Acknowledgements

- Statistical Relational Learning (SRL) and AI (StarAI) are a synthesis of ideas of many individuals who have participated in various SRL/StarAI events, workshops and classes.

- Thanks to all of you!
General Take-Away Message

- Graphs are not enough
- We need logic

Roadmap

1. Motivation

2. Statistical Relational Learning / AI: a short overview

3. Markov Logic Networks
MOTIVATION

[Hermann Rorschach († Nov 8, 1884; † April 2 1922)]

Rorschach Test

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
Etzioni’s Rorschach Test for Computer Scientists

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications

Moore’s Law?

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
Storage Capacity?

Number of Scientific Publications?
Number of Facebook Users?

Number of Web Pages?
So, Tasks Are Often Structural

- Objects are not just feature vectors
  - They have parts and subparts
  - Which have relations with each other
  - They can be trees, graphs, etc.
- Objects are seldom i.i.d.
  - (independent and identically distributed)
  - They exhibit local and global dependencies
  - They form class hierarchies (with multiple inheritance)
  - Objects’ properties depend on those of related objects
- Deeply interwoven with knowledge

How do computer systems deal with structural problems?
(First-order) Logic handles Structures

- Main theoretical foundation of computer science
- General language for describing complex structures and knowledge: trees, graphs, hierarchies, etc.
- Inference algorithms (satisfiability testing, resolution, theorem proving, etc.)

More compact knowledge representation. Consider e.g. classical examples such as chess or wumpus:

\[ \text{FOL} \ll \text{PL} \ll \text{atomic} \]

Tasks are also often Statistical

- Information are ambiguous
- Our information is always incomplete
- Our predictions are uncertain

How do computer systems deal with uncertainty?
Probability handles Uncertainty

- Mixture models
- Hidden Markov models
- Bayesian networks
- Markov random fields
- Maximum entropy models
- Conditional random fields

... 

So, will traditional (U)AI scale?
Propositional vs. Relational Data

- Traditional work in robotics, machine learning and knowledge discovery assume data instances form a single table.

- Traditional statistical models assume independence among instances (rows) and find associations among the values of multiple variables within a single instance.

- Relational models assume dependence among instances in different rows and tables and find associations among these values.

Let’s consider a simple relational domain: Reviewing Papers

- The grade of a paper at a conference depends on the paper’s quality and the difficulty of the conference.
  - Good papers may get A’s at easy conferences
  - Good papers may get D’s at top conference
  - Weak papers may get B’s at good conferences
  - ...
(Reviewing) Bayesian Network

**Random Variables**

<table>
<thead>
<tr>
<th>Qual</th>
<th>Diff</th>
<th>P(Grade)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>low</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>low</td>
<td>middle</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>low</td>
<td>high</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[ P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | X_{i-1}, \ldots, X_1) \]

**Direct Influence**

\[ P(\text{Qual}) \]

<table>
<thead>
<tr>
<th>low</th>
<th>middle</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.5</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\[ P(\text{Diff}) \]

<table>
<thead>
<tr>
<th>low</th>
<th>middle</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The real world, however, ... ... has **inter-related objects**

These ‘instance’ are not independent!
Statistical Relational Learning and AI

Let's deal with uncertainty, objects, relations, and learning jointly.

The study and design of intelligent agents that act in noisy worlds composed of objects and relations among the objects.
Why the Tutorial?

- **A very active, multi-disciplinary research area**
  - Involves all sub-disciplines of AI: reasoning and acting under uncertainty, knowledge representation, constraint satisfaction, machine learning, ...
  - Unfortunately, can be hard to follow: *they all speak a different language*

- **A success story**
  - Often outperforms state-of-the-art
  - Novel ways of using the structure for faster and/or more robust solutions
  - Growth path for (U)AI in general
STATISTICAL RELATIONAL LEARNING / AI: A SHORT OVERVIEW

Applications to Date

- Natural language processing
- Information extraction
- Entity resolution
- Link prediction
- Collective classification
- Social network analysis
- Robot mapping
- Activity recognition
- Scene analysis
- Computational biology
- Probabilistic Cyc
- Personal assistants
- Etc.
Information Extraction

Parag Singla and Pedro Domingos, “Memory-Efficient Inference in Relational Domains” (AAAI-06).


