Exploring Unknown Universes
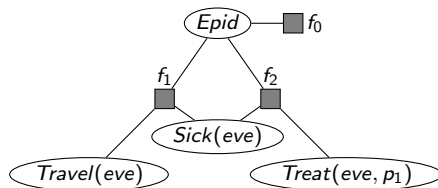# in Probabilistic Relational Models

Tanya Braun, Ralf Möller

Institute of Information Systems
University of Lübeck

December 4, 2019

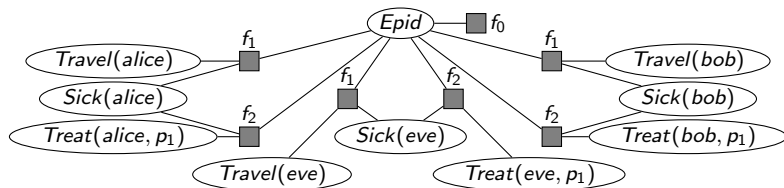# Probabilistic Graphical Models
Factor Graph F: Semantics

Compact encoding of full joint distribution

$$P_F = \frac{1}{Z} \prod_i f_i$$

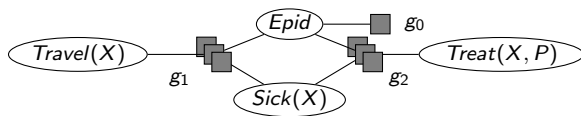# Probabilistic Graphical Models
Factor Graph F: Semantics

Compact encoding of full joint distribution

$$P_F = \frac{1}{Z} \prod_i f_i$$
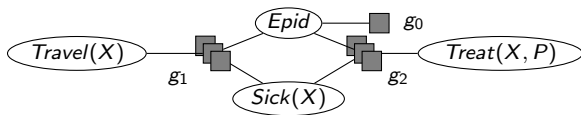
Probabilistic Relational and Lifted Models

Parfactor Graph G: Distribution Semantics (Sato 1995)

Probabilistic Relational and Lifted Models

Parfactor Graph G: Distribution Semantics (Sato 1995)
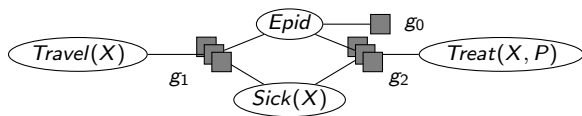

Compact encoding of full joint distribution
with tractable inference in domain sizes

$$P_G = \frac{1}{Z} \prod_{f \in gr(G)} f$$

Probabilistic Relational and Lifted Models

Parfactor Graph G: Distribution Semantics (Sato 1995)

Compact encoding of full joint distribution
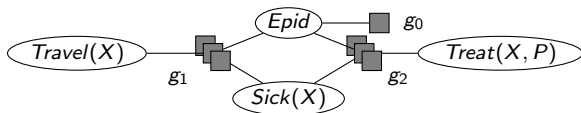with tractable inference in domain sizes

$$P_G = \frac{1}{Z} \prod_{f \in gr(G)} f$$

⇓

Requires knowing domains of parameters (X, P) in g_i's
(i.e., the universe is known)

Known Universe

Parfactor Graph G: Distribution Semantics (e.g., Poole 2003, Taghipour 2013)

Universe

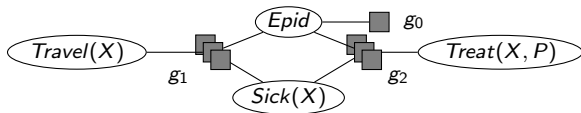General domains of parameters, e.g.,

$$dom(X) = \{alice, eve, bob\}$$

Constraint C_i for g_i as (subsets of) combinations of domains, e.g.,

$$C_2 = ((X, P), dom(X) \times dom(P))$$

Distribution semantics apply given a universe

Problem: **Un**known Universe

Parfactor Graph G: Distribution Semantics

Ramifications

General domains of parameters are **empty**,

$$dom(X) = \{\cancel{alice}, \cancel{eve}, \cancel{bob}\}$$

Constraint C_i for g_i as combinations of domains **empty**

$$C_2 = ((X, P), \cancel{dom(X) \times dom(P)})$$

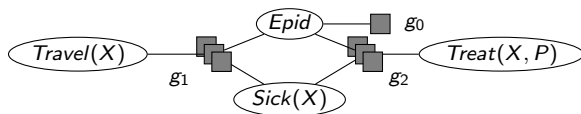Semantics do not apply $\rightarrow$ Lifted algorithms do not work

Conference Contribution

Semantics for Lifted Models with Unknown Universes

- Template model + constraint program
  - $\rightarrow$ Local distributions as templates
  - $\rightarrow$ Abstract description of constraints
- Domain worlds
  - $\rightarrow$ Describe possible universes
  - $\rightarrow$ Inputs to constraint program
- Generate a set of possible models
  - $\rightarrow$ Within universe: distribution semantics apply
  - $\rightarrow$ Reasoning over possible models

  Enable tractable inference with lifted algorithms again

Template Models

Template Model

Local distributions as templates with empty constraints, e.g.,

$$C_2 = ((X, P), \perp)$$

Template Models

Template Model

Local distributions as templates with empty constraints, e.g.,

$$C_2 = ((X, P), \bot)$$

What do we do about empty constraints?

