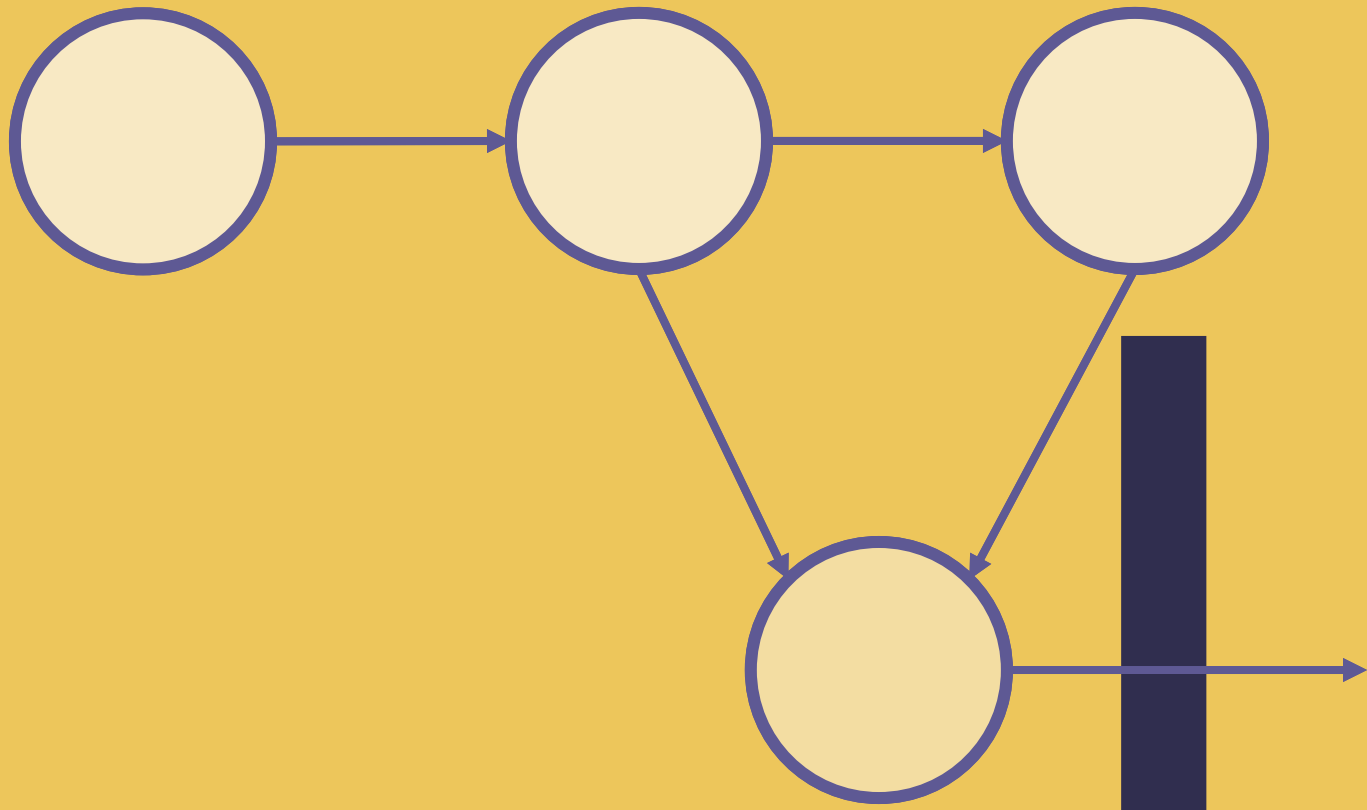