### Segmentation

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parag Singla</td>
<td>&quot;Memory-Efficient Inference in Relational Domains&quot; (AAAI-06).</td>
<td></td>
</tr>
</tbody>
</table>

### Entity Resolution

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parag Singla</td>
<td>&quot;Memory-Efficient Inference in Relational Domains&quot; (AAAI-06).</td>
<td></td>
</tr>
</tbody>
</table>

Relations are at the heart of entity resolution.
Gene Localization

- Predict the localization of a given gene in a cell among 15 distinct positions
- Relations important as sequence similarity does not help

Relational Kernels better than Hayashi et al.‘s KDD Cup 2001 winning approach

Semantic Labeling of 3D Scan Data

- Neighbouring pixels/voxels have the same semantic label

Relations as constraints
Relational approaches outperform traditional ranking approaches.

Social Recommendation / Collaborative Filtering

- Predict whether a user **likes** a movie given attributes of users and movies, as well as known ratings and **complex link structures**.

Relational approaches outperform set-based recommendation systems.
What is the world talking about?

Relational approaches estimate better low-dimensional embeddings

Topic Models
How do you spend your spare time?

YouTube like media portals have changed the way users access media content in the Internet. Every day, millions of people visit social media sites such as Flickr, YouTube, and Jumpcut, among others, to share their photos and videos, while others enjoy themselves by searching, watching, commenting, and rating the photos and videos; what your friends like will bear great significance for you.

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications

How do you efficiently broadcast information?

Lifted inference faster than belief propagation
Heart diseases and strokes – cardiovascular disease – are expensive for the world

According to the World Heart Federation, cardiovascular disease cost the European Union EURO169 billion in 2003 and the USA about EURO310.23 billion in direct and indirect annual costs. By comparison, the estimated cost of all cancers is EURO146.19 billion and HIV infections, EURO22.24 billion.

Electronic Health Records
A New Opportunity for AI to Save our Lifes

Mining EHR is a non trivial problem!

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
Atherosclerosis is the cause of the majority of Acute Myocardial Infarctions (heart attacks)

Plaque in the left Coronary artery

What are Relations?

- There are several types of relations and in turn there are several views on what (statistical) relational learning is

1. **Relations provide additional correlations/regularization**
   - If two words occur frequently in the same context (page, paragraph, sentence, …) then there must be some semantic relation between them

2. **Often extensional (data) only, for one relation**
   - Covariance function, distance functions, kernel functions, graphs, tensors, random walks with restarts...
What are Relations?

3. Relations are symmetries/redundancies in the model
   - E.g. lifted inference based on bisimulation

4. Hypergraph representations of data
   - Multiple (extensional) relations
   - Random walks with restarts as similarity measure or to produce path features

5. Full-fledged relational (or logical) knowledge as considered in this tutorial
   - Multiple (often typed) relations
   - Intensional formulas (often Horn clauses)

\[ \text{ancestor}(X,Z) \land \text{parent}(Z,Y) \Rightarrow \text{ancestor}(X,Y) \]

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
Key Dimensions with some prototypes

Directed: Probabilistic Relational Models (PRMs)
Bayesian logic Programs (BLPs)

∀x author(x, p) ∧ smart(x) ⇒ high_quality(p)
∀x high_quality(p) ⇒ accepted(p)

Macro for conditional probability table

<table>
<thead>
<tr>
<th>high_quality, smart</th>
<th>high_quality, smart</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.9, 0.1)</td>
<td>yes</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Rule Graph

Deterministic background knowledge

Predicates

Probabilistic rule

Placeholders

Atoms

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications

[Getoor et al. 2002; Kersting, De Raedt 2007]
Inference on BN constructed by instantiating the rules/ macros using back- or forward chaining

But what happens if instead we have author(bob,p1)?

So, we can deal with a variable number of objects. The induced BN depends on the domain elements and the background knowledge we have.

Directed: Aggregate Dependencies
Directed: Aggregate Dependencies

| A | P(HQ | A) |
|---|---|
| t | 0.9 | 0.1 |
| f | 0.2 | 0.8 |

Still, the induced model is assumed to be acyclic

Option 1: Relational Dependency Networks (RDNs)

\[
\begin{align*}
\forall x \text{ author}(x, p) \land \text{smart}(x) & \Rightarrow \text{high}_\text{quality}(p) \\
\forall x \text{ high}_\text{quality}(p) & \Rightarrow \text{accepted}(p) \\
\forall x, y \exists p \text{ author}(x, p) \land \text{author}(y, p) & \Rightarrow \text{co-author}(x, y)
\end{align*}
\]

