An Inference Model for Semantic Entailment in Natural Language

Rodrigo de Salvo Braz **Roxana Girju** **Vasin Punyakanok**
Dan Roth **Mark Sammons**
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL, 61801, USA
{braz, girju, punyakan, danr, mssammon}@cs.uiuc.edu

Semantic entailment is the problem of determining if the meaning of a given sentence entails that of another. This is a fundamental problem in natural language understanding that provides a broad framework for studying language variability and has a large number of applications. We present a principled approach to this problem that builds on inducing re-representations of text snippets into a hierarchical knowledge representation along with a sound inferential mechanism that makes use of it to prove semantic entailment.

1 Introduction

Semantic entailment is the task of determining, for example, that the sentence: "*WalMart defended itself in court today against claims that its female employees were kept out of jobs in management because they are women*" entails that "*WalMart was sued for sexual discrimination*".

Determining whether the meaning of a given text snippet *entails* that of another or whether they have the same meaning is a fundamental problem in natural language understanding that requires the ability to abstract over the inherent syntactic and semantic variability in natural language [Dagan and Glickman, 2004]. This challenge is at the heart of many high level natural language processing tasks including Question Answering, Information Retrieval and Extraction, Machine Translation and others that attempt to reason about and capture the meaning of linguistic expressions.

Research in natural language processing in the last few years has concentrated on developing resources that provide multiple levels of text analysis (both syntactic and semantic), resolve various context sensitive ambiguities, and identify abstractions (from syntactic categories like POS tags to semantic ones like named entities), and the text relational structures.

Indeed, several decades of research in natural language processing and related fields have made clear that the use of deep structural, relational and semantic properties of text is a necessary step towards supporting higher level tasks. However, beyond these resources, in order to support fundamental tasks such as inferring semantic entailment between two texts snippets, there needs to be a unified knowledge representation of the text that **(1)** provides a hierarchical encoding of the structural, relational and semantic properties of the given text, **(2)** is integrated with learning mechanisms that can be used to induce such information from raw text, and **(3)** is equipped with an inferential mechanism that can be used to support inferences with respect to such representations. Just resorting to general purpose knowledge representations – FOL based representations, probabilistic representations or hybrids – along with their corresponding general purpose inference algorithms is not sufficient.

We have developed an integrated approach that provides solutions to all challenges mentioned above. We formally define the problem of *semantic entailment* for Natural Language and present a computational approach to solving it, that consists of a hierarchical knowledge representation language into which we induce appropriate representations of the given text and required background knowledge, along with a sound inferential mechanism that makes use of the induced representation to determine entailment. Our inference approach is formalized as an optimization algorithm that we model as an integer linear programming problem. The preliminary evaluation of our approach is very encouraging and illustrates the significance of some of the key contributions of this approach.

1.1 General Description of Our Approach

Given two text snippets S (source) and T (target) (typically, but not necessarily, S consists of a short paragraph and T, a sentence) we want to determine if $S \models T$, which we read as "S *entails* T" and, informally, understand to mean that *most people would agree that the meaning of* S *implies that of* T. Somewhat more formally, we say that S *entails* T when some representation of T can be "matched" (modulo some meaning-preserving transformations to be defined below) with some (or part of a) representation of S, at some level of granularity and abstraction.

The approach consists of the following components:

KR: A Description Logic based hierarchical knowledge representation, EFDL, into which we re-represent the surface level text representations, augmented with induced syntactic and semantic parses and word and phrase level abstractions.

KB: A knowledge base consisting of syntactic and semantic rewrite rules, written in EFDL.

Subsumption: An extended subsumption algorithm which determines subsumption between EFDL expressions (representing text snippets or rewrite rules). "Extended" here means that the basic unification operator is extended to support several word level and phrase level abstractions.

First a set of machine learning based resources are used to induce the representation for S and T. The entailment algorithm then proceeds in two phases: (1) it incrementally gen-

