Frame semantics and the derivational process in English

Formal languages and Programming Languages

Jeremy Dohmann

Contact Information:
Harvard University
Email: dohmann@college.harvard.edu

Abstract

This investigation is the marriage of several topics in linguistics, cognitive science, programming languages, and intelligent agent planning levied in the pursuit of an all-purpose discourse planning architecture. The chief tasks of discourse planning in NLP are representation of knowledge, coalescing primitive knowledge into more complex knowledge, organizing and ordering concepts, and finally, producing coherent text from a planned out semantic structure.

This paper sketches out a particular approach to this problem being developed between Harvard University and Charles River Analytics called Tree abstraction metalanguage (Taml). Taml is a functional programming language in the early design stages which tackles the first three tasks in discourse planning laid out above.

We will first discuss Taml's knowledge representation system, inspired primarily by Marvin Minsky's seminal paper"A Framework for Representing Knowledge" (1974).

Next we will turn our attention to the theory and practice behind our system of tree rewriting. Taml's approach to building up complex concepts from knowledge primitives is built on functional programming language paradigms as found in languages like OCaml and Haskell.

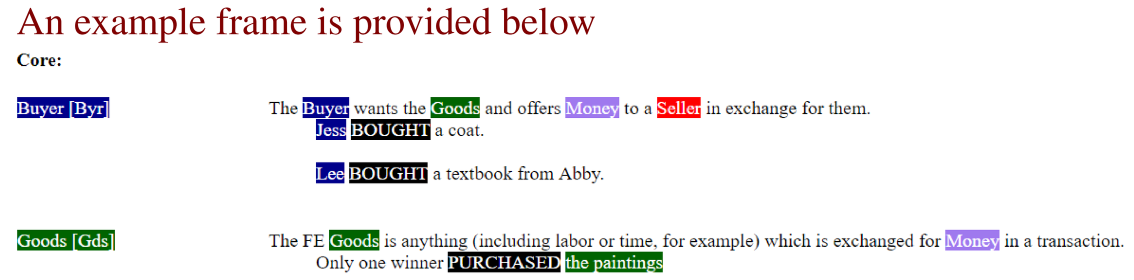
Next we present an overview of how contemporary approaches to knowledge-intensive intelligent agents bears on the reasoning tasks inherent in Taml's rule application selection process. Taml's preliminary reasoning architecture is built on the supposition that during discourse planning, organization is primarily driven by our goals (e.g. clarify XYZ, discuss XYZ, show the relationship between XYZ) and many stylistic choices are driven by our priorities (e.g. emphasize individual X, or process Y, or outcome Z).

Finally, we will briefly discuss the application of these principles to Speech Act Theory and J.R. Ross' Performative Sentence Hypothesis.

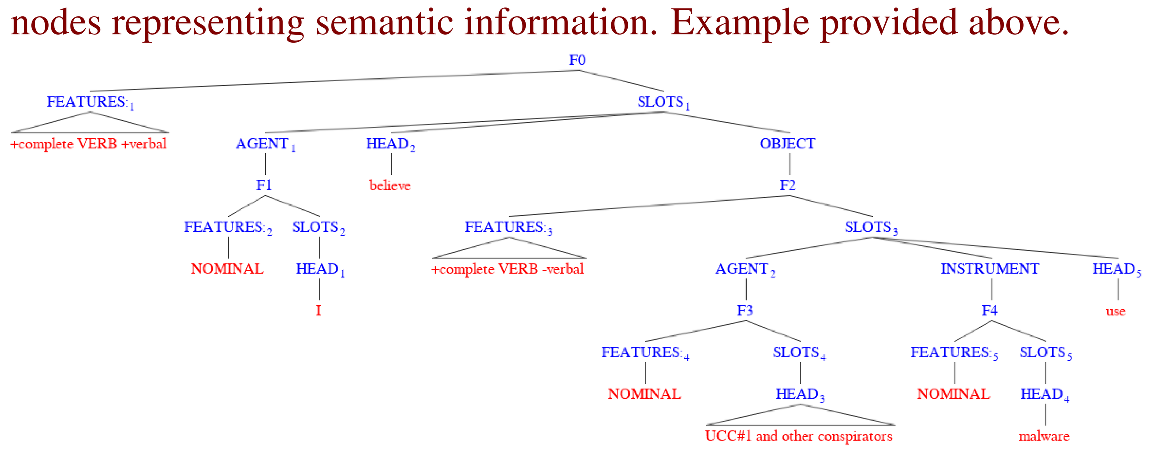
Frames

We choose to represent semantic information as frames. Frames have a long history in cognitive science stretching back to Minsky (1974). Frames, conceptually, adopt the belief that knowledge is stored by humans in terms of stereotypical, symbolic abstractions wherein contexts are associated with certain expectations, and furthermore that this system of representing expectations of a given context is compositional. This theory was originally applied to computer vision but it has since been the focus of several projects in NLP including FrameNet, where human annotated frames provide an ontology of the different verbs and nouns in English, the deep roles they have associated with them, and their relations to other contexts.