Constraint Programs

Constraint Program

Generate specific constraints for template models if given domain,
e.g., using probabilistic Datalog:

```
element_of_C2(X,Y1) :- linked(X,Y1,Y2).
element_of_C2(X,Y2) :- linked(X,Y1,Y2).
linked(X,Y1,Y2) :- instance_of_X(X) & pair(Y1,Y2).
0.7 pair(t1,t2).  0.2 pair(t2,t3).  0.1 pair(t1,t3).
```
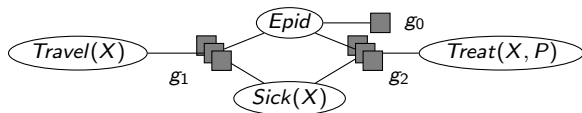
Constraint Programs

Constraint Program

Generate specific constraints for template models if given domain, e.g., using probabilistic Datalog:

```
element_of_C2(X,Y1) :- linked(X,Y1,Y2).
element_of_C2(X,Y2) :- linked(X,Y1,Y2).
linked(X,Y1,Y2) :- instance_of_X(X) & pair(Y1,Y2).
0.7 pair(t1,t2).  0.2 pair(t2,t3).  0.1 pair(t1,t3).
```

Where do we get a given domain from?

Domain Worlds

Domain World

Specify or generate possible domains, e.g., for X

Distribution-based Semantics

Semantics for Lifted Models with Unknown Universes

- Template model:
  - $\rightarrow$ Empty constraints
- Constraint program:
  - $\rightarrow$ Generate constraints given a domain world
- Domain worlds:
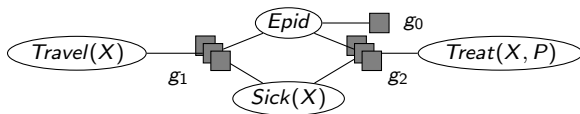 - $\rightarrow$ Generate possible domain worlds

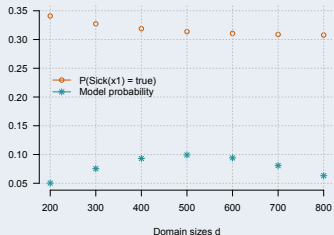**Generate a set/distribution of possible models**
$$\Downarrow$$
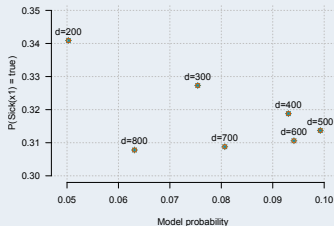Distribution semantics apply again

New Queries Emerging

Exploration

Model and query probabilities

with respect to domain size
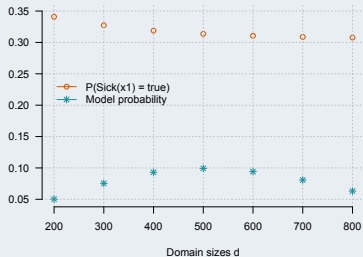

for skyline query

New Queries Emerging
Model Checking

Behaviour between models

E.g., does the probability of

- an individual being sick decrease with larger domains?

- an epidemic decrease with more treatments?

- an epidemic rise if more people travel?

Analysis: Probabilistic Inference with Unknown Universes

Syntactic components

- Template model
- Constraint program
- Domain worlds

Set of possible worlds

- $\rightarrow$ Distribution: Expected values
- $\rightarrow$ Runtime efforts increase
- $\rightarrow$ Possibilities for new queries arise

Constraint meta-programming

- $\rightarrow$ Build oracle for algorithms

Transfer learning

- $\rightarrow$ Decoupling from specific domain

Conference Contribution

Semantics for Lifted Models with Unknown Universes

- Template model + constraint program
  - $\rightarrow$ Local distributions as templates
  - $\rightarrow$ Abstract description of constraints
- Domain worlds
  - $\rightarrow$ Describe possible universes
  - $\rightarrow$ Inputs to constraint program
- Generate a set of possible models
  - $\rightarrow$ Within universe: distribution semantics apply
  - $\rightarrow$ Reasoning over possible models

Enable tractable inference with lifted algorithms again