THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Form-Independent Meaning Representation for Eventualities

Citation for published version:
Steedman, M 2019, Form-Independent Meaning Representation for Eventualities. in R Truswell (ed.), *The Oxford Handbook of Event Structure.* Oxford University Press, New York, pp. 605-623.
https://doi.org/10.1093/oxfordhb/9780199685318.013.3

Digital Object Identifier (DOI):
10.1093/oxfordhb/9780199685318.013.3

Published In:
The Oxford Handbook of Event Structure

CHAPTER 21

FORM-INDEPENDENT MEANING REPRESENTATION FOR EVENTUALITIES

MARK STEEDMAN

NATURAL language text such as that found on the web or in newspapers can nowadays be efficiently parsed, including building meaning representations or logical forms, with somewhat usable accuracy, at speeds of hundreds or even thousands of sentences a second. And Google is parsing everything we type at it. Nevertheless, we still don't have real question-answering (QA), so that we can ask a question such as 'Is the president in Washington today?', have it mapped to an equivalent query, and get a precise answer. Such an answer could in principle be based on a semantic net or knowledge graph of eventualities, continually built and updated by semantic parsers reading the newspapers. Instead, we are still presented with a bunch of snippets from pages whose words and linkages may or may not answer our question when we ourselves do the reading.

The central problem in using parsers to answer questions from unrestricted text like this is that the answer to our question is very likely to be there somewhere, but that it is almost certainly in a form which is not the same as that suggested by the form of our question. For example, the question 'Is the president in Washington?' is in fact answered by the statement in today's paper that 'The president has arrived at the White House'. However, understanding this requires inferences that 'having arrived' at a place at a time entails 'being at' that place at that time, that being at the White House entails being in Washington, and so on. We ourselves draw all of these inferences effortlessly when we read the latter sentence. However, the standard logical form for our question is something like *present* (*in washington president*), while that of the text is *present* (*perfect* (*arrived whitehouse president*))

Of course, the commonsense knowledge that links the two statements can be hand-engineered for specialized domains, in this case by the use of resources such as named-entity linkers, ontologies, and gazeteers, and inference rules linking *arriving* with *being there*. However, there is simply too much of it to hand-engineer the open domain.

The chapter begins by briefly reviewing some early attempts to build such representations by hand. It then compares the two main alternative contemporary approaches to the discovery of hidden meaning-representations for relation-denoting content words. Section 21.3 then examines the extension of one of these approaches to the discovery of latent episodic relations such as temporal sequence and causality between such terms, and examines some extensions and limitations of the approach. A brief concluding section considers some broader implications for the theory of meaning, and its implications for practical tasks like question answering.

21.1 Decompositional lexical semantics

Linguists, starting with the Generative Semanticists of the late '60s, have tried for many years to build a form-independent semantics. The following are various attempts to specify the meaning of the sentence 'Bugs kill plants' in terms of semantically primitive relations like causation and change (see Travis, this volume, for discussion):

(1) Montague (1973): $\forall x[bug'x \Rightarrow \exists y[plants'(y) \wedge kill'y\, x]]$
McCawley (1968):
$[_S\, \mathit{CAUSE\ BUGS}\ [_S\, \mathit{BECOME}\ [_S\, \mathit{NOT}\ [_S\, \mathit{ALIVE\ PLANTS}]]]]$
Dowty (1979): $[\mathit{CAUSE}\ [\mathit{DO\ BUGS}\ \varnothing]\ [\mathit{BECOME}\ \neg[\mathit{ALIVE\ PLANTS}]]]$
Talmy (2000): *Bugs ARE-THE-AUTHOR-OF[plants RESULT-TO-die]*
Van Valin (2005): $[\mathit{do'}(\mathit{bugs'}, \varnothing)]\ \mathit{CAUSE}\ [\mathit{BECOME}\ [\mathit{dead'}(\text{plants}')]]$
Goddard (2010): *BUGS* do something to *PLANTS*; because of this, something happens to *PLANTS* at the same time; because of this, something happens to *PLANTS*' body; because of this, after this *PLANTS* are not living anymore.

Other related representations are graphical, such as that of Schank (1972), in which the left–right arrow $\longleftrightarrow$ represents the subject dependency, while the double up-arrow $\uparrow\uparrow$ represents the causal dependency of the plants' death upon the ACT of the bugs (cf. Langacker 2008):

(2)
$$\begin{array}{ccc} \text{bugs} & \longleftrightarrow & \text{do} \\ & \uparrow\uparrow & \\ \text{plants} & \longleftrightarrow & \text{die} \end{array}$$

21.2 Decomposing temporality

In a similar vein, Reichenbach (1947) identified three temporal entities underlying the semantics of the tensed verb group. These were: S—the *speech time*, or the time of the speech act itself; R—the *reference time*, or the time referred to; and E—the *event time*, the time of the eventuality identified by the main verb.

These entities may all be temporally disjoint, or may coincide or overlap. For example, in the case of the pluperfect in (3a), S is the time of utterance, R is at the time before S that we are talking about, and E is at a time before R:

(3) a. My luggage had arrived.
 b. We leave at dawn.
 c. He is driving to London.

In (3b), on the other hand, R the reference time is in the future, after S, and R and E coincide at dawn. In (3c), S and R coincide.

Tense—past in (3a), the futurate present in (3b), and the simple present of the progressive in (3c)—defines the relation between R and S as precedence in (a), succession in (b) and as identity in (c).[1]

Reichenbach seems to have conceived of S, R, and E as undifferentiated monolithic intervals and their relations as purely temporal. However, the relation between R and E, which is defined by progressive and perfect Aspect, respectively marked in English by the auxiliary verbs *be* and *have*, is not a purely temporal relation between intervals.