An example frame is provided below



We have slightly extended and reinterpreted the FrameNet style as follows. We represent frames as trees (complete, directed acyclic graphs, with n-1 edges) with edge labels representing slot names (correlates of deep roles in generative syntax), and string sets attached to nodes representing semantic information. Example provided above.



The spec given above was designed to be the minimal structural representation needed to describe the content, properties, relationships, and compositionality of natural language semantics.

Functions

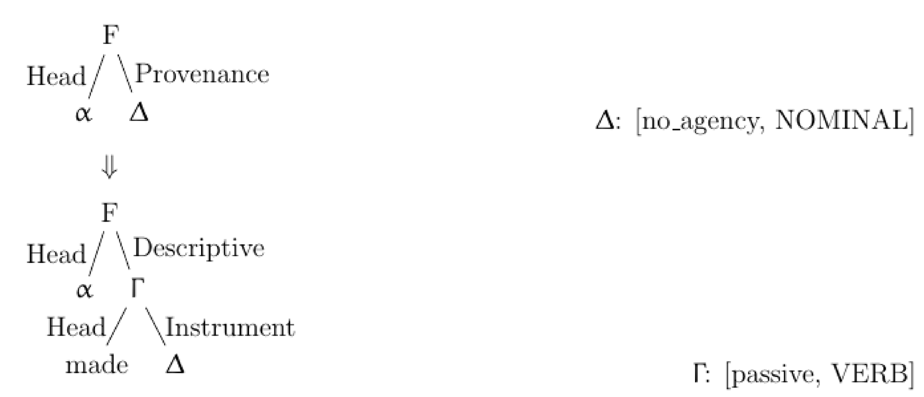
We posit that a discourse planner consists of a way to represent content-based knowledge (e.g. what is this document about and what do we know about the subject), a way to represent conventionalized knowledge about how to transform and combine semantic information (i.e. the subject of this section), and a way to reason over the previous two forms of information in order to accomplish linguistic tasks.

Thus our decision to make semantic transformation the focal point of our language design embodies the belief that there is a high degree of regularity of the kinds of constructions we encounter in the process of writing text. Ultimately, these similar constructions can be thought of as originating from some primitive, context-dependent knowledge (the content) as well as a procedural grasp on how to how exactly to combine information (this is both logical and stylistic). Thus we suspect that the high degree of regularity across discourse reflects the fact that ultimately, the writing process consists of the routine application of transformations that we have learned and ritualized over time.

To this end, we allow users to explicitly state the processes via which simple statements get coalesced, interposed, contrasted, and paraphrased in terms of functions. Functions are lightweight transformational signatures indicating what kinds of semantic argument they can be applied to, what the resulting semantic output ought to be, and when and where these functions are"appropriate" to apply.

An example for a graphical template of a function is given below:

6. NP with passive verbal modifier showing instrument



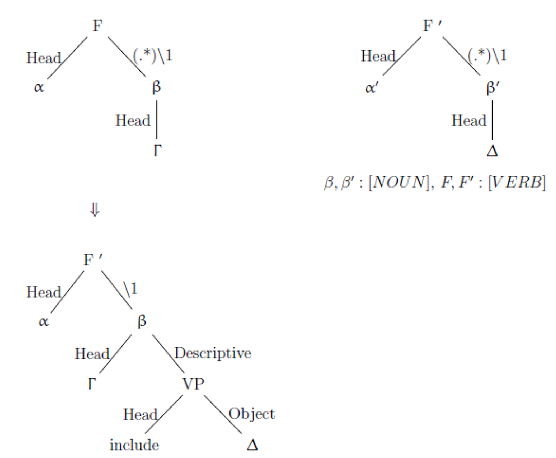
e.g. "malware in the C programming language" → "Malware made with the C programming language"

This graphical representation shows that we describe a function permitting the transformation of a simple referential description (i.e. there exists some malware made with C) into an NP with a passive VP describing how it was made. Note that the description of this rule includes what the input must look like (it must have Head slot, a Provenance slot, it must be a NOMINAL, etc.) and what the resulting output looks like. Furthermore, in order to specify the resulting functional transform, we permit naming and binding of frame elements in the original frame.

Thus this example shows that we can represent the process by which we highlight the origin or provenance a given NP by producing a construction relating the referent and the provenance information associated with it.

A more complex example is given below in which we describe the process by which any description containing a some process occurring to an NP can be further narrowed if we know there exists another NP that is a subtype of the original NP. Thus we can narrow our scope from saying "The man sold pets" to "The man sold pets, including parrots".

12. Replacing non-head type in verb frame with type + subtype/variant qualification



e.g. "The FBI collected samples of malware_i" AND "The FBI collected media_center.exe." AND "media_center.exe is a variant of malware_i." → "The FBI collected samples of malware_i, including media_center.exe."

