Computational Linguistic Approaches to Temporality

(Draft 2.2, 4 April 2011

Mark Steedman

1 Introduction

Temporality in computational linguistics and natural language processing can be considered from two aspects. One concerns the use of linguistic and philosophical theories of temporality in computational applications. The other concerns use of computational theory in its own right to define new kinds of theories of dynamical systems including natural language and its temporal semantics. The latter influence is at least as important as the former.

2 Linguistic Contributions to Computational Linguistics

As in the case of nominal expressions in natural language, we should be careful to distinguish temporal *semantics*, or the question of what kinds of objects and relations temporal categories denote, from the question of temporal *reference* to particular times or events that the discourse context affords.

It is useful to draw a further distinction within the semantics between temporal *ontology*, or the types of temporal entity that the theory entertains, such as instants, intervals, events, states, or whatever, temporal *quantification* over such entities, and the temporal *relations* over them that it countenances, such as priority or posteriority, causal dependence, and the like.

2.1 Temporal Semantics: Ontologies, Quantifiers, and Relations

Applications such as information retrieval (IR), and question-answering (QA), have made surprisingly little use of the riches offered by linguistic temporal semantics and temporal logic. The reason is the inextricable entanglement of the temporal categories with everyday knowledge. As other chapters of this handbook show, categories like tense, mood, and aspect are confounded with non-temporal relations such as causality, teleology, counterfactuality, evidentiality, and the like, to an extent that makes firm boundaries to temporal semantics hard to draw.

For example, the question

(1) Have you met Miss Jones?

does not define a temporal relation between the present time and an event of meeting Miss Jones, as theories of temporal reference founded on Reichenbach (1947) might seem to suggest, but rather asks whether the state of affairs that is *consequent upon* such an event—roughly speaking, *knowing Miss Jones*—is in force (Moens and Steedman, 1988). One may be able to answer such questions in the affirmative even if one has no recollection of the event in question, nor any idea when it might have been, or even lacks the capacity for such recollection, as in the case of certain agnosias.

Of course, one may infer that meeting Miss Jones must have preceded the present, for this state of affairs to hold—but that is an entailment of the relation between cause and effect, rather than temporal sequence as such.

Similarly, if a search engine offers (2A) in answer to a query (2Q), in order to answer the question correctly, a question answerer must understand the textual entailment of A that, although one might have expected Swatman to win, in the event he did not:

(2) Q: Did John Swatman win the British Open Gold Medal?

A: In 1980 at 16 years of age he fought his way to the final in the under 60 Kg category, and *was winning the contest when he was forced to withdraw through injury*.

That is, the progressive denotes a state of affairs that would normally bring about a win, rather than a temporal relation to an actual event of winning.

Such inferences are extremely specific to the particular content that is involved. Thus, the temporal extent of the state of *having met Miss Jones* is generally (as the song says) only bounded by the lifetime of the participants. However, if the following question is asked, the relevant consequent state is bounded by our knowledge of the digestive process to a few hours:

(3) Has the patient had anything to eat?

Similarly, if the question is:

(4) Has the patient had a tetanus shot recently?

—then the answer depends on specific knowledge of the length of time the consequent protection typically lasts, which is a few years.

Human beings are remarkably good at such associative inference, which they seem to achieve quite effortlessly. However, the problem of formulating such knowledge in computational terms, and carrying out similar inference by machine, using the standard logics that have been designed for the purpose (Prior, 1967; McDermott, 1982; Allen, 1983), is very hard, although Prior's work provided one of the foundations for the computational dynamic logics discussed in section 3. Pratt and Francez (2001) formulate temporal generalized quantifiers for such a framework, and Pratt-Hartmann (2005) proves complexity properties of this system.

There have been attempts to design limited logics with better search properties, which have tended to trade under the name of "temporal database query languages", and there have been attempts to design natural language user-interfaces or "front ends" for such systems, drawing on linguistically-informed semantics, notably the ontologies of Vendler (1967) and followers (e.g. Bruce, 1972; Ritchie, 1979; Moens, 1987; Hinrichs, 1988; Palmer et al., 1993; Crouch and Pulman, 1993; White, 1994; Dorr and Olsen, 1997; Androutsopoulos et al., 1998; Dorr, 2007).

Such ontologies typically distinguish event-types according to a number of dimensions including $\pm$durativity and $\pm$telicity, and distinguish states as of type progressive, consequent, iterative, habitual, and so on. Several of these systems constitute recursive mereologies, or part-whole hierarchies, sometimes embodying a notion of type-*coercion* or *overloading*, whereby aspects and adverbial modifiers compositionally add layers of temporal predication such as preparation, initiation, iteration, and culmination, and the like, without any limit on depth of embedding, as in:

(5) It took me two years to be able to play "Young and Foolish" in less than thirty seconds for up to an hour at a time.

Further examples of work of this linguistically informed kind are Pustejovsky 1991a; Kameyama et al. 1993; Hitzeman 1997; Narayanan 1997. However, such attempts at general-purpose solutions to the problem of temporal query have not been widely used, and typically involve serious investments of time in knowledge-engineering by hand.

Instead, early natural language query systems for temporally rich domains, such as Sager et al. 1994 tended to hard-wire the requisite knowledge into collections of very specific hand-built inference rules representing actions directly, in ad-hoc (although nonetheless revealing) ways, in the tradition of AI action representations systems, to which we will return in section 3.

The CYC project (Lenat, 1995) is an attempt in the same tradition to build a knowledge representation of this kind by hand on a very large scale, and with general applicability including temporal reasoning, based on a rather general ontology and a number in the millions of specific rules of inference, many of which represent relations among events and states. Specialized versions of CYC have been developed for supporting effective inference in several narrower "closed" domains, many of them industrially significant, and a research version is freely available.