The effect of the progressive auxilliary *be* is rather to turn the event into a progressive state. The identity of that state depends on the type of the event. Events are anatomized by Moens and Steedman (1988), following Vendler (1957), as falling into four aspectual types or *Aktionsarten*, as follows (see chapters by Mittwoch, Verkuyl, Thomason, Ramchand, and Travis, in this volume):

(4)

Name	Type	Example	Grammatical test
Accomplishments	+telic, +durative	drive to London	#for an hour/in an hour/#at dawn
Achievements	+telic, −durative	arrive in London	#for an hour/in an hour/at dawn
Activities	−telic, +durative	drive	for an hour/#in an hour/#at dawn
Points	−telic, −durative	start/stop driving	#for an hour/#in an hour/at dawn

[1] More accurately, past tense defines the reference time as *other than* the situation of utterance, since past tense is also a marker of counterfactual modality, as in its use in counterfactual conditionals (Isard 1974).

The symbol '#' on a test such as combination of 'drive to London' with 'at dawn' means that the combination is impossible *without a change in the type of the event*—in this case, to something like 'start to drive to London'. (It is important to remember this point, because almost any of these combinations is possible with such 'coercions' to different event types.)

The effect of simple past tense on these event types is simply to identify the entire extent of the eventuality time E with the anterior reference time. For the telic accomplishments and achievements, this entails that the goal of the event—in this case being in London—was achieved.

(5) a. He drove to London.
 b. He arrived in London.
 c. He drove.

The effect of the progressive auxiliary is to turn the core eventuality into a *progressive state*. The type of this state is determined by the above eventuality types, as in the following examples:

(6) a. He was driving to London.
 b. He was arriving in London.
 c. He was driving.

The progressive of an activity (6c) says that R is anterior to S and that the eventuality E is a progressive state of *him driving* holding at a past R, where the start and stop points of E are undefined.

The progressive of an accomplishment (6a) is almost identical to (6c). The progressive state is *his driving with the goal of being in London*, and it holds at the anterior reference time. However, it is not entailed that the goal was achieved: it is perfectly consistent to continue 'but the car broke down and he never got there'.

The progressive of an achievement (6b) says that the progressive state holding at R was not the *arriving* but an inferrable activity that would normally result in arrival, such as *his driving the last part of the route to London*.

Thus the three examples in (6) have rather similar truth conditions. One might think of this as the progressive auxiliary turning everything into the nearest related activity. Often this is what Moens and Steedman called the *preparatory activity*, but it may also be iteration of the core event, as in 'I am seeing a doctor.'

Since most states do not have associated preparatory activities, and nor can they iterate, they can only combine with the progressive by rather extreme coercions to events. Thus the following seems to refer to *repeatedly showing that you know the answer whenever asked*:

(7) #I am knowing the answer (these days).

The perfect auxiliary has a similar effect of mapping events onto states. The states in question are what Moens and Steedman called the *consequent state* of the core event (cf. Portner 2003).

(8) a. I have driven to London.
b. I have arrived in London.
c. #I have driven.

Thus, the perfect of an accomplishment (a) and an achievement (b) are true just in case the consequent state (in this case, *my being in London*) hold at the (present) reference time. The perfect of an activity is only acceptable to the extent that the activity has an accessible consequent state, such as *my probably still remembering how to drive*.[2]

Many questions about such representations were never satisfactorily resolved, such as whether the representation should be 'decompositional', as in the above cases, or 'procedural' (Woods 1968), or based instead on 'meaning postulates' or rules of entailment (Fodor *et al.* 1975).

Nevertheless, all such formalisms have the attraction of being potentially language-independent, together with the considerable advantage of being immediately compatible with inference using first-order logical operators such as negation. Thus, one could in principle deduce an answer to the question 'Are the plants alive?' from the text 'The bugs killed the plants', or the equivalent in another language, or use such meaning representations to support machine translation. However, such semantics was confined to small fragments, and remained somewhat language-specific (Dorr *et al.* 2010).

Related attempts at a decompositional semantics have been more recently realized semi-automatically as computational lexical resources, including WordNet (Fellbaum 1998), FrameNet (Baker *et al.* 1998), VerbNet/PropBank (Hwang *et al.* 2010), BabelNet (Navigli and Ponzetto 2012), Abstract Meaning Representations (AMR, Banarescu *et al.* 2012), and the relations over named entities of the Google Knowledge Graph (Singhal 2012).

However, such hand-built semantic resources are invariably incomplete, in the sense that they leave out many relations, usually because such resources are built (consciously or unconsciously) for human users, and omit many essential entailments that humans find too obvious to ever need to state.

For example, at the time of writing, the FrameNet entry for the verb 'arrive' tells us a great deal about the verb 'arrive', but omits the information that the consequent state or *result* of the *theme* arriving at the *goal* is that the former is situated *at* the latter, which is what the relation *perfect′* in the logical form for the text in our running example needs to access in order to know that the text does actually answer the question.[3]

[2] Of course, in context, other coercions than those suggested here may be possible.
[3] https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Arriving (accessed 27 August 2018).

Of course, this paricular lacuna would be easy enough to fix, but there are many more (such as that *not being at the goal already* is a precondition of *arriving*). It is hard to believe that such resources will ever be complete enough to support our hypothetical question-answerer.

21.3 Decompositional primitives as 'hidden'

This realization prompts the following thought: why not let parsing and machine learning do the work of completing the semantics instead, using the 'machine reading' approach of Etzioni *et al.* (2007) and Mitchell *et al.* (2015) to mine 'hidden' or latent entailment relations such as that between *arriving* and *being at* a place?