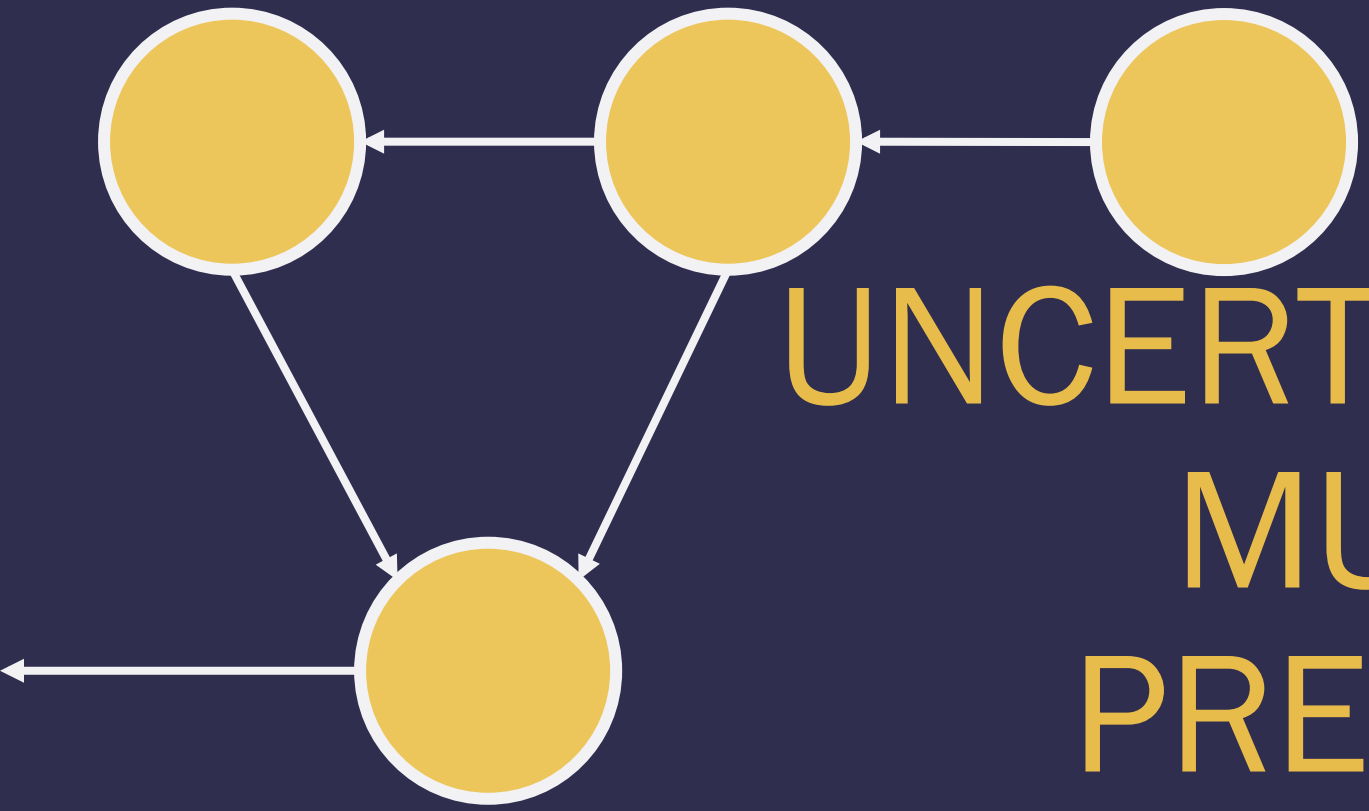


