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:

$\text{man}(M) \wedge \text{drive}(D) \wedge \text{agent}(D, M) \wedge \text{convertible}(C) \wedge \text{patient}(D, C)$

Priors:

$\forall x.\ \text{man}(x)$	| -1
$\forall x.\ \text{drive}(x)$	| -1
$\forall x.\ \text{convertible}(x)$	| -1
$\forall x.\ \text{guy}(x)$	| -1
$\forall x.\ \text{own}(x)$	| -1
$\forall x.\ \text{nice}(x)$	| -1
$\forall x.\ \text{car}(x)$	| -1
$\forall x,y.\ \text{agent}(x, y)$	| -1
$\forall x,y.\ \text{patient}(x, y)$	| -1

Everything is False by default

Rules:

$\forall x.\ \text{man}(x) \rightarrow \text{guy}(x)$	| w1
$\forall x.\ \text{drive}(x) \rightarrow \text{own}(x)$	| w2
$\forall x.\ \text{convertible}(x) \rightarrow \text{nice}(x) \wedge \text{car}(x)$	| w3

Query:

$\neg\ \exists x,y,z.\ \text{guy}(x) \wedge \text{own}(y) \wedge \text{agent}(y, x) \wedge \text{nice}(z) \wedge \text{car}(z) \wedge \text{patient}(y, z)$

Negating the query because universally quantified formulae are easier to ground

Modified Closed-world Assumption (MCW)

Ground atoms are **false except** the ones **reachable** form the Evidence

Evidence:

$\text{man}(M) \wedge \text{drive}(D) \wedge \text{agent}(D, M) \wedge \text{convertible}(C) \wedge \text{patient}(D, C)$

Priors:

Not needed. The modified closed-world assumption plays their role.

Ground Rules:

$\text{man}(M) \rightarrow \text{guy}(M)$	| w1
$\text{man}(D) \rightarrow \text{guy}(D)$	| w1
$\text{man}(C) \rightarrow \text{guy}(C)$	| w1
$\text{drive}(M) \rightarrow \text{own}(M)$	| w2
$\text{drive}(D) \rightarrow \text{own}(D)$	| w2
$\text{drive}(C) \rightarrow \text{own}(C)$	| w2
$\text{convertible}(M) \rightarrow \text{nice}(M) \wedge \text{car}(M)$	| w3
$\text{convertible}(D) \rightarrow \text{nice}(D) \wedge \text{car}(D)$	| w3
$\text{convertible}(C) \rightarrow \text{nice}(C) \wedge \text{car}(C)$	| w3

TRUE: evidence
UNKNOWN: reachable from evidence
FALSE: otherwise

Ground Query:

$\neg$ { $\text{guy}(M) \wedge \text{own}(M) \wedge \text{agent}(M, M) \wedge \text{nice}(M) \wedge \text{car}(M) \wedge \text{patient}(M, M)$ }

....

$\neg$ { $\text{guy}(M) \wedge \text{own}(D) \wedge \text{agent}(D, M) \wedge \text{nice}(C) \wedge \text{car}(C) \wedge \text{patient}(D, C)$ }

....

$\neg$ { $\text{guy}(C) \wedge \text{own}(C) \wedge \text{agent}(C, C) \wedge \text{nice}(C) \wedge \text{car}(C) \wedge \text{patient}(C, C)$ }

Resulting MLN with MCW

Much smaller than without MCW

Ground Rules:

$\text{guy}(M)$	| w1
$\text{own}(D)$	| w2
$\text{nice}(C) \wedge \text{car}(C)$	| w3

Ground Query:

$\neg$ ($\text{guy}(M) \wedge \text{own}(D) \wedge \text{nice}(C) \wedge \text{car}(C)$)

MCW-Reachability is NOT Graph-Reachability

Evidence: g(C1), h(C2)

Rules: $\forall x,y.\ \text{g}(x) \vee \text{h}(y) \vee \text{i}(x, y)$ | w

Ground Rules:

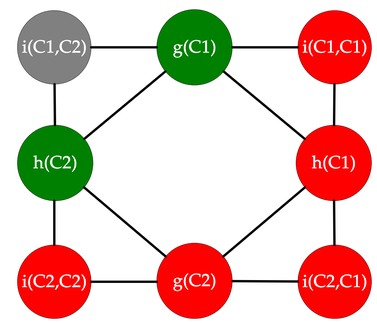
$\text{g}(C1) \vee \text{h}(C1) \vee \text{i}(C1, C1)$ |w
$\text{g}(C1) \vee \text{h}(C2) \vee \text{i}(C1, C2)$ |w
$\text{g}(C2) \vee \text{h}(C1) \vee \text{i}(C2, C1)$ |w
$\text{g}(C2) \vee \text{h}(C2) \vee \text{i}(C2, C2)$ |w

MCW-reachability:

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: $\text{result}(\text{“dummyConst”})$

New inference method with Query formula

$\Pr(Q \mid R) = Z(R \cup \{(Q, \infty)\}) \div Z(R)$ = ratio between Z 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

System	Accuracy	CPU Time	Timeouts
mln	57%	2min 27sec	96%
mln+qf	69%	1min 51sec	30%
mln+mcw	66%	10sec	2.5%
mln+qf+mcw	72%	7sec	2.1%

Conclusion

The MCW significantly reduces size of the ground network and makes inference tractable.

Inference with query formula is faster and more accurate.

References

[MLN] M. Richardson and P. Domingos. Markov logic networks. Machine Learning 2006

[SampleSearch] V. Gogate and R. Dechter. Samplesearch: Importance sampling in presence of determinism. Artificial Intelligence 2011