New Rules for Domain Independent Lifted MAP Inference

Happy Mittal, Prasoon Goyal
Dept. of Comp. Sci. & Engg., I.I.T. Delhi

Vibhav Gogate
Dept. of Comp. Sci., UT Dallas

Parag Singla
Dept. of Comp. Sci. & Engg., I.I.T. Delhi

Introduction

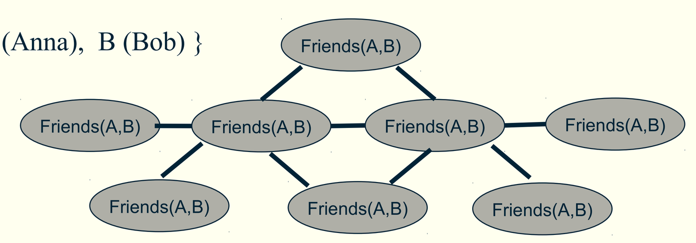
Markov Logic Networks (MLN) [Richardson & Domingos, 2006]

- A Markov Logic Network (MLN) is a set of pairs (F,w) where
  - F is a formula in first order Logic
  - w is a real number (weight of formula)
- When a world violates a formula, it becomes less probable, but not impossible.
- Together with a set of constants, it defines a Markov network.

$1.5 \;\; \forall x \; Smokes(x) \Rightarrow Cancer(x)$

$1.1 \;\; \forall x, y \; Friends(x,y) \Rightarrow (Smokes(x) \Leftrightarrow Smokes(y))$

Constants = { A (Anna), B (Bob) }



- **Join Probability :**

$$P(y) = \frac{1}{Z}\exp\left(\sum_k w_k f_k(y)\right) = \frac{1}{Z}\exp\left(\sum_i w_i n_i(y)\right)$$

Weight of Feature *i* — Feature *k* — No. of true groundings of formula i

- **MAP Inference :** Find most likely state of world given evidence

$$\arg\max_y P(y \mid x) = \arg\max_y \frac{1}{Z_x}\exp\left(\sum_i w_i n_i(x,y)\right) = \arg\max_y \sum_i w_i n_i(x,y)$$

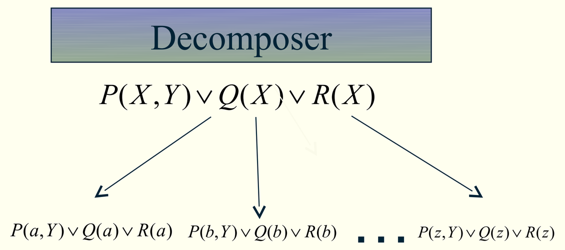
Query — Evidence

- ❑ This is just the **weighted MaxSAT** problem
- ❑ Use weighted SAT solver (e.g., MaxWalkSAT [Kautz et al. 97])

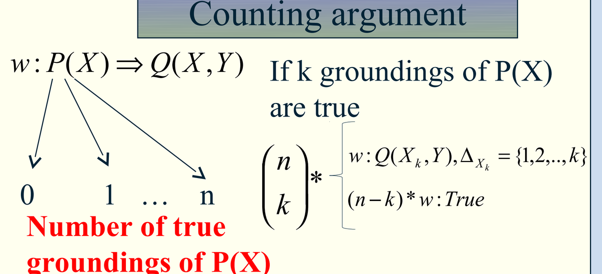
Lifted MAP Inference

- Naive way to solve MAP inference is to ground whole theory and run any weighted SAT solver
- Ground network can be very large, having millions of random variables and features.
- Lifted inference techniques can exploit the structure.
- Existing lifted inference techniques exploit following two properties

Decomposer



Counting argument



- Consider an MLN :

  $w1 : Parent(X,Y) \wedge Friend(Y,Z) \Rightarrow Knows(X,Z)$

  $w2 : Friend(X,Y) \wedge Friend(Y,Z) \Rightarrow Friend(X,Z)$
- No Decomposer or counting argument can be applied here
- Existing techniques will resort to partial grounding.
- We exploit specific properties of MAP, and show that MAP inference for above two formulae is domain independent i.e. its complexity doesn't depend on domain size of the variables in MLN theory.

