolute time coordinate indicates an individual time interval, whereas each relative time coordinate indicates a set of time intervals. For example, "January 29" refers to a set of calendar days, one for each calendar year.

Addition of a duration value to a time coordinate, subtraction of a duration value from a time coordinate, and difference of two time coordinates (a duration value) use standard mixed-based arithmetic. However, difference is not defined between relative time coordinates that span the end of February because of the possibility of a leap day. For example, "2 March – 28 February" may be either 2 days or 3 days.

Comparison of absolute time coordinates is performed according to the ordering of the corresponding time intervals on the Time Axis. We define comparison of relative time coordinates as within the boundaries of the "next larger" time scale. Thus we think of February as being before March even though February of a later year follows March of an earlier year. We specify such comparisons as converting the two time coordinates to a common finite time scale and comparing their indices on that scale. Some relative time coordinate comparisons require special logic because of leap days. For example, the comparison "day-of-year 59 is before March 1" is true in leap years and false in common years. We handle this by converting time coordinates that include months to *time sets* that represent multiple possible equivalent time points on the target time scale. For example, we convert "March 1" to the time set {day-of-year 59, day-of-year 60} to represent the two possible time points that "March 1" could be on the Gregorian year-of-days scale. Then we compare the results of the conversion. This scheme permits comparison of any two relative time coordinates that fill either of these criteria: (a) are at least two days apart; or (b) do not span the end of February.

6. Temporal Concepts for Situations

The raison d'etre for the Date Time Vocabulary is to associate situations and events with time intervals. The ontology uses the general concept *occurrence* to refer to things that happen. It does not distinguish events from states or activities. The underlying philosophy follows Davidson [24] in that an occurrence is a thing in the universe of discourse and it may be an argument in a relation. In particular, we define predicates that relate occurrences to time, following the approach that is typically used in formalizing natural language. The alternative approach used in DOLCE [18] is to make time an explicit role in each "occurrent" relation.

Everything that happens in a world of interest happens at some point(s) in time in that world. Having only time intervals of non-zero particular duration, DTV assumes that each occurrence has some associated time interval, called its *occurrence interval*, over which it is continuously a part of the state of the world. The occurrence interval may be in the past, or the future, or overlap the "current" time interval. We assume that time and change over time is a part of possible worlds, rather than a factor that distinguishes possible worlds, and that each possible world populates its conceptual Time Axis with its own occurrences. In particular, future time is a part of the world, not a modality.

We refer to the conceptualization of a situation – a category of possible occurrences – as a *situation model*. Occurrences – the things that actually happen in a world of interest – are said to *exemplify* situation models. For example, the situation model "weekly staff meeting" is exemplified by multiple individual meetings, each of which is an occurrence. A situation model can fail to occur – have no exemplifying occurrences; and it can "recur" – have more than one distinguishable occurrence. In many cases, two occurrences of the same situation may be distinguished by having distinct occurrence intervals. Ultimately distinguishing occurrences of a situation model depends on properties chosen by the viewer. An occurrence is a situation that is viewed as ongoing from inception to termination.

The fundamental axioms for associating time with occurrences are based on the primitive predicate "occurrence *occurs throughout* time interval". It means only that the occurrence is ongoing in every time interval that is part of the time interval. For example, World War II *occurs throughout* 1944; it did not *occur throughout* 1945. In this section we use the present tense in the historical examples. The interpretation of tense is discussed in the next section.

An occurrence is said to *occur within* a time interval if it occurs throughout a time interval that is part of that time interval. E.g., World War II occurs within 1945. Importantly, if an occurrence does not occur within a given time interval, there is no time interval during that time interval in which the occurrence is part of the state of the world.

An occurrence is said to *occur for* (or *occur over*) exactly one time interval – its *occurrence interval*. The occurrence starts happening at the start of the occurrence interval and stops happening at the end of the occurrence interval. E.g., World War II occurred for the time interval from September 1, 1939 to August 14, 1945 (assuming the time scale granularity is *day*).

Lacking the time instant concept, we define two other predicates: An occurrence *starts with* a time interval *t* if either *t* starts the occurrence interval or the occurrence interval starts *t* (where *starts* is the Allen relation described earlier). Similarly, an occurrence *finishes with* a time interval *t* if either *t* finishes the occurrence interval or the occurrence interval finishes *t*.

Occurrences are partially ordered by their occurrence intervals. That is, an occurrence *A occurs before* another occurrence *B* if the occurrence interval of *A* is before the occurrence interval of *B*. Occurrences *overlap* if their occurrence intervals overlap; they *start jointly* if one occurrence interval starts the other, and they *finish jointly* if one occurrence interval finishes the other. These relationships allow us to use relative time in describing relationships among events.