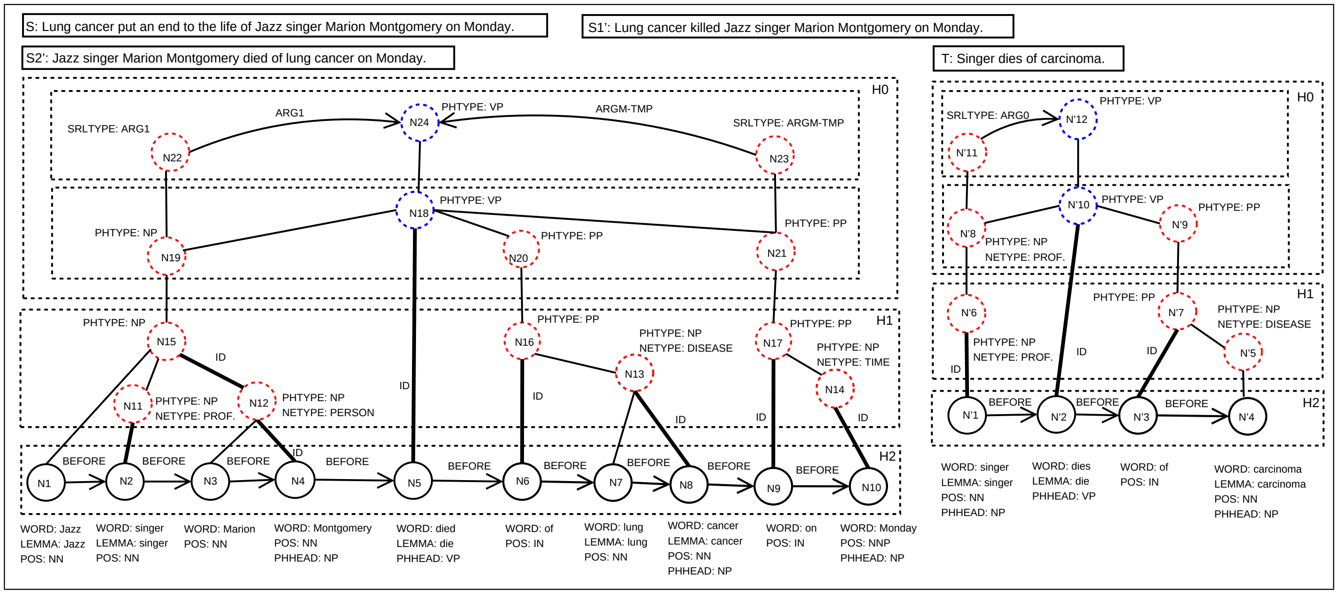

Figure 1: Example of *Re-represented Source & Target* pairs as concept graphs. The original source sentence S generated several alternatives including S'_1 and the sentence in the figure (S'_2). Our algorithm was not able to determine entailment of the first alternative (as it fails to match in the extended subsumption phase), but it succeeded for S'_2. The dotted nodes represent phrase level abstractions. S'_2 is generated in the first phase by applying the following chain of inference rules: #1 (genitives): "Z's W $\longrightarrow$ W of Z"; #2: "X put end to Y's life $\longrightarrow$ Y die of X". In the extended subsumption, the system makes use of WordNet hypernymy relation ("*lung cancer*" IS-A "*carcinoma*") and NP-subsumption rule ("*Jazz singer Marion Montgomery*'" IS-A "*singer*"). The rectangles encode the hierarchical levels (H_0, H_1, H_2) at which we applied the extended subsumption. Also note that in the current experiments we don't consider noun plurals and verb tenses.

erates re-representations of the original representation of the source text S and (2) it makes use of an (extended) subsumption algorithm to check whether any of the alternative representations of the source entails the representation of the target T. The subsumption algorithm mentioned above is used in both phases in slightly different ways.

Figure 1 provides a graphical example of the representation of two text snippets, along with an sketch of the extended subsumption approach to decide the entailment.

Along with the formal definition developed here of semantic entailment, our knowledge representation and algorithmic approach provide a novel solution that addresses some of the key issues the natural language research community needs to address in order to move forward towards higher level tasks of this sort. Namely, we provide ways to represent knowledge, either external or induced, at multiple levels of abstractions and granularity, and reason with it at the appropriate level.

2 Experimental Evaluation

Data. We tested our approach on the PASCAL challenge data set (http://www.pascal-network.org/Challenges/RTE/). As the system was designed to test for semantic entailment, the PASCAL data set is ideally suited, being composed of 276 source - target sentence pairs, indicating whether the source logically entails the target. The set is split into various tasks: CD (Comparable Documents), IE (Information Extraction), MT (Machine Translation), PP (Prepositional Paraphrases), QA (Question Answering), and RC (Reading Comprehension).

In Table 1 we show the system's performance. The baseline is a lexical-level matching based on bag-of-words representation with lemmatization and normalization (LLM).

Perform.	Overall [%]	Task [%]						
		CD	IE	IR	MT	PP	QA	RC
System	64.8	74.0	35.0	62.0	87.5	63.8	84.0	49.0
LLM	54.7	64.0	50.0	50.0	75.0	55.2	50.0	52.9

Table 1: System's performance obtained for each experiment on the Pascal corpora and its subtasks.

Acknowledgement

We thank Dash Optimization for the free academic use of their Xpress-MP software. This work was supported by the Advanced Research and Development Activity (ARDA)'s Advanced Question Answering for Intelligence (AQUAINT) Program, NSF grant ITR-IIS-0085980, and ONR's TRECC and NCASSR programs.

References

[Dagan and Glickman, 2004] I. Dagan and O. Glickman. Probabilistic textual entailment: Generic applied modeling of language variability. In *Learning Methods for Text Understanding and Mining*, Grenoble, France, 2004.