However, common sense reasoning about the everyday world, of the kind that is needed to capture the simple natural language examples with which this piece began, remains a very difficult task. Attempts to apply the publicly available version in tasks like text-based inference have not proved very successful (Mahesh et al., 1996), The general suspicion is that this is because such hand-built resources are both too high-level in terms of ontology, and too small-scale in comparison to what a mixture of animal evolution and social learning has put in our own heads.

In reaction to this realization concerning the limits of hand-built knowledge resources of many kinds, including ontologies such as WordNet (Fellbaum, 1998b), FrameNet (Baker et al., 1998), and VerbNet (Kipper et al., 2008), there is renewed interest recently in building larger resources of this kind by clustering on collocations in large volumes of text (Lin and Pantel, 2001), or by searching such text corpora for string-based proxies for ontological relations (Webber et al., 2002). The proposals of Banko and Etzioni (2007); Etzioni et al. (2007) and Mitchell et al. (2009) for "machine reading" to create similar knowledge resources are related.

A more radical proposal of this kind is to use more directly associative knowledge representations such as associative memory models and semantic networks whose nodes directly correspond to individuals of various types, and whose arcs represent relations between them. Harrington and Clark (2009) have proposed a method using a wide coverage parser to parse unlabeled text and constructing a semantic network an orders of magnitude larger than CYC, using spreading network activation to bound the complexity of network access and update.

The biggest obstacle to such ambitious plans is the low reliability of wide-coverage parsers. Because of the large search spaces involved, the best-performing parsers use parsing models (and often grammars) acquired by "supervised" machine learning from human-annotated corpora such as the one million word Penn *Wall Street Journal* Treebank. The models are acquired by deriving probabilities (or feature weights) from the frequency with which components of derivations are found in the corpus, in order to choose the derivation that most closely resembles the training material.

The most successful parsers frequently exploit "head dependency" models, in which probabilities or weights are computed on the basis of frequency of co-occurrence of relations between particular words, such as the noun "days" acting as the subject of verb "elapsed". Such parsers, while performing better than the hand-built alternative, still have dependency recovery rates of only around 90% (Collins, 2003). Since word dependencies are closely related to semantic predicate-argument relations, there is a danger that the structures the parsers deliver will be too errorful for this process to deliver useful semantic networks.

Error analysis suggests that the reason the parsers are so weak is that 1M words of fairly arbitrarily selected annotated newspaper text is not enough to give us a grammar or a parsing model comparable to what we have in our own heads. There is considerable work going on to develop unsupervised methods for parser induction from unannotated text, and semisupervised methods for using unlabeled text to generalize the treebank parsers.

The relative success of supervised-learned parsers using head dependency models trained on human-labeled data might seem to suggest a quite different and much bolder solution to the common-sense reasoning bottleneck in temporal semantics. The content-dependency of the extent of the consequent state denoted by the perfect on the nature of the core event—*meeting Miss Jones* versus *eating something* versus *having a tetanus shot*—is reminiscent of the way in which those parsing models rank parses by assigning higher probability to a head-word dependency that occurs frequently in the training data than one that appears rarely or not at all. Head-word dependency parsing models work because they approximate a mixture of semantics and real world knowledge that underlies frequent collocations.

One can therefore consider as a thought experiment the idea of approximating a similar mixture underlying the interpretation of tenses, moods, and aspects, by having human annotators annotate texts about events like *meeting Miss Jones* with the implicit consequent states like *knowing Miss Jones* and preparatory processes like *traveling for the purpose of fulfilling an appointment with Miss Jones*, together with their temporal extents, and learning a model that would allow a machine to answer questions like *Was Mr. Smith meeting Miss Jones when he had the accident?* and *had he met Miss Jones?*.

Of course, such an experiment is completely unrealistic, both in terms of the possibility of obtaining reliable annotations, and in terms of the amount of annotated data that would be required for effective machine learning, let alone in terms of the limitations of the learning techniques themselves when faced with an essentially AI-complete problem. However, a scaled-down version of this idea is being attempted in the related area of temporal reference, to which we now turn.

2.2 Temporal Reference

By temporal reference is meant the anchoring of temporal descriptions to specific clock-times, or to other events in an established narrative. The simple tenses—the past, present, and in English the bare infinitival future—are temporally referential, in the sense that their underlying Reichenbachian reference time R must stand in an inferable temporal/causal relation to some time or event that is either already give discourse-given, as in (6a), or provided by a modifier that is itself temporal/causally anchored, as in (6b,c);

(6) a. It was the night they raided Minsky's. I met Miss Jones.
  b. I met Miss Jones the night they raided Minsky's.
  c. I met Miss Jones when they raided Minsky's.
  d. I met Miss Jones soon after they raided Minsky's.

Webber (1988) points out that such temporal anchoring processes resemble definite noun-phrase reference in allowing "bridging" reference to inferred referents, as when the mention of a car supports reference to "the driver" (Clark and Marshall, 1981). Such inferences are knowledge-dependent in the same way as the temporal semantic interpretations considered in the last section. Thus, in the following example, the fact that we know that throwing us in jail *followed* the raid, and that coming with a warrant *preceded* it, is a matter of world-knowledge about preparations for, and consequences of, such events:

(7) The night they raided Minsky's,

  a. they threw us all in jail.

  b. they came with a warrant

Such discourse is also characterized by shifting temporal focus: once we have decided the relation of throwing in jail to the raid, it may in turn act as anchor for further events such as phone calls, which may act as anchors for events of obtaining bail, and so on, under similar conditions of script-like general knowledge about goal-directed activities (Schank, 1975).

As in the case of early work on temporal semantics, for suitably closed domains we may be able to finesse explicit reference. Wiebe et al. (1998) describe a domain-specific rule-based approach to temporal reference resolution in the sense of time-stamping for a corpus of scheduling dialogs consisting of exchanges like the following:

(8) a. Would you like to meet Wednesday, August 2nd?

  b. No, how about Friday at 2?