Notations & Preliminaries

Binding Relation : variables X and Y are related if :

- They appear in same position of a predicate P, or
- X,Z and Y,Z are related

w1 P(X) v Q(X,Y), w2 P(Z) v Q(U,V), w3 R(W)

Equivalence Class : Binding relation splits variables into equivalence classes :

E1 : {X, Z, U}, E2 : {Y, V}, E3 : {W}

Single Occurrence EQ Class :

- No two variables of same equivalence class appear together in any formula.
- **E1** is **not** single occurrence., **E2**, **E3** are single occurrence

Single Occurrence MLN : Each equivalence class is single occurrence.

Normal MLN :

1)No constants in any formula

2)Variables X and Y at same position of predicate implies dom(X) = dom(Y)

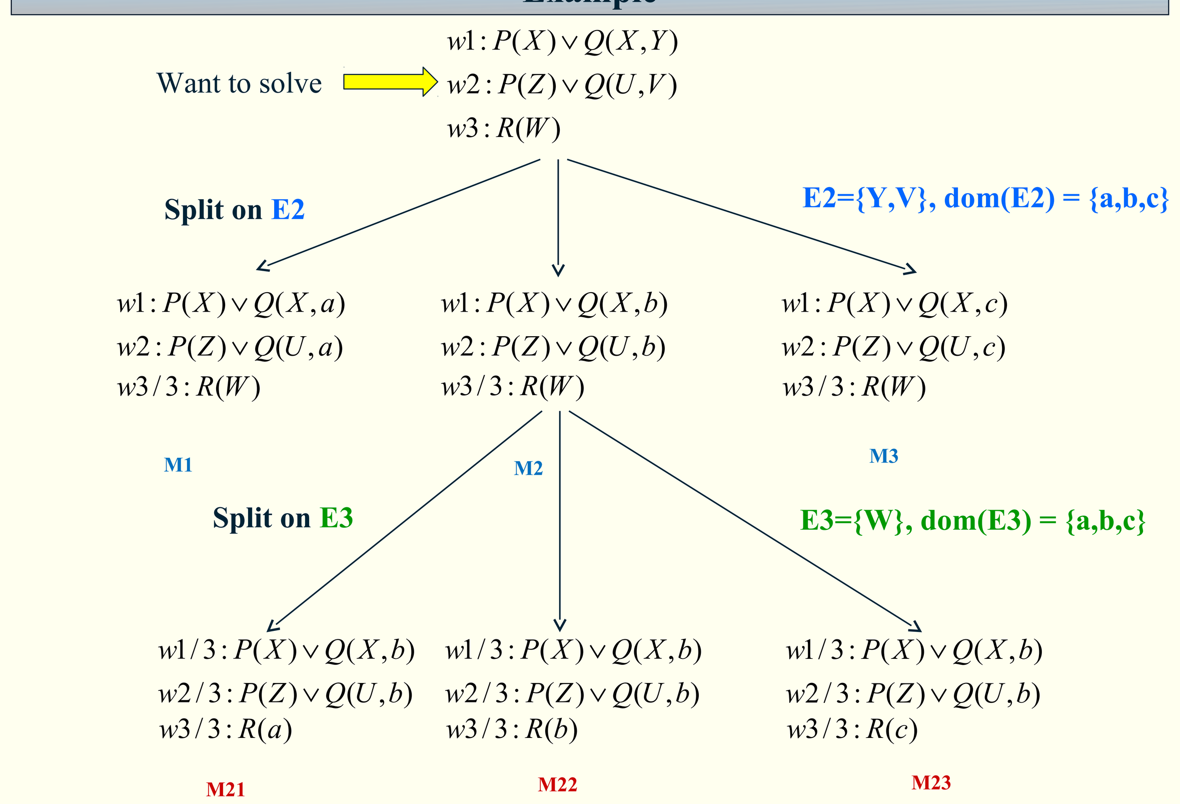
$S(X) \vee S(Y); \Delta_X = \Delta_Y = \{a,b\}$ **Normal**

$S(X) \vee S(Y); \Delta_X = \{a\}, \Delta_Y = \{b\}$ **Not Normal**

Any MLN can be converted into normal form in polynomial time

First Rule

Example



Solve any one of the M21, M22, M23 to get original solution.