AAAI-16 Tutorial

CP-nets

- Thomas E. Allen
- Judy Goldsmith
- Francesca Rossi

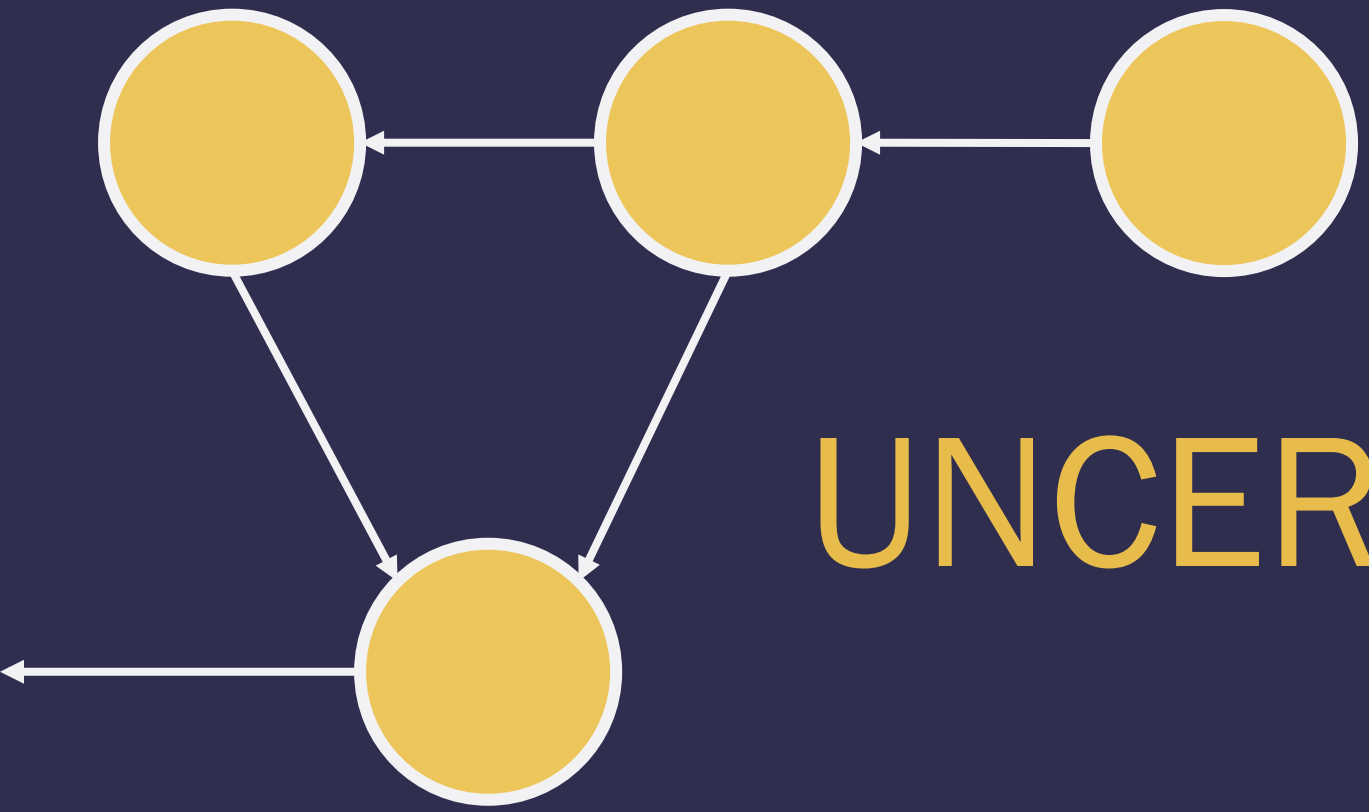


UNCERTAINTY AND MULTI-AGENT PREFERENCES

Francesca Rossi

Motivation

- Noise, uncertainty, missing data
  - ➔ probabilities in CP-nets (PCP-nets)
- Preferences coming from multiple sources
  - ➔ voting theory to aggregate several CP-nets
  - ➔ PCP-nets to model the aggregation



UNCERTAINTY IN CP-NETS

Uncertainty in CP-nets

- CP-nets require that we know exactly what we prefer
  - *Although we may leave out some items, or say we are indifferent to some of them*
- What if we are not really sure that we prefer Paris to Rome?
  - *Or we want to model the percentage of people with a certain preference?*
- CP-nets were inspired by Bayesian nets
  - *Replacing probabilities by preferences*
- To allow for uncertainty, we add probabilities back
  - *Instead of a total order, a probability distribution over all possible total orders*