This work uses a graph-structured stack as a focus model that allows non-adjacent antecedent anchors in complex dialogs. Filatova and Hovy (2001) offer a related approach to time-stamping event clauses in the more open domain of newspaper stories, including relations of anteriority and priority. Both papers evaluate on held-out data—that is, additional human-labeled data that have not been used for training.

Developed as part of the ACE (automatic content extraction) initiative hosted at the Linguistic Data Consortium at the University of Pennsylvania, the TIMEX2 annotation scheme (Ferro et al., 2005) has been used to annotate corpora such as the ACE 2005 corpus (around 600 documents), which has been used for training and evaluating automatic temporal expression recognition and normalization (TERN) programs using a mixture of small numbers of hand-written rules and machine learning (e.g. Ahn et al. 2007)

The TimeML temporal mark-up language (Pustejovsky et al., 2003a; Verhagen et al., 2009) is a reformulation of TIMEX2 that has been extended to cover events, temporal relations, and certain kinds of state, and used for annotation of the Timebank corpus of 186 news reports (Pustejovsky et al., 2003b). Pan et al.

(2006) have extended the Timebank annotation to include estimated upper and lower bounds on the temporal extents of TimeML temporal expressions, with reasonable inter-annotator agreement. Chambers et al. (2007) present a temporal relation classifier for six relations trained on the original Timebank corpus, reporting 72% accuracy when these relations are collapsed to the simplest binary classification *before/after*.

Mazur and Dale (2010) criticize both ACE and Timebank for the brevity of the documents that they include, and the limitations on the complexity of the temporal reference that they support. They point out that most temporal expressions in these corpora can be interpreted relative to a single temporal focus or "anchor", defined as the dateline of the report, rather than involving the kinds of shifting focus characteristic of extended discourse and narrative. They offer an alternative WikiWars corpus comprising 22 much more extended Wikipedia articles on the major wars of human history, containing around 2700 TIMEX2annotated temporal expressions.

It is possible in principle that the typical extent of events could be learned from such data, and used to improve TERN-style temporal reference programs of the kind discussed earlier. While TimeML does not mark consequent states (of the kind crucial to the interpretation of (1), *Have you met Miss Jones?*) as such, it does mark "signal words" such as modals and auxiliary verbs, so it is even possible in principle that the typical temporal extent that should be considered in answering questions like (3) (*Has the patient eaten anything?*) and (4) (*Has the patient had a tetanus shot recently?*) could be learned.

Not surprisingly, nothing as ambitious as this has been attempted so far. As Lapata and Lascarides (2006) point out, these corpora are quite small in comparison with the Penn *Wall Street Journal* treebank (the TIMEX2-annotated English portion of ACE 2005 is around 26K words, while TimeBank is around 69K.) Given the sparse nature of these data, and the sheer difficulty in many cases of annotating temporal relations reliably, it is unclear whether supervised learning with human-labeled data can succeed practically on this problem, although, as Lapata and Lascarides point out, the labeled corpora remain valuable as gold-standards against which unsupervised methods can be evaluated.

In reaction to these resource limitations for supervised learning, there has been considerable research into unsupervised methods for training such classifiers using unsupervised methods based on wide-coverage parsing of unlabeled text. Chklovski and Pantel (2004) learn verb subcategorization frames and semantic relations between them including temporal relations, the latter chosen on the model of Fellbaum 1998a. Lapata and Lascarides (2006) have used such methods successfully to automatically extract a restricted class of specifically temporal relations, by parsing unlabeled text using a wide-coverage statistical parser trained on the Penn treebank, in search of main and subordinate clauses linked by temporal connectives such as "after", "while", and "until", evaluating against the human-labeled Timebank corpus (see above). Chambers and Jurafsky (2009) show how script-like narrative chains involving shared participants can be mined using similar unsupervised methods, evaluating in comparison to the related but non-narrative relations in the hand-built FrameNet corpus (Baker et al., 1998), as well as by a novel "narrative Cloze" procedure.

Automatically identifying temporal semantics and temporal reference remains an extremely hard problem, to which linguistic semantics provides only part of a solution which we do not seem very close to attaining. The Recognizing Text Entailment (RTE) task (Dagan et al., 2006) attempts to provide a standard test-set of pairs of text passages of the kind delivered by real information retrieval and machine translation programs, and questions or "hypotheses" which the text may or may not answer in the positive or negative. Many of the examples involve temporal reference, such as the following:

(9) T: Bush returned to the White House late Saturday while his running mate was off campaigning in the West.
H: Bush left the White House. (RTE example no. 1960:PP)

(10) T: De La Cruz's family said he had gone to Saudi Arabia a year ago to work as a driver after a long period of unemployment.
H: De la Cruz was unemployed. (RTE example no. 1030:RC)

The question of whether **T** entails **H** is in both cases dependent upon the temporal referent of the latter. If in (9), it is taken as Saturday relative to the dateline of **T** then the latter entails that **H** is false. If it is taken as sometime prior to that Saturday, then the entailment is true. Similarly, in (10), the text **T** says that at the time the family spoke, the time De la Cruz went to employment in SA was *after* being unemployed. If the reference time of the hypothesis **H** is the time the family spoke, then either the entailment is false, or there is no entailment (because we are not actually told how long the employment lasted). Thus it seems that the RTE task examples considerably underspecify the task of temporal reference (Beigman-Klebanov and Beigman 2010). Linguistic semantics will certainly continue to be crucial to solving these hard computational problems, but it is not in itself a sufficient solution.

3 Computational Contributions to Linguistic Temporal Semantics

The temporal semantics of both human languages and programming languages can be thought of as logical languages predicating relations over a model (in the logicians' narrow sense of the term) that can be visualized as in figure 1.

