

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none"> I. Beltagy Stephen Roller Pengxiang Cheng Katrin Erk Raymond J. Mooney 	authors	<ul style="list-style-type: none"> Iz Beltagy Stephen Roller Pengxiang Cheng K. Erk R. Mooney 	DUPLICATES	375
	title	Representing Meaning with a Combination of Logical and Distributional Models	title	Representing Meaning with a Combination of Logical and Distributional Models		
	publication_date	2015-05-26 06:19:18+00:00	publication_date	2015-05-26 00:00:00		
	source	SupportedSources.ARXIV	source	SupportedSources.SEMANTIC_SCHOLAR		
	journal	None	journal	Computational Linguistics		
	volume		volume	42		
	doi		doi	10.1162/COLI_a_00266		
	urls	<ul style="list-style-type: none"> http://arxiv.org/pdf/1505.06816v5 http://arxiv.org/abs/1505.06816v5 http://arxiv.org/pdf/1505.06816v5 	urls	<ul style="list-style-type: none"> https://www.semanticscholar.org/paper/d799ab952c0f677d4526c485cc14555b7265ecc6 		
	id	id675663874911795781	id	id3338067326014052050		
	abstract	NLP tasks differ in the semantic information they require, and at this time no single semantic representation fulfills all requirements. Logic-based representations characterize sentence structure, but do not capture the graded aspect of meaning. Distributional models give graded similarity ratings for words and phrases, but do not capture sentence structure in the same detail as logic-based approaches. So it has been argued that the two are complementary. We adopt a hybrid approach that combines logic-based and distributional semantics through probabilistic logic inference in Markov Logic Networks (MLNs). In this paper, we focus on the three components of a practical system integrating logical and distributional models: 1) Parsing and task representation is the logic-based part where input problems are represented in probabilistic logic. This is quite different from representing them in standard first-order logic. 2) For knowledge base construction we form weighted inference rules. We integrate and compare distributional information with other sources, notably WordNet and an existing paraphrase collection. In particular, we use our system to evaluate distributional lexical entailment approaches. We use a variant of Robinson resolution to determine the necessary inference rules. More sources can easily be added by mapping them to logical rules; our system learns a resource-specific weight that corrects for scaling differences between resources. 3) In discussing probabilistic inference, we show how to solve the inference problems efficiently. To evaluate our approach, we use the task of textual entailment (RTE), which can utilize the strengths of both logic-based and distributional representations. In particular we focus on the SICK dataset, where we achieve state-of-the-art results.	abstract	NLP tasks differ in the semantic information they require, and at this time no single semantic representation fulfills all requirements. Logic-based representations characterize sentence structure, but do not capture the graded aspect of meaning. Distributional models give graded similarity ratings for words and phrases, but do not capture sentence structure in the same detail as logic-based approaches. It has therefore been argued that the two are complementary. We adopt a hybrid approach that combines logical and distributional semantics using probabilistic logic, specifically Markov Logic Networks. In this article, we focus on the three components of a practical system:1 1) Logical representation focuses on representing the input problems in probabilistic logic; 2) knowledge base construction creates weighted inference rules by integrating distributional information with other sources; and 3) probabilistic inference involves solving the resulting MLN inference problems efficiently. To evaluate our approach, we use the task of textual entailment, which can utilize the strengths of both logic-based and distributional representations. In particular we focus on the SICK data set, where we achieve state-of-the-art results. We also release a lexical entailment data set of 10,213 rules extracted from the SICK data set, which is a valuable resource for evaluating lexical entailment systems.2		
	versions		versions			