There are two active approaches to this problem. The first treats the meaning of a content word as a location in a high-dimensional vector space. The dimensions of this space can initially be thought of as all the other content words of the language, with distances along those dimensions corresponding to counts of the occurrences of those words in the immediate neighbourhood of the word in question. However, this is a space of such high dimensionality and such sparse occupancy that its dimensionality must in practice be reduced. The reduction must be such as to preserve the Euclidean property of the original space to some degree of tolerance. Closeness in the space then represents relatedness in meaning (although relatedness tends to include antonymy as well as synonymy).

The attraction of such representations is that one can accomplish the composition of words into phrase- and sentence-level meanings using linear-algebraic operations like vector addition and multiplication (Church and Hanks 1989, Smolensky 1990, Landauer and Dumais 1997, Lin 1998, Baroni and Zamparelli 2010, Grefenstette and Sadrzadeh 2011, Padó and Lapata 2007, Mikolov *et al.* 2013, Bordes *et al.* 2013, Mitchell and Steedman 2015, Guu *et al.* 2015, Neelakantan *et al.* 2015, Weir *et al.* 2016, *passim*).

Vector-based 'embeddings' representing all the contexts a word has been encountered in can be trained by unsupervised methods over vast amounts of text, and can be very useful for disambiguating unseen words. In particular, when used as features to tune a supervised parsing model, they can be very effective in deciding which seen events in the supervised model most resemble unseen events in unseen text (Henderson 2003, Henderson *et al.* 2008, Chen and Manning 2014, Lewis and Steedman 2014*a*,*c*, Dyer *et al.* 2015, 2016).

However, for the same reason, it is questionable whether we can think of vectors as *meaning representations*. In particular, it remains unclear how to make such representations compatible with the logical operators such as negation, conjunction, and disjunction that are crucial to tasks such as question-answering.

21.3.1 Combined distributional and formal semantic representations

An alternative approach, advocated by Moro and Navigli (2012), Navigli and Ponzetto (2012), Nakashole *et al.* (2012), and Grycner and Weikum (2014), Grycner *et al.* (2015) among others, shows that relational ontologies, including multilingual ones, can be built by mining text concerning recognizable named entities.

Lewis and Steedman (2013*a*,*b*, 2014*b*) and Lewis (2015) propose to combine distributional and formal semantics by mining text concerning typed named entities such as the person named by *Mr. Obama* and the office named by *President* for consistent directional entailments using so-called *distant supervision* (Mintz *et al.* 2009), making strong assumptions concerning the entailment relations between predications that we frequently see made about sets of entities of the same type. For example, if when we read about a person *being elected* to an office, we often also read about them *running for* that office (but not vice versa), we may hypothesize that the former entails the latter. Typing is necessary, because distinct relations are sometimes homonymous, as with the *born in* relation, which denotes distinct relations between people and places on the one hand, and people and times on the other. Such candidate entailments will therefore be probabilistic and noisy, and are inherently distributional (for example, *the president* is sometimes a person and sometimes an office). But Lewis and Steedman (2014*b*) follow Berant *et al.* (2015) in exploiting the transitivity of entailment to make cleaner entailment graphs out of the candidate entailments, using various techniques to refine the entailment graph.[4]

For example, the typed named-entity technique is applied to (errorfully) estimate local probabilities of entailments using an asymmetric similarity measure such as Weeds precision (Weeds and Weir 2003), giving data that might look like the following simplified example for pairs of people and things *xy*, where ⇒ means 'probabilistically entails' (cf. Lewis and Steedman 2014*b*):

(9) a. $p(buy\,x\,y \Rightarrow acquire\,x\,y) = 0.9$
b. $p(acquire\,x\,y \Rightarrow own\,x\,y) = 0.8$
c. $p(acquisition\,(of\,x)\,(by\,y) \Rightarrow own\,x\,y) = 0.8$
d. $p(acquire\,x\,y \Rightarrow acquisition\,(of\,x)\,(by\,y)) = 0.7$
e. $p(acquisition\,(of\,x)\,(by\,y) \Rightarrow acquire\,x\,y) = 0.7$
f. $p(buy\,x\,y \Rightarrow own\,x\,y) = \mathbf{0.4}$

[4] It is important in what follows to distinguish the entailment graph that is used to identify paraphrase clusters and entailment relations from the knowedge graph, which represents all the knowledge in some body of text such as the Web. Of course, some of the mentions of entities in the text will involve pronominal and other forms of definite reference. However, pilot experiments with Reginald Long using the Stanford coreference sieve suggest that, in newspaper text at least, such coreference is involved in only around 20% of recoverable relations.

g. $p(buy\,x\,y \Rightarrow buyer\,(of\,x)\,y) = 0.7$
h. $p(buyer\,(of\,x)\,y \Rightarrow buy\,x\,y) = 0.7$
i. $p(inherit\,x\,y \Rightarrow own\,x\,y) = 0.7$
(etc.)

These local entailment probabilities are used to construct an entailment graph shown in Figure 21.1, with the global constraint that *entailment graphs must be closed under transitivity* (Berant *et al.* 2011).

Thus, (9f) will be correctly included, despite low observed frequency, because it is supported by the transitivity of entailment, while other low-frequency spurious local entailments will be dropped.

'Cliques' within the entailment graphs—that is, groups of relations that all mutually entail each other such as *acquire, acquisition-of*, and are therefore paraphases—can be collapsed to a single cluster relation identifier, such as rel_2 in Figure 21.1.