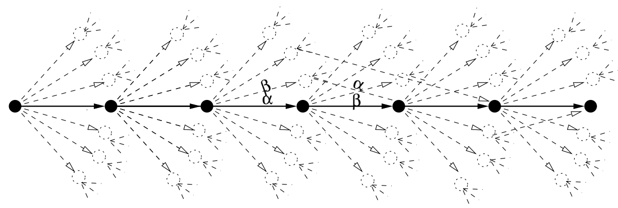

Figure 1: The S4 model

The figure depicts a Kripke or S4 model, in which nodes represent possible states of the world (only a few of which are depicted, and which should be thought of as complex structures, consisting of a number of propositional "fluents", or facts subject to change), and directed arcs represent events α, β, etc. that transform one state into another (of which few are depicted also). We may want to distinguish some particular sequence of states and events as actual or historical: those might be the ones in solid black.

This structure is not "there," in the mind or the computer. It is not something that can ever be built—for one thing, it is infinitely extending. Rather, it describes the space of possibilities that we or a machine inhabit, and to a very limited extent can think about by searching it to some limited depth.

What we and other animals *do* have in our heads (as do machines, if we program them with that capability, or allow them to acquire it for themselves) is a finite but extendable set of rules that describe the events that change one state to another, some of which are probabilistically under our control.

These rules (together with some computational resources that must include a (possibly simulated) push-down automaton) are what allows us and some other animals to see small portions of the eternal world, and to construct *plans* or sequences of actions that (with any luck) will take us to the more desirable possible worlds (or at least to the ones that we can find by searching the forest of crossing destinies to some very limited depth).

Most of the semantic theories discussed in the present volume assume such a model, implicitly or explicitly, and can be seen as addressing the question of the precise content of the states, and the nature of the events that take us from one state to another.

For example, the theories differ as to whether they take intervals as the basic temporal primitive, and regard events as durative, or whether they take instants as primitive and intervals as composite. Under the first view, a Vendlerian Activity like *running* would be represented as a transition, with a temporal and spatial extent. Under the second view, an Activity would be regarded as a progressive fluent, or property of a state, with the states that it characterizes being accessed via instantaneous incipitative events of *beginning running* and abandoned via terminative events of *stopping running*. (Vendler and his followers seem equivocal between these two interpretations.) Under the latter interpretation, the instantaneous incipitative and terminative events themselves correspond to Vendlerian Achievements, associated with further changes in fluents corresponding to consequent states, such as *running* and *having stopped running*. Vendlerian Accomplishments like *running to the bus-stop* are then the composition of an Activity of *running with the goal of being at the bus-stop*, the terminative Achievement of *stopping running* and the culminative achievement of *reaching the bus-stop*, which in turn initiates its own consequent state of *being at the bus-stop*.

When scaled to practical problems of planning in realistic worlds, such models are clearly going to be very complex. In deciding which of the many theories they allow we should adopt, we will be guided not only by the usual questions of soundness in representing temporal knowledge, but also by questions of efficiency for the purpose of searching for plans, which we can think of as *proofs* in a logic of change.

It is in this connection that theoretical computer science can be of help to linguistics. Computation can also be modeled as a space of states and operators that change state, using logics of change, as Prior (1967) noted himself.

In computer science, issues of constructivity (that proofs of attainability of a state are always accompanied by an algorithm for actually getting there) and efficiency (that useful proofs can actually be found with affordable resources) are always paramount. As a result, theoretical computer science has been the main engine driving progress in the use of temporal logics since the time of Prior.

With these ends in mind, computer scientists have made a very important observation about logics of change as they apply to programs and human reasoning. That is that the kind of changes that we are interested in are *localized*, affecting only a very few among the vast number of facts or fluents that define the current state in the model. Thus, *assigning a value to a register* affects that register and no other aspect of the state of the machine. Similarly, *drinking an ice-cold beer* affects the beer, and the consumer, in predictable ways, but leaves unaffected a miriad other facts that hold in the situation of the action, such as the weather, the color of the walls, and the current popularity ratings of the president of the United States. This suggests that events are to be defined in terms of *partial* descriptions of situations. (There are some events, like detonating H-bombs, that change practically everything. However, these too are only useful to the extent that they can be defined in terms of simple partial descriptions—for example, using a universal quantifier.)

This insight has been captured in a number of ways, both informal and formal. In linguistic theory, an early version of the idea surfaced in Lewis's (1973) idea of "inertia worlds," which he defined in terms of similarity between actual and hypothetical worlds, in order to provide a semantics for counterfactuals, which Dowty (1991) used explain the imperfective paradox. (Fine, 1975, criticized this definition using examples involving H-Bombs and the like, for which similarity between the inertia world and the actual world doesn't seem to work. Lascarides (1991) showed that inertia worlds cannot sensibly be defined other than in terms of the progressive itself.)

In Artificial Intelligence, the idea is usually identified with the STRIPS representation for actions for the purpose of automatically constructing plans (Fikes and Nilsson, 1971), although the idea seems to have been arisen more than once (cf. PLANNER, Hewitt, 1969). STRIPS actions are represented in terms of three elements: a list of *preconditions*, that is, facts or "fluents" which must hold for the action to be possible; a list of *deletions*, or fluents that cease to be true when the action occurs, and a list of *additions*, or fluents that become true in the aftermath of the action. Additions and deletions are modeled by database update, so the property that every fact that is not explicitly mentioned in the rule remains as it is, holds by default.