PROBABILISTIC CP-NETS: PCP-NETS

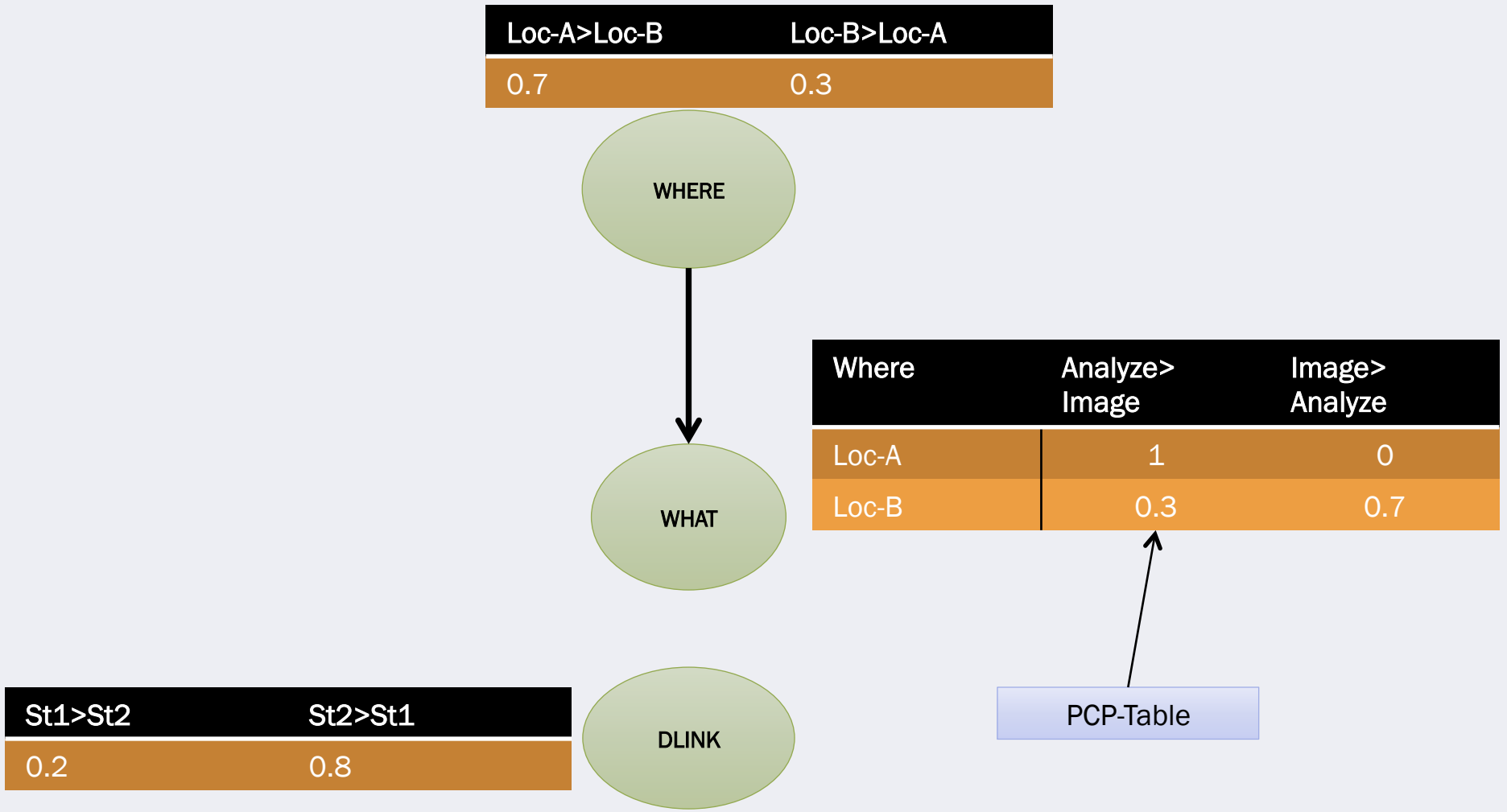
Adding probabilities to CP-nets

- Add probabilities to cp-statements
  - *associating to each feature a PCP-table indicating for each combination of values of the parents a probability distribution over the orderings on the domain of the feature*
- → Probabilistic CP nets (PCP-nets)