Run approximate Gibbs sample
Relational Dependency Networks

Option 2: Markov Logic Networks

Suppose we have constants: alice, bob and p1

\[
\begin{align*}
&1.5 \quad \forall x \text{ author}(x, p) \land \text{smart}(x) \Rightarrow \text{high}_\text{quality}(p) \\
&1.1 \quad \forall x \text{ high}_\text{quality}(p) \Rightarrow \text{accepted}(p) \\
&1.2 \quad \forall x, y \text{ co_author}(x, y) \Rightarrow (\text{smart}(x) \leftrightarrow \text{smart}(y)) \\
&\infty \quad \forall x, y \exists p \text{ author}(x, p) \land \text{author}(y, p) \Rightarrow \text{co_author}(x, y)
\end{align*}
\]

[Richardson, Domingos MLJ 62(1-2): 107-136, 2006]
Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications

Key Dimensions with some prototypes

directed
- PRMs
- CLP(BN)
- IHRM

undirected
- PRISM
- MLNs
- IHRM

macro
- BLOG
- RBNs
- RDN

proofs
- LPAD
- SLPs
- BLOG

- ProbLog
- IHRM
- RDN
- PRISM
- MLNs
- BLOG
- RBNs
- ICL

- PRMs
- CLP(BN)
- IHRM

- BUGS
- RBNs
- ICL

- MLNs
- BLPs
- PRMs

- PRMs
- CLP(BN)
- PRMs

Label of a clause/fact $c$ is the probability that $c$ belongs to the target program; Facts/clauses independent of each other

- Defines a distribution over programs $P(L|Program) = \prod_{\alpha \in L} p_\alpha \prod_{\alpha \notin L} (1 - p_\alpha)$

$P(path(x_gene,disease2)) =$ sum of probs of all programs that entail the query

- Exponentially many subprograms! To avoid explosion, consider proofs/paths only + store them in a BDD in order to count correctly

$P(path(x_gene,disease2)) =$ sum of probs of all programs that entail the query

$P(path(x_gene,disease2)) =$ sum of probs of all programs that entail the query

- $P=0.1*0.66*0.39$ + $P=(1-0.1)*0.66*0.39$ + $P=0.1*0.66*(1-0.39)$

- $0.10 :: \text{edges}(x\_gene, disease2)$
- $0.66 :: \text{edge}(x\_gene, disease1)$
- $0.39 :: \text{edges}(disease1,disease2)$

- path($X,Y$) :- $\text{edge}(X,Y)$
- path($X,Y$) :- $\text{edges}(X,Z)$, path($Z,Y$)

- path($X,Y$) :- $\text{edge}(X,Y)$
- path($X,Y$) :- $\text{edges}(X,Z)$, path($Z,Y$)
Many other approaches!!

- **directed**
  - PRMs
  - BLNs
  - LPAD
  - ProbLog
  - MLNs
  - RBNs
  - IHRM
  - SLPs

- **undirected**
  - PRISM
  - LPAD
  - RBNs
  - IHRM

- **macro**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG
  - SLGs

- **parametric**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG
  - RBNs

- **non-parametric**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG

- **proofs**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG
  - RBNs

- **CWA**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG
  - RBNs

- **OWA**
  - MLNs
  - BLPs
  - PRMs
  - IHRM
  - BLOG
  - RBNs

And actually they span the whole AI spectrum

- Relational topic models
- Mixed-membership models
- Relational Gaussian processes
- Relational reinforcement learning
- (Partially observable) MDPs
- Systems of linear equations
- Kalman filters
- Declarative information networks

No, this is very much like in the early days of UAI!

So, should we worry about the soup?
The early days of UAI

- Maximum entropy inference
- Odds-likelihood updating
- Dempster-Shafer Belief Functions
- Mycin's Certainty Factors
- Bayesian Networks
- Expert-rating
- Decision-theoretic metrics
- Belief Maintenance System
- Prospector
- Fuzzy Set Theory
- Probabilistic Logic
- Incidence Calculus

References:
- D. Hunter. Uncertainty Reasoning Using Maximum Entropy Inference. UAI-85
- S. Ursic. Generalizing Fuzzy Logic Probabilistic Inferences. UAI-86
- D. Hunter. An Empirical Comparison of Three Inference Methods. UAI-84

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications

This soup boiled down to Graphical Models as intermediate representation

Distributions can naturally be represented as **Factor Graphs**

- There is an edge between a circle and a box if the variable is in the domain/scope of the factor

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
Factor Graphs from Graphical Models

$$p(x) = p(x_1)p(x_2) \frac{p(x_3|x_1, x_2)}{p(x_1)}$$

$$f_a(x_1) = p(x_1)$$

$$f_b(x_2) = p(x_2)$$

$$f_c(x_1, x_2, x_3) = p(x_3|x_1, x_2)$$

$$\psi(x_1, x_2, x_3) = \psi(x_1, x_2, x_3)$$

Similar “boiling down” process is going on in SRL!