For example, the earlier example of *my drinking an ice-cold beer* might be represented by the following triple, in which the variable *x* is implicitly existentially quantified:

(11)	preconditions:	*beer(x), ice-cold(x), here(x), here(me), thirsty(me)*
	delete:	*beer(x), here(x), ice-cold(x), thirsty(me)*
	add:	*high(me)*

This rule says that when I drink an ice-cold beer, it ceases to be, while I, although I still exist, stop being thirsty and start being high. Whatever else holds in the current state remains unchanged. If we want to model a counterfactual situation such as *If I had not drunk an ice-cold beer* we can reverse the rule and work out that it would be pretty much the same apart from my state and that of the beer. Since our representation is in terms of actions rather than possible worlds, notions of similarity between worlds don't come into it: if we want to include actions like detonating an H-bomb in our plans we can do so. (Simplifying a bit, the latter action deletes everything, so it is easy to work out that counterfactuals like *If someone had detonated an H-Bomb, you wouldn't be here* are true.)

A number of important lessons were learned from work using STRIPS-like action representations. First, it isn't at all easy to represent even the simplest temporal knowledge domains consistently, especially if you want to be able to extend the domains by freely adding new actions. For example, if you represent the fact that some boxes and a truck are *in Edinburgh* as ground facts in a database, then your action of loading boxes *on trucks* should delete the ground fact that the boxes are *in Edinburgh*, and add a ground fact that they are *on the truck*. If the *move* action for trucks is defined in the obvious way, as deleting the ground fact of the truck being *in Edinburgh* and adding one of it being *in London* (say), this stratagem will ensure consistency in reasoning about where the boxes are when the truck moves. (Alternatively, you could define the move action as deleting the location at the origin of any objects that are *on* the moving object and asserting their location at the destination.) It is easy to make mistakes defining domain knowledge like this.

The other lesson learned from STRIPS is that, if you want to do any kind of temporal reasoning over the representation, or more generally need to represent co-occurrent actions, then you have to represent durative events like trucks moving as composed of an instantaneous incipitative event that introduces a progressive fluent, and a terminative event removing it (cf. Kowalski and Sergot, 1986). This ensures that if the truck is *dry* at the start of its journey, and it starts to rains while the truck is moving, the database will not end up saying inconsistently that the truck is both *wet* and *dry* at its destination. (This observation seems to suggest that instants and not intervals should be taken as the primitive elements in any model theory for systems of this kind—see Allen and Hayes, 1989 for a dissenting opinion.)

Event calculi of this kind underlie some very powerful planning programs, which compete at an annual competition on shared tasks of considerable complexity. The STRIPS idea was enshrined in a standard notation by McDermott et al. (1998), and extended in a series of papers culminating in Fox and Long (2006) to cover a rich ontology of time-stamped event types for purposes of the competition.

STRIPS representations were initially derided by logicians for their non-monotonicity. However, they provide a very natural expression for change of state, particularly when events are instantaneous and discrete, as they are in digital computers. Hoare Logic (Hoare, 1969) is founded on a very similar idea of "triples" $P\{S\}Q$, where P is a set of preconditions for a program statement S and Q describes the resulting state.

Pratt (1976) and Harel (1984) extend Hoare Logic to Dynamic Logic (DL) for the purpose of proving correctness of programs. DL combines a modal logic with the algebra of regular events or finite-state machines. DL is a multimodal one, in which the $\Box$ and $\Diamond$ modalities are relativized to particular event-types. For example, if a program or command α computes a function F over the integers, then we may write the following:

(12) a. $n \geq 0 \Rightarrow [\alpha](y = F(n))$
b. $n \geq 0 \Rightarrow \langle\alpha\rangle(y = F(n))$

The meaning of (12a) is that in any state in which $n \geq 0$, executing α always results in a state where $y = F(n)$. The meaning of (12b) is that in any state in which $n \geq 0$, executing α sometimes results in a state where $y = F(n)$. Although our knowledge of action is inherently nondeterministic, as far as reasoning about the world goes, we usually reason as if we could predict outcomes, even if we attach a probability of success, so we will mostly be dealing with the $[\alpha]$ modalities.

The α may be sequences $\alpha;\beta;\ldots$. They may also include loops or iteration, as in the following representation for Piaget's "primary circular reaction" of sucking in infants (1936):

(13) $[\textbf{while}\ \mathit{hungry}\ \ \textbf{do}\ \mathit{suck}]$

Dynamic Logic is usually applied to pure functional programming languages, without update. If we want to apply it for representing actions and change we may want to extend to make it "resource sensitive", by extending it to include Girard's (1995) linear implication operator, written $\multimap$ (pronounced "lolly"), to yield "linear dynamic" versions of event calculi such as STRIPS and its descendants.

For example, we might represent the earlier naive version of the *move* action as follows:

(14) $\mathit{affords}(\mathit{move}(x,\mathit{loc}_2)) \wedge \mathit{at}(x,\mathit{loc}_1) \multimap [\mathit{move}(x,\mathit{loc}_2)]\mathit{at}(x,\mathit{loc}_2)$

This formula means that if you can move and are at a location, and you move to another location, you stop being *at* the first place and start being *at* the other place. We adopt a convention that only *ground* fact (that is, the ones actually explicit in the data-base) are deleted or added, so the rule doesn't define whether the new situation supports inferrable facts like $\mathit{affords}(\mathit{move}(x,\mathit{loc}_2)$. If we define the latter in terms of a ground facts, using standard implicature as follows, to say that you can't move to a place if you are at that place, then it will not:

(15) $\neg \mathit{at}(x,\mathit{loc}) \Rightarrow \mathit{affords}(\mathit{move}(x,\mathit{loc}))$

(The predicate *affords* for preconditions is used in homage to Gibson's (1979) notion of affordance of actions by situations, which lies at the heart of effective action representation.)

