Efficient Markov Logic Inference for Natural Language Semantics

Islam Beltagy, Raymond Mooney

The University of Texas at Austin

Abstract

- Using Markov Logic Networks (MLN) to represent Natural Language Semantics results in complex inference problems involving large ground network and complex formulae.
- We address this problem through:

**MCW**: A modified closed-world assumption (MCW) that removes unnecessary ground atoms, which significantly reduces the size of the ground network

**QF**: Inference algorithm that utilizes SampleSearch to compute probabilities of complete formulae not just individual ground atoms
- Evaluation: on the recognizing textual entailment (RTE) task

Modified Closed-World Assumption (MCW)

**MLN for an RTE problem**

**Why is this MLN difficult??**

**Evidence:**

\[
\text{man(M)} \land \text{drive(D)} \land \text{agent(D, M)} \land \text{convertible(C)} \land \text{patient(D, C)}
\]

**Priors:**

- \( \forall x. \text{man(x)} \) \[ \leftarrow \]
- \( \forall x. \text{drive(x)} \) \[ \leftarrow \]
- \( \forall x. \text{convertible(x)} \) \[ \leftarrow \]
- \( \forall x. \text{man(x)} \) \[ \leftarrow \]
- \( \forall x. \text{own(x)} \) \[ \leftarrow \]
- \( \forall x. \text{nice(x)} \) \[ \leftarrow \]
- \( \forall x. \text{car(x)} \) \[ \leftarrow \]
- \( \forall x,y. \text{agent(x, y)} \) \[ \leftarrow \]
- \( \forall x,y. \text{patient(x, y)} \) \[ \leftarrow \]

**Rules:**

- \( \forall x. \text{man(x)} \rightarrow \text{guy(x)} \) \[ \leftarrow \]
- \( \forall x. \text{drive(x)} \rightarrow \text{own(x)} \) \[ \leftarrow \]
- \( \forall x. \text{convertible(x)} \rightarrow \text{nice(x)} \land \text{car(x)} \) \[ \leftarrow \]

**Query:**

\[ \neg \exists x, y, z. \text{guy(x)} \land \text{own(y)} \land \text{agent(y, x)} \land \text{nice(z)} \land \text{car(z)} \land \text{patient(y, z)} \]

- Negating the query because universally quantified formulae are easier to ground

**Resulting MLN with MCW**

**Ground Rules:**

- \( \text{man(M)} \rightarrow \text{guy(M)} \) \[ \leftarrow \]
- \( \text{man(D)} \rightarrow \text{guy(D)} \) \[ \leftarrow \]
- \( \text{man(C)} \rightarrow \text{guy(C)} \) \[ \leftarrow \]
- \( \text{drive(M)} \rightarrow \text{own(M)} \) \[ \leftarrow \]
- \( \text{drive(D)} \rightarrow \text{own(D)} \) \[ \leftarrow \]
- \( \text{drive(C)} \rightarrow \text{own(C)} \) \[ \leftarrow \]
- \( \text{convertible(M)} \rightarrow \text{nice(M)} \land \text{car(M)} \) \[ \leftarrow \]
- \( \text{convertible(D)} \rightarrow \text{nice(D)} \land \text{car(D)} \) \[ \leftarrow \]
- \( \text{convertible(C)} \rightarrow \text{nice(C)} \land \text{car(C)} \) \[ \leftarrow \]

**Ground Query:**

\[ \neg (\text{guy(M)} \land \text{own(D)} \land \text{nice(C)} \land \text{car(C)}) \]

**MCW-Reachability is NOT Graph-Reachability**

**Evidence:**

- \( \text{g(C1)} \land \text{g(C2)} \)

**Rules:**

- \( \forall x, y. \text{g(x)} \land \text{h(y)} \lor \text{i(x, y)} \) \[ \leftarrow \]

**Ground Rules:**

- \( \text{g(C1)} \lor \text{h(C1)} \lor \text{i(C1, C1)} \) \[ \leftarrow \]
- \( \text{g(C1)} \lor \text{h(C2)} \lor \text{i(C1, C2)} \) \[ \leftarrow \]
- \( \text{g(C2)} \lor \text{h(C1)} \lor \text{i(C2, C1)} \) \[ \leftarrow \]
- \( \text{g(C2)} \lor \text{h(C2)} \lor \text{i(C2, C2)} \) \[ \leftarrow \]

**MCW-reachability:**

- \( \text{i(C1, C2)} \)

**Graph-reachability:**

- all ground atoms

Query Formula (QF): inference with complex queries Q

**Standard work-around:**

- Extra rule: \( Q \leftrightarrow \text{result(\text{"dummyConst"})} \)

**Query:**

- result(\text{"dummyConst"})

**New inference method with Query formula**

\[
Pr(\{Q\mid R\}) = \frac{Z(\{Q\mid R\})}{Z(\{R\})} = \text{ratio between } Z \text{ of the ground network of the MLN with and without Q added as a hard rule.}
\]

**Estimate Z using SampleSearch. Why?**

Evaluation: 10,000 RTE pairs

<table>
<thead>
<tr>
<th>System</th>
<th>Accuracy</th>
<th>CPU Time</th>
<th>Timeout</th>
</tr>
</thead>
<tbody>
<tr>
<td>mln</td>
<td>57%</td>
<td>2min 27sec</td>
<td>96%</td>
</tr>
<tr>
<td>mln+qf</td>
<td>69%</td>
<td>1min 51sec</td>
<td>30%</td>
</tr>
<tr>
<td>mln+mcw</td>
<td>66%</td>
<td>10sec</td>
<td>2.5%</td>
</tr>
<tr>
<td>mln+qf+mcw</td>
<td>72%</td>
<td>7sec</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

Conclusion

- The MCW significantly reduces size of the ground network and makes inference tractable.
- Inference with query formula is faster and more accurate.

References
