# A Short Introduction to Probabilistic Soft Logic

Angelika Kimmig, Stephen H. Bach, Matthias Broecheler,
Bert Huang and Lise Getoor

NIPS Workshop on Probabilistic Programming 2012

http://psl.umiacs.umd.edu

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]
- Dependencies as weighted first order rules

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]
- Dependencies as weighted first order rules
- Support for similarity functions and aggregation

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]
- Dependencies as weighted first order rules
- Support for similarity functions and aggregation
- Linear (in)equality constraints

# Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]
- Dependencies as weighted first order rules
- Support for similarity functions and aggregation
- Linear (in)equality constraints

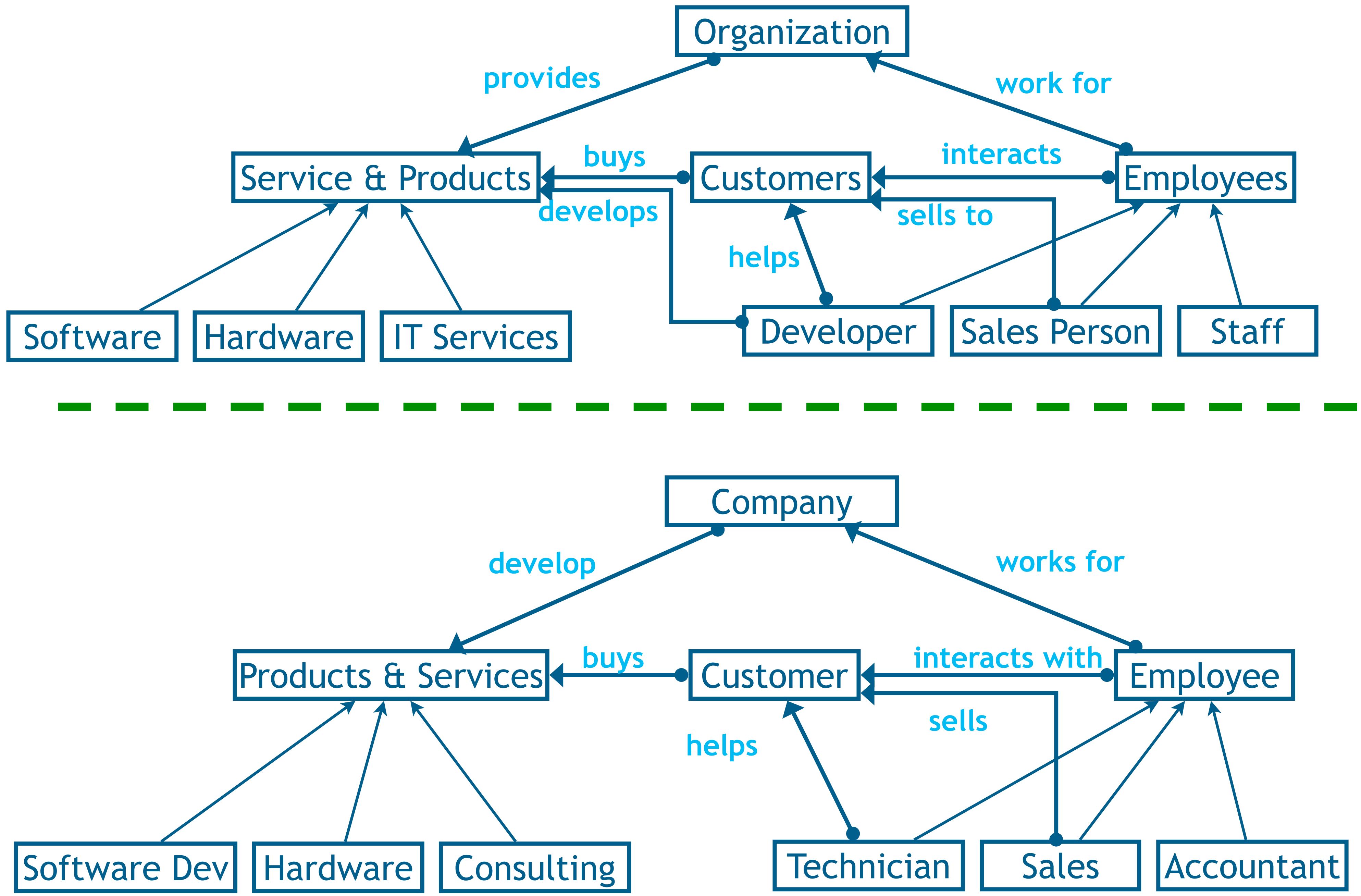
new approach
NIPS 12

Efficient MPE inference: continuous convex optimization

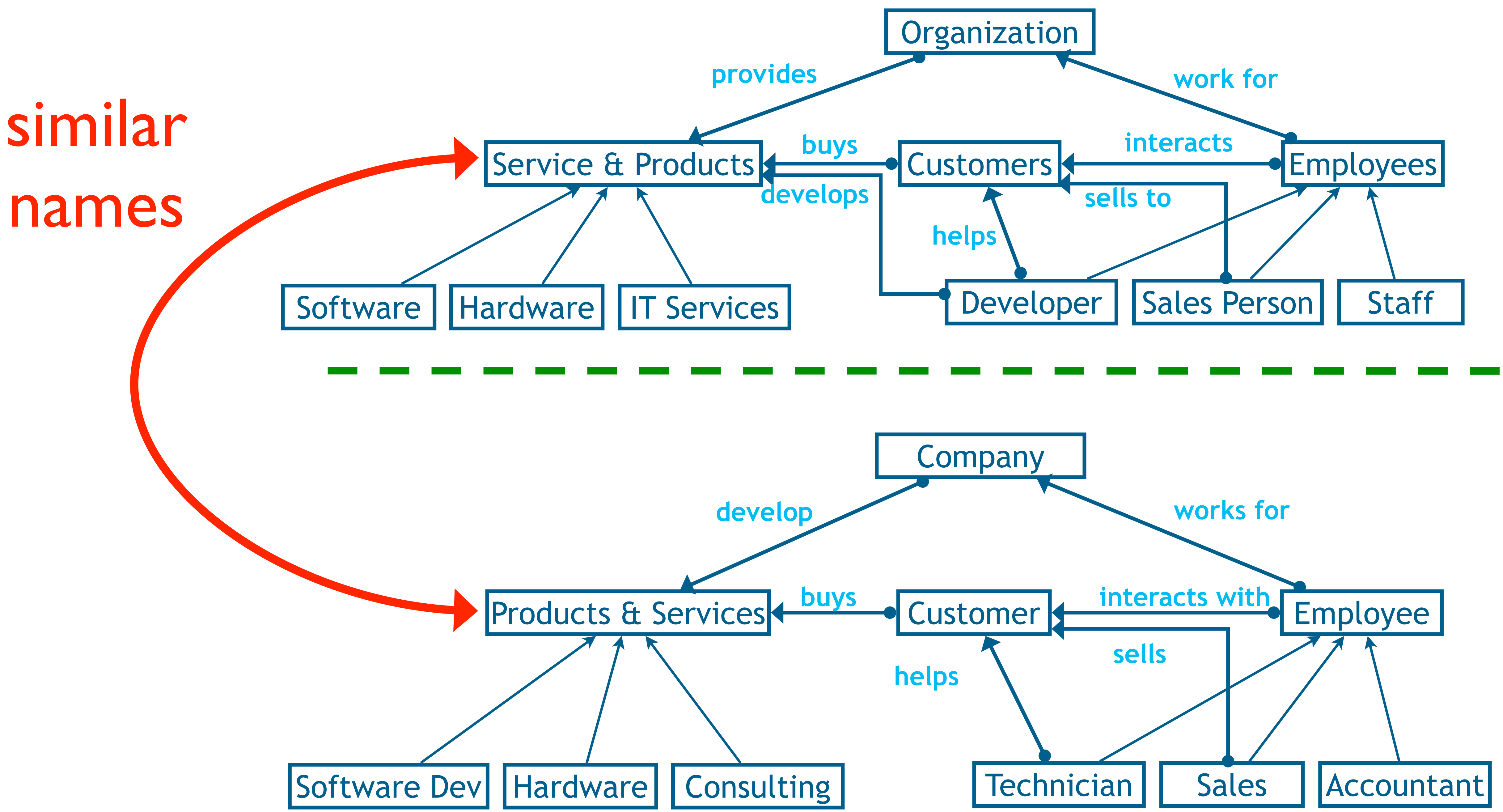
# Applications

- Collective classification
- Ontology alignment
- Entity resolution
- Link prediction
- Trust in social networks
- Social group modeling
- Personalized medicine
- ...