Boiled-Down SRL Alphabet Soup

- Given a relational model in your language of choice, a set of constants and a query, compile everything into an intermediate representation
  - Factor graphs
  - BDDs, Arithmetic Circuits, d-DNNFs, ...
  - Weighted CNFs

- Run inference
Rules + Potential: Logically Parameterized Factors

- ∀X. φ₁(popular, attends(X))
- ∀X. φ₂(attends(X), series)

Atoms represent a set of random variables

Parfactors parameterized factors

Logical Variables parameterize RV

There can also be constraints to logical variables such as X=/=UAI11

Rules + Weights: Weighted CNF

- Weighted MAX-SAT as mode finding for log-linear distributions
- Each configuration has a cost: the sum of the weights of the unsatisfied (ground) clauses.
- An infinite cost gives a ‘hard’ clause.
- Goal: find an assignment with minimal cost.
Common Approach to Inference

Inference within a graphical model

Relational Model

Grounding

Factor Graph:

\[ \begin{array}{c}
X & Y & Z \\
\alpha & \beta \\
\end{array} \]

Weighed CNF

\[ \begin{align*}
& w_1 \neg x \lor \neg x \lor y \\
& w_2 \neg x \lor z
\end{align*} \]

Weighted SAT

Lifted Inference

- Inference in first-order logic is not “ground” but lifted
  - handling whole sets of indistinguishable objects together

Resulting lifted approaches are often faster, more compact and provide more structure for optimization

Let’s exploit such symmetries in order to speed up inference
Lifted Inference

Graph-based
- Exact
  - FOVE, ...
  - [Poole 03, de Salvo Braz 06, 07, Milch 07, Kisyinski 09, 09b, Choi 10, 11, Sen 08, Taghipou 12]
- Approximate
  - [Gogate 10, 11, Jha 10, Van-den-Broeck 11]
  - Deterministic
    - [Singla 08, Kersting 09, 10, Sen 09, Nath 10, Hadjii 10, 11, Ahmadi 10, 12, 13, Riedel 08, Mladenov 12, Bui 13, Noessner 13, Van-den-Broeck 12, 13]
  - Sampling
    - [Milch 06, Poon 08, Zettlemoyer 07, Gogate 12, Niepert 12]
  - Interval
    - [de Salvo Braz 09]

Search
- Knowledge Compilation, Probabilistic Theorem Proving, ...
  - [Gogate 10, 11, Jha 10, Van-den-Broeck 11]
  - [Shavlik 09, Mihalkova 09]

Preprocessing
- MC-SAT, Lifted MCMC, Lifted Importance Sampling, ...
  - [Kersting, ECAI 2012, FAIA Track]
Lifted Belief Propagation

These are so-called fractional automorphisms computable in $O(n^2 \log n)$.

$AS = SA$ where $S$ is a doubly-stochastic and not a permutation matrix like for automorphisms.

An instance of the Weisfeiler-Lehman algorithm for (approx.) testing graph isomorphism.

Lifted Inference

What are the symmetries we are exploiting? Can we characterize them mathematically?

<table>
<thead>
<tr>
<th>Lifted Model</th>
<th>Ground Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP, VE, JT, ...</td>
<td>small/fast</td>
</tr>
<tr>
<td>(C,G-)FOVE, Probabilistic Theorem Proving,...</td>
<td>large/slow</td>
</tr>
<tr>
<td>LRCR</td>
<td></td>
</tr>
</tbody>
</table>
Lifted Inference

LBP, CBP, ...

LRCR

GI-complete! LBP does not compute them!

What are the symmetries we are exploiting? Can we characterize them mathematically?

(C,G-)FOVE, Probabilistic Theorem Proving, ...

Lifted Model

Automorphisms
[Bui et al. UAI13]

Fractional Automorphisms
[LBP and Co]

Ground Model

BP, VE, JT, ...

Hierarchy of Lifted Inference

LBP, CBP, ...

LRCR

Sherali-Adams Tightening

(C,G-)FOVE, Probabilistic Theorem Proving, ...

Lifted Model

Fractional Automorphisms
[LBP and Co]

Automorphisms
[Bui et al. UAI13]

But lifted inference actually covers the whole AI spectrum

BP, VE, JT, ...