If we want to avoid ramification problems arising from unexpected events like rain, as in the earlier example, then we need to recast the representation in terms of the instantaneous and stative components described there. For example:

(16) a. $\neg at(x, loc) \Rightarrow affords(start(move(x, loc)))$

b. $affords(start(move(x, loc_2))) \wedge at(x, loc_1)$
$\multimap [start(move(x, loc_2))] moving(x, loc_1, loc_2)$

(17) a. $moving(x, loc_1, loc_2) \Rightarrow affords(stop(move(x, loc_2)))$

b. $affords(stop(move(x, loc_2)))) \wedge at(x, loc_3) \wedge moving(x, loc_1, loc_2)$
$\multimap [stop(move(x, loc_2))] at(x, loc_3)$

Equipped with such rules, practical planning programs can search possible futures by progressing the database breadth-first to some limited depth (say, by iterative deepening, Korf, 1985), and build and execute plans to reach desirable states by search and composing actions in very complicated domains involving multiple actions and objects.

At some point, it may be thought desirable to timestamp everything in such representations. However, the causal structure implicit in the representation will often define the simplest relations of temporal antecedence and aspectual state without explicit indexing to clock-times. For example a simple history of starting to move from loc_1 to loc_2 followed by stopping doing so at a different place loc_3 will contain the information necessary to answer the question "Was x moving to loc_2 when she stopped at loc_3?" Such calculi therefore appear to offer a transparent and efficient representation for the concepts implicit in most current linguistic theories of temporal semantics for natural language, and have obvious relevance for purposes of linguists and philosophers of language.

Systems related to dynamic and nonmonotonic logics in application to natural language semantics are described by van Benthem (1991); Blackburn et al. (1994); Barwise and Seligman (1997)and Fernando (2011). Dynamic and nonmonotonic formalisms invoking or capturing real-world knowledge and related to the computational calculi outlined in this section have been applied to elegant effect in linguistic theories of temporality by Dowty (1986); Sperber and Wilson (1986); Pustejovsky (1991b); Moltmann (1991); Lascarides (1991); Asher (1992); Kamp and Reyle (1993); ter Meulen (1995); Glasbey (2004); Ramchand (1997); Piñon (1997); Pianesi and Varzi (1999); Stone and Hardt (1999); van Lambalgen and Hamm (2005); Bittner (2007); Truswell (2007), among others.

4 Further Reading

The computational literature on temporality and representation of causal action and its applications is overwhelming, and I am painfully aware of having been forced to pass over entirely or, worse, to treat very superficially, a great deal of important and relevant work. The following sources are offered as a means of entry to a more extensive literature.

Vardi 2008 provides an exceptionally readable summary of temporal and dynamic logic from Prior to the present from a computer science perspective. Mani et al. 2005 is an indispensable collection of mainly computational readings including several of the papers discussed above, and much other work that deserves attention. Virtually all of the more recent literature on computational approaches is accessible from the web, and in particular from the similarly indispensable ACL Computational Linguistics Anthology, at `http://www.aclweb.org/anthology-new/`. References to the broader literature on tense and aspect in natural language can be found in the Annotated Bibliography of Contemporary Research in Tense, Grammatical Aspect, Aktionsart, and Related Areas at `http://www.scar.utoronto.ca/~binnick/TENSE/`

Acknowledgments

I am grateful to Robert Binnick and Alex Lascarides for commenting on the draft. The work was supported by EU ERC Advanced Fellowship 249520 GRAMPLUS and EU FP7 IP Xperience.

REFERENCES

Ahn, D., van Rantwijk, J., and de Rijke, M. (2007). A cascaded machine learning approach to interpreting temporal expressions. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pages 420–427, Rochester, NY. ACL.

Allen, J. (1983). Maintaining knowledge about temporal intervals. *Communications of the Association for Computing Machinery*, 26:832–843.

Allen, J. and Hayes, P. (1989). Moments and points in an interval-based temporal logic. *Computational Intelligence*, 5:225–238.

Androutsopoulos, I., Ritchie, G., and Thanisch, P. (1998). Time, tense and aspect in natural language database interfaces. *Natural Language Engineering*, 4:211–228.

Asher, N. (1992). A default truth conditional semantics for the progressive. *Linguistics and Philosophy*, 15:469–598.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The berkeley framenet project. In *Proceedings of the 17th international conference on Computational linguistics*, pages 86–90, Montreal, Quebec, Canada. ACL.

Banko, M. and Etzioni, O. (2007). Strategies for lifelong knowledge extraction from the web. In *Proceedings of the 4th International Conference on Knowledge Capture*, pages 95–102, New York. SIGART, ACM.

Barwise, J. and Seligman, J. (1997). *Information Flow*. Cambridge Tracts in Theoretical Computer Science 44. Cambridge University Press, Cambridge.

Beigman-Klebanov, B. and Beigman, E. (2010). Some empirical evidence for annotation noise in a benchmarked dataset. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 438–446, Los Angeles, CA. ACL.

Bittner, M. (2007). Aspectual universals of temporal anaphora. In Rothstein, S., editor, *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, pages 348–385. John Benjamins, Amsterdam.

Blackburn, P., Gardent, C., and de Rijke, M. (1994). Back and forth through time and events. In Gabbay, D. and Ohlbach, H.-J., editors, *Temporal Logic*, pages 225–237. Springer, Berlin.

Bruce, B. (1972). A model for temporal references and its application in a question answering program. *Artificial Intelligence*, 3:1–25.

Chambers, N. and Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2*, pages 602–610, Suntec, Singapore. ACL.

Chambers, N., Wang, S., and Jurafsky, D. (2007). Classifying temporal relations between events. In *Proceedings of the 45th Annual Meeting of the ACL: Interactive Poster and Demonstration Sessions*, pages 173–176, Prague. ACL.

Chklovski, T. and Pantel, P. (2004). Verbocean: Mining the web for fine-grained semantic verb relations. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 33–40, Barcelona. ACL.

Clark, H. and Marshall, C. (1981). Definite reference and mutual knowledge. In Joshi, A., Webber, B., and Sag, I., editors, *Elements of Discourse Understanding*, pages 10–63. Cambridge University Press, Cambridge.