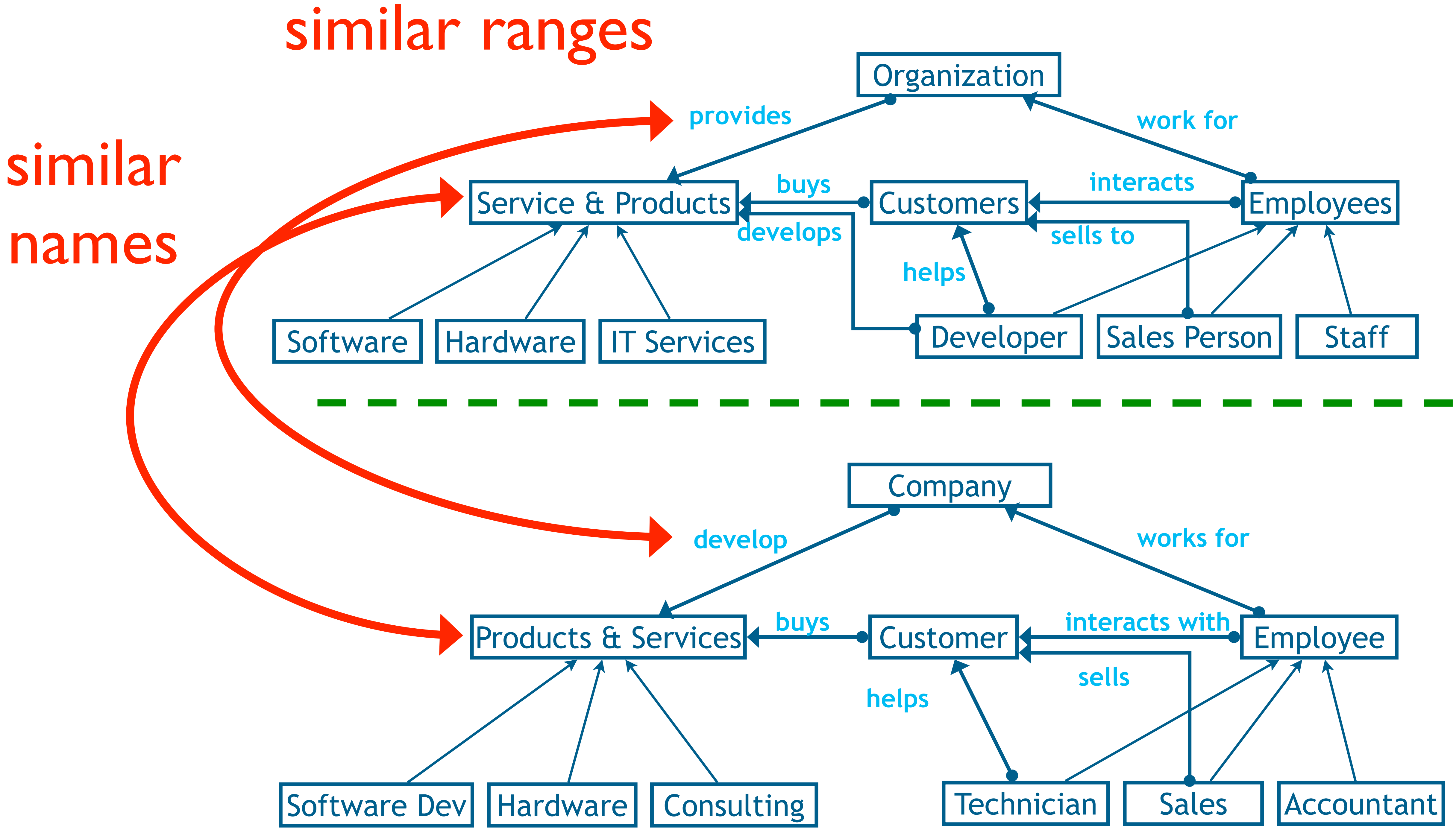
# Ontology Alignment



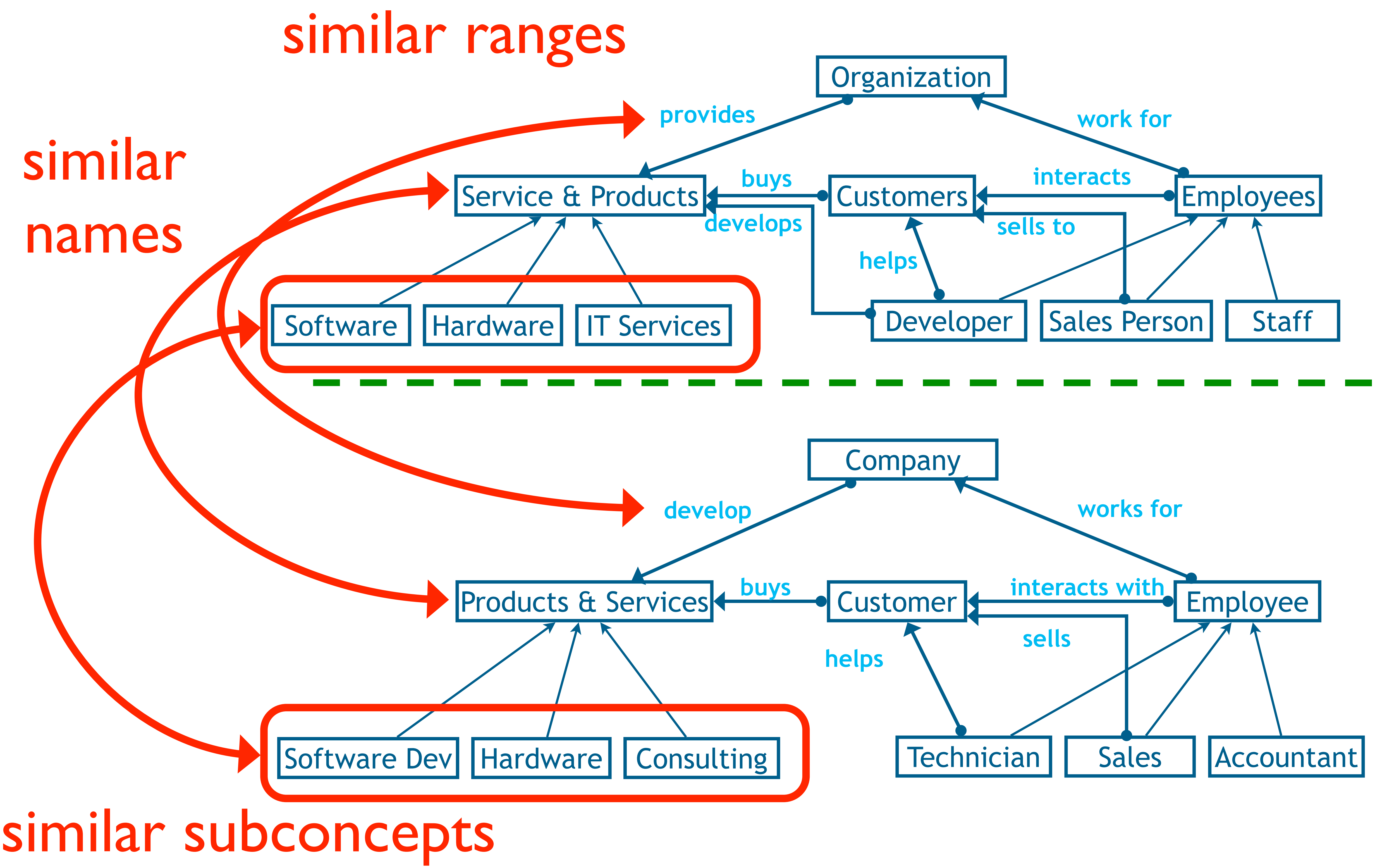
# Ontology Alignment



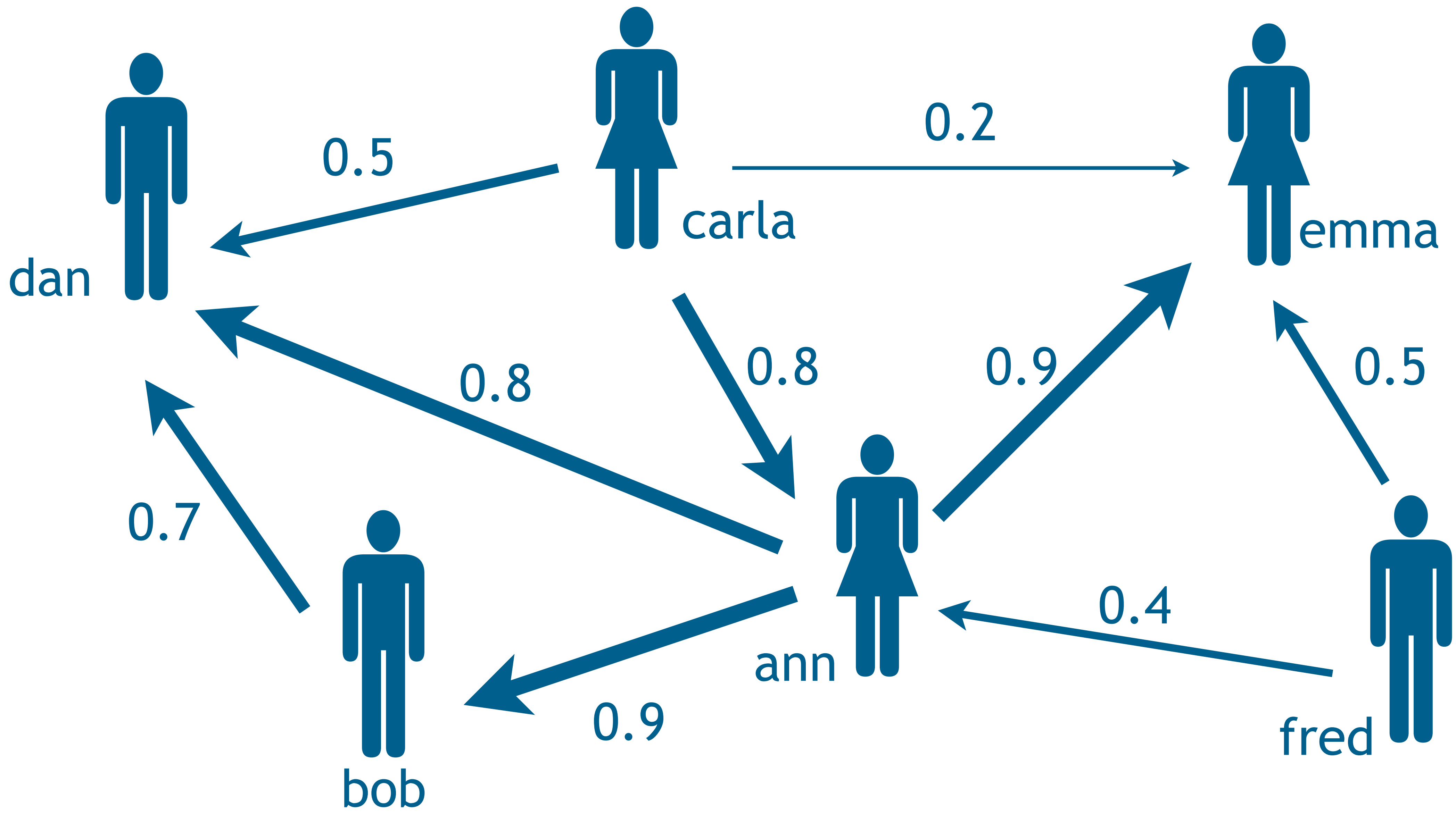
# Ontology Alignment



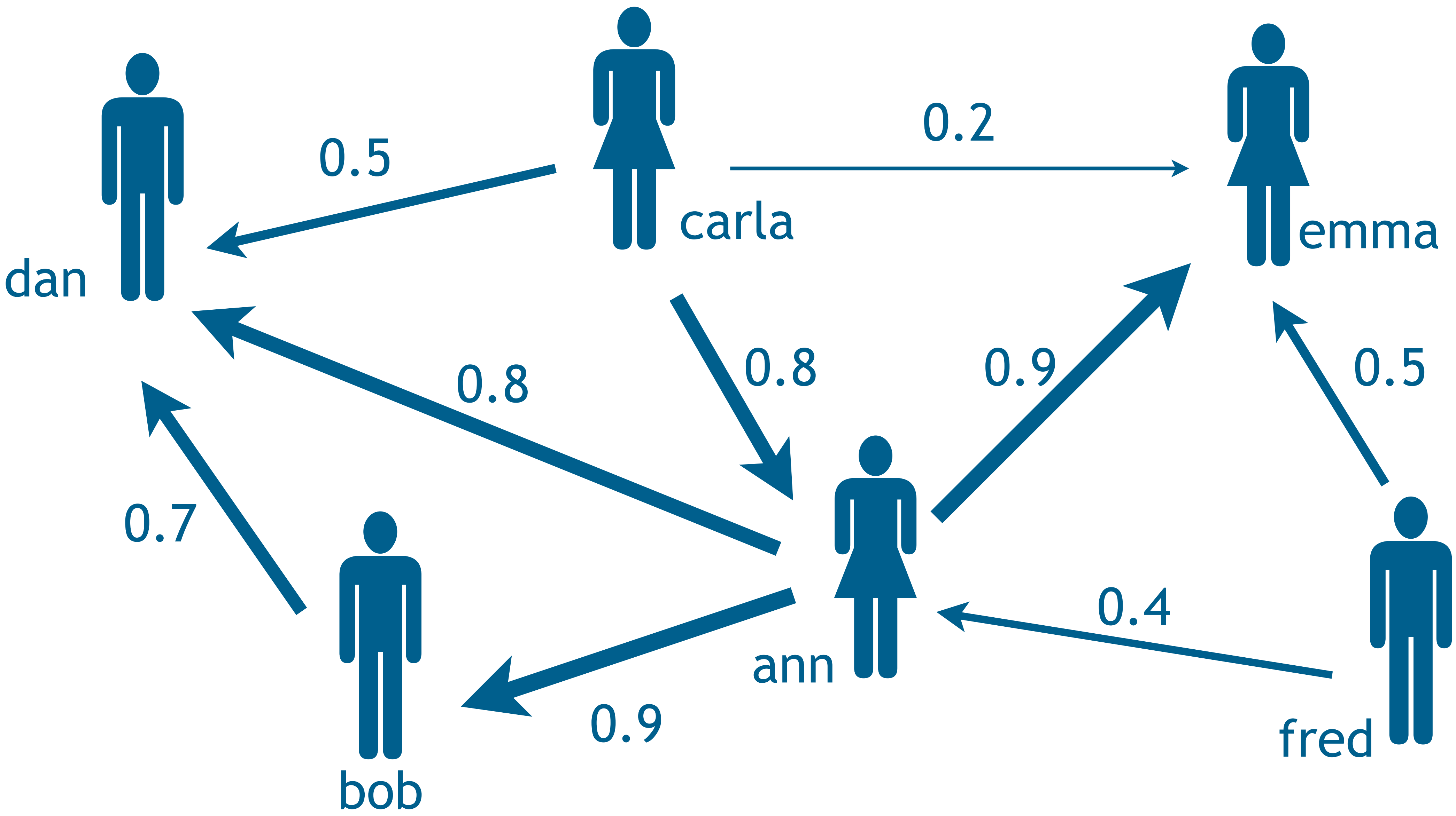
# Ontology Alignment



# Trust Modeling

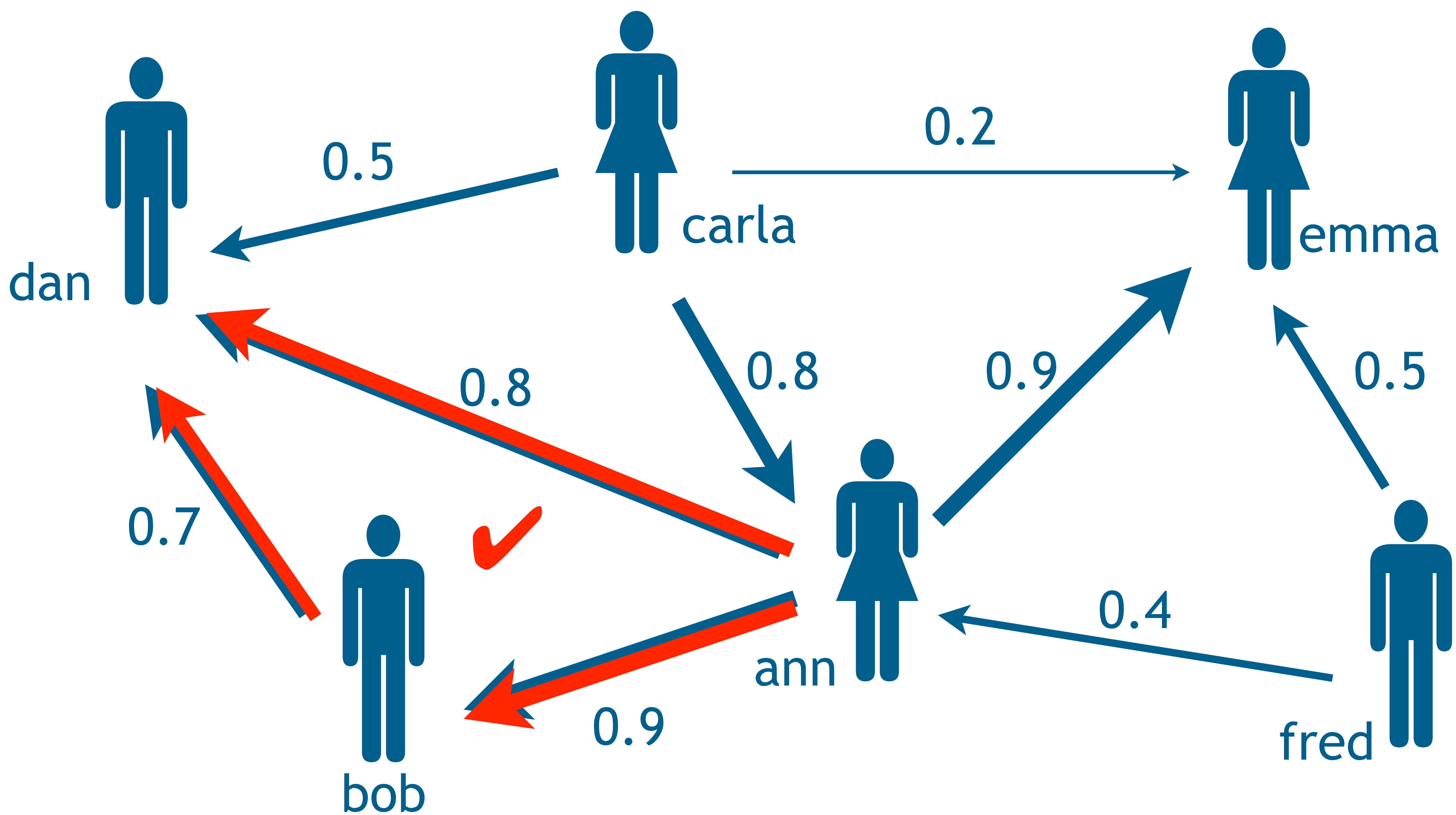


# Trust Modeling



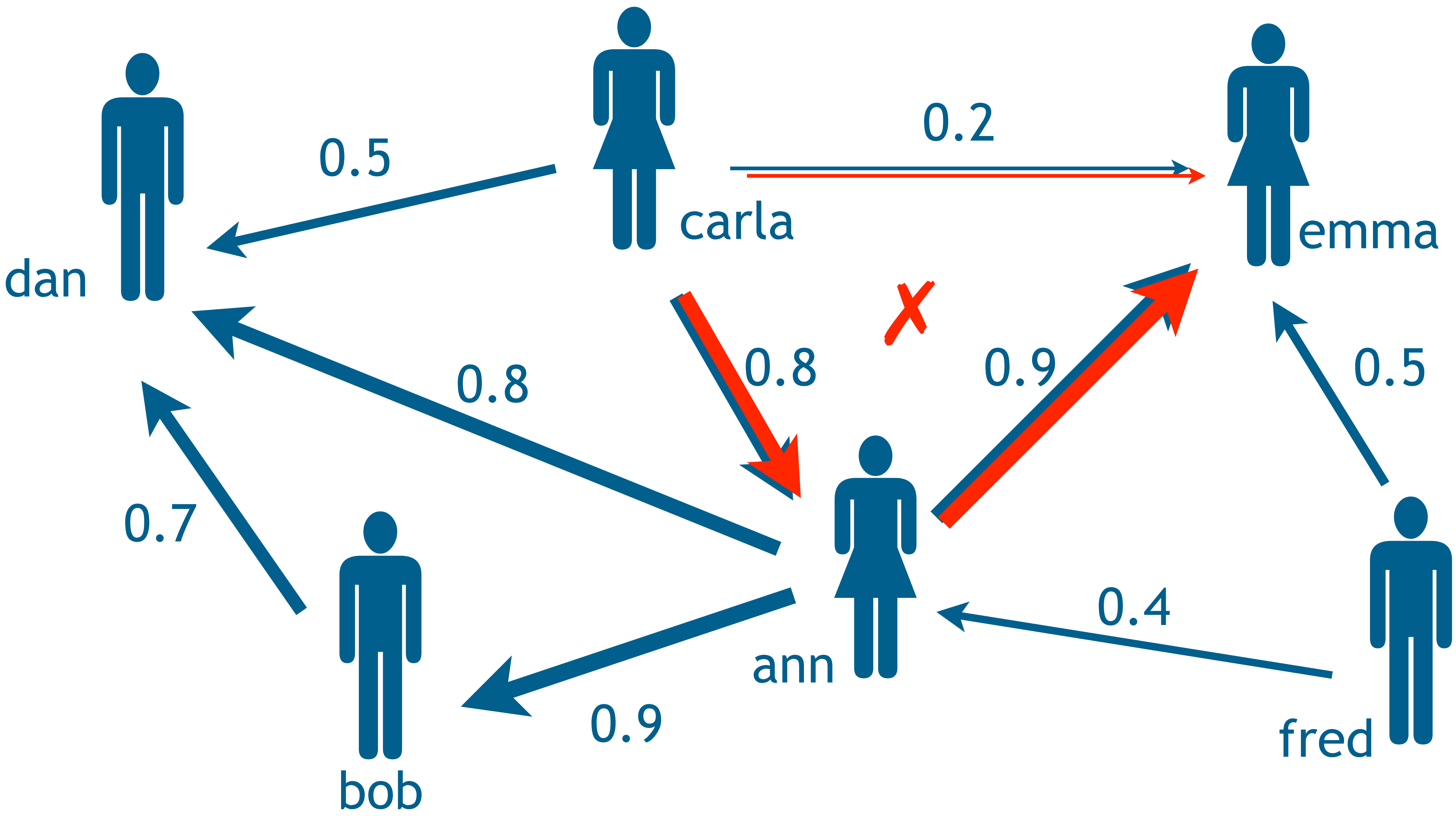

$$trusts(X,Y) \land trusts(Y,Z) \rightarrow trusts(X,Z)$$

# Trust Modeling



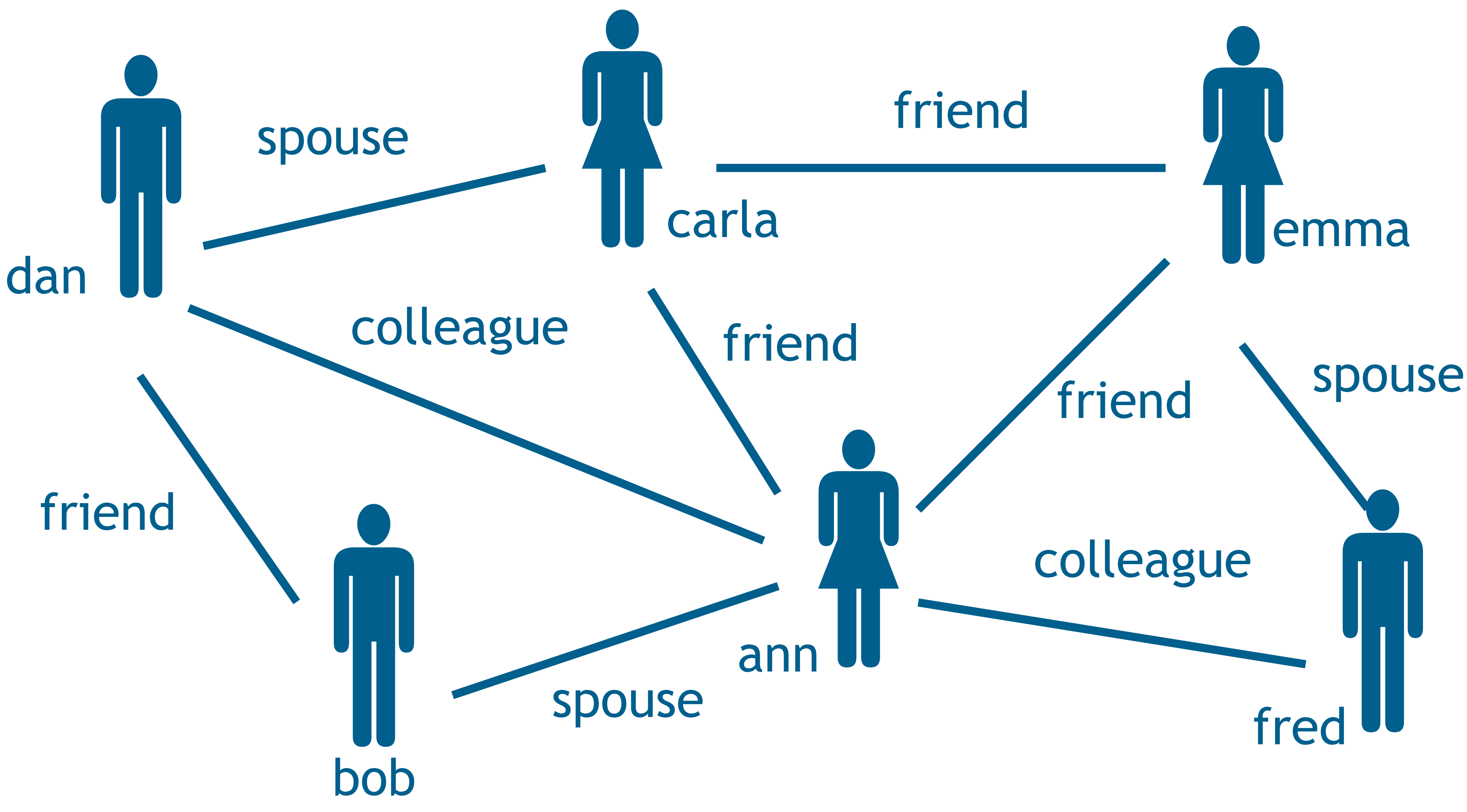

$$trusts(X,Y) \wedge trusts(Y,Z) \rightarrow trusts(X,Z)$$