[Mladenov, Kersting, 2013 submitted]
[Mladenov, Kersting, Ahmadi AISTATS 2012]

**Lifted Linear Programming**

\[
\begin{align*}
\max_{x, y, z} & \quad (c^T x + y + 1z) \\
\text{s.t.} & \quad \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ -z \end{bmatrix} \leq \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}
\end{align*}
\]

Now, run WL to compute cluster matrix B

Solve using any LP solver

---

[Grohe, Mladenov, Kersting, Selman 2013, submitted; Van-den-Broeck 13, Noessner 13]

**Dimensionality Reduction**

Now, run WL to compute cluster matrix B

Solve using any LP solver
Relaxed ILPs with symmetries present (ILPs are due to Margot) for computing maximum cardinality binary error correcting codes, edge colourings, minimum dominating sets in Hamming graphs, and Steiner-triple systems, among others. [Grohe, Kersting, Mladenov, Selman, 2013, under submission] http://wpweb2.tepper.cmu.edu/fmargot/lpsym.html

Empirical Illustration: Combinatorial Problems

Illustration: AI Problems

Grid MDP with the same reward in each corner

Standard MAP-LP using the smoker-friends MLN with the Frucht and McKay graphs to encode the social network as well as (attracting) Ising models
The current lifted frontier

Already in 1848, Louis Pasteur recognized “Life as manifested to us is a function of the asymmetry of the universe”

PTP vs. symmetries? fractional autom. vs. autom.? distributed lifting, asymmetric & approximate lifting, variational lifted inference, lifted tree decompositions, temporal models, open world, inferning, lifted QPs and SDPs, ...

Exponential speed-ups achievable but actually any speed-up is a success

ILP = Machine Learning + Logic Programming
[Muggleton, De Raedt JLP96]

Find a set of general rules
- mutagenic(X) :- atom(X,A,c), charge(X,A,0.82)
- mutagenic(X) :- atom(X,A,n), ...

Examples E
- pos(mutagenic(m1))
- neg(mutagenic(m2))
- pos(mutagenic(m3))

Background Knowledge B
- molecule(m1)
- atom(m1,a11,c)
- atom(m1,a12,n)
- bond(m1,a11,a12)
- charge(m1,a11,0.82)
- ...

- molecule(m2)
- atom(m2,a21,o)
- atom(m2,a22,n)
- bond(m2,a21,a22)
- charge(m2,a21,0.82)
- ...

Pedro Domingos, Kristian Kersting
Combining Probability and Logic: Languages, Algorithms and Applications
**Example ILP Algorithm: FOIL**
[Quinlan MLJ 5:239-266, 1990]

\[
\begin{align*}
\text{mutagenic}(X) & :- \text{atom}(X,A,n), \text{charge}(A,0.82) \\
\text{mutagenic}(X) & :- \text{atom}(X,A,c), \text{bond}(A,B)
\end{align*}
\]

\[
\vdash \text{true}
\]

\[
\text{Coverage} = 0.5, 0.7
\]

\[
\vdash \text{atom}(X,A,c)
\]

\[
\text{Coverage} = 0.5, 0.7
\]

\[
\vdash \text{atom}(X,A,n)
\]

\[
\text{Coverage} = 0.6, 0.3
\]

\[
\vdash \text{atom}(X,A,f)
\]

\[
\text{Coverage} = 0.4, 0.6
\]

**Vanilla SRL Approach** [De Raedt, Kersting ALT04]

\[
\begin{align*}
\text{mutagenic}(X) & :- \text{atom}(X,A,n), \text{charge}(A,0.82) \\
\text{mutagenic}(X) & :- \text{atom}(X,A,c), \text{bond}(A,B)
\end{align*}
\]

\[
=0.882
\]

- Traverses the hypotheses space a la ILP
- Replaces ILP’s 0-1 covers relation by a “smooth”, probabilistic one [0,1]

\[
\begin{align*}
\text{cover}(e, H, B) & = P(e|H, B) \\
\text{cover}(E, H, B) & = \prod_{e \in E} \text{cover}(e, H, B)
\end{align*}
\]
Boosted Statistical Relational Mining

\[ \text{Data} + \text{Loss fct} + \text{Initial Model} \]

\[ \text{Final Model} \]

### Algo | Likelihood
--- | ---
Boosting | 0.810
MLN | 0.730

To predict \( \text{Fine}(X) \):

- \( \text{speed}(X, S), S > 120 \)
- \( \text{job}(X, \text{politician}) \)
- \( \text{Count}_{Y, \text{knows}(X, Y)} > 0 \)
- \( \text{job}(Y, \text{politician}) \)

MARKOV LOGIC