In practice, a few occurrences have names, e.g., World War II, but the majority of occurrences are referred to as instances of situation models. That is, an occurrence is considered to be a thing that corresponds to a conceptualization of a situation, even when it is the unique instance. The general form of a reference to a category of occurrences – a situation model – can be a simple predicate, e.g., "war", or a combination of qualifying predicates, e.g., "wars that were fought by the U.S.", or a sentence/proposition that characterizes those occurrences, e.g., "Germany invades France" or "the Convention du Mètre is signed". As a consequence, most propositions that state relationships between occurrences and time have the form: "situation model *occurs throughout/within/for* time interval", which is defined as "there exists an occurrence of that situation model that occurs throughout/within/for that time interval". E.g., "Germany invades France in 1870" is formulated as "The situation "Germany invades France" occurs within calendar year 1870", and the meaning is that an occurrence of that situation model occurred within calendar year 1870. In the IKRIS Knowledge Language (IKL) [31] [32], a dialect of CL, this is:

```
(forall ((y calendar_year)) (if (indicates "1870" y)
  (forall ((t time_interval)) (if (y t)
    (exists ((s occurrence)) (and
      (proposition_describes_occurrence
            (that (invades Germany France)) s)
      (occurs_within s t) )) )) ))
```

That is, if t is the time interval that instantiates the calendar year indicated by 1870, then an occurrence of "Germany invades France" occurs within t. "y" is a variable that refers to a time point as a predicate.

Reasoning about occurrences and time requires a logic that supports nominalization of propositions, so as to support "proposition describes occurrence". We use IKL "that" to nominalize propositions.

7. Logical Formulation of Tense and Aspect

The previous section described the relationships between situations and explicit time intervals. In addition, it is necessary for many purposes to address relative time, and, in particular, the relationship between the time of an event and the time of an utterance.

Natural languages express relative temporal information in the tense and aspect of verbs. A tense is a verb form that relates the time of the action of the verb to the time of utterance. An aspect is a verb form that indicates whether the situation associated with the verb is in process or completed. English indicates the present and past tenses via verb forms (e.g., *run* and *ran*, *plan* and *planned*), and the future using the modal verb "will", sometimes with the copula "be". There are two aspects in English: the progressive or continuing aspect (typically using the auxiliary verb *be* with a form ending in *-ing*, as in *I was reading a book*) and the perfect or perfective (expressing completed action, typically using the auxiliary verb *have* with a past participle, as in *I have read the book*). We provide guidance on how to formulate the semantics of such sentences, enabling the capture of a wide variety of temporal distinctions. Thirteen different tense/aspect example combinations are given in the Date-Time specification. When using DTV, every sentence has a temporal interpretation.

Tense is represented in our vocabulary by the time concepts *past*, *present* (*current time* or *now*) and *future*. These are called "indexical times". The indexical verbs – *is past*, *is current*, *is future* – characterize time intervals in relationship to a reference time interval that is determined by context. *is past* means a time interval that is *before* the reference time interval; *is current* means "now", formally, a time interval that *includes* a time interval that *is past* and that *includes* another time interval that *is* not *past*; *is future* means a time interval that *includes* no time interval that *is in the past*. Indexical nouns, such as *today* and *next week* refer to time intervals using those characteristics.

In most cases, the reference time interval is the "now" time of the utterance, or the "now" time of the reasoning/decision process, indicated by the *current time* in DTV. For example, "next month" is an indexical in the statement "The construction must be finished by next month". In such cases, a time of utterance can be set for an entire corpus (e.g., a CLIF module) by specifying the value of the DTV symbol *current_time*.

On the other hand, "the clerk must check that the vehicle has been returned" refers to a time interval relative to the time the clerk is checking.

English does not have a "timeless" tense, so we borrow existing verb forms to formulate the base atemporal situation from the original sentence. It is customary, but not necessary, to use verb forms in the simple present tense for this purpose. This base situation is then restricted to occur at the time prescribed by the verb. For an example, consider the past tense sentence "Joyce wrote *Ulysses.*" Its atemporal base situation is formulated using the verb concept "person *writes* book" as "Joyce writes *Ulysses.*" This base is then restricted to occur in the past. Here is a logical formulation of "Joyce wrote *Ulysses,*" in IKL:

```
(person Joyce) (book Ulysses) (time_interval current_time)
(cl:comment "specify the utterance time, the publication
             date"
    (indicates "19220202" current_time))
(exists ((o occurrence))
  (and
    (proposition_describes_occurrence
         (that (person_writes_book Joyce Ulysses)) o)
    (cl:comment "past tense restriction"
         (occurrence_is_past o)) ))
```

The DTV predicate "proposition *describes* occurrence" is a one-to-many relation between a proposition and its occurrences, and the predicate "occurrence *is past*" means that the occurrence happened before "current time", where "`current_time`" is the putative utterance time. The example gives it a value, in order to de-index [32] the indexical reference. The IKL example means "Joyce wrote *Ulysses* before February 2, 1922." Alternatively, an occurrence can be related to a time interval which itself is said to be in the past relative to a reference time.