# Trust Modeling

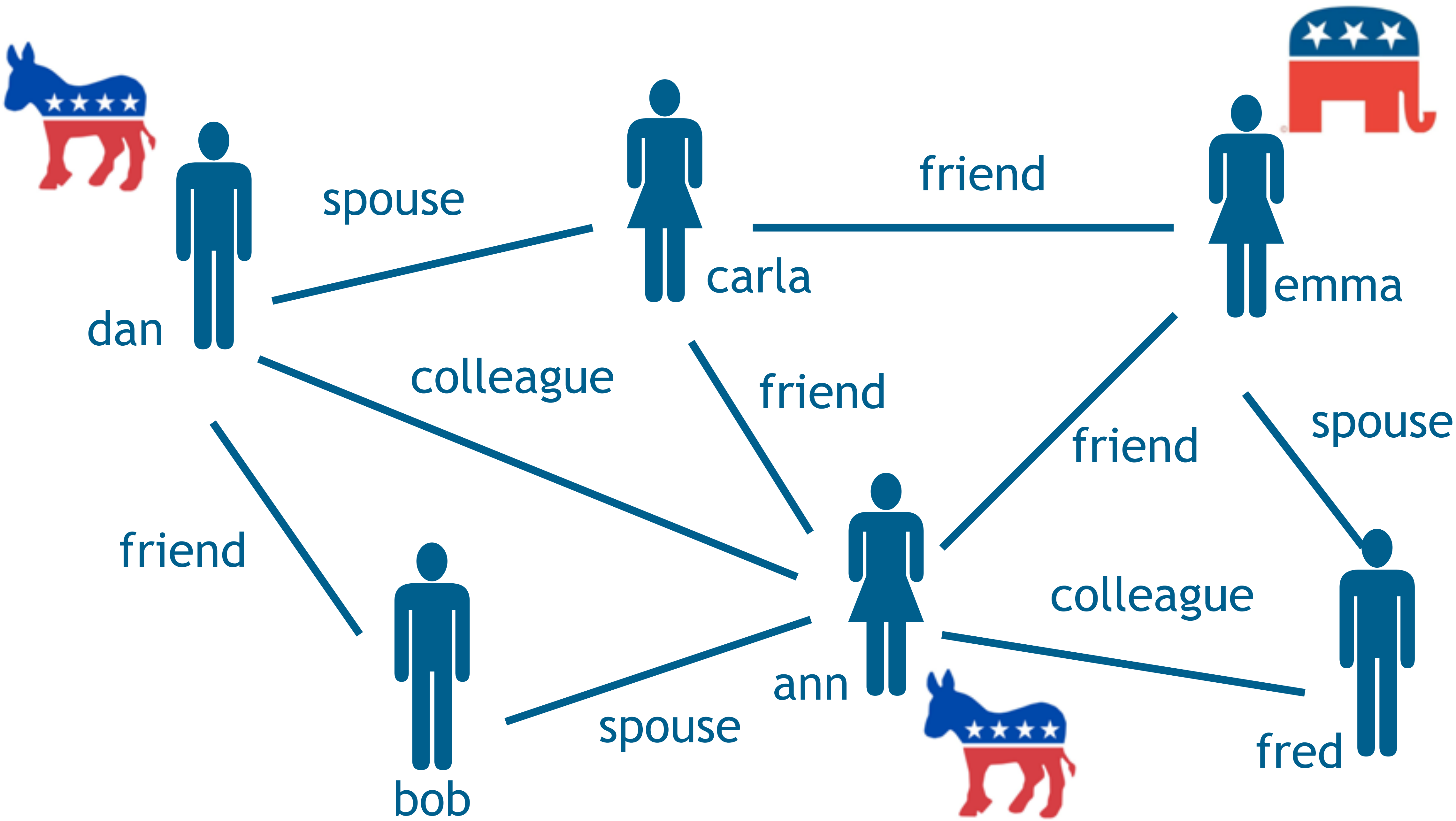



$$trusts(X,Y) \wedge trusts(Y,Z) \rightarrow trusts(X,Z)$$

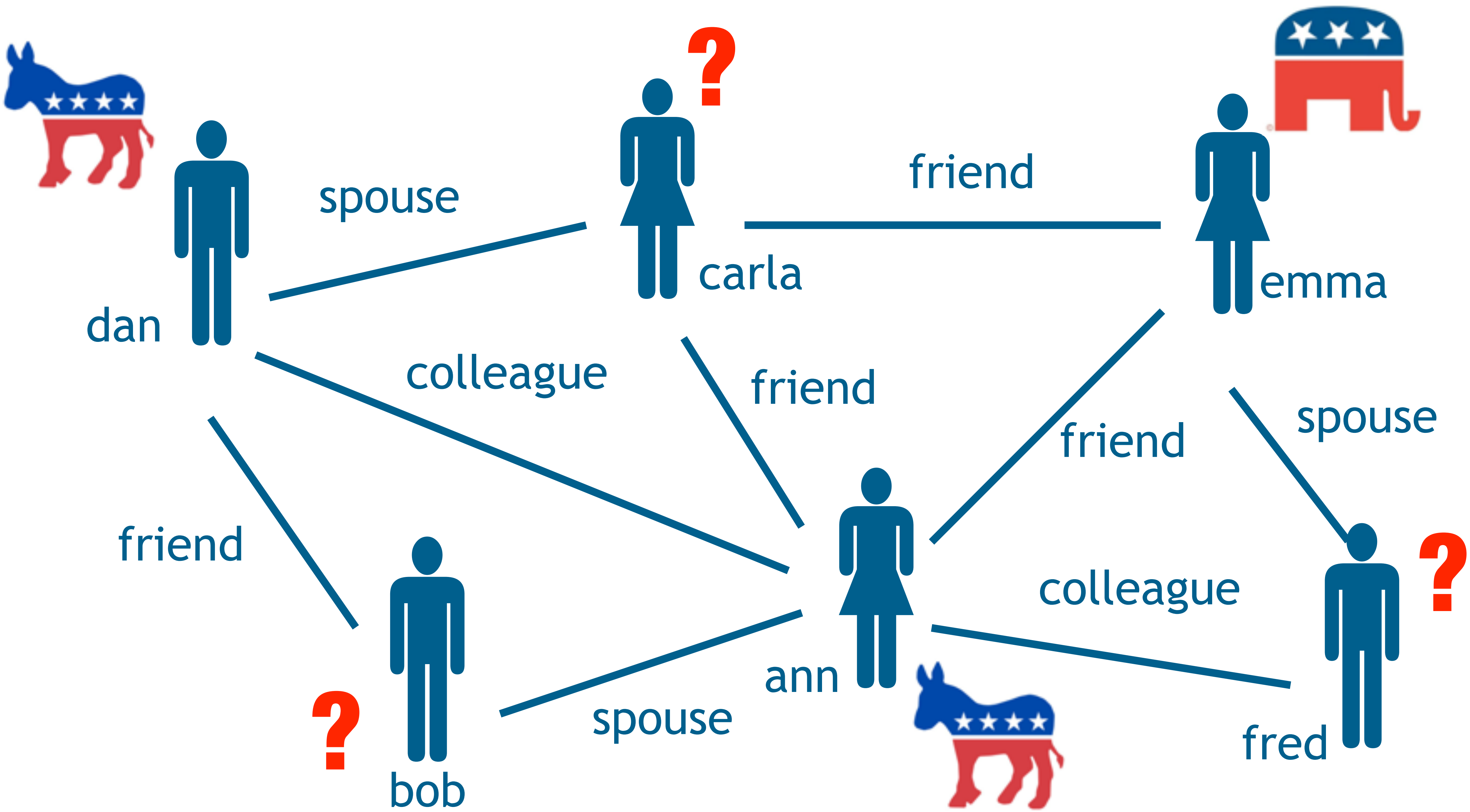
# Voter Opinion Modeling



# Voter Opinion Modeling

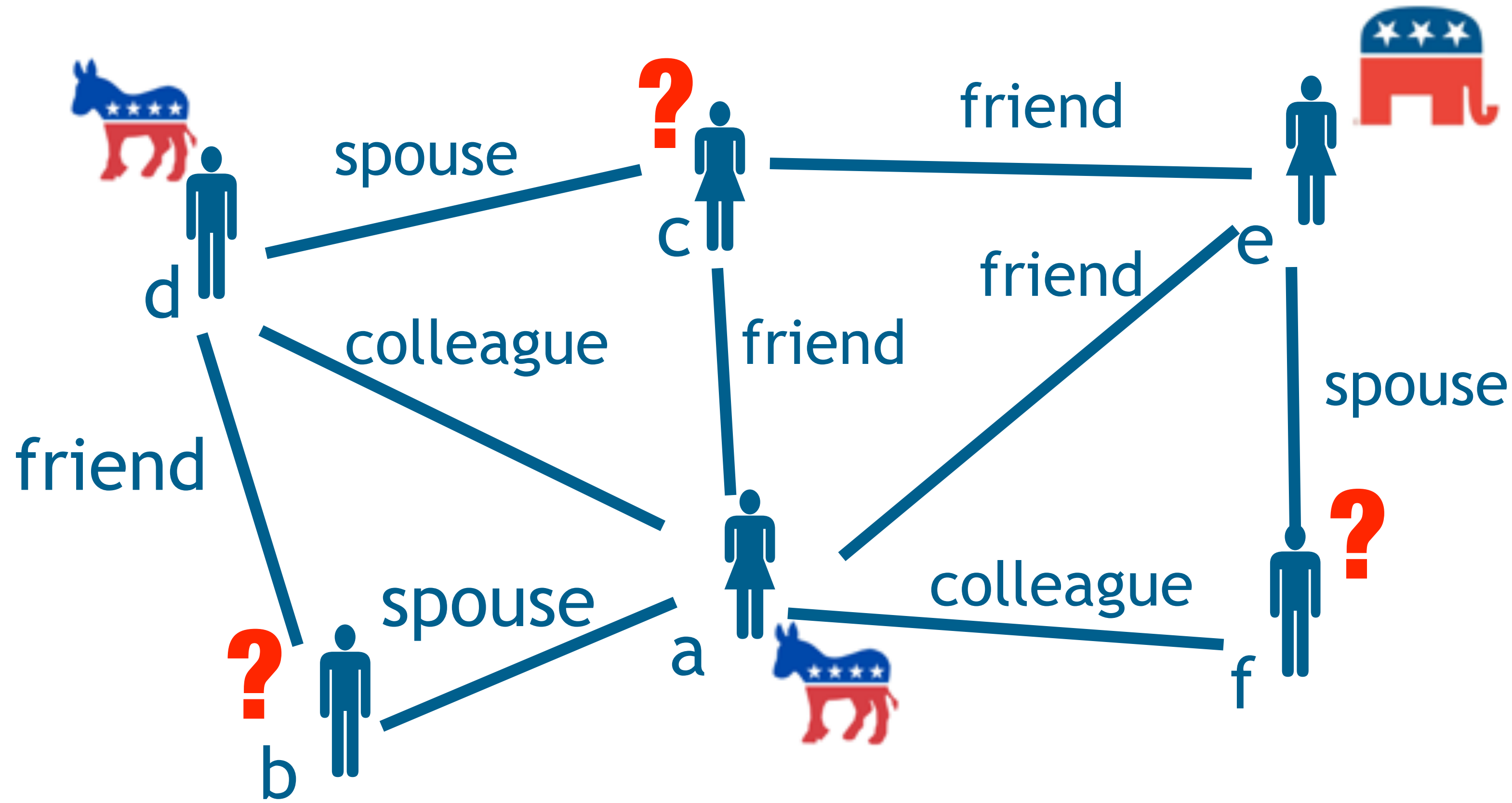


# Voter Opinion Modeling



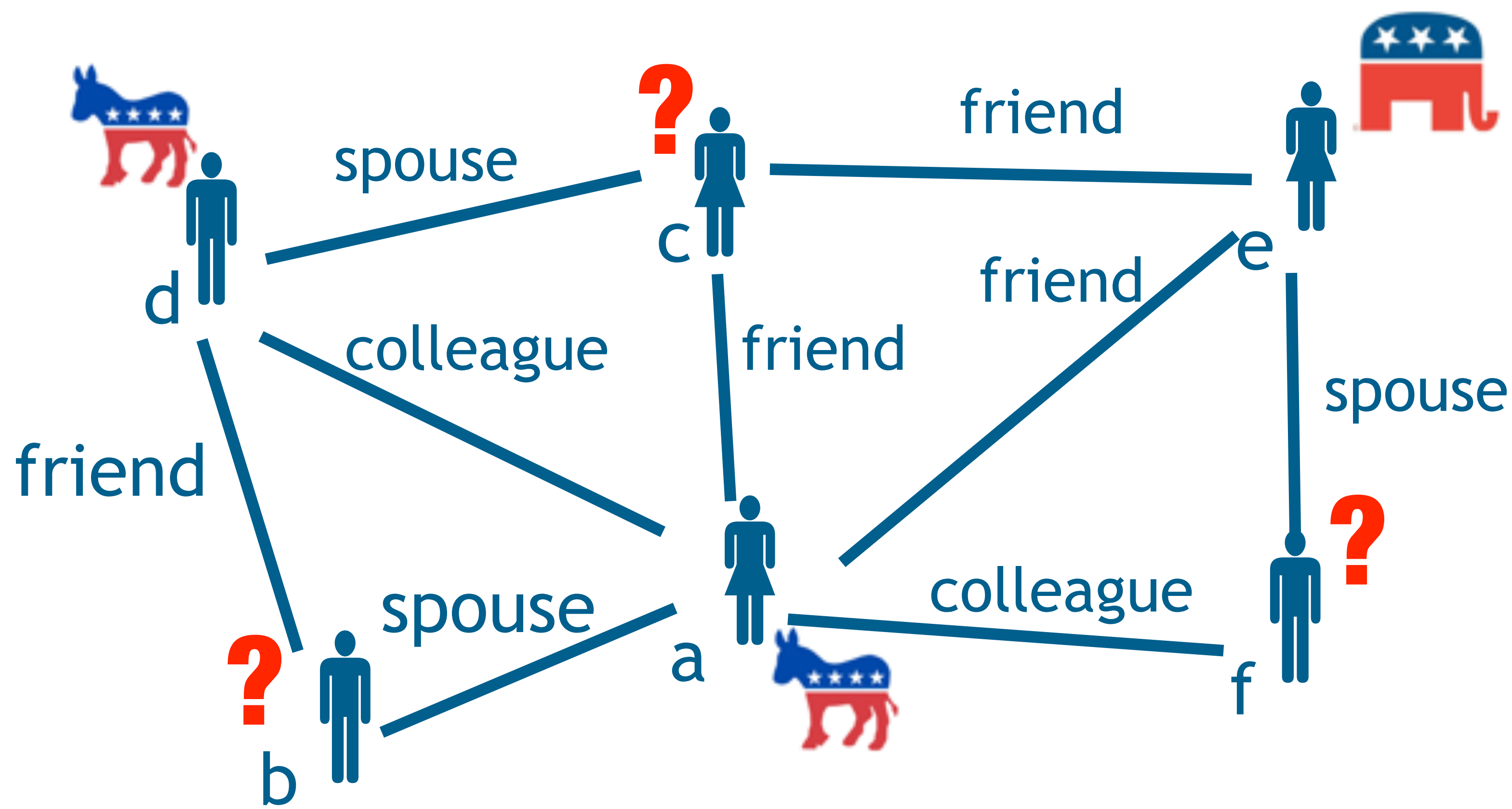
# PSL Program

friend(carla,emma)=0.9
friend(bob,dan)=0.4
spouse(ann,bob)=1.0
prefers(ann, )=0.8
...



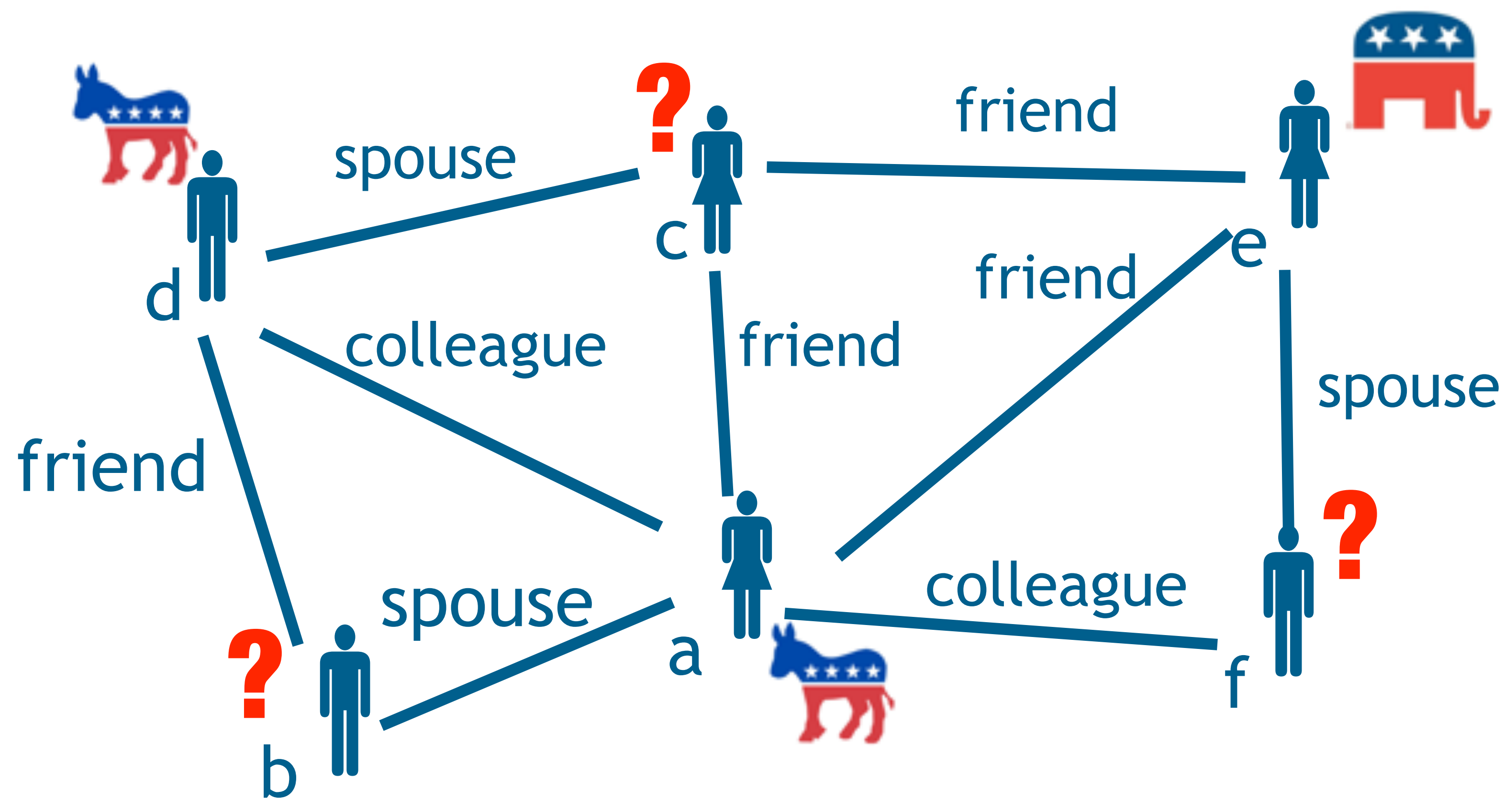
# PSL Program

friend(carla,emma)=0.9
friend(bob,dan)=0.4
spouse(ann,bob)=1.0
prefers(ann, )=0.8
...