On the basis of this graph of entailments, we can take the categorial lexicon used by the parser to identify the original text-dependent local entailments, and transform it into something better-adapted to question-answering, by replacing the form-dependent Montague-style predicates by paraphrase cluster identifiers for their meaning. For example, some lexical items related to the *buying* entailment graph in Figure 21.1 will now look something like the following:

(10) own := $(S\backslash NP)/NP : \lambda x \lambda y.rel_1\,x\,y$
inherit := $(S\backslash NP)/NP : \lambda x \lambda y.rel_4\,x\,y$
acquire := $(S\backslash NP)/NP : \lambda x \lambda y.rel_2\,x\,y$
buy := $(S\backslash NP)/NP : \lambda x \lambda y.rel_3\,x\,y$
buyer := $N/PP_{of} : \lambda x \lambda y.rel_3\,x\,y$

In order to answer a question such as 'Did Verizon acquire Yahoo?', which denotes relation rel_2, we retrieve from the entailment graph all relations which either entail or are entailed by rel_2. If we can derive either from raw text, or from some knowledge graph representing the information in the text, rel_2 or any of the relations entailing rel_2

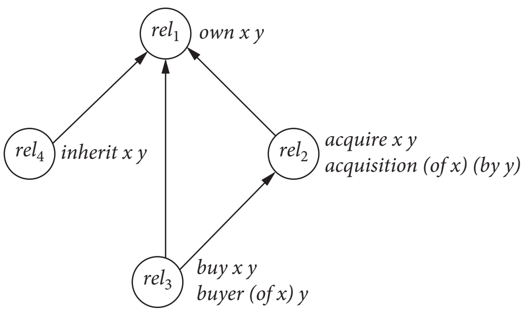

FIGURE 21.1 A simple entailment graph for property relations between people and things.

(here, rel_3), we can answer the question affirmatively. If on the other hand we can derive the negation of any of the entailments of rel_2 (here, rel_1), then the answer is negative.

The clustered clique or paraphrase identifiers such as rel_1 play much the same role in the entailment-based semantics as semantic *features* such as ALIVE did in the decompositional theories of semantics in (1), whereas the relations over such paraphrase cluster identifiers in the entailment graph correspond to Carnapian/Fodorian *meaning postulates*. Such logical forms immediately support correct inference under negation, such as that *bought* in the text entails *acquired* and *doesn't own* therefore entails *didn't buy*.

An example of open-domain questions succesfully answered from unseen text using these techniques is shown below. (Following Poon and Domingos 2008, the questions were artificially generated by replacing arguments in parsed web text with a dummy *wh*-question element 'What' to generate pseudo-*wh*-questions, which were then answered on the basis of unseen text of the same genre. See Lewis and Steedman 2013*a*, Lewis 2015 for further details and experiments.)

(11)

Question	**Answer**	**From unseen sentence:**
What did Delta merge with?	Northwest	The 747 freighters came with Delta's acquisition of Northwest
What spoke with Hu Jintao?	Obama	Obama conveyed his respect for the Dalai Lama to China's president Hu Jintao during their first meeting
What arrived in Colorado?	Zazi	Zazi flew back to Colorado. . .
What ran for Congress?	Young	. . .Young was elected to Congress in 1972

21.3.2 An application to machine translation

It should be apparent at this point that we can collect local entailments between expressions in languages other than English, provided that we can recognize and type the named entities in the language concerned, and align their types with the English ones. Lewis and Steedman (2013*b*) report an extension of the paraphrase/entailment semantics to French, and apply it to the task of reordering MOSES (Koehn *et al.* 2007) phrase-based statistical machine translations from French sentences to English. The bilingual semantics is evaluated by parsing the top 50 English translations into language-independent meaning representations and reordering them according to how well they preserve the multilingual entailment-based meaning obtained by parsing the original French. Where this process prefers a translation that is different from Moses' own top-ranked translation, bilingual judges are asked which they prefer. In 39% of cases where there is a difference, the judges prefer the reranked alternative, compared to the 5% of cases in which they prefer the Moses 1-best. (Many of the remaining 56% of cases in which there was no preference are ones in which the difference between the candidates was a matter of a syntactic attachment which was not available to the judges from mere presentation of the strings.)

An example of a successful reordering of Moses SMT translations is the following:

(12)

Source:	Le Princess Elizabeth **arrive à Dunkerque** le 3 août 1999
SMT 1-best:	The Princess Elizabeth **is to manage to Dunkirk** on 3 August 1999.
Reranked 1-best:	The Princess Elizabeth arrives at Dunkirk on 3 August 1999.

See Lewis and Steedman (2013*b*) for detailed results and further experiments.

21.4 Meaning representation for eventualities

It is natural to ask what kinds of semantic information can be mined in this way. This section considers a variety of open problems in the semantics of content words concerning eventualities.

21.4.1 Temporality and causality

If the text that we are mining is datelined, as news material usually is, then we should be able to work out that the entailments associated with *people* visting *places* in the graph in Figure 21.2 are temporally (or rather, causally) ordered, and that *being there* is the result of *arriving*, and therefore an entailment of *having arrived*, as in the example with which this chapter began. Certain finer distinctions, such as that between the present and the simple futurate, can be drawn on the basis of temporal modifiers, as in *visits/is visiting Hawai'i next week*, whose automatic extraction has been investigated by Chambers *et al.* (2014), using supervised learning over labelled resources such as TimeBank (UzZaman *et al.* 2013).

We may also expect to find entailments stemming from inceptive and conclusive aspectual 'coercions' involving light verbs like 'start', of the kind discussed by Moens and Steedman (1987, 1988) and Pustejovsky (1995), such as progressive *is visiting*, and compounds like *start a visit*, *finish a vacation*, and the like.