Indexicals play an important role in temporal expressions, as demonstrated in the example. Many business rules are stated in terms of relative times using indexical time intervals. Indexical referents can be grounded in time by giving effective dates of rules as explicit temporal restrictions in the rule statements. Date-Time provides more than 30 common time indexicals, including *current day*, *last week*, *next month*, *future year*, etc. The list is not exhaustive, but is easily extended by users.

The formulation of aspect is similar. The perfective aspect is formulated in our vocabulary by using the predicate "situation *is accomplished*". The progressive aspect is formulated using "situation *is continuing*". The "situation" is an occurrence that is described by an atemporal proposition derived from the original sentence, as described above. This occurrence formulation is then used with an aspect predicate to formulate the aspect, and also with a temporal predicate, which may be explicit or indexical, to formulate the time. The overall formulation is a conjunction that combines the aspect formulation and the temporal formulation.

For example, "Joyce was writing Ulysses" can be formulated:

```
(person Joyce) (book Ulysses) (time_interval current_time)
(cl:comment "specify the utterance time, the publication
             date"
    (indicates "19220202" current_time))
(exists ((o occurrence))
  (and
    (proposition_describes_occurrence
         (that (person_writes_book Joyce Ulysses)) o)
    (cl:comment "past tense restriction"
```

```
            (occurrence_is_past o))
        (cl:comment "performative aspect restriction"
            (occurrence_is_continuing o))
    ))
```

8. Conclusion

We believe this work contributes in the following ways:

- Covering temporal concepts in breadth: relationships between time and situations, continuous and discrete time, temporal indexicals, schedules, and language tense and aspect.
- Providing linguistic forms for the noun concepts and relations, so that they can be used in business rules stated in a Controlled Natural Language.
- Targeting business use, but also providing formal definitions in UML, CLIF and eventually OWL, so that implementations of the rules can rely upon these formal definitions.
- Demonstrating how vocabularies defined in SBVR can map to multiple software languages and knowledge representation formalisms.

We created a vocabulary of formalized English descriptions of time concepts, and a matching vocabulary for software models that incorporate time concepts, both of which are supported by a rigorous ontology. These provide a sound semantics for the use of these concepts in formulating rules for behaviors of humans and machines.

This effort exposed a significant tension between three goals that conflict in part:

- The SBVR focus is on the business perspective – "a vocabulary that is conceptualized optimally for the way people think and communicate about things in their organizations using natural language ...." [33].
- The Common Logic and OWL focus is on machine interpretation and reasoning: machine processing (by automated reasoners) is possible, based on a well-grounded formal representation.
- The UML focus is on implementation.

We managed this conflict by carefully distinguishing related concepts, by describing the relationships among concepts precisely and in detail, and by providing extensive supporting rationale, notes, and examples. We believe the result finds a reasonable compromise among these three goals.

Acknowledgements

The Date-Time Vocabulary is the work of a team, including the authors, and Don Baisley, Roger Burkhart, Cory Casanave, Donald Chapin, Paul Haley, John Hall, Terry Halpin, Elisa Kendall, Sjir Nijssen, Ron Ross, Markus Schacher, and Ravi Sharma.

References

[1] Object Management Group (OMG), *Date-Time Vocabulary*, available at http://www.omg.org/spec/DTV/.
[2] Object Management Group (OMG), *Semantics of Business Vocabularies and Business Rules* (SBVR), http://www.omg.org/spec/SBVR/.