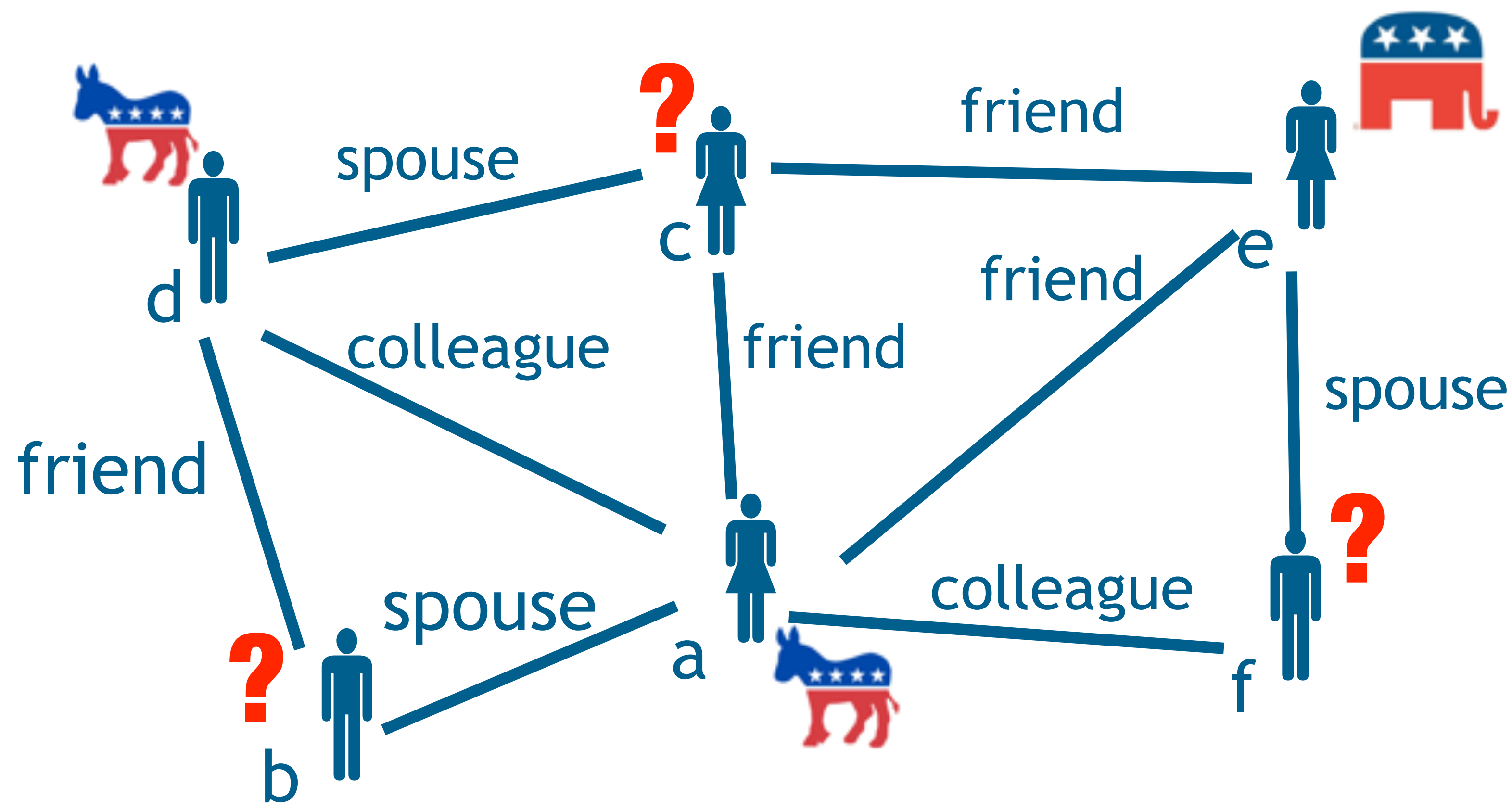

0.3: lives(A,S) ∧ majority(S,P) → prefers(A,P)
0.8: spouse(B,A) ∧ prefers(B,P) → prefers(A,P)
0.1: similarAge(B,A) ∧ prefers(B,P) → prefers(A,P)
0.4: prefers(A,P) → prefersAvg({A.friend},P)
partial-functional: prefers

# Constraints




partial-functional: prefers

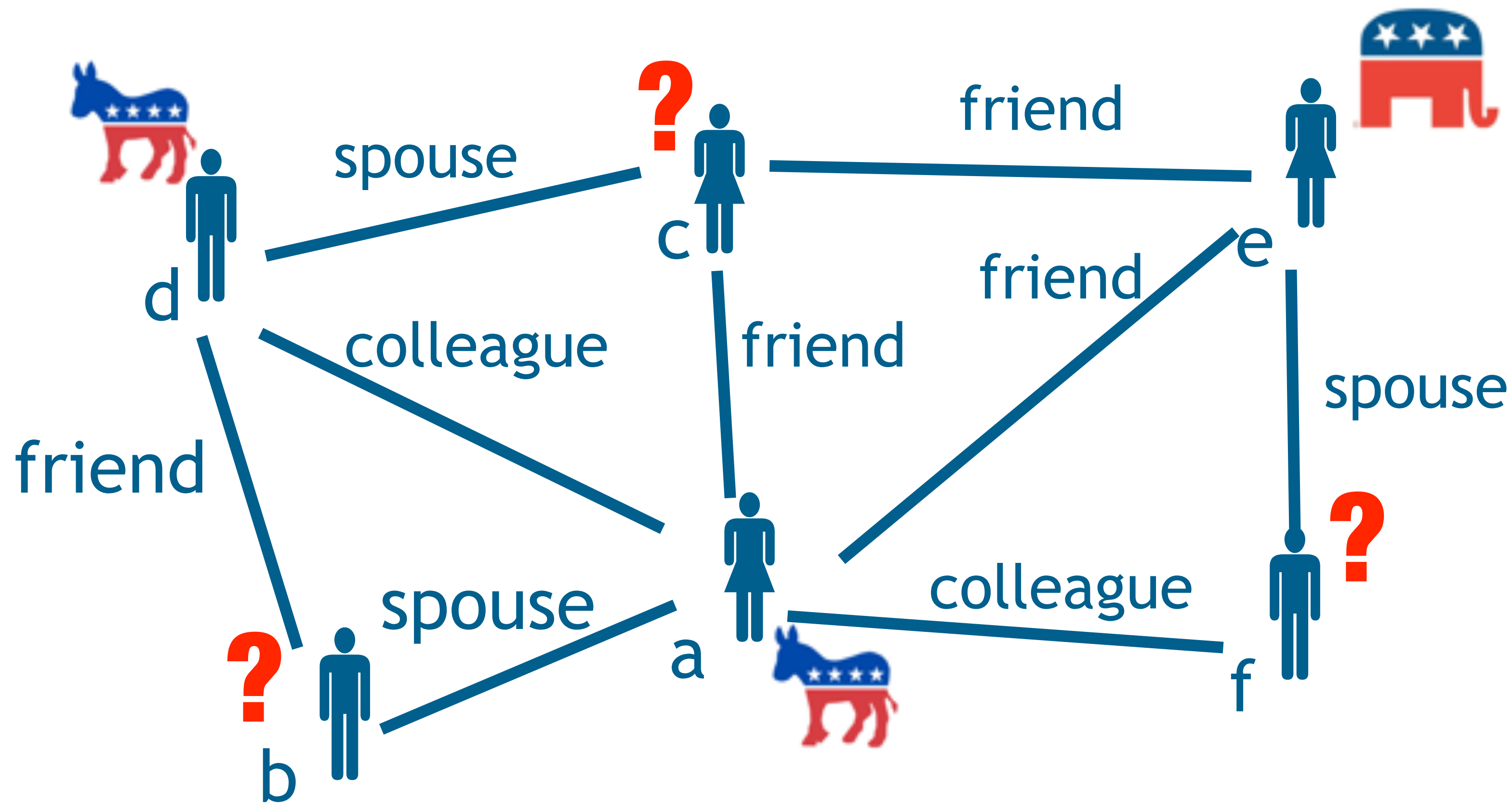
# Constraints




partial-functional: prefers

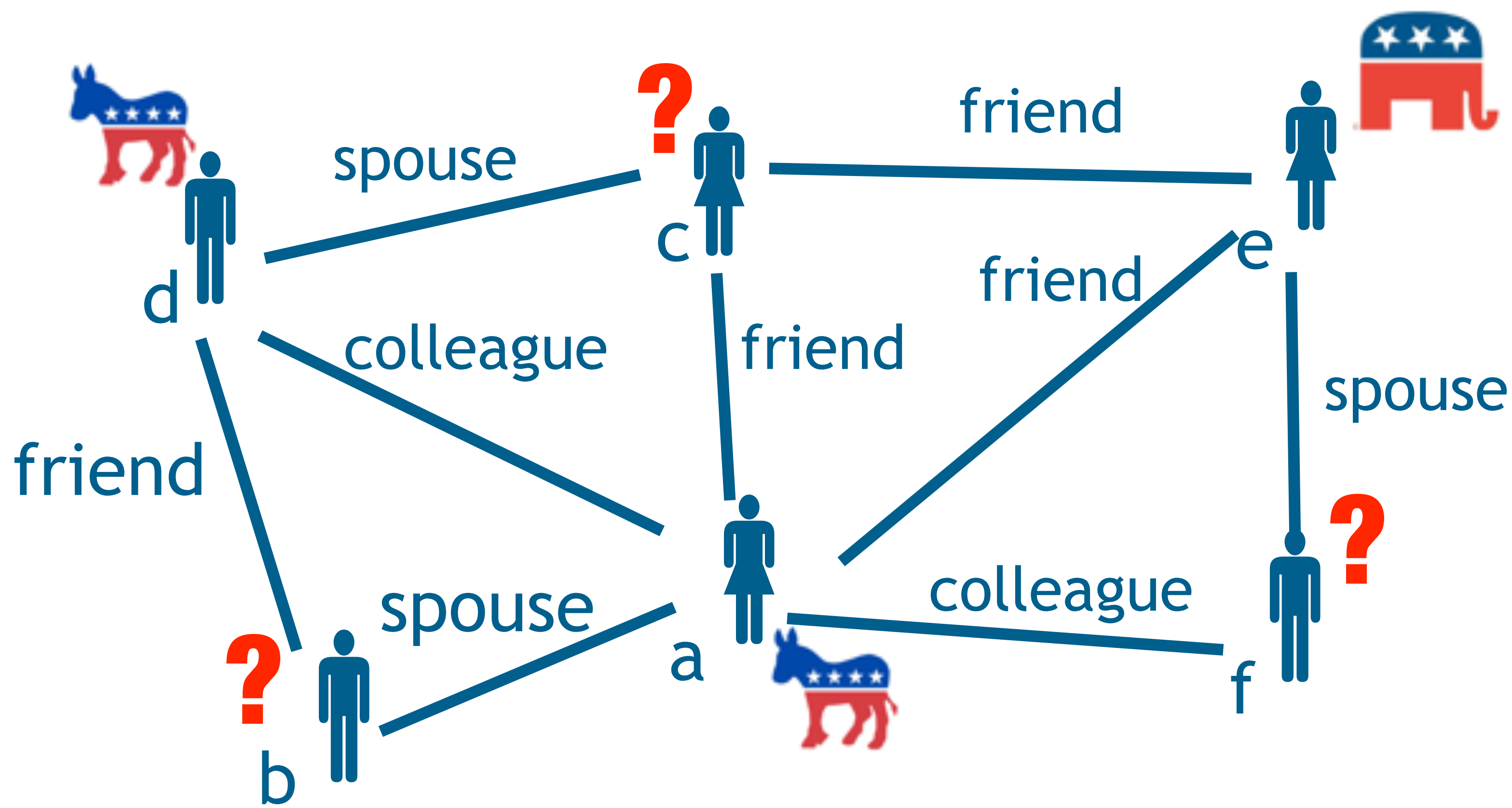
prefers(A, )+prefers(A, ) $\leq$ 1.0

# Local Rules



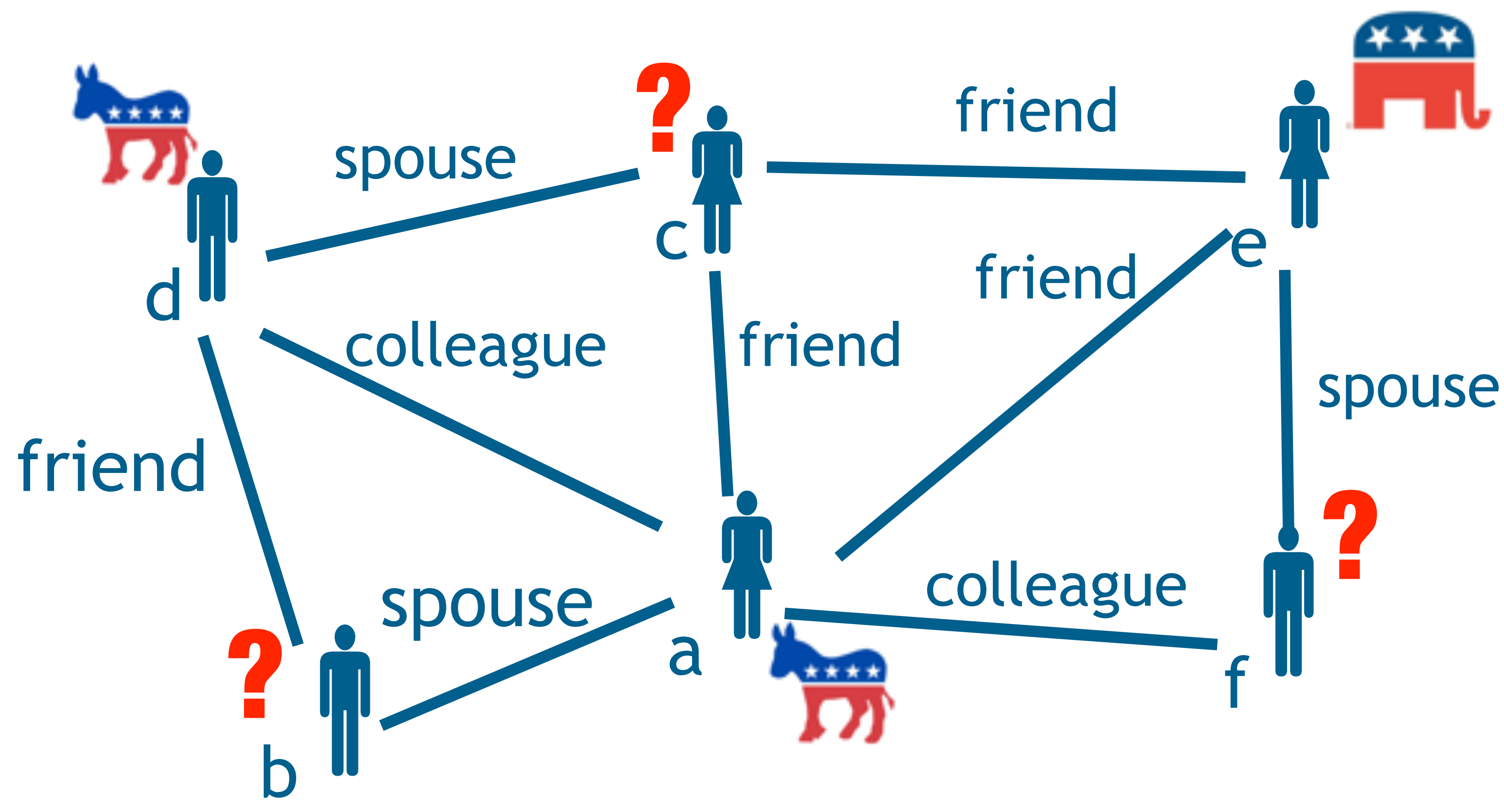

$$0.3: \text{lives}(A,S) \wedge \text{majority}(S,P) \rightarrow \text{prefers}(A,P)$$

# Propagation Rules




0.8: spouse(B,A) ∧ prefers(B,P) → prefers(A,P)