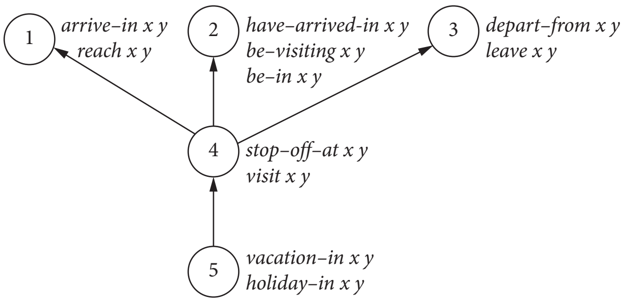

FIGURE 21.2 A temporal entailment graph for people visiting places.

These finer distinctions in varieties of entailment between relations can be discovered automatically from data like the following concerning particular pairs of named entities, from the University of Washington NewsSpike corpus (Zhang and Weld 2013), all of which have the same dateline of 4 February 2013.

(13) {"arg1":"OBAMA","arg2":"MINNEAPOLIS","sentences": [{"relation-phrase":"be in","tokens":
["Obama","is","in","Minneapolis","to","push","for","tougher",
"gun","laws","and","highlight","some","of","the","things","the",
"city","has","done","to","try","and","reduce","gun","violence",
"as","Mayor","R.T.","Rybak","and","some","of","his",
"counterparts", "across","the","country","try","to","put",
"direct","pressure", "on","firearms","makers","."],
"a1":[0,1],"a2":[3,4],"v":[1,3],"fromArticleId":371037},
{"relationphrase":"head to","tokens":
["Obama","heads","to","Minneapolis","to","sell","gun","plan",
"."], "a1":[0,1],"a2":[3,4],"v":[1,3],"fromArticleId":369952},
{"relationphrase":"be visit","tokens":
["Monday",",","Obama","is","visiting","Minneapolis","to", "dis-cuss","his","plan","to","battle","gun","violence","."],
"a1":[2,3],"a2":[5,6],"v":[3,5],"fromArticleId":433846}],...}

In such data, we find that statements that so-and-so *is visiting*, *is in*, and the perfect *has arrived in* such and such a place, occur in stories with the same dateline, whereas *is arriving*, *is on her way to*, occur in preceding stories, while *has left, is on her way back from, returned*, etc. occur in later ones. We also use the TimeBase/TimeML supervised-trained event and time-ordering system CAEVO to handle time-adverbials and order events (Pustejovsky *et al.* 2003*a*,*b*, Chambers *et al.* 2014).

This information provides a basis for inference that *visiting* entails *being in*, that the latter is the consequence of *arriving*, and that *arrival* and *departure* coincide with the beginning and end of the progressive state of *visiting*.

In order to capture the semantics behind these intuitions, we follow Moens and Steedman (1988), Hornstein (1990), Smith (1991), Steedman (1997), Portner (2003), and Fernando (2015) in assuming a neo-Reichenbachean semantics of tense, aspect, and modality.

Our event calculus is instant- and state-based (Steedman 1982, Kowalski and Sergot 1986, Copley and Harley 2015, and Copley's chapter in this volume), not interval- or event-based as in Dowty (1979), Allen (1983), Bach (1986*a*)—cf. Galton (1990).

We could use it as the input to a neo-Reichenbachian semantics of temporality, via (handbuilt) lexical entries for auxiliary verbs and other closed-class words (Steedman 1977, 1982, 1997, 2012*a*, Webber 1978, Moens and Steedman 1988, White 1994, Pustejovsky 1995, Filip 2008, Fernando 2015, *passim*), where the prime on *consequent-state′* etc. indicates that it is the identifier of a cluster of linguistic forms that are paraphrases:

(14)
$$\begin{aligned}
&\text{has} := (S\backslash NP)/VP_{pastpl}\\
&\qquad : \lambda p_E \lambda y.consequent\text{-}state'\ p_E\, y\, \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{will} := (S\backslash NP)/VP_{b}\\
&\qquad : \lambda p_E \lambda y.P_{prior}\ (p_E\, y) \Rightarrow imminent\text{-}state'\ p_E\, y\, \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{must} := (S\backslash NP)/VP_{b}\\
&\qquad : \lambda p_E \lambda y.P_{posterior}\ (p_E\, y) \Rightarrow imminent\text{-}state'\ p_E\, y\, \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{is} := (S\backslash NP)/VP_{prespl}\\
&\qquad : \lambda p_E \lambda y.progressive\text{-}state'\ p_E\, y\, \mathbf{R} \wedge \mathbf{R} = \mathbf{S}
\end{aligned}$$

However, we have already noted the absence of a source where we can look up the (possibly multiple) consequent states of relations like *arriving somewhere*, so logical forms like the above are rather vacuous. It would be more straightforward to write them simply as the paraphrase cluster identifiers *have′*, *will′*, *must′*, *be′*, etc., as follows, and to let the entailment graph do the rest of the work—for example, the work of saying that *having arrived somewhere* entails *being there*, and that the epistemic modal 'will' in 'That will be the postman' entails that the prior probability of the postman's presence is high, whereas epistemic 'must' in 'That must be the postman' entails that the posterior probability based on some further evidence is high:

(15)
$$\begin{aligned}
&\text{has} := (S\backslash NP)/VP_{ptpl}\\
&\qquad : \lambda p_E \lambda y.have'\ (p_E\, y\, \mathbf{E})\, \mathbf{R} \wedge \mathbf{E} < \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{will} := (S\backslash NP)/VP_{b}\\
&\qquad : \lambda p_E \lambda y.will'\ (p_E\, y\, \mathbf{E})\, \mathbf{R} \wedge \mathbf{E} > \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{must} := (S\backslash NP)/VP_{b}\\
&\qquad : \lambda p_E \lambda y.must'\ (p_E\, y\, \mathbf{E})\, \mathbf{R} \wedge \mathbf{E} > \mathbf{R} \wedge \mathbf{R} = \mathbf{S}\\
&\text{is} := (\mathrm{S}\backslash \mathrm{NP})/\mathrm{VP}_{\mathrm{prespl}}\\
&\qquad : \lambda p_E \lambda y.be'\ (p_E\, y\, \mathbf{E})\, \mathbf{R} \wedge \mathbf{E} \supset \mathbf{R} \wedge \mathbf{R} = \mathbf{S}
\end{aligned}$$

(Like the first-order logical operators such as quantifiers and negation, the tense operators relating **S** and **R**, and the further information in the modals can be handcoded into the lexicon, either via the morphology or semi-automatically for unanalysed verbs.)

