Semantic Parsing using Distributional Semantics and Probabilistic Logic

Islam Beltagy, Katrin Erk, Raymond Mooney
The University of Texas at Austin

Abstract

Goal: a semantic parser that is not constrained by a fixed formal ontology and purely logical inference
1) Formal language: First Order Logic
2) Ontology: use Distributional Semantics to generate the relevant part of a "graded" on-the-fly ontology
3) Inference: Probabilistic Logic inference

Tasks:
1) Recognizing Textual Entailment (RTE) using Markov Logic Networks (MLNs)
2) Semantic Textual Similarity (STS) using Probabilistic Soft Logic (PSL)

Inference
- Given input sentences and the generated ontology, Inference draws conclusions and answers queries
- Standard theorem provers are insufficient because the ontology is "graded" not binary
- Probabilistic logic solves this problem because it accepts weighted first order logic formulas
- Given the input sentences, and the rule base, we build a probabilistic logic program whose solution is the answer to the target task
- A probabilistic Logic program is an evidence set E, set of rules R and a query Q
- Inference computes Pr(Q|E, R)
- The probabilistic logic frameworks we use are MLNs for the RTE task and PSL for the STS task

Markov Logic Networks (MLNs)
- Given a set of weighted first-order logic formula, MLNs construct complex undirected graphical model
- Weighting the rules is a way of softening them compared to hard logical constraints.
- MLNs define a probability distribution over possible worlds, where a world’s probability increases exponentially with the total weight of the logical clauses that it satisfies.

Probabilistic Soft Logic (PSL)
- Another probabilistic logic framework, with emphasize on efficient inference
- Atoms have continuous truth values in interval [0,1] (in contrast with boolean atoms in MLNs)
- Logical operators are replaced with Łukasiewicz logic:
  I(¬ ℓ1) = 1 – I(ℓ1)
  I(ℓ1 ∨ ℓ2) = max {0, I(ℓ1) + I(ℓ2) – 1}
  I(ℓ1 ∧ ℓ2) = min {1, I(ℓ1) + I(ℓ2) }
- MPE inference is a linear program (100 times faster than MLN in our experiments)

System Architecture

Task1: RTE using MLNs
- Task: Given T, H, find if T Entails, Contradicts or not related (Neutral) to H
- Inference: compute Pr(T|H, RB), Pr(¬T|H, RB)
- Entailment decision = F (Pr(T|H, RB), Pr(¬T|H, RB)) where F is trained
- Computational overhead: reduce size of the ground network by removing unnecessary ground atoms [MLN, SL]

Task2: STS using PSL
- Task: Given S1, S2, find their semantic similarity score
- Adapt PSL for the STS task: replace PSL conjunction with an average, and change the grounding accordingly
- Inference: compute Pr(S1|S2, RB), Pr(S2|S1, RB)
- Similarity score = F (Pr(S1|S2, RB), Pr(S2|S1, RB)) where F is trained

Evaluation

System | SICK-RTE (acc) | SICK-STS (corr)
--- | --- | ---
Distribution-only | 0.60 | 0.65
Logic-only | 0.71 | 0.68
Logic+Distribution | 0.73 | 0.70

Conclusion
We propose a Semantic Parser that does not require a fixed ontology.
Ontology is generated from distributional information, and tasks are performed using Probabilistic Logic Inference.

References

System