Collins, M. (2003). Head-driven statistical models for natural language parsing. *Computational Linguistics*, 29:589–637.

Crouch, R. and Pulman, S. (1993). Time and modality in a natural language interface. *Artificial Intelligence*, 63:265–304.

Dagan, I., Glickman, O., and Magnini, B. (2006). The PASCAL recognising text entailment challenge. In Joaquin Quiñonero Candela, Ido Dagan, B. M. and dAlché Buc, F., editors, *Machine Learning Challenges*, volume 3944 of *Lecture Notes in Computer Science*, pages 177–190. Springer, Berlin.

Dorr, B. (2007). Exploiting aspectual features and connecting words for summarization-inspired temporal relation extraction. *Information Processing and Management*, 43:1681–1704.

Dorr, B. J. and Olsen, M. (1997). Deriving verbal and compositional lexical aspect for NLP applications. In *Proceedings of the 8th Conference of the European Chapter of the Association for Computational Linguistics*, pages 151–158, Madrid. ACL.

Dowty, D. (1986). The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy*, 9:37–62.

Dowty, D. (1997/1991). *Word Meaning in Montague Grammar*. Reidel, Dordrecht, 2nd edition.

Etzioni, O., Banko, M., and Cafarella, M. (2007). Machine reading. In *Proceedings of AAAI Spring Symposium on Machine Reading*, Menlo Park, CA. AAAI Press.

Fellbaum, C. (1998a). Semantic network of English verbs. In Fellbaum, C., editor, *WordNet: an Electronic Lexical Database*, pages 69–104. MIT Press, Cambridge, MA.

Fellbaum, C., editor (1998b). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Fernando, T. (2011). Constructing situations and time. *Journal of Philosophical Logic*, 40:371–396.

Ferro, L., Gerber, L., Mani, I., Sundheim, B., and Wilson, G. (2005). Tides 2005 standard for the annotation of temporal expressions. Technical report, MITRE Corporation, Boston, MA.

Fikes, R. and Nilsson, N. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208.

Filatova, E. and Hovy, E. (2001). Assigning time-stamps to event-clauses. In *Proceedings of the Workshop on Temporal and Spatial Information Processing, held at ACL 2001*, pages 104–111, Toulouse. ACL.

Fine, K. (1975). Critical notice of lewis' "counterfactuals". *Mind*, 84:451–458.

Fox, M. and Long, D. (2006). Modelling mixed discrete-continuous domains for planning. *Journal of Artificial Intelligence Research*, 27:235–297.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton-Mifflin Co., Boston, MA.

Girard, J.-Y. (1995). Linear logic: Its syntax and semantics. In Girard, J.-Y., Lafont, Y., and Regnier, L., editors, *Advances in Linear Logic*, London Mathematical Society Lecture Notes 222, pages 1–42. Cambridge University Press, Cambridge.

Glasbey, S. (2004). Event structure, punctuality, and 'when'. *Natural Language Semantics*, 12:191–211.

Harel, D. (1984). Dynamic logic. In Gabbay, D. and Guenthner, F., editors, *Handbook of Philosophical Logic*, volume 2, pages 497–604. Reidel, Dordrecht.

Harrington, B. and Clark, S. (2009). ASKNet: Creating and evaluating large scale integrated semantic networks. *International Journal of Semantic Computing*, 2:343–364.

Hewitt, C. (1969). Planner: a language for proving theorems in robots. In *Proceedings of the 1st International Joint Conference on Artificial Intelligence (IJCAI'69)*, pages 295–301, Washington, DC. AAAI.

Hinrichs, E. (1988). Tense, quantifiers, and contexts. *Computational Linguistics*, 14:3–14.

Hitzeman, J. (1997). Semantic partition and temporal adverbials. *Natural Language Semantics*, 5:87–100.

Hoare, C. A. (1969). An axiomatic basis for computer programming. *Communications of the Association for Computing Machinery*, 30:576–583.

Kameyama, M., Passonneau, R., and Poesio, M. (1993). Temporal centering. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 70–77, Columbus, OH. ACL.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic*. Kluwer, Dordrecht.

Kipper, K., Korhonen, A., Ryant, N., and Palmer, M. (2008). A large-scale classification of English verbs. *Language Resources and Evaluation*, 42:21–40.

Korf, R. (1985). Depth-first iterative deepening: An optimal admissible tree search. *Artificial Intelligence*, 27:97–109.

Kowalski, R. and Sergot, M. (1986). A logic-based calculus of events. *New Generation Computing*, 4:67–95.

Lapata, M. and Lascarides, A. (2006). Learning sentence-internal temporal relations. *Journal of Artificial Intelligence Research*, 27:85–117.

Lascarides, A. (1991). The progressive and the imperfective paradox. *Synthèse*, 87(6):401–447.

Lenat, D. (1995). Cyc: a large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38.

Lewis, D. (1973). *Counterfactuals*. Harvard University Press, Cambridge, MA.

Lin, D. and Pantel, P. (2001). Induction of semantic classes from natural language text. In *Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data-Mining (KDD-01)*, pages 317–322, San Francisco.

Mahesh, K., Nirenburg, S., Cowie, J., and Farwell, D. (1996). An assessment of Cyc for natural language processing. Technical Report MCCS-96-302, New Mexico State University.

Mani, I., Pustejovsky, J., and Gaizauskas, R., editors (2005). *The Language of Time*. Oxford University Press, Oxford.

Mazur, P. and Dale, R. (2010). WikiWars: A new corpus for research on temporal expressions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 913–922, Cambridge, MA. ACL.

McDermott, D. (1982). A temporal logic for reasoning about processes and actions. *Cognitive Science*, 6:101–155.