By treating *have arrived* as a distinct relation, and a node in its own right in the entailment graph, we will then be able to learn by our standard machine-reading process that it entails *being there* at the reference time.

Similarly, the following are some potentially learnable lexical entries for implicative verbs (Karttunen 1971, 2012):

(16) try := $(S\backslash NP)/VP_{to} : \lambda p_E \lambda y.try' \, (p_E \, y) \, y \, \mathbf{E}$
manage := $(S\backslash NP)/VP_{to} : \lambda \mathrm{p_E} \lambda \mathrm{y}.manage' \, (\mathrm{p_E} \, \mathrm{y}) \, \mathrm{y} \, \mathbf{E}$
fail := $(S\backslash NP)/VP_{to} \lambda p_E \lambda \mathrm{y}.fail' \, (\mathrm{p_E} \, \mathrm{y}) \, \mathrm{y} \, \mathbf{E}$
stop := $(S\backslash NP)/VP_{ing} \lambda p_E \lambda \mathrm{y}.stop' \, (\mathrm{p_E} \, \mathrm{y}) \, \mathrm{y} \, \mathbf{E}$

Let us assume that the entailment graph built using the procedure outlined earlier includes the following, in which '⊨' denotes directional entailment in the graph:[5]

(17) a. $win'_{person,game} \models play'_{person,game}$
b. $\neg win'_{person,game} \models play'_{person,game}$
c. $fail'_{person,event} \models \neg event \wedge try'_{person,event}$
d. $\neg fail'_{person,event} \models event \wedge try'_{person,event}$
e. $stop'_{person,event} \models event$
f. $\neg stop'_{person,event} \models event$
etc.

(That is, our machine reading can be expected to detect latent 'presuppositional' relations between both *winning* and *not winning* and *playing* on the one hand, and *failing/not failing* and *trying* on the other.)

In the latter case, learning that *failing* to do something entails *not* doing it relies on the assumption that the text will elsewhere include explicit negation of the outcome. This assumption seems reasonable, since *trying* to do something creates an *a priori* likelihood of doing it, which is a precondition for the felicitous use of negation (Freud 1925).

21.4.2 Presupposition as entailment

The above subsumption of presupposition to entailment is akin to Wilson's (1975) account of logical presupposition in terms of entailment. However, a number of criticisms of this assumption have been raised, whose relevance to the present purpose is briefly reviewed below.

[5] Negation ¬ here is clearly nonclassical as in intuitionistic or relevance logics (Fitting 1969, Anderson and Belnap 1975).

The entailment account is at first glance also broadly consistent with Karttunen's account of the behaviour of presupposition under various kinds of embedding predicate. Karttunen divided such predicates into three categories, called 'plugs', 'holes', and 'filters', according to whether they blocked all presuppositions from emerging from an intensional context, or allowed all of them to so emerge, or blocked only *some* presuppositions. The 'plugs' constituted a large class of propositional attitude verbs such as 'say', and were so called because, while the following seems to entail that Frank believes Mary used to smoke, it doesn't seem to commit the speaker to that belief:

(18) Frank said that Mary had stopped smoking.

The 'holes' were a narrower, 'factive', class of propositional attitude verbs, such as *know*, which *did* commit the speaker to any presuppositions of the complement. Thus, either of the following seems to commit the speaker to the belief that Mary used to smoke:

(19) a. Frank knows that Mary has stopped smoking.
b. Frank doesn't know that Mary has stopped smoking.

It seems reasonable to assume that factivity can also be captured by the same process of mining text for coincidences across multiple predications over multiple sets of entities of the same type.

The 'filters' were the natural language equivalent of logical connectives, such as 'if... then...'. Whereas in (20a) the construction acts like a hole to the definite presupposition of the consequent that *John has children*, in (20b), where the antecedent includes that same presupposition, it acts like a very strong plug, and the presupposition does not emerge.

(20) a. If it is past eight o'clock, then John's kids are asleep.
b. If John has kids, then John's kids are asleep.

Similar considerations apply to disjunction and conjunction. Neither of the following presupposes that John has chidren:

(21) a. (Either) John has no kids, or John's kids are asleep.
b. (Both) John has kids, and John's kids are asleep.

The behaviour of the filters should clearly be made to follow from the semantics of the connectives themselves (Karttunen 1974, Heim 1983). For example, the facts in (20) would follow if $P \Rightarrow Q$ were equivalent to $\neg P \vee (P \wedge Q)$, rather than the standard Philonean $\neg P \vee Q$, as is the case for the semantics proposed in Steedman (2012*b*).

However, two further kinds of example have caused some people, including Wilson herself, to question the simple identification of presupposition with entailment. The first arises from the fact that a presupposition can be 'accommodated' or assented to by a hearer who didn't know that it held before it was made, provided only that it is consistent with everything that they do know. For example, if someone tells me to 'fetch me the envelope that is in the drawer' under conditions where I have no prior knowledge of such an envelope, then unless I already know that the drawer is empty, I will simply add such a referent to my model of the situation, and proceed to plan the requested action as if I'd known about the referent all along. Frequently, this will require updates to my world knowledge, as for the following from van der Sandt (1992), for which I will probably have to add the knowledge that every scuba-diver has a regulator (whatever *that* is):

(22) If a scuba-diver comes, they will bring their regulator.