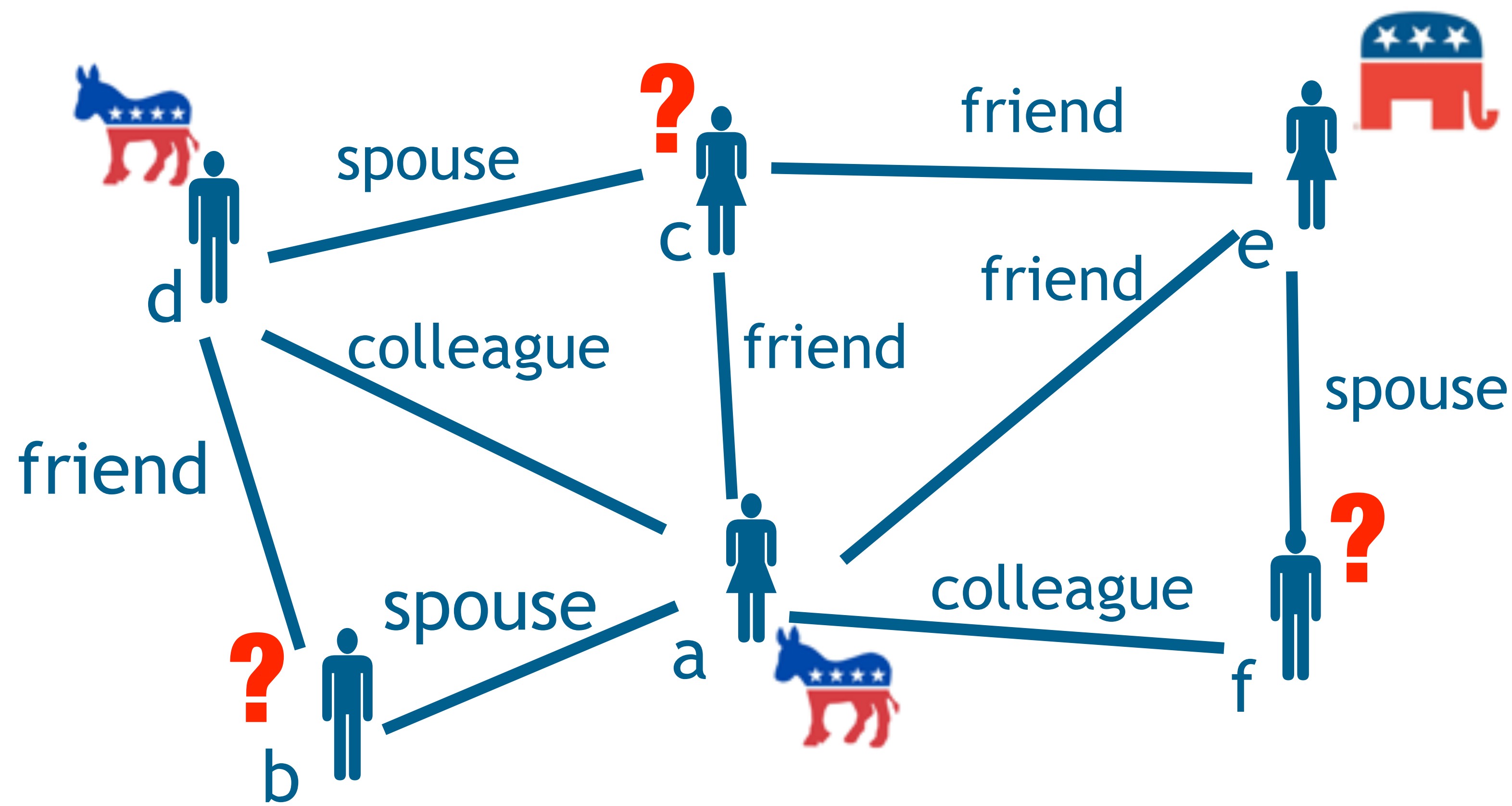
# Similarity Rules




0.1: similarAge(B,A) ∧ prefers(B,P) → prefers(A,P)

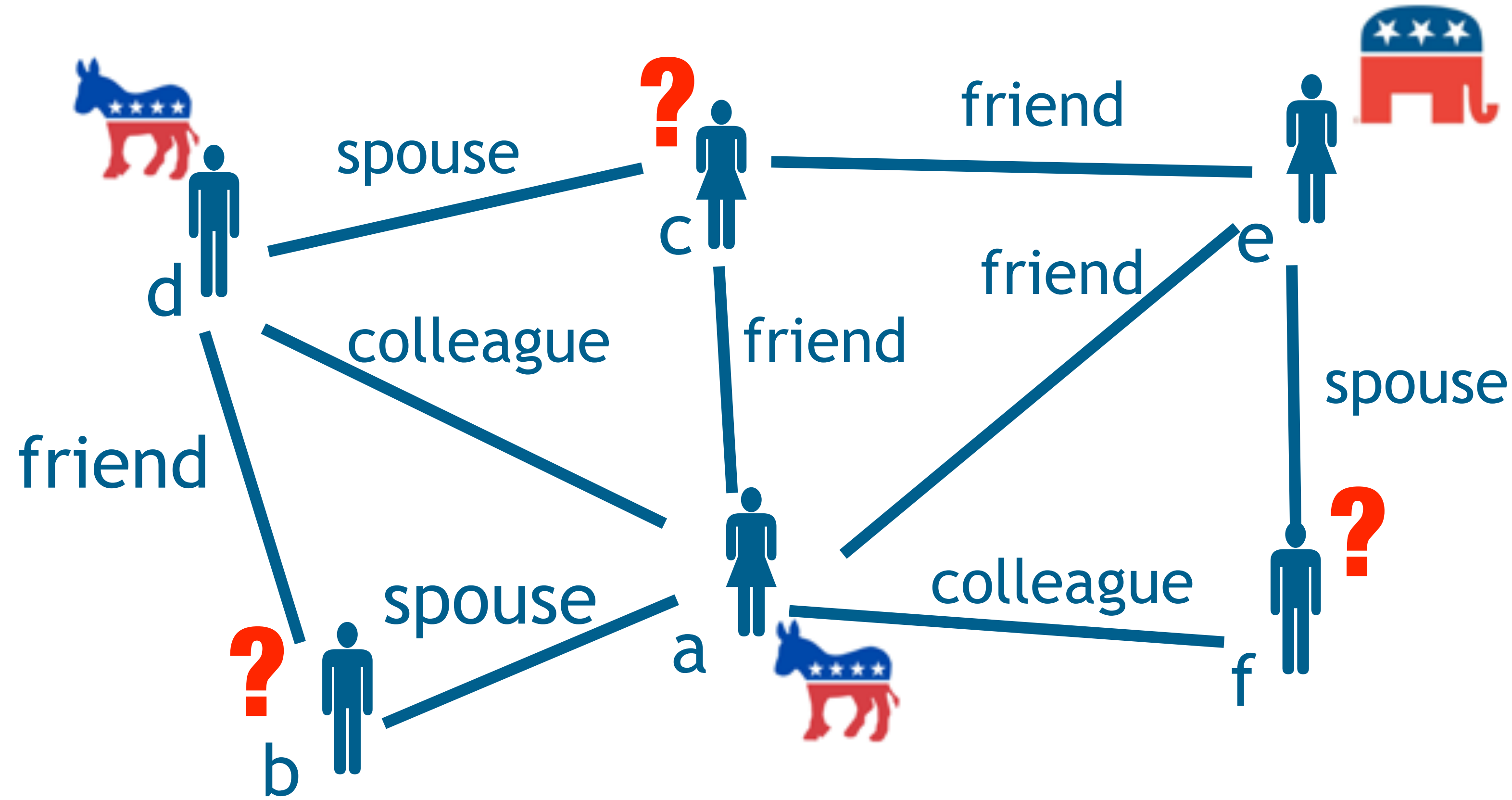
Similarity function with range [0,1]

# Sets




0.4: prefers(A,P) → prefersAvg({A.friend},P)

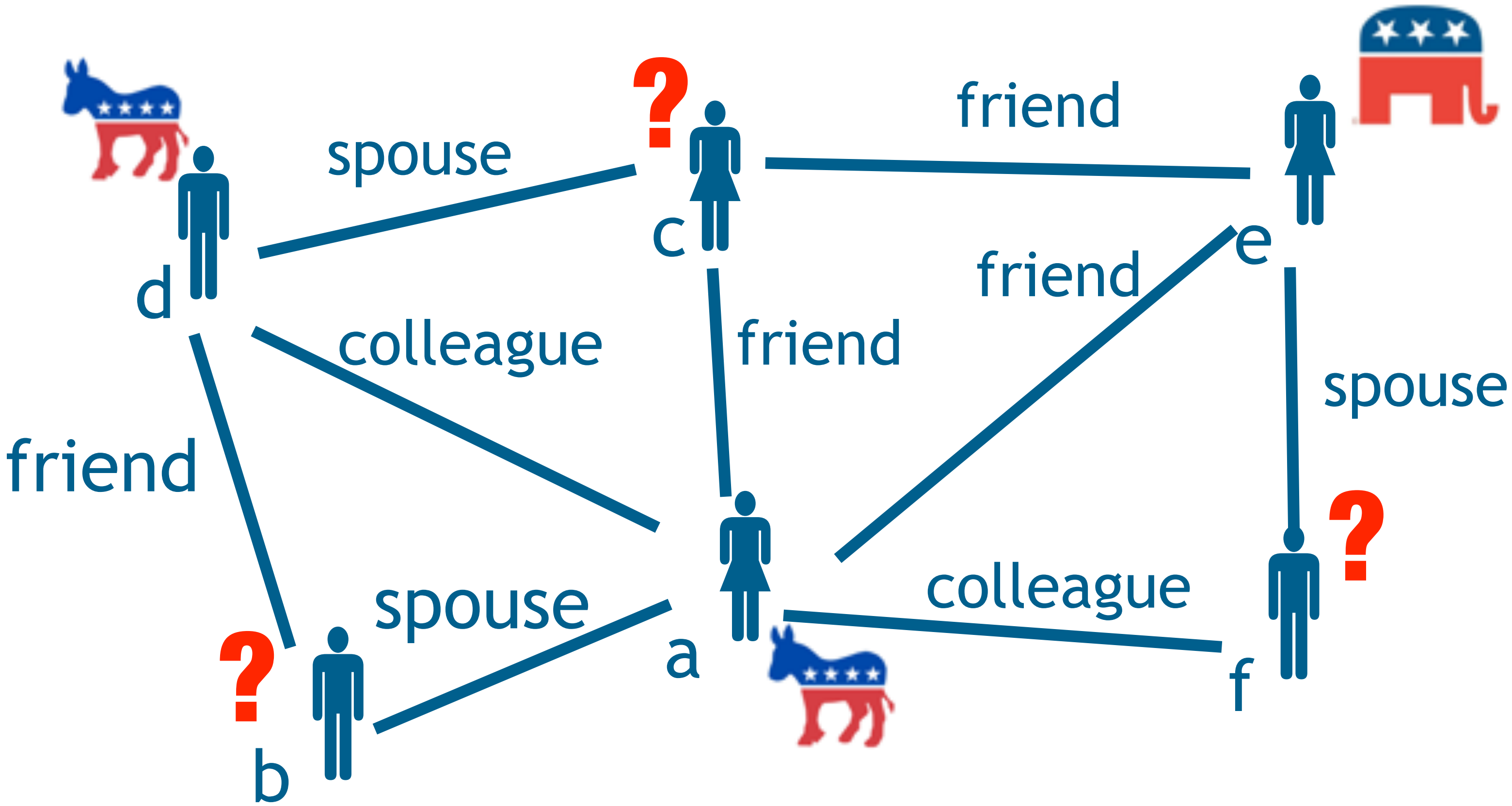
# Sets




all X such that friend(A,X)

0.4: prefers(A,P) → prefersAvg({A.friend},P)

# Sets




all X such that friend(A,X)

0.4: prefers(A,P) → prefersAvg({A.friend},P)

truth value := average truth value of prefers(X,P)

# PSL Program

- Ground atoms = random variables
- Soft truth value assignments
- Assignment satisfying more rules more likely
- Constraints to rule out unwanted assignments

# Probabilistic Model

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

# Probabilistic Model

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

# Probabilistic Model

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

Set of rule groundings

# Probabilistic Model

$$f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

Rule's weight

Set of rule groundings

# Probabilistic Model

Ground rule's distance from satisfaction given I

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

Rule's weight

Set of rule groundings

# Probabilistic Model

Ground rule's distance from satisfaction given I

∈{1,2}

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

Rule's weight

Set of rule groundings

# Probabilistic Model

Ground rule's distance from satisfaction given I

∈{1,2}

$$f(I) = \frac{1}{Z} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

Interpretation

Rule's weight

Set of rule groundings

Normalization constant

$$Z = \int_{J \in \mathcal{I}} \exp\left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(J))^k \right)$$

# Distance from Satisfaction

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

# Distance from Satisfaction

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

*body* → *head* satisfied

⇔

truth value of *body* ≤ truth value of *head*

# Distance from Satisfaction

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

*body* → *head* satisfied

⇔

truth value of *body* ≤ truth value of *head*

Lukasiewicz t-norm

$$I(v_1 \wedge v_2) = \max\{0, I(v_1) + I(v_2) - 1\}$$
$$I(v_1 \vee v_2) = \min\{I(v_1) + I(v_2), 1\}$$
$$I(\neg l_1) = 1 - I(v_1)$$

# Distance from Satisfaction

similarAge(bob,ann) $\wedge$ prefers(bob, ) $\rightarrow$ prefers(ann, )

0.8 0.5 0.5

# Distance from Satisfaction

similarAge(bob,ann) ∧ prefers(bob, ) → prefers(ann, )

0.8 0.5 0.5

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

$$I(v_1 \wedge v_2) = \max\{0, I(v_1) + I(v_2) - 1\}$$

# Distance from Satisfaction

similarAge(bob,ann) ∧ prefers(bob, ) → prefers(ann, )

0.8 0.5 0.5

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

$$I(v_1 \wedge v_2) = \max\{0, I(v_1) + I(v_2) - 1\}$$

max{0, 0.8+0.5−1} ≤ 0.5 d=0.0

# Distance from Satisfaction

similarAge(bob,ann) $\wedge$ prefers(bob, ) $\rightarrow$ prefers(ann, )

0.8 0.5 0.2

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

$$I(v_1 \wedge v_2) = \max\{0, I(v_1) + I(v_2) - 1\}$$

max{0, 0.8+0.5−1} ≤ 0.5 d=0.0

max{0, 0.8+0.5−1} ≤ 0.2 d=0.1

# Distance from Satisfaction

similarAge(bob,ann) ∧ prefers(bob, ) → prefers(ann, )

0.8 0.9 0.2

$$d_r(I) = \max\{0, I(body) - I(head)\}$$

$$I(v_1 \wedge v_2) = \max\{0, I(v_1) + I(v_2) - 1\}$$

max{0, 0.8+0.5−1} ≤ 0.5 d=0.0

max{0, 0.8+0.5−1} ≤ 0.2 d=0.1

max{0, 0.8+0.9−1} ≤ 0.2 d=0.5

# Tasks

# Tasks

- MPE Inference

  prefers(bob, ) $\geq$ prefers(bob, ) ?

# Tasks

- MPE Inference

  prefers(bob, ) ≥ prefers(bob, ) ?

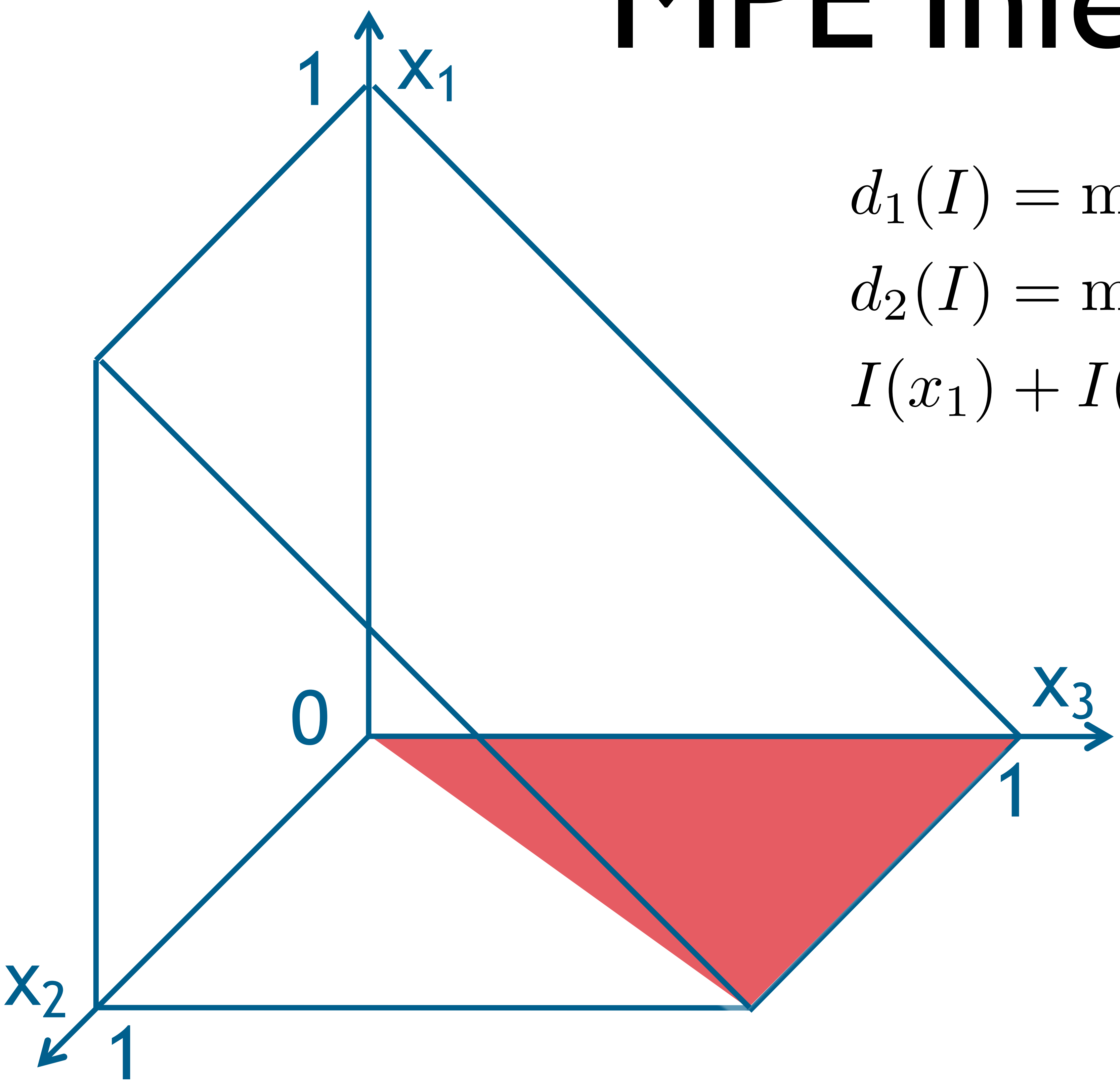
- Computing Marginals

  P(prefers(bob, ) ≥ 0.8) ?