[Bigot, Zanuttini, Fargier, Mengin 2013]

[Cornelio, Goldsmith, Mattei, Rossi, Venable 2013]

PCP-net: example



PCP-nets and induced CP-nets

- A PCP-net models a probability distribution over a collection of CP-nets (induced CP-nets)
- Each induced CP-nets is obtained from the PCP-net by choosing
  - *a specific ordering on the domain value for each parent assignment*
- An induced CP-net will have a subset of edges of the PCP-net

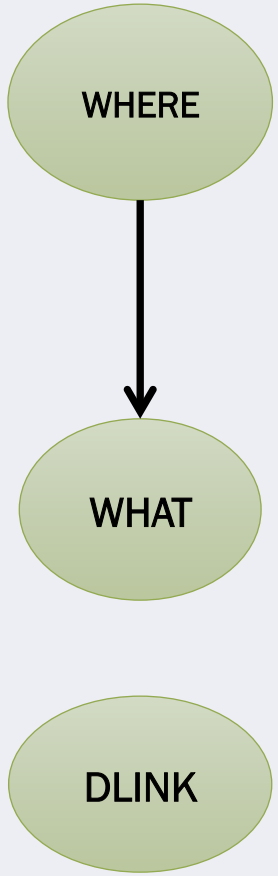
Induced CP-nets: example

PCP-net

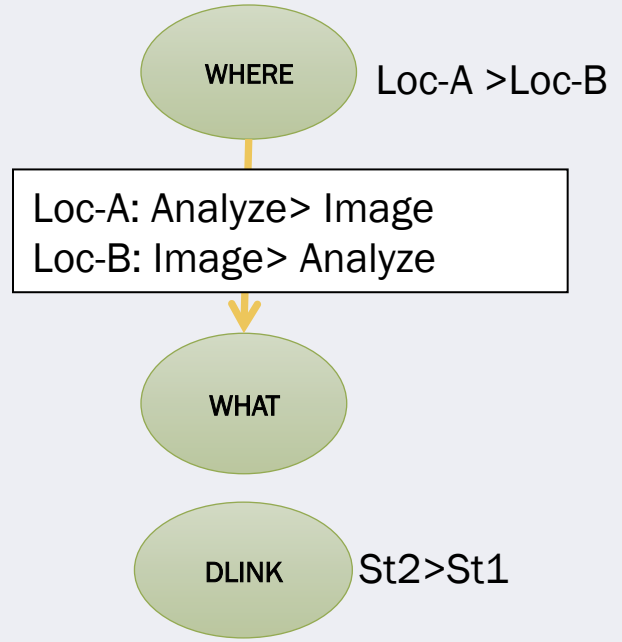
Loc-A>Loc-B	Loc-B>Loc-A
0.7	0.3

Where	Analyze> Image	Image> Analyze
Loc-A	1	0
Loc-B	0.3	0.7

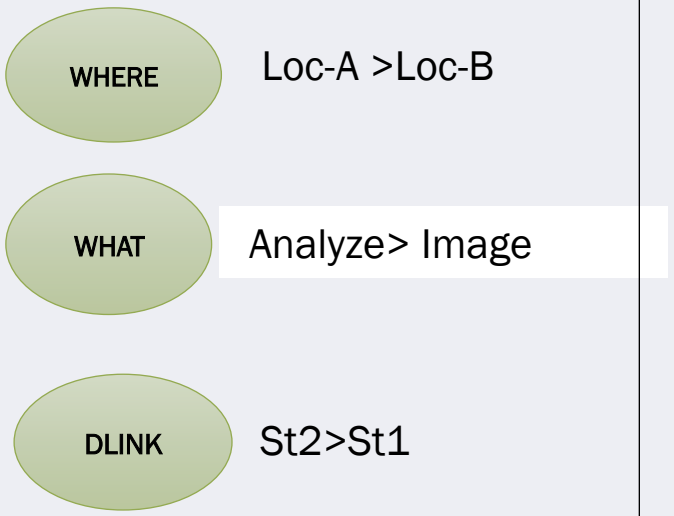
St1>St2	St2>St1
0.2	0.8



An induced CP-net



Another Induced CP-net



Probability of induced CP-nets

- Each induced CP-net has an associated probability obtained from the PCP-net by taking the product of:
  - *the probability of the chosen orderings*
- A PCP-net defines a probability distribution over the set of induced CP-nets

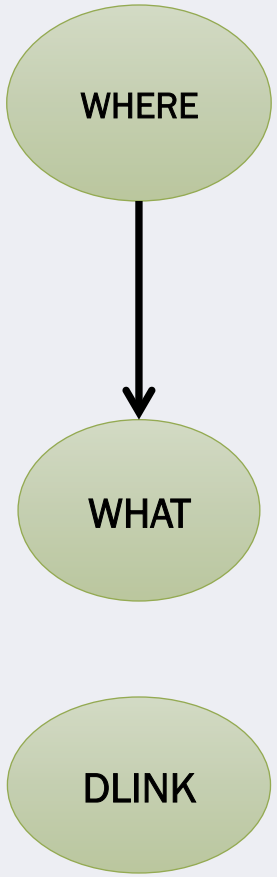