Clearly, in terms of the entailment account, so long as we allow the model to be dynamic or 'updateable', this dynamism simply corresponds to modifying the model so that it is consistent with the entailment asserted by the speaker. The only complication arises when the prior state of knowledge of the hearer is *not* consistent with the speaker's entailment, when either the whole utterance must be rejected, or the prior knowledge must be modified, in a process of 'consistency maintenance' or 'belief revision', according to which the model of the common ground of the discourse is changed.

The second class of presuppositional effects that has caused people to question the entailment account arises from the possibility of 'cancelling' or negating a presupposition, as in utterances like the following in response to a positive presupposition-carrying assertion:

(23) a. Jane hasn't stopped smoking, because Jane has never smoked!
b. The present King of France hasn't died! There is no King of France.
c. It isn't true that John knows Mary is angry, because Mary is not angry!

However, there is clearly something odd about these utterances that takes them beyond the scope of sentential semantics. These utterances are clearly speech acts of *contradiction involving more than one speaker*, as when speaker A says 'Mary is angry' and speaker B says 'No she isn't'. We would not want to say that B has thereby uttered an inconsistency. The fact that the syntactic form of the first clause in (23) seems to assign it a truth value. This is just the vagueness of everyday language concerning the difference between truth/falsity and the speech acts of assertion/denial.

The failure to recognize the extra-semantic nature of accommodation and presupposition-cancellation has engendered a certain amount of confusion in the literature on presuppositions. For example, the following alternative to (20b) does not seem to /presuppose/ that John has children, but rather to *assert* it. This fact does not follow from our explanation of presupposition projection from the conditional in terms of non-Philonean implication, which would generate (b):

(24) a. If John's kids are asleep, then John has kids.
b. $\neg(\mathit{asleep} \wedge \mathit{has}) \vee (\mathit{asleep} \wedge \mathit{has}) \wedge \mathit{has}$

Clearly, this utterance only makes sense if *someone other than the speaker* has said 'John's kids are asleep'—a property of such utterances that Fillmore (1969) called 'semiquotation'—and the speaker is *affirming their accommodation* of the other's presupposition.

21.4.3 An application

The most ambitious application of such a semantics would be to build a knowledge graph or semantic network in which the nodes would be entities such as *Barack Obama*, *Sherlock Holmes*, and *Aluminium*, and the arcs would be relations and eventualities of the kind we have been looking at, mined from text. Such networks would be very large, on the same order of magnitude as social networks such as Facebook, with nodes numbering in the billions, calling for techniques like 'spreading activation' or the modern equivalent to limit the complexity of querying and updating them (Harrington and Clark 2009).

The advantage of building such semantic networks using the semantic representations proposed in this chapter would be that they could be queried in natural language, via a semantic parser building form- and language-independent meaning representations of the same kind as those in the semantic network itself, avoiding the severe problems of representational impedance mismatch that make the problem of natural language query of databases and knowledge graphs such as the Google knowledge graph so difficult (Reddy *et al.* 2014).

Building knowledge graphs from natural language semantic parses remains a challenging problem which has not yet been solved. However, we might think of the process of building and interrogating such a graph in terms of thought-experiments like the following.

Suppose that, in the course of continually updating the knowledge graph, the semantic parser encounters the following sentence in the text it is mining:

(25) Watford has failed to win the Cup.

The parser assigns the following Fodorian meaning representation, simplified as usual for present purposes:

(26) $have'\,(fail'\,(win'\,cup'\,watford')\,watford'\,\mathbf{E})\,\mathbf{R} \wedge \mathbf{R} = \mathbf{S}$

At this point, the program should inspect the entailment graph for this relation, and prepare to add to the knowledge graph arcs corresponding to this relation and all its entailments, namely the following:

(27) a. $\neg win'\,cup\,watford'\,\mathbf{E} \wedge \mathbf{E} < \mathbf{R}$
b. $try'\,(win'\,cup'\,watford')\,watford'\,\mathbf{E} \wedge \mathbf{E} < \mathbf{R} \wedge \mathbf{R} = \mathbf{S}$

c. $\mathit{play}'\ \mathit{watford}'\ \mathbf{E} \wedge \mathbf{E} < \mathbf{R} \wedge \mathbf{R} = \mathbf{S}$
d. $\mathit{losers}'\ \mathit{watford}'\ \mathbf{R} \wedge \mathbf{R} = \mathbf{S}$

If any of those relational arcs are already present in the graph—as when it has already been read that *Watford tried to win the cup*—or if any of them are inconsistent—as when it has already been read that *Watford won the cup*—then they all should not be added just yet. On the other hand, if they are not there already, they should be added to the nodes for the two entities involved.

For example, if it is already known that (27c) *Watford played* at some time E in the past, then only (27a,b,d) need be added.

If on the other hand the graph says that *Watford didn't play* then some process of 'consistency maintenance' must be entered, possibly via a dialogue with the source of the new information, for which the inconsistency of playing and not playing suggests the appropriate opening is 'But Watford didn't play', followed by a discussion of the grounds for the disagreement. (This process would be akin to what in the earlier discussion of the presupposition literature was referred to as 'cancelling' the presupposition.)

If, on the other hand, the sentence to be used for knowledge-graph update is the following, then the entailment graph (17) dictates that arcs be added corresponding to *not failing, winning, trying to win, playing*, and *being a winner*, equivalent to the utterance 'Watford managed to win the cup' (Karttunen 2007, 2012):

(28) Watford didn't fail to win the cup.

This underlines the importance of treating the meaning representations of *has Xed*, *didn't X*, and so on as nodes in the entailment graph in their own right, rather than as formulæ involving logical negation and modality.