Algorithm 1

- reduce(MLN M)
  - Initialize M' = M
  - for each single occurrence EQ class E:
    - if (isSingleOccurrence(E)) then
      - M' = reduceEQ(M',E)
    - end if
  - end for
  - return M'
- F_E : Formulas having variables in E
- F_{-E} : Formulas not having variables in E

- reduceEQ(MLN M, class E)
  - M' = {}, size = dom(E)
  - dom(E) = $\{a_1^E\}$ //any constant in E
  - for each formula f ∈ F_E
    - Add (f,w) to M'
  - end for
  - for each formula f ∈ F_{-E}
    - Add(f,w/size) to M'
  - end for
  - return M'

Theorem

Theorem : MAP Inference problem in an MLN theory M can be reduced to MAP inference problem in reduced theory M_s, where domain of single occurrence equivalence class X has been reduced to a single constant.

Proof :

1. Set of weighted constraints in M is union of weighted constraints in sub-theories $M_1, M_2, \ldots, M_r$:
   - For formula containing variables of equivalence class X : partial groundings are split across sub-theories.
   - For formula not containing variables of equivalence class X : weight divided equally among all sub-theories, hence total weight is original weight in M.
2. MAP assignments across sub-theories are same up to constant renaming :
   - Follows from symmetry of sub-theories.
3. MAP solution X^{MAP} for M can be obtained from any of the sub-theory M_j
   - Create MAP solution for M as follows :
     - For predicate not containing variable of class X :
       - Use solution from MAP of M_1 .
     - For predicate containing variable of class X :
       - Read off any of its partial grounding's assignment in any theory M_j.
4. X^{MAP} is indeed a MAP solution for M
   - Suppose it is not, then let X^{ALT} exists such that $W_M(X^{ALT}) > W_M(X^{MAP})$
   - But then $\sum_j W_{M_j}(X_j^{ALT}) > \sum_j W_{M_j}(X_j^{MAP})$
   - Which means there is some sub theory j, whose X_j^{MAP} is not a MAP solution,

Domain Independence

- An inference procedure is **domain independent** if its time complexity is independent of the domain size of the variables.
- MAP inference in single occurrence MLN is domain independent.
- Proof : Since MLN is single occurrence,
  - ❑ Successively reduce domain of each single occurrence equivalence class to be a constant.
  - ❑ Since MLN is single occurrence, all variables' domain is reduced to constant.
  - ❑ Hence domain independent MAP inference.

Second Rule for Lifting MAP

- **Extreme assignment :** Given an assignment, predicate P is at extreme, if all the groundings of P take same value (either true or false)
- **Theorem :** If Predicate P is single occurrence in M i.e. if each of the equiv. class of its argument variables is single occurrence, then P has an extreme assignment in X^{MAP} .
- **Proof :**
  - Reduce the MLN theory M into M' using Algorithm 1.
  - Since P is single occurrence, its domain is reduced to constant.
  - MAP solution of M can be read off from M', hence every grounding of P gets identical values in MAP solution of M
- **Corollary :** A single occurrence MLN admits a MAP solution which is at extreme.
- Consider an MLN : w1 Parent(X,Y) ^ Friends(Y,Z) => Knows(X,Z)
- Single occurrence MLN, hence each predicate has an extreme assignment in MAP.
- Now add a transitivity formula w2 Friends(X,Y) ^ Friends(Y,Z) => Friends(X,Z)
- **Tautology at extremes :** A formula f is tautology at extremes if all of its groundings are satisfied at any of the extreme assignments of its predicates.
- Now MLN is not single occurrence.
- But transitivity formula is **tautology at extreme.**
- In the theory

  w1 Parent(X,Y) ^ Friends(Y,Z) => Knows(X,Z)

  ~~w2 Friends(X,Y) ^ Friends(Y,Z) => Friends(X,Z)~~
- Friends is at extreme, hence MAP solution for above theory is same for theory containing transitive formula.