Induced CP-nets: example

PCP-net

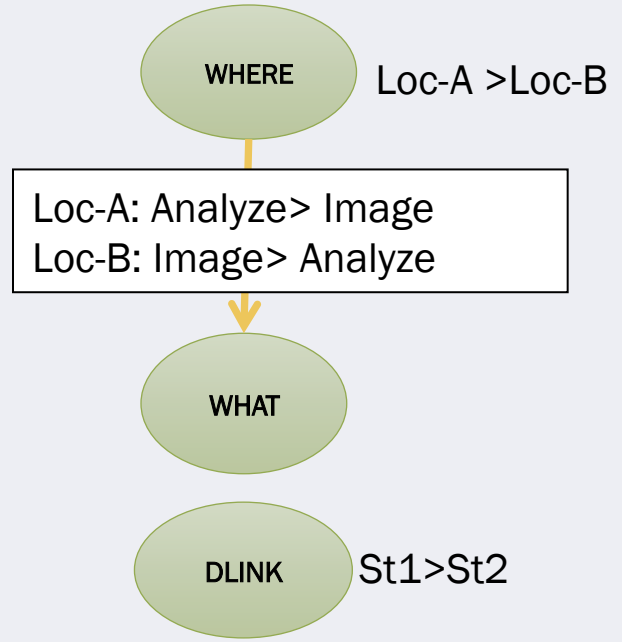
Loc-A>Loc-B	Loc-B>Loc-A
0.7	0.3

Where	Analyze> Image	Image> Analyze
Loc-A	1	0
Loc-B	0.3	0.7

St1>St2	St2>St1
0.2	0.8



Induced CP-net with prob 0.098



Induced CP-net with prob 0.042

Reasoning with PCP-nets: optimality

- What is an optimal outcome in a PCP-net?
  - *The most probable optimal outcome*
    - the outcome with the highest probability defined as the sum of the probabilities of the induced CP-nets with that outcome as optimal
  - *The optimal outcome of the most probable induced CP-net*

  (sometimes not the same)
- Both polynomial if low connectivity in the dependency graph
  - *Bounded number of parents for each feature*
  - *Bounded induced width of dependency graph*

Reasoning with PCP-nets: dominance

- Given two outcomes O1 and O2, return the probability that O1 is preferred to O2
  - *Difficult even in acyclic PCP-nets and in tree PCP-nets*

PCP-NETS FOR AGGREGATING CP-NETS

RPI, March 17, 2015

CP-net preference aggregation

- Since PCP