Finally, we would like to permit user-defined predicates to be defined in the signature of a transformation, thus ensuring that correct application of the rule can be carried out when a.) the inputs match the pattern in the signature and b.) when the user-defined predicate is satisfied. This will permit us to draw in real-world semantic knowledge that operates at a level above the mere structural description given by patterns.

The above example, for instance, may have a semantic predicate that Δ must be a subtype of Γ (information that could be mined from a WordNet-type ontology.

As has been covered so far, Taml's spec will focus on user-defined transformations. We would like to be robust enough to handle descriptions of the "correctness" of a transformation (i.e. what pattern the inputs must match against, how to bind the inputs, and what to do with the bound constituent components) and the appropriateness. Whereas "correctness" has the scope of patterns and predicates, "appropriateness" is the purview of a reasoning agent.

Reasoning

There are a number of different paradigms in AI concerning themselves with symbolic reasoning approaches to decision-making in intelligent agents. Some of the biggest names are BDI (Belifes, Desires, Intentions), GOMS (Goals, Operators, Methods, and Selections), and Soar. The architecture of these three approaches generally uses some pipeline from input, to beliefs, to goals and plans, and finally, to actions. Furthermore, they each approach these different subtasks in different ways and they break them down further into varying levels of granularity. For instance, Soar uses a common semantic representation for inputs, outputs, and beliefs, represents goals and actions similarly as atomic "proposed operations" or "primitive operations" and maintains the balance between these things in real time by allowing beliefs to change, allowing goals to be satisfied by atomic operations and allowing plans to be revisited following updates to beliefs and intentions.

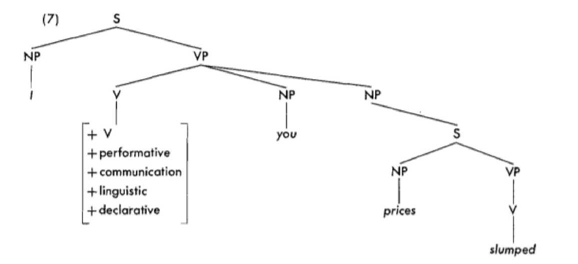
This schema bears a striking similarity to that proposed in Taml. Namely, the elementary unit of action in our reasoning system is the"transformation", which is analogous to the"operators" and"operations" used in other reasoning languages. Our project bundles what the literature variously refers to as"desires", goals","intentions" into similarly titled"goals" and"priorities". We speculate that some elements of"goals" will be conventionalized, compiled knowledge. For instance, when introducing the beginning of a counterexample, it is a conventionalized"goal" to explain to rehash previous points in the form of a simplified strawman. Another example of a precompiled"goal" would be the basic argumentative structure associated with a given genre. Thus an indictment may have several ordered goals such as"describe accused","offer allegations","present evidence from least convincing to most convincing" (which may have subgoals relating the substructure of the argument to the semantic relationships between pieces of evidence), and"discuss motives" (which may itself have stipulations relating entailments between motives and evidence observed). Finally, there may be goals unique to the circumstances which could be as broad as"give background on cybercrime" or as specific as"discuss monetary consequences of crime".

If"goals" concern themselves primarily with topics and organization,"priorities" concern themselves primarily with phrasing. For instance, there may be a genre-based priority to"emphasize role of criminal" which may cause the discourse planner to prioritize active over passive voice. Other examples of"priorities" could be,"emphasize John Smith","emphasize theft instead of bribery", etc.

Performative Sentence Hypothesis

Finally, we would like to show that our framework can be used to explain a phenomenon in English known as the"Performative Sentence Hypothesis".

The"performative sentence hypothesis", posited by Ross (1970) suggests all declarative sentences originate inside of performative speech acts. For example the sentence"prices slumped" originates in a deep structure as follows, which clarifies the fact that implicitly, declarative sentences asserting S are really speech acts in which the speaker asserts to the audience that S.



Ross (1970) and Rutherford (1970) use this hypothesis to account for the observation that much of written/spoken discourse consists not of content but rather metacommentary on the act of discourse itself. For example "Prices slumped, in case you're wondering", isn't meant to convey a conditional relationship between S="prices slumped" and P="you're wondering". Rather it serves as a metacommentary where P is a remnant of the since-deleted performative. Thus implicitly the sentence means "I tell you S in case P".

The interesting takeaway here is that a huge amount of the coherence of a discourse comes from the parts of the utterance concerned not with the content S, but rather with the metacommentary P left behind from the deletion of the implicit performative.

Thus we can reframe the metacommentary residue appearing in discourse to be attempts by a reasoning agent to explain its behavior, contextualize the discourse (e.g. relationship between speaker and listener), or clarify its intentions.

Thus a major goal of the Taml reasoning agent would be to account for coherence-oriented metacommentary in discourse as a built-in part of the rule selection process, wherein the reasoning agent attempts to clarify its actions in the context of its goals, its priorities, and the relationships its teasing out in the text.