# Tasks

- MPE Inference
  prefers(bob, ) $\geq$ prefers(bob, ) ?
- Computing Marginals
  P(prefers(bob, ) $\geq$ 0.8) ?
- Weight Learning
- Structure Learning

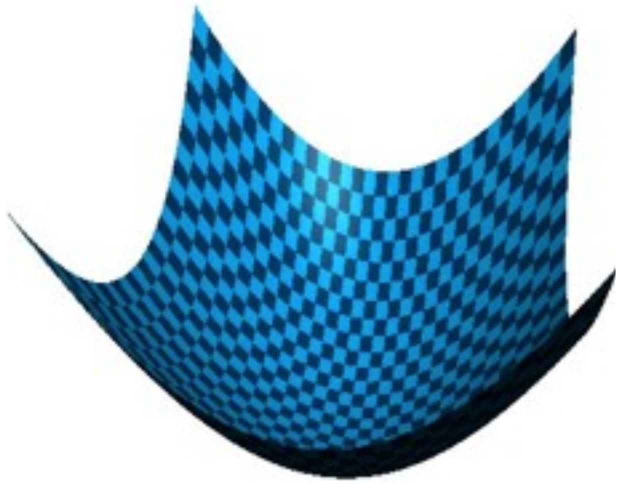
# Geometric Intuition: MPE Inference



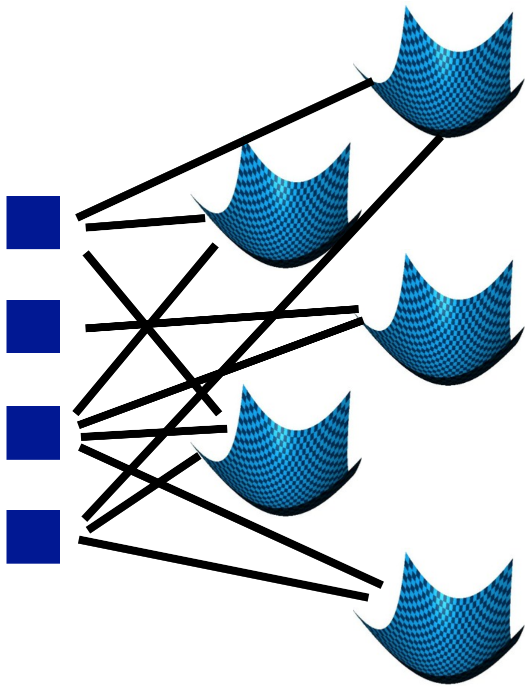

$$d_1(I) = \max\{0, I(x_1) - I(x_2)\}$$
$$d_2(I) = \max\{0, I(x_2) - I(x_3)\}$$
$$I(x_1) + I(x_3) \leq 1$$

most likely interpretations given $x_1$=0

# MPE Inference

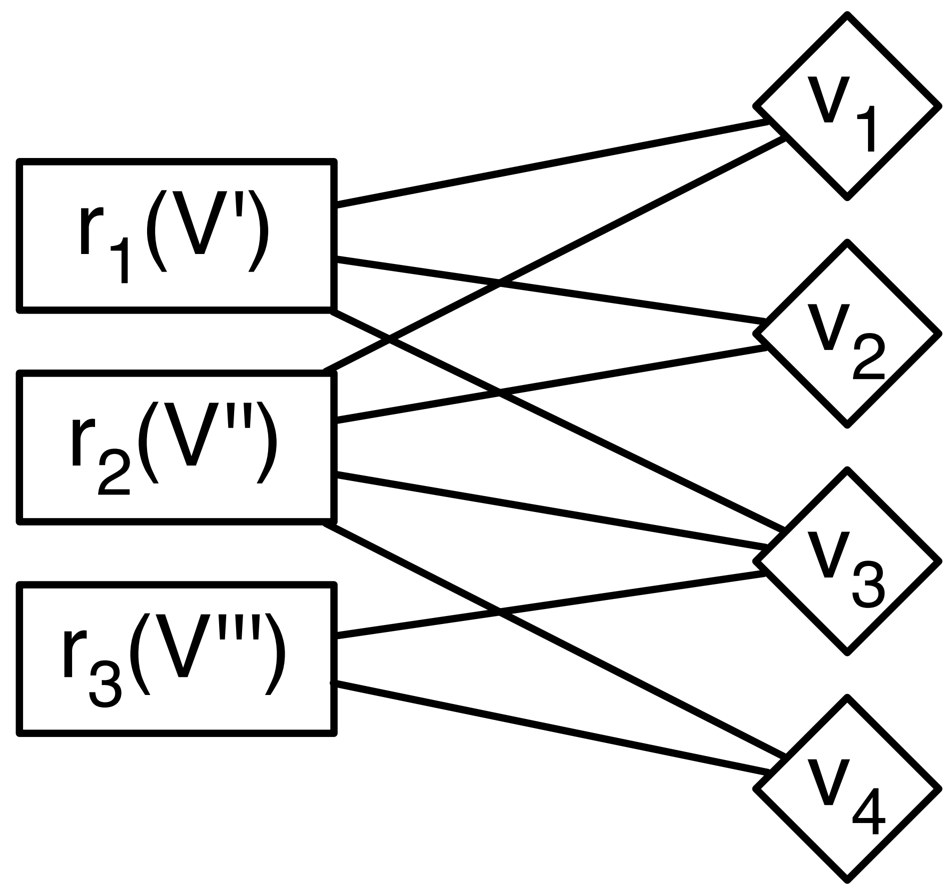


- Convex optimization problem
- New solver [Bach et al, NIPS 12]
  - Consensus optimization
  - Linear time in practice
  - Closed form solutions for subproblems



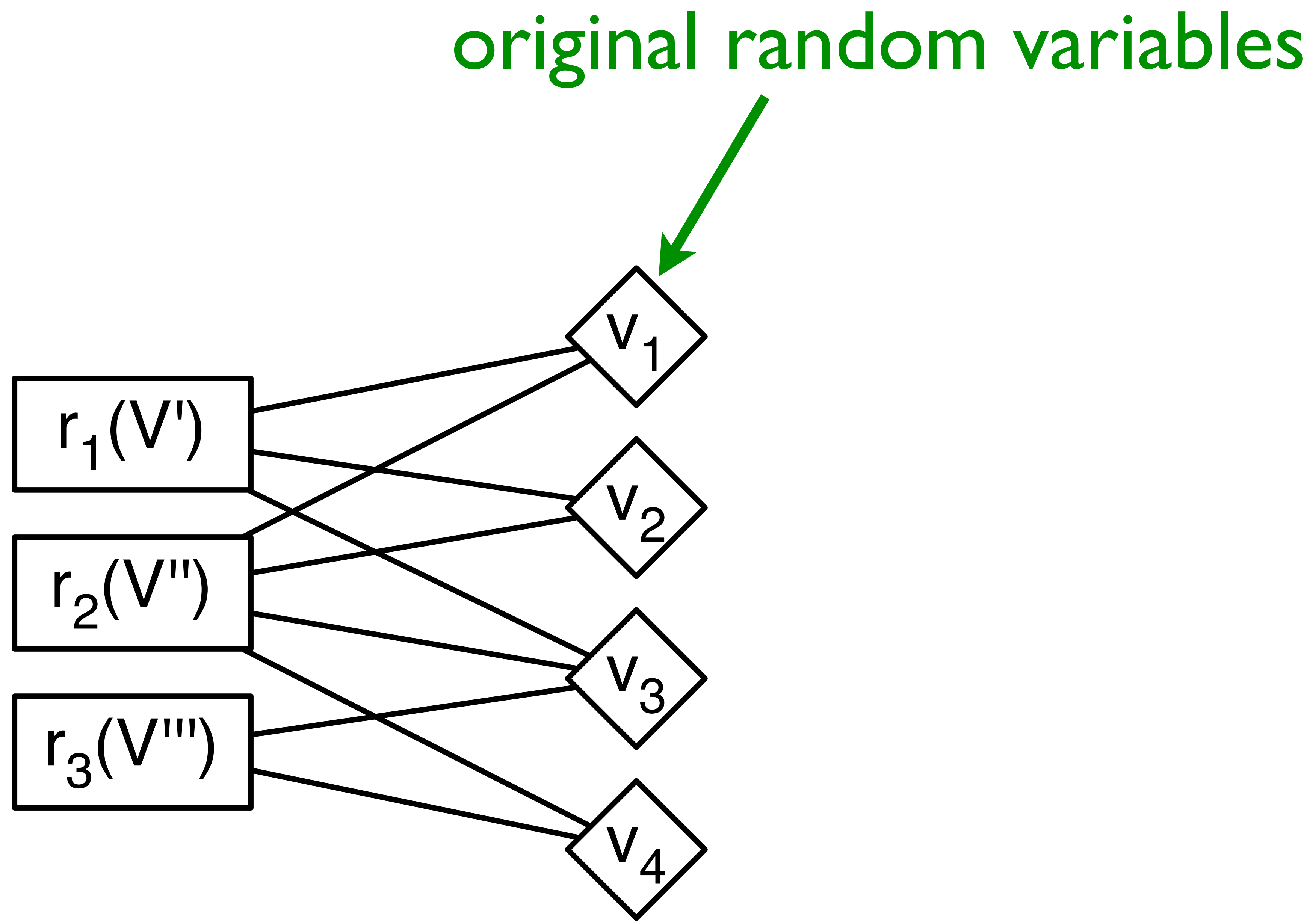
# Consensus Optimization

[Bach et al, NIPS 12]



# Consensus Optimization

[Bach et al, NIPS 12]

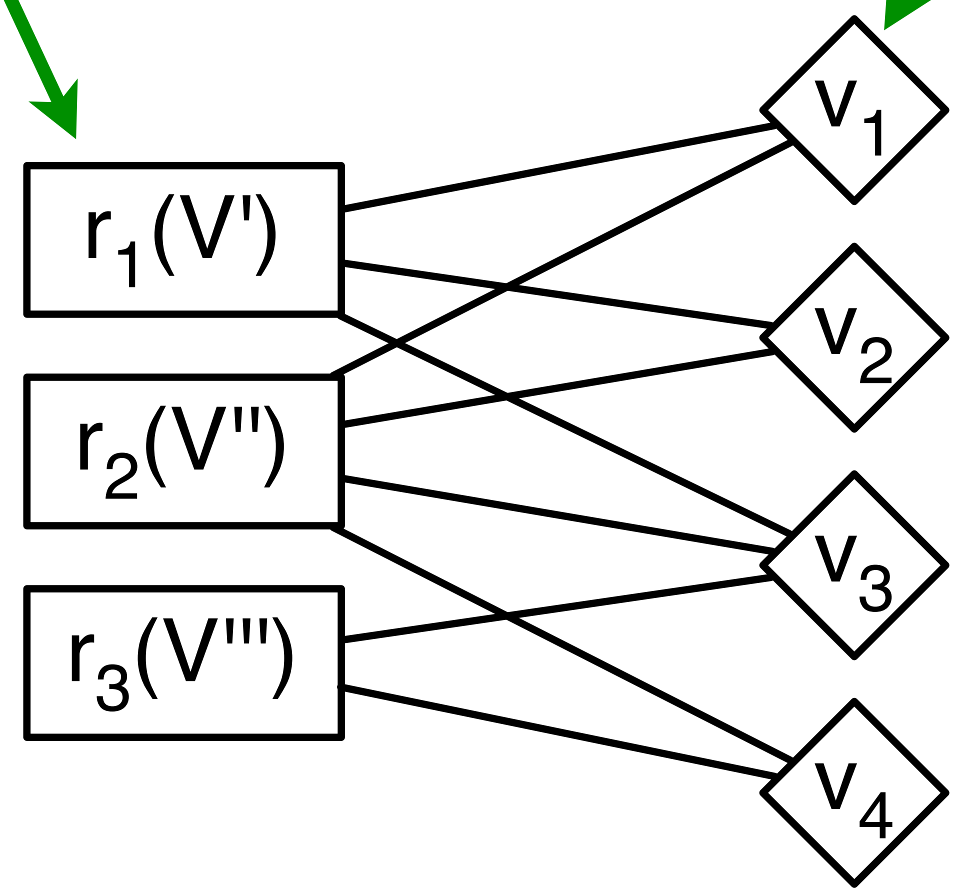


# Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables



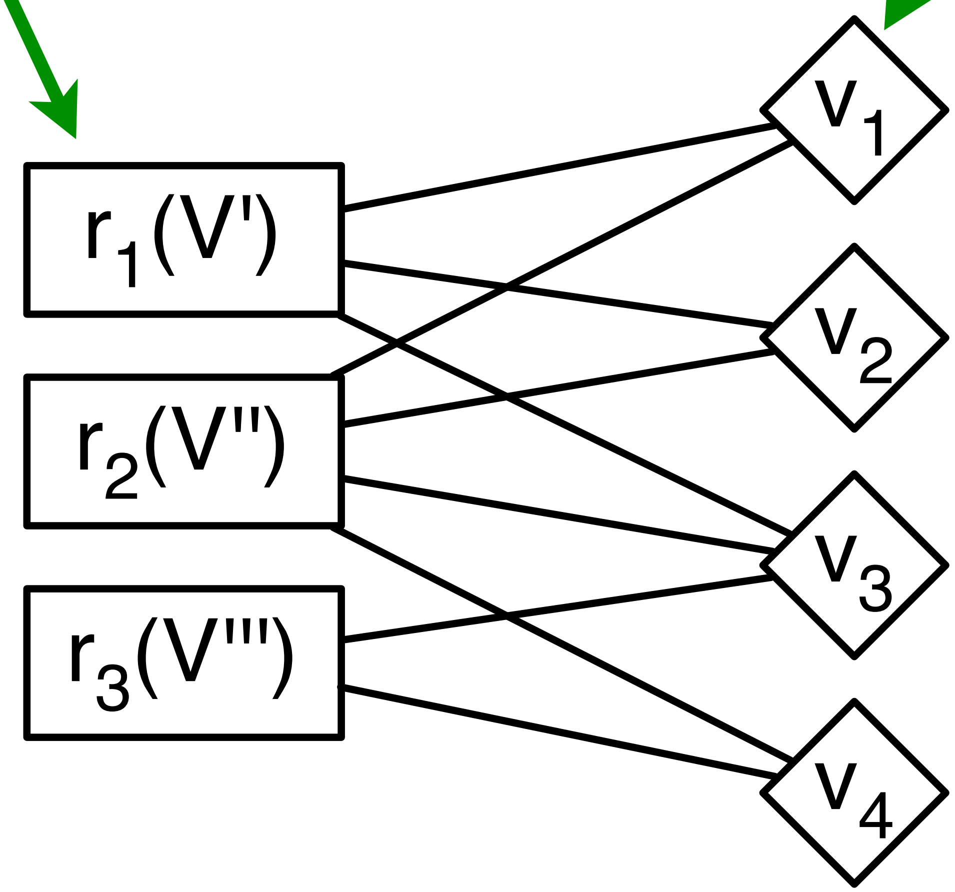
# Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables

optimize truth values & agreement with original variables per rule



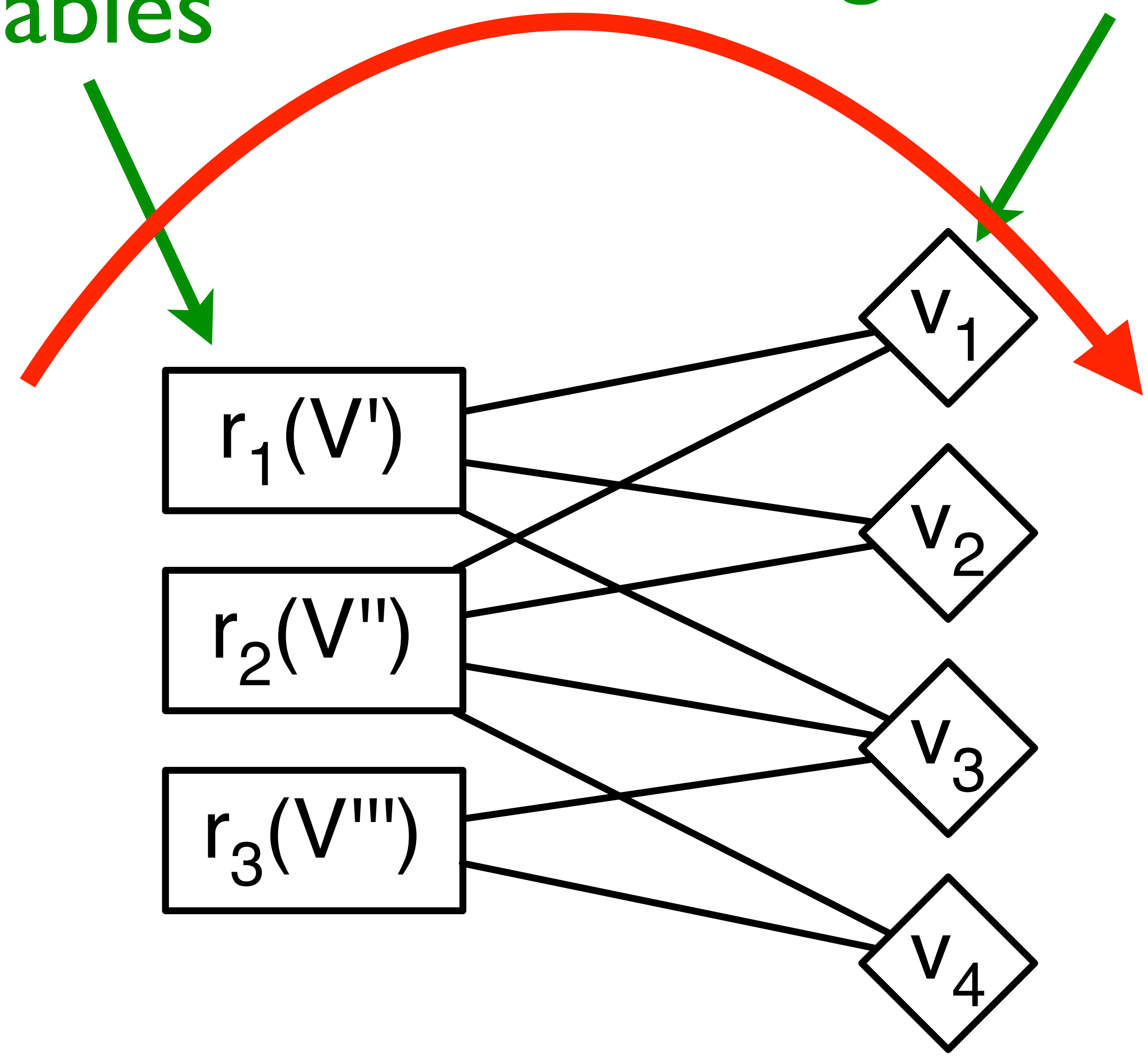
# Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables

optimize truth values & agreement with original variables per rule



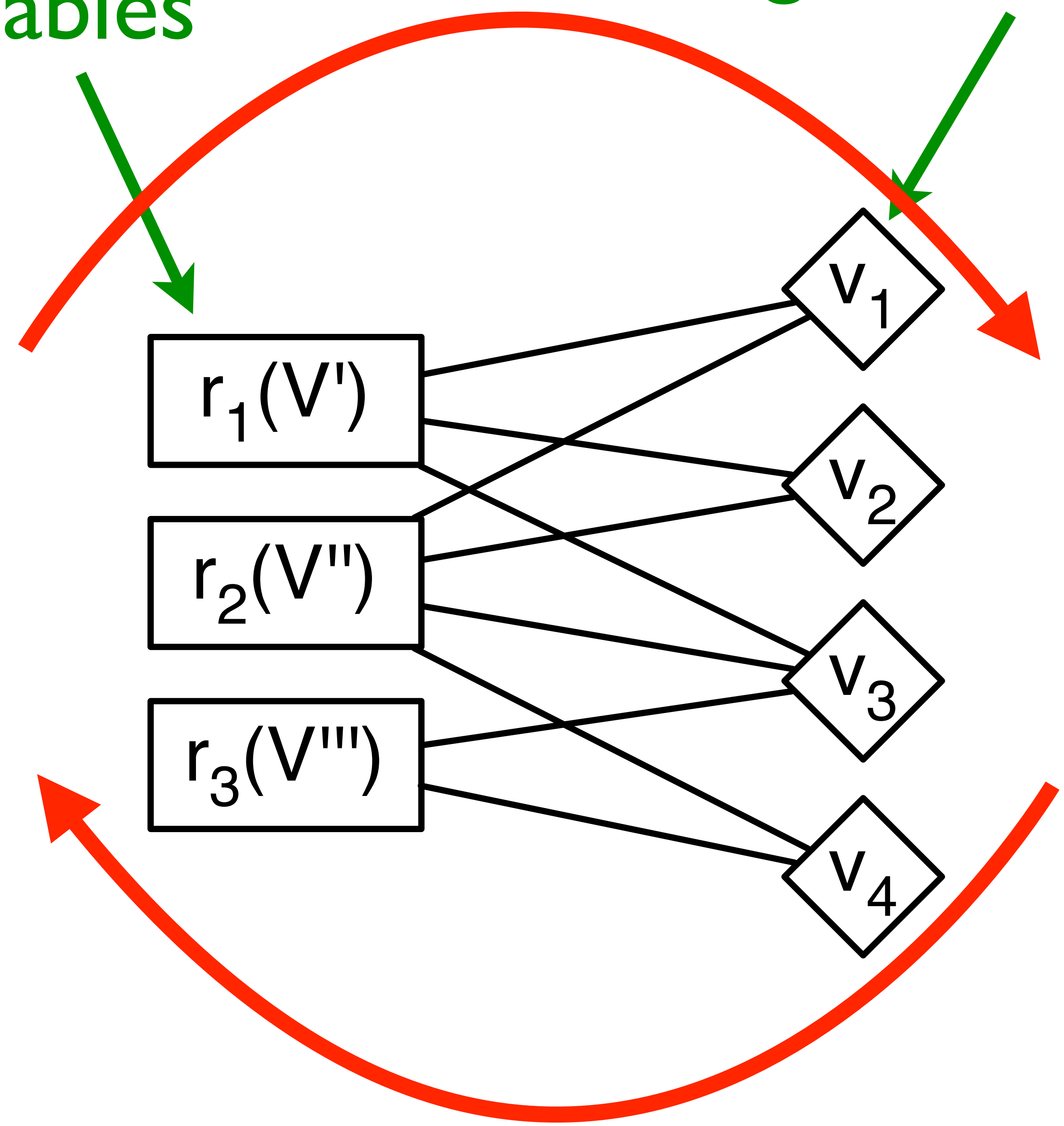

update variables to average of copies

# Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables




optimize truth values & agreement with original variables per rule

update variables to average of copies

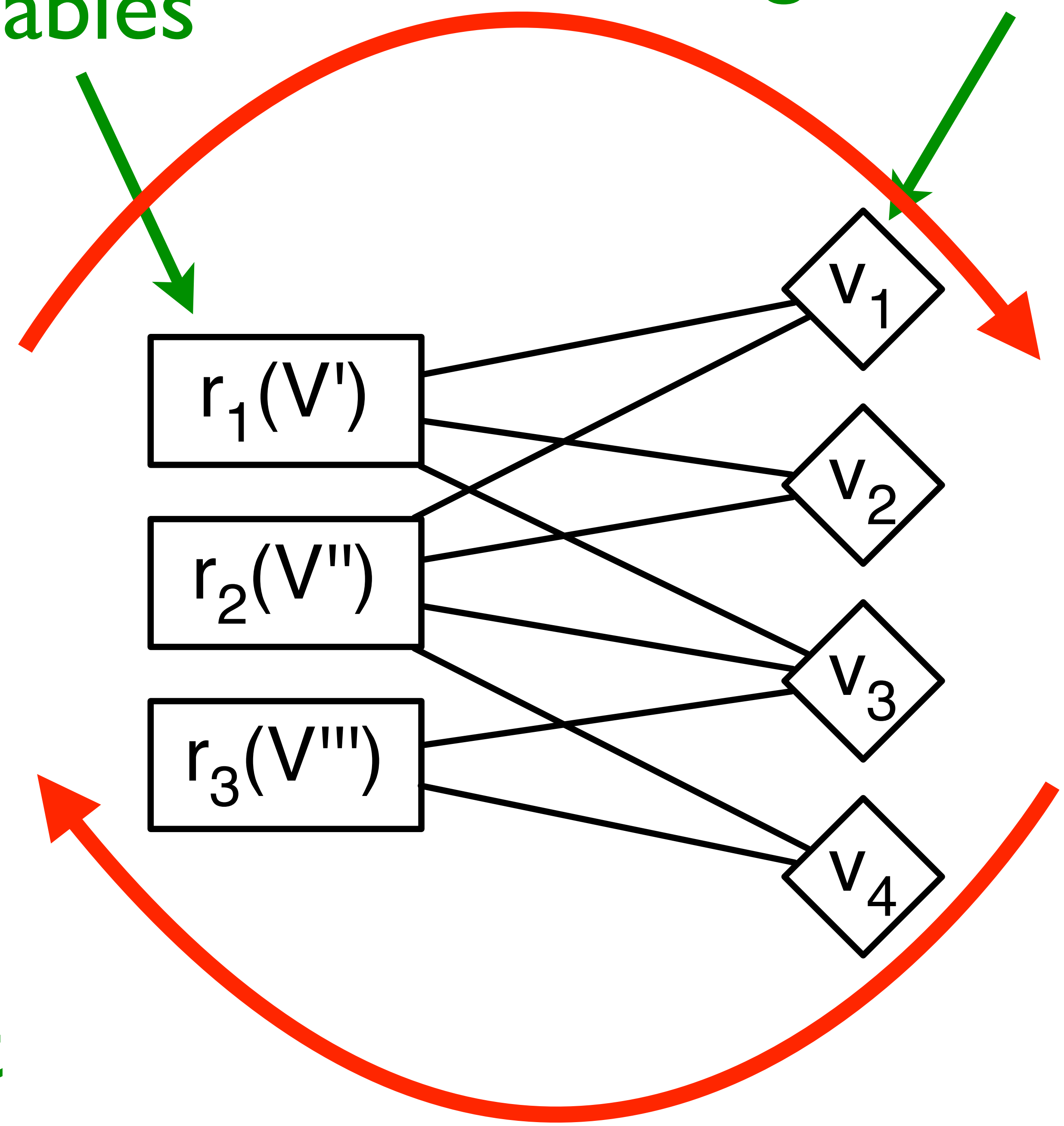
# Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables

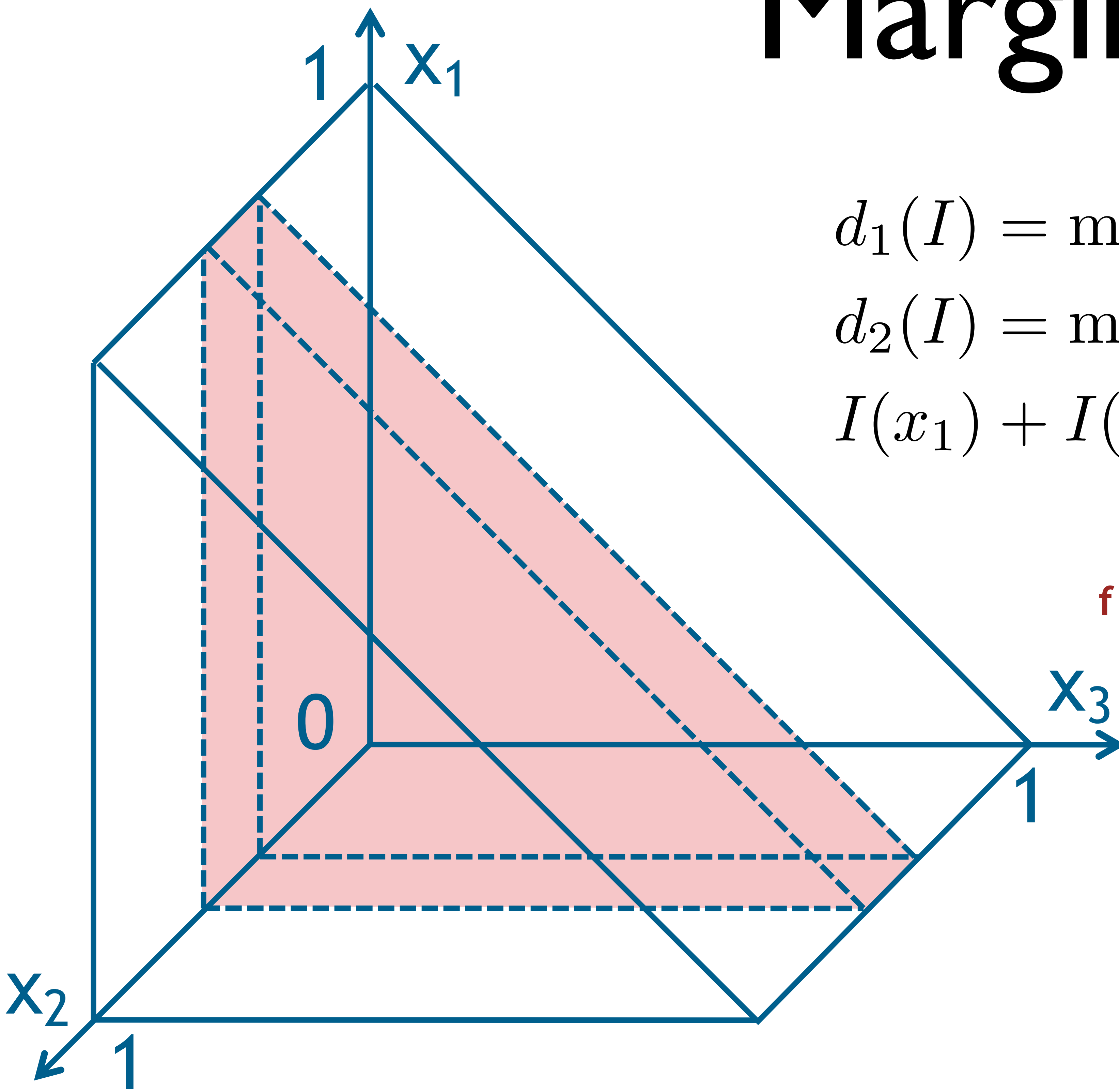
optimize truth values & agreement with original variables per rule



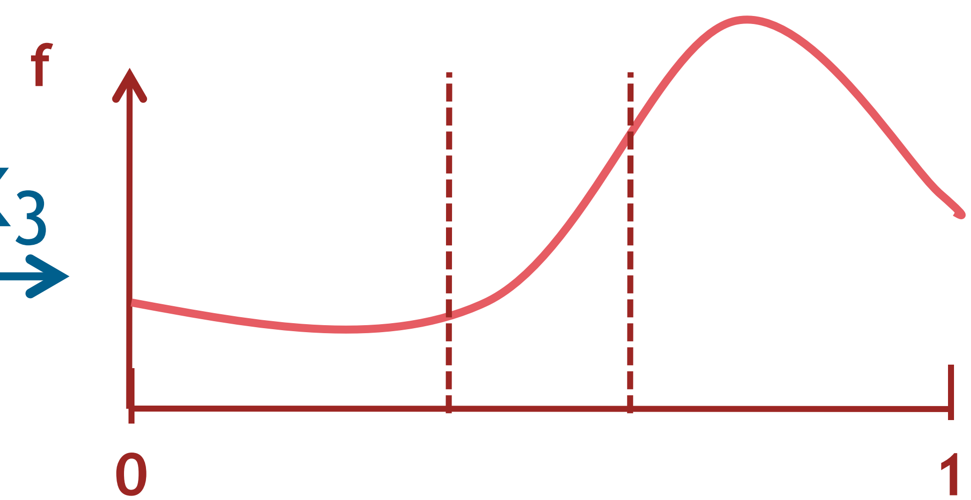

update variables to average of copies

new: fast solutions

# Geometric Intuition: Marginals




$$d_1(I) = \max\{0, I(x_1) - I(x_2)\}$$
$$d_2(I) = \max\{0, I(x_2) - I(x_3)\}$$
$$I(x_1) + I(x_3) \leq 1$$

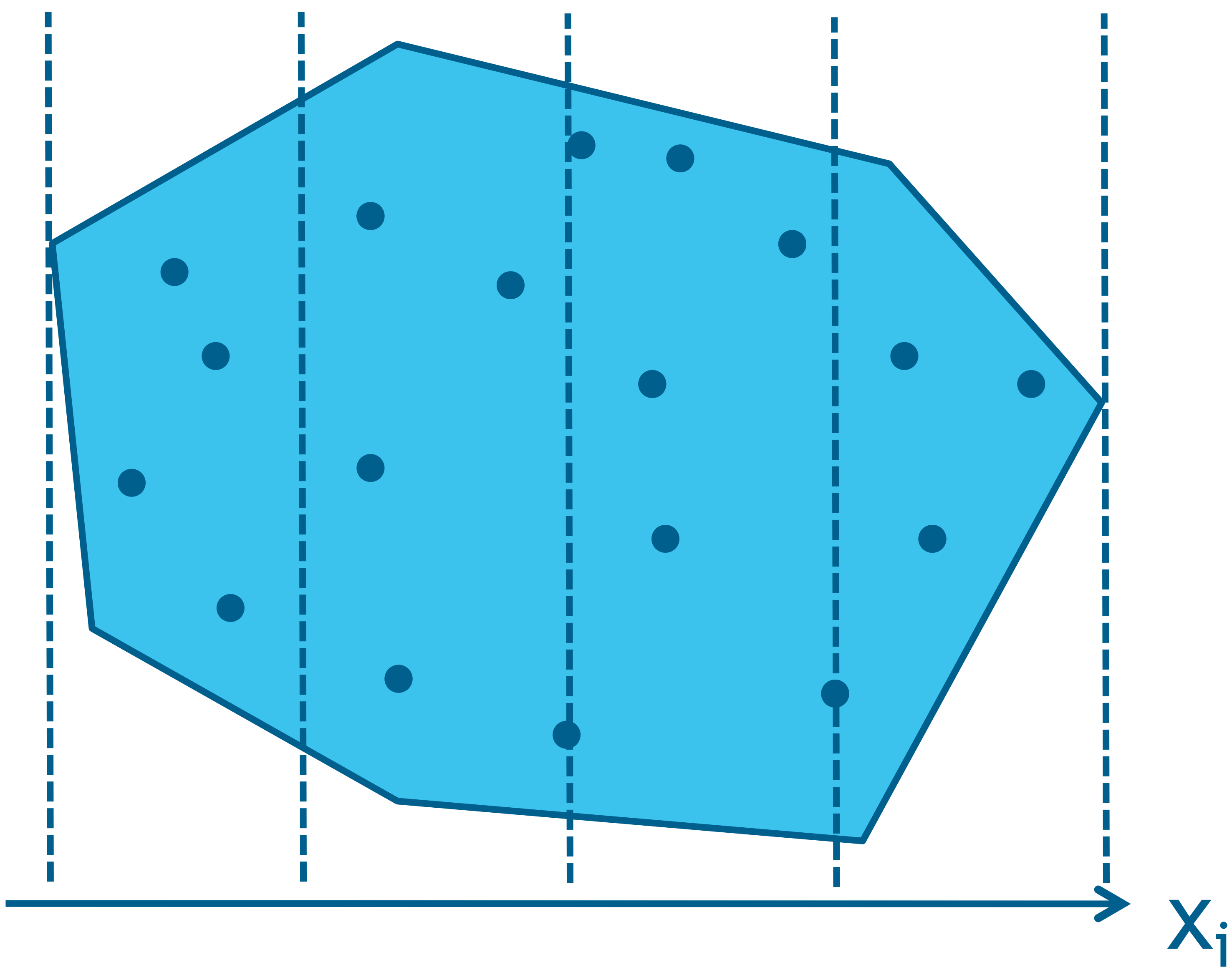



**$P(0.4 \leq x_2 \leq 0.6)$**

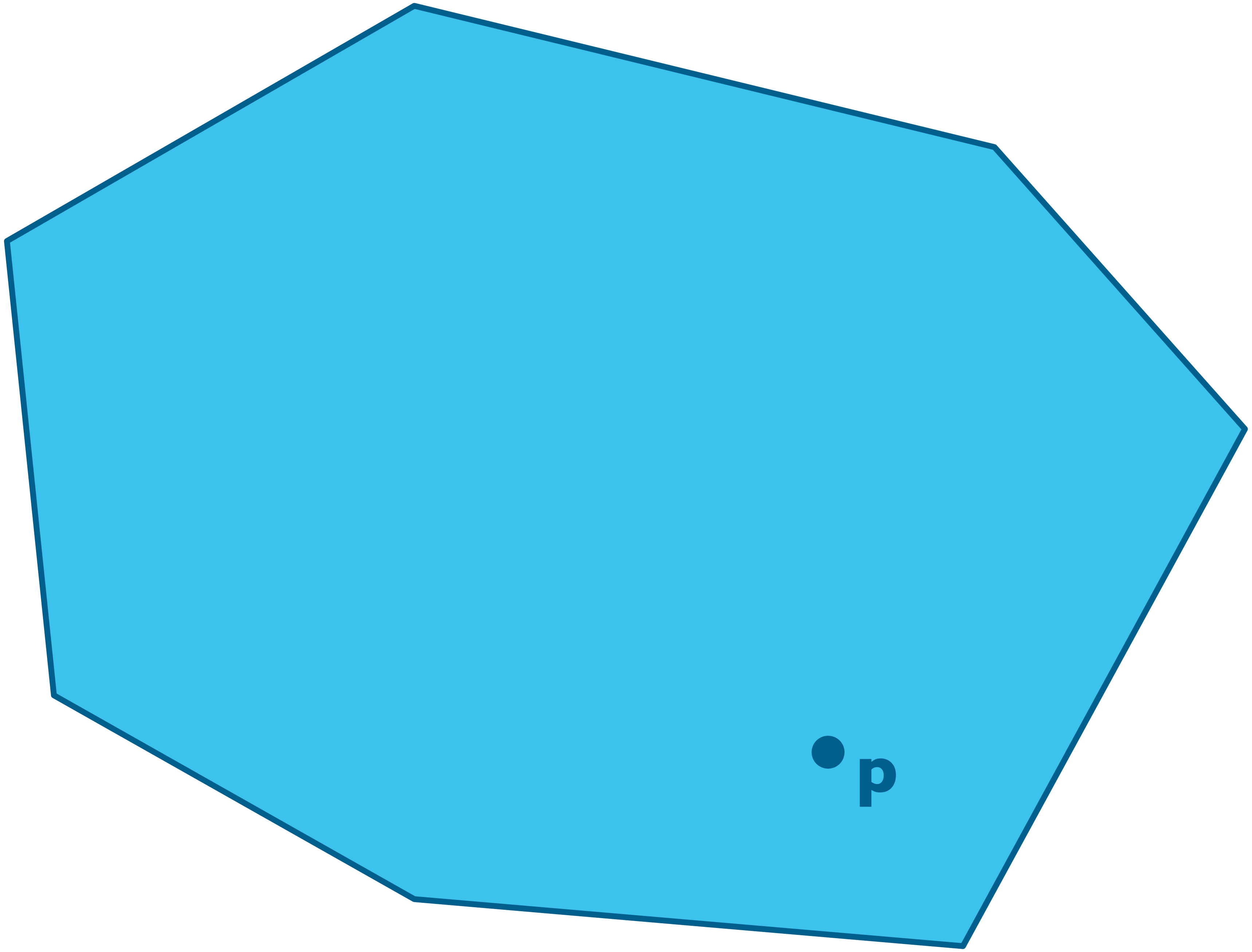
# Computing Marginals

[Broecheler and Getoor, NIPS 10]