Second Rule : If an MLN M contains set of formulas F which are tautologies at extreme, and all predicates in F are single occurrence in remaining theory, then MAP inference over M can be reduced to MAP inference over smaller theory M', in which F has been removed, and all single occurrence predicates have been propositionalized.

Corollary : Let M be an MLN theory. Let M' be a single occurrence theory (with variable domains identical to M) obtained after removing a subset of formulas in M which are tautologies at extremes. Then, MAP inference in M is domain independent

- Algorithm 2 below gives a procedure to identify largest set of tautologies at extreme such that all variables in them are single occurrence with respect to remaining theory.

Algorithm 2

- **getSingleOccurTautology(MLN M)**
  - Initialization :
    - F_{te} = getAllTautologyAtExtremes(M)
    - F' = remaining formulas (F - F_{te})
    - done = *False*
  - while (not done)
    - done = *True*
    - EQ = getSingleOccurrVars(F')
    - for each formula in F_{te} :
      - if there is a variable in f not in EQ :
        - Add f into F' ; done = *False*
  - return F – F'

getAllTautologyAtExtremes(MLN M)

- /* iterate over all the formulas in M and return the subset of formulas which are tautologies at extremes */

isTautologyAtExtreme(formula f)

- f' = f
- for all unique predicates P in f' :
  - ReplaceByNewPropPred(P,f')
- return isTautology(f')

- Necessary and sufficient condition for a clausal formula to be a tautology at extremes is to have both positive and negative occurrence of same predicate symbol.

Subsumption of Sarkhel et. al [2014]'s rule

- Sarkhel et. al[2014] show that non-shared MLNs (with no self-joins) have a MAP solution at the extreme.
- **Theorem :** If an MLN theory is non-shared and has no self-joins, then M is single occurrence.
- **Proof :**
  - Non-shared MLN => every eq. class's variables appear in single predicate at a particular position.
  - No self join => No two variables of same equivalence class in same formula, because if they did, they will have to appear in different positions of same predicate, or in different predicates, which is a contradiction

Subsumption of Decomposer rule

- **Theorem :** Let M be an MLN theory and let X be an equivalence class of variables. If X is a decomposer for M, then X is a single occurrence in M.
- **Proof :** Let X_D be a decomposer, then no two variables X,Y ∈ X_D can appear in same formula, because if they did, they have to appear in all predicates at same position, which is not possible

Experiments, Conclusion and Future work

Experiments

- Compared our performance with
  - Direct grounding of MLN theory
  - Sarkhel et al. [2014]'s non-shared MLN approach
- Used ILP based solver Gurobi [http://gurobi.com] as base solver
- Notation :
  - GRB : purely grounded version
  - NSLGRB : Sarkhel et al.[2014]'s non-shared MLN approach
  - SOLGRB : our approach (single occurrence lifted GRB)
- 3 benchmark MLNs used for experiments :
  - Information Extraction (IE), Friends & Smokers (FS), and Student

Dataset	No.of formulas	Single Occurrence	Taut. At extreme	Example formulas
IE	7	No (Mix)	No	Token(t,p,c)=>InField(p,f,c) !Equal(f1,f2) ^ InField(p,f1,c) => !InField(p,f2,c) etc
FS	5	No (Mix)	Yes	S(x) => C(x) F(x,y) ^ F(y,z) => F(x,z) etc
Student	3	Yes	No	Teaches(t,c) ^ Takes(s,c) => JobOffers(s, company) etc.

- For each algorithm, we report :
  - **Time :** Time to reach the optimal as the domain size is varied from 25 to 1000.
  - **Cost :** Cost of the unsatisfied clauses as the running time is varied for a fixed domain size (500).
  - **Theory size :** Ground theory size as the domain size is varied.