[3] Object Management Group (OMG), *Unified Modeling Language* (UML®), http://www.omg.org/spec/UML/.
[4] Object Management Group (OMG), Object Constraint Language (OCL), http://www.omg.org/spec/OCL/.
[5] International Standards Organization/International Electrotechnical Commission (ISO/IEC) 24707, *Common Logic* (CL), http://standards.iso.org/ittf/PubliclyAvailableStandards/c039175_ISO_IEC_24707_2007%28E%29.zip.
[6] World Wide Web Consortium, *Web Ontology Language 2* (OWL2), http://www.w3.org/2007/OWL/wiki/OWL_Working_Group.
[7] Enterprise Data Management Council (EDMC) and Object Management Group (OMG), *Financial Industry Business Ontology* (FIBO). Forthcoming.
[8] International Standards Organization (ISO) 8601, *Data elements and interchange formats – Information interchange – Representation of Dates and Times*, Third edition, December 1, 2004.
[9] International Standards Organization (ISO), ISO/IEC 80000-3, *Quantities and units -- Part 3: Space and time,* 2006.
[10] World Wide Web Consortium (W3C), *XML Schema Part 2: Datatypes Second Edition*, 28 October 2004, http://www.w3.org/TR/xmlschema-2/.
[11] Bureau International des Poids et Mesures (BIPM), *The International System of Units* (SI), 8th edition, 2006.
[12] International Standards Organization/International Electrotechnical Commission (ISO/IEC), JCGM 200:2008, *International Vocabulary for Metrology – Basic and General Concepts and Associated Terms* (VIM), 3rd edition.
[13] International Standards Organization (ISO), ISO 18026, *Information technology — Spatial Reference Model (SRM)*, 2009.
[14] World Wide Web Consortium (W3C), *An OWL Ontology of Time*, 27 September 2006, http://www.w3.org/TR/owl-time/.
[15] Hobbs, Jerry, *An OWL Ontology of Time*, July 2004, http://www.isi.edu/~hobbs/time/owl-time-july04.txt.
[16] Pan, Feng, *Representing Complex Temporal Phenomena for the Semantic Web and Natural Language*, PhD Thesis, 2007, http://www.isi.edu/~hobbs/time/pub/pan-phdthesis.pdf.
[17] Niles, I., and Pease, A., *Towards a Standard Upper Ontology,* proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Chris Welty and Barry Smith, eds, Ogunquit, Maine, October 17-19, 2001.
[18] Masolo, C., et al., *The WonderWeb Library of Foundational Ontologies Preliminary Report*, Institute for Cognitive Sciences and Technology, National Research Council (Italy), 2003, http://www.loa.istc.cnr.it/DOLCE.html.
[19] Allen, James, *Maintaining Knowledge about Temporal Intervals*, Communications of the ACM, Volume 26, Issue 11, November 1983, pp. 832-843, http://portal.acm.org/citation.cfm?id=358434.
[20] Hayes, Pat, *A Catalog of Temporal Theories*, University of Illinois Technical Report UIUC-BI-AI-96-01, 1995-1996, http://www.ihmc.us/users/phayes/docs/TimeCat96.pdf.
[21] The New Oxford Dictionary of English.
[22] Sowa, John, *Laws, Facts, and Contexts,* 2003, www.jfsowa.com/pubs/laws.htm.
[23] Parsons, Terence, *Events in the Semantics of English: a study in subatomic semantics*, MIT Press, 1990, ISBN 0-262-16120-6, www.humnet.ucla.edu/humnet/phil/faculty/tparsons/Event%20Semantics/download.htm.
[24] Davidson, Donald, *The Logical Form of Action Sentences*. In: Nicholas Rescher (ed.), *The Logic of Decision and Action*, Pittsburgh: The University Press, pp. 81-95. (1967).
[25] Menzel, Christopher, *Actualism*, The Stanford Encyclopedia of Philosophy (Fall 2011 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2011/entries/actualism/.
[26] Ontolog Forum, *A Quantity and Unit of Measure Ontology-based Standard* (QUOMOS), http://ontolog.cim3.net/cgi-bin/wiki.pl?UoM_Ontology_Standard.
[27] Hodgson, R., and Keller, P., *QUDT – Quantities, Units, Dimensions, and Data Types in OWL and XML*, qudt.org.
[28] Object Modeling Group (OMG), *Quantities, Units, Dimensions, Values* (QUDV), http://www.omgwiki.org/OMGSysML/doku.php?id=sysml-qudv:quantities_units_dimensions_values_qudv.
[29] Casati, Roberto, and Varzi, Achille C, *Parts and Places*, MIT Press, 1999.
[30] Internet Engineering Task Force (IETF), RFC 5905: *Network Time Protocol Version 4: Protocol and Algorithms Specification* (NTP), http://www.ietf.org/rfc/rfc5905.txt.
[31] Hayes, Pat, and Menzels, Chris, *IKL Specification Document*, http://www.ihmc.us/users/phayes/IKL/SPEC/SPEC.html.
[32] Hayes, Pat, *IKL Guide*, http://www.ihmc.us/users/phayes/IKL/GUIDE/GUIDE.html.
[33] Chapin, Donald, private communication.