21.5 Other varieties of entailment

More broadly, we expect to see type-based entailments involving (frozen) metaphors of the kind discussed by Lakoff (1994*a*,*b*), such as that between a government *attacking* a disease and it *trying to prevent* the disease. (This will simply be a differently typed sense of the verb *attack* from the one involving pairs of individuals or countries.)

On the other hand we shall not expect to capture Gricean conversational implicature the same way: if I ask for bread, and you tell me that there is a bakery around the corner, I will infer that you think the bakery is open. But that is not a matter of entailment. We do not expect to be able to detect such implicatures by the present method.

Similarly, we shall see entailments depending on type-based coercions of verbs like *start* and *finish* to distinct inchoative and cuminative events for the characteristic affordances of entities like *sandwiches* and *novels* (Pustejovsky and Bouillon 1995).

We also expect to see entailments betwen verbs and their nominalizations. Thus, a corpus that talks about *Rome destroying Carthage* may also include mentions of *Rome's destruction of Carthage*, *Rome's destroying Carthage* and vice versa. These forms will therefore be reduced to the same paraphrase cluster relation in their lexical logical forms.

While gerundive transitive nominalizations like the latter *Rome's destroying Carthage* are completely productive, the derived nominals like *Rome's destruction of Carthage* are notoriously idiosyncratic. Thus we have *John grows potatoes* and gerundive *John's growing potatoes*, and derivational *the growth of John's potatoes*, but not **John's growth of potatoes* (Chomsky 1970, Marantz 1997). While there seem to be semantic regularities underlying the availability or otherwise of derivational nominals for transitives (Dowty 1982, Grimshaw 1990, Levin and Rappaport Hovav 1995), they depend on the same features of the semantics that we are treating as hidden, so it seems appropriate to discover them piecemeal in this fashion.

We shall also expect to see that multi-word relations like *write a book about* have detectable entailments like *know about*, whereas syntactically identical relations like *destroy a book about* do not, and therefore do not end up with a distinct clustered entailment semantics in their own right. This fact may explain the well-known but otherwise puzzling sensitivity of the corresponding preposition-stranding constructions to semantic and pragmatic factors (Bresnan 1982, Takami 1992):

(29) a. Who did you write a book about?
b. #Who did you destroy a book about?

There are of course limits to what we can expect text-mining of this kind to discover. It is also important in answering such questions as 'Is the president in Washington?' to understand that *nothing can be in more than one place at a time*, so that the question can be answered in the negative on the basis of a text saying that 'The president is in Hawai'i'. This knowledge is probably too banal to be mentioned in text, ever.

However, this is not the kind of knowledge we want to put into the lexical entailments of *being there*. It is the kind of nonlinguistic knowledge that we share with other animals. (My cat seemed quite clear on the point that things could only be in one place at a time.) It is the kind of knowledge that we hard-wire into the knowledge representation for our robot planning systems, as in the STRIPS planner (Fikes and Nilsson 1971), perhaps along lines suggested in Steedman (2002) or van Lambalgen and Hamm (2005). Among other things, such systems build in such knowledge as that, if you move something, then you don't change the containment relations with things that it may contain. (This has the pleasing consequence that if my bicycle is on a train in London, and the train goes to Edinburgh, I know that my bicycle is in Edinburgh and not in London without ever having to invoke an axiom that things can only be in one place at a time.) A representation of this kind is probably built in to the natural logical 'language of mind' that underpins language acquisition and language evolution in the same way, so it should be built into the Reichenbachian representation proposed above, rather than making such banal entailments explicit in the lexicon.

21.6 Conclusion

A decompositional semantics of this kind, acquired by machine-learning of latent semantic primitives from text, builds commonsense entailments into logical form itself, so that they can be derived directly from sentential meanings, rather than by theorem-proving search. It can be seen as a lexicalized implementation of Carnap's (1952) 'meaning postulates', which he proposed to make the repository of the knowledge that someone *being a bachelor* entails them *being male* and *unmarried*, and which Fodor *et al.* (1975), and J.D. Fodor (1977: 152–5) advocate as a basis for lexical semantics over the feature-based account of Katz and Postal (1964). Indeed, our lexicon is entirely Fodorian, in the sense that many lexical logical forms are atomic primitives. The difference in the present approach is that *bidirectional* entailments or paraphrases are reduced to semantically primitive typed relational clusters like those in Figure 21.1, representing a hidden version of the features in such decompositional accounts, as in the hybrid decompositional/meaning-postulate approach proposed by Lakoff (1970). The approach can also be seen as a practical implementation of Wittgenstein's famous (1953: ¶43) identification of the meaning of (content) words with *usage* ('Gebrauch'):

> For a *large* class of occasions of use of the word 'meaning'—though not for *every* occasion of its use—this word can be defined thus: the meaning of a word is its usage in the language.[6]

Such a usage-based semantics of content-words, if refined considerably further than it has been so far, especially by the use of multilingual data, might ultimately approach the hidden conceptual language to which the child must have access in order to hang language-specific grammar onto it during first language acquisition, as required by Fodor (1975).

Acknowledgements

The research was supported in part by ERC Advanced Fellowship GA 742137 SEMANTAX, ARC Discovery grant DP160102156, a Google Faculty Award, and a Bloomberg L.P. Gift Award. The author thanks Tomas Kober, Alex Lascarides, Sander Bijl De Vroe, Bonnie Webber, the reviewers, and the editor for their comments on the draft.

[6] 'Man kann für eine *große* Klasse von Fällen der Benützung des Wortes "Bedeutung"—wenn auch nicht für *alle* Fälle seiner Benützung—dieses Wort so erklären: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.'