Conclusion

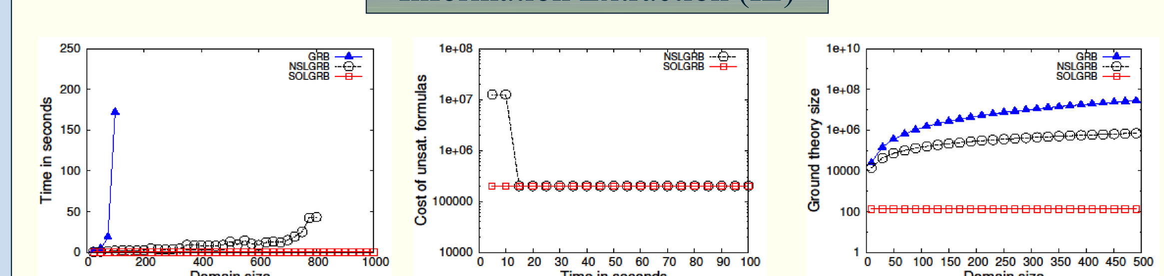
- Presented two new rules for lifting MAP inference, applicable to wide variety of MLN theories.
- MAP inference becomes domain independent in single occurrence MLN.
- Rules used as pre-processing step to generate a reduced theory.

Future work

- How to effectively combine our rules with existing lifting rules ?
- Ex :
  - w1 : S(X) V R(X)
  - w2 : S(Y) V R(Z) V T(U)
- Application of Binomial rule before single occurrence would lead to larger savings

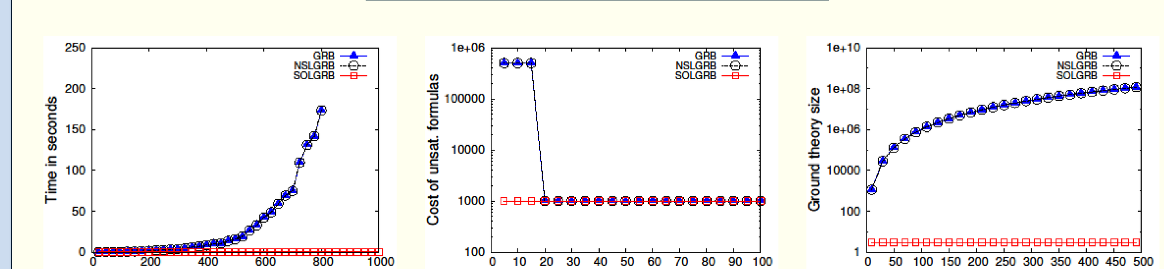
Results

Information Extraction (IE)



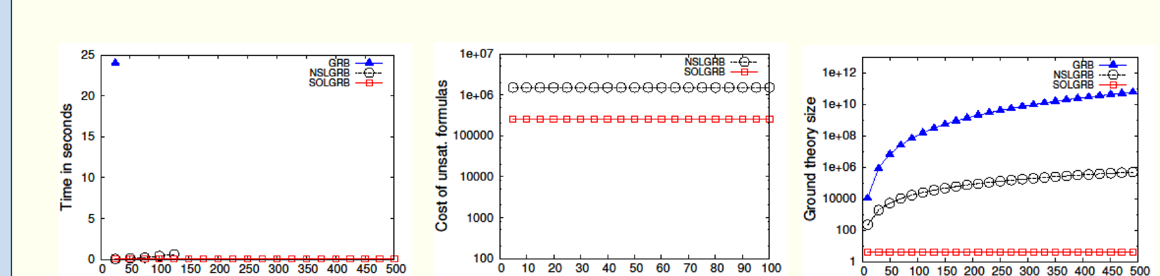
Time taken vs Domain size — Cost at domain size 500 — # of Groundings vs Domain size

Friends & Smokers (FS)



Time taken vs Domain size — Cost at domain size 500 — # of Groundings vs Domain size

Student Network



Time taken vs Domain size — Cost at domain size 500 — # of Groundings vs Domain size

Observations :

- In IE and FS domain, SOLGRB reaches the optimal instantaneously for all domain sizes. In comparison, time taken by GRB and NSLGRB to reach optimal increases with increasing domain size.
- Size of ground theory with varying domain size remains constant in all networks in SOLGRB, whereas it increases polynomially in GRB & NSLGRB