McDermott, D., Ghallab, M., Howe, A., Knoblock, C., Ram, A., Veloso, M., Weld, D., and Wilkins, D. (1998). PDDL the planning domain definition language. DCS TR-1165, Yale Center for Computational Vision and Control, New Haven, CT.

Mitchell, T., Betteridge, J., Carlson, A., Hruschka, E., and Wang, R. (2009). Populating the semantic web by macro-reading Internet text. In *Proceedings of the 8th International Semantic Web Conference (ISWC 2009)*, Karlsuhe. Semantic Web Science Association.

Moens, M. (1987). *Tense, Aspect, and Time Reference*. PhD thesis, Edinburgh.

Moens, M. and Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14:15–28.

Moltmann, F. (1991). Measure adverbials. *Linguistics and Philosophy*, 14:629–660.

Narayanan, S. (1997). Talking the talk *is* like walking the walk: A computational model of verbal aspect. In *Proceedings of the Nineteenth Conference of the Cognitive Science Society, Stanford, August*, Mahwah, NJ. Lawrence Erlbaum Associates.

Palmer, M., Passonneau, R., Weir, C., and Finin, T. (1993). The kernel text understanding system. *Artificial Intelligence*, 63:17–68.

Pan, F., Rutu, M., and Hobbs, J. (2006). Extending timeml with typical durations of events. In *Proceedings of the Worshop on Annotating and Reasoning about Time and Events, held at the 44th Annual Meeting of the Association for Computational Linguistics*, pages 38–45, Sydney. ACL.

Piñon, C. (1997). Achievements in an event semantics. In *Proceedings of the 7th Conference on Semantics and Linguistic Theory*, pages 276–293, Ithaca, NY. Cornell University.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Delachaux et Niestle, Paris. translated 1953 as *The Origin of Intelligence in the Child*, Routledge and Kegan Paul.

Pianesi, F. and Varzi, A. (1999). The context-dependency of temporal reference in event semantics. In Bouquet, P., Brezillon, P., and Serafini, L., editors, *Modeling and Using Context: Proceedings of the 2nd International and Interdiciplinary Conference*, pages 507–510. Springer, Berlin.

Pratt, I. and Francez, N. (2001). Temporal prepositions and temporal generalized quantifiers. *Linguistics and Philosophy*, 24:187–222.

Pratt, V. (1976). Semantical considerations on floyd-hoare logic. In *Proceedings of the 17th Symposium on Foundations of Computer Science*, pages 109–121.

Pratt-Hartmann, I. (2005). Temporal prepositions and their logic. *Artificial Intelligence*, 166:1–36.

Prior, A. (1967). *Past, Present and Future*. Clarendon Press, Oxford.

Pustejovsky, J. (1991a). The generative lexicon. *Computational Linguistics*, 17:409–441.

Pustejovsky, J. (1991b). The syntax of event structure. *Cognition*, 41:47–81.

Pustejovsky, J., Casteño, J., Ingria, R., Saurí, R., Gaizauiuskas, R., Setzer, A., Katz, G., and Radev, D. (2003a). Timeml: Robust specification of event and temporal expressions in text. In *Proceedings of the 5th International Workshop on Computational Semantics*, Tilburg. AAAI. as TR-SS-03-07.

Pustejovsky, J., Hanks, P., Saurí, R., See, A., Day, D., Ferro, L., Gaizauiuskas, R., Lazo, M., Setzer, A., and Sundheim, B. (2003b). The timebank corpus. In *Proceedings of the Conference on Corpus Linguistics*, pages 647–656. University of Lancaster.

Ramchand, G. (1997). *Aspect and Predication*. Oxford University Press, Oxford.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. University of California Press, Berkeley.

Ritchie, G. (1979). Temporal clauses in English. *Theoretical Linguistics*, 6:87–115.

Sager, N., Lyman, M., Bucknall, C., Nhan, N., and Tick, L. (1994). Natural language processing and the processing of clinical data. *Journal of the American Medical Informatics Association*, 1:142–160.

Schank, R. (1975). The structure of episodes in memory. In Bobrow, D. and Collins, A., editors, *Representation and Understanding*, pages 237–272. Academic Press, New York.

Sperber, D. and Wilson, D. (1986). *Relevance*. Harvard, Cambridge, MA.

Stone, M. and Hardt, D. (1999). Dynamic discourse referents for tense and modals. In Bunt, H. and Muskens, R., editors, *Computing Meaning*, volume 1, pages 301–319. Springer, Berlin.

ter Meulen, A. (1995). *Representing Time in Natural Language: the Dynamic Interpretation of Tense and Aspect*. MIT Press, Cambridge, MA.

Truswell, R. (2007). *Locality of Wh-Movement and the Individuation of Events*. PhD thesis, University College London.

van Benthem, J. (1991). *Language in Action*. North Holland, Amsterdam.

van Lambalgen, M. and Hamm, F. (2005). *The Proper Treatment of Events*. Blackwell, Oxford.

Vardi, M. (2008). From Church and Prior to PSL. In Grumberg, O. and Veith, H., editors, *25 Years of Model Checking*, volume 5000 of *LNCS*, pages 150–171, Berlin. Springer.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Moszkowitcz, J., and Pustejovsky, J. (2009). The tempeval challenge: Identifying temporal relations in text. *Language Resources and Evaluation*, 43:161–179.

Webber, B. (1988). Tense as discourse anaphor. *Computational Linguistics*, 14:61–73.

Webber, B., Gardent, C., and Bos, J. (2002). Position statement: Inference in question answering. In *Proceedings of the International Conference on Language Resources and Evaluation*, Las Palmas. LREC.

White, M. (1994). *A Computational Approach to Aspectual Composition*. PhD thesis, University of Pennsylvania.

Wiebe, J., O'Hara, T., Öhrström, T., and McKeever, K. (1998). An empirical approach to temporal reference resolution. *Journal of Artificial Intelligence Research*, 9:247–293.