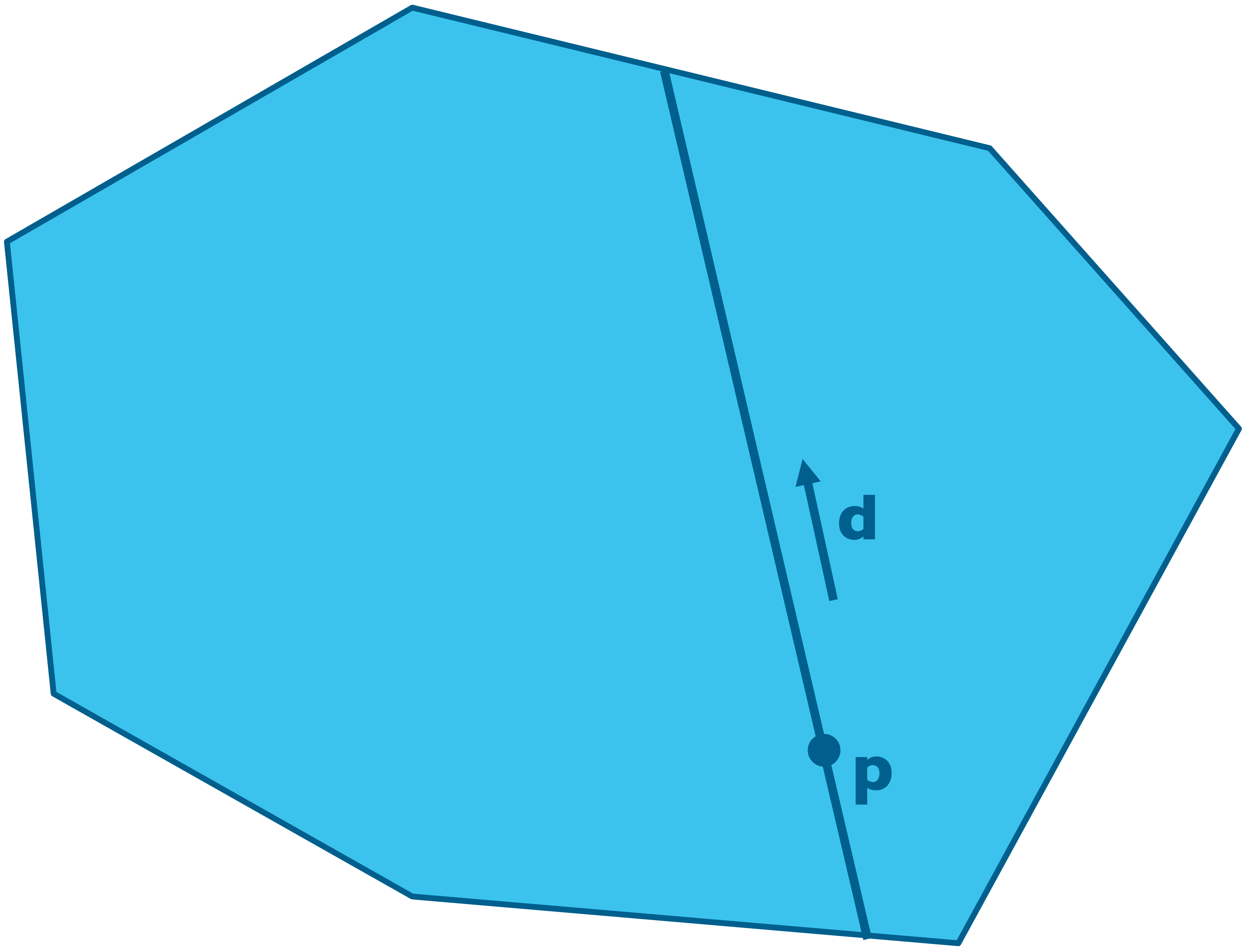
Histogram sampling
using hit-and-run Monte Carlo scheme



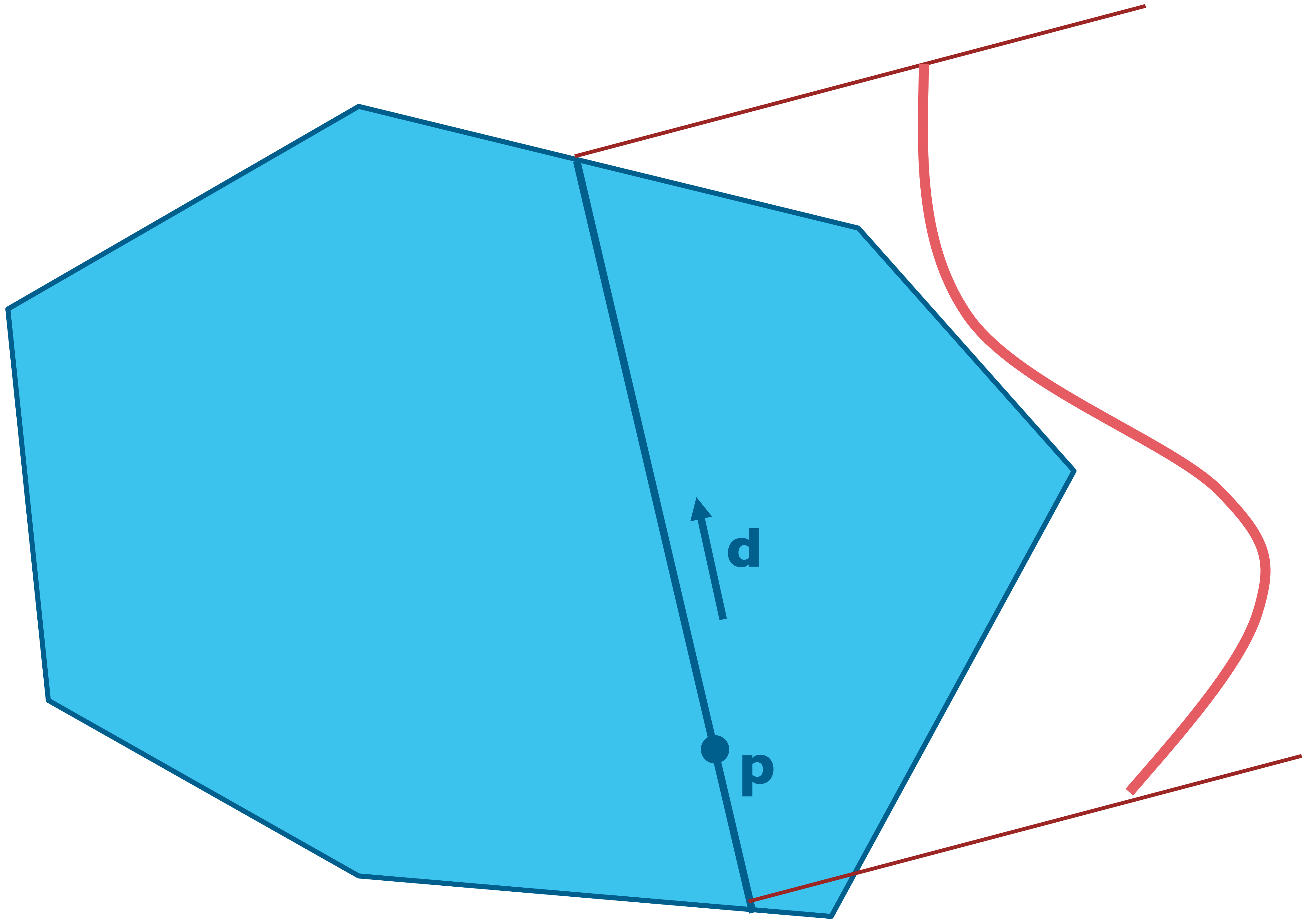
# Hit-and-Run



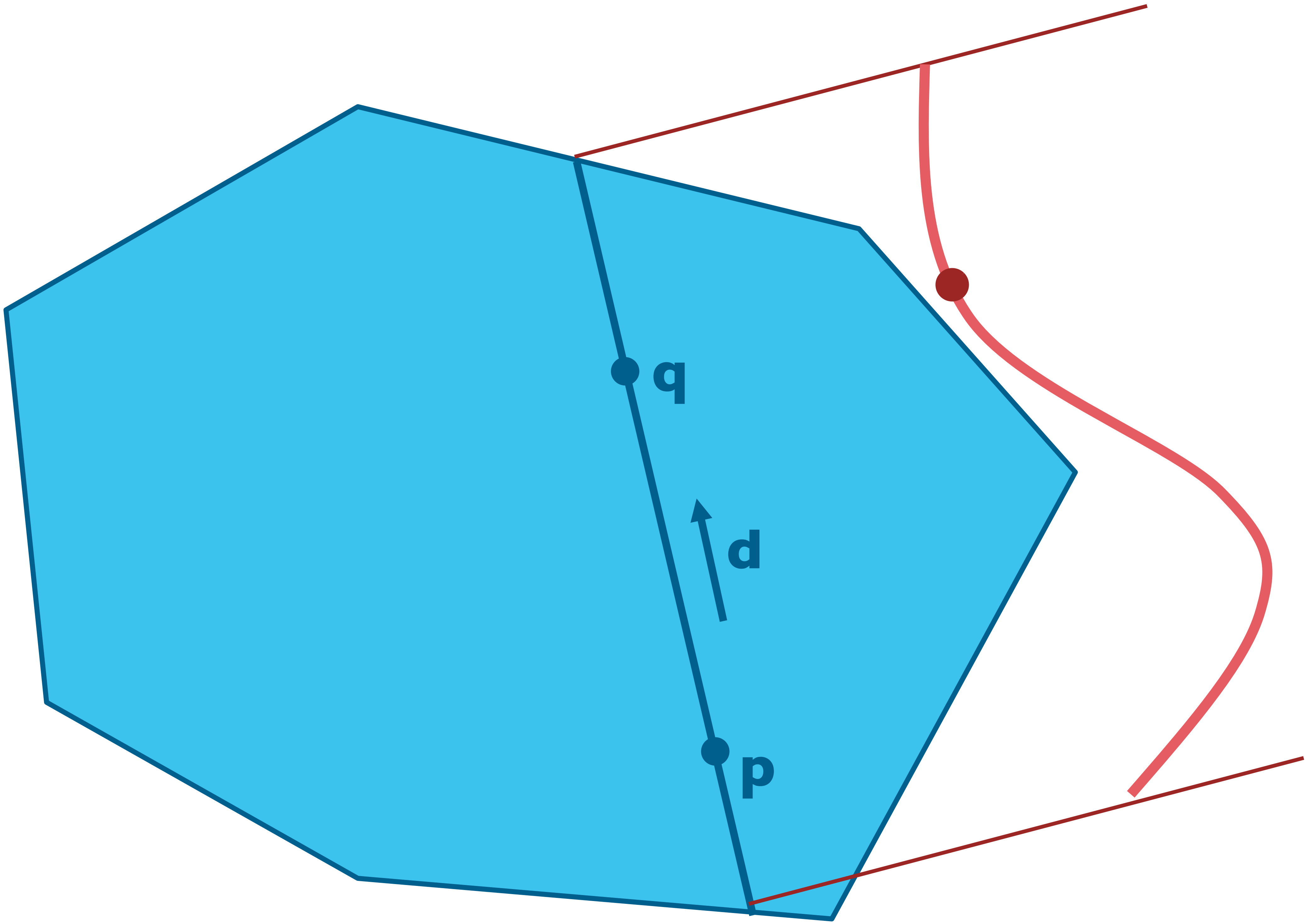
# Hit-and-Run



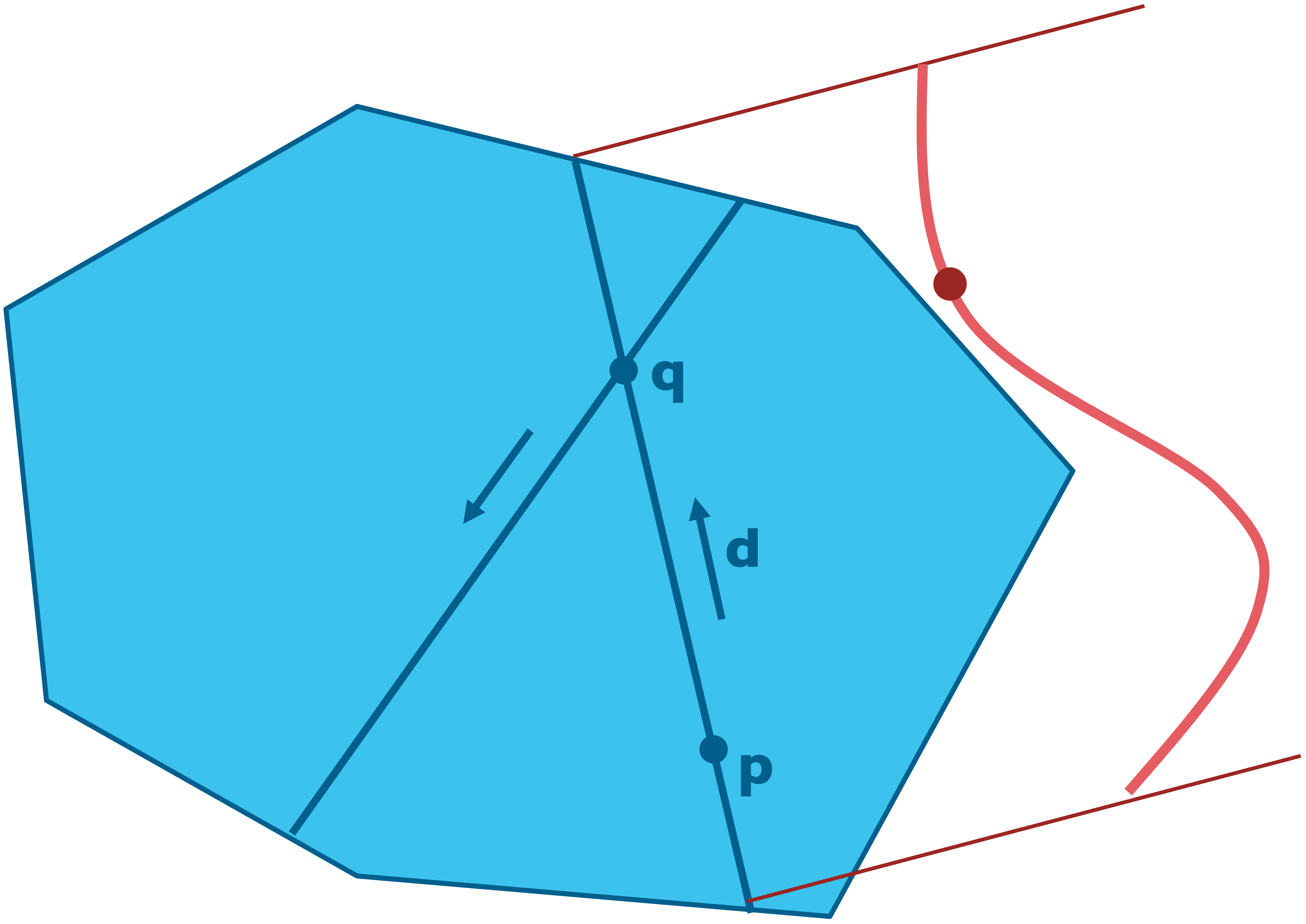
# Hit-and-Run



# Hit-and-Run



# Hit-and-Run



# Probabilistic Soft Logic (PSL)

Declarative language to specify graphical models

- Logical atoms with soft truth values in [0,1]
- Dependencies as weighted first order rules
- Support for similarity functions and aggregation
- Linear (in)equality constraints

Inference

- MPE: consensus optimization
- Marginals: hit-and-run histogram sampling

# Thank you!

http://psl.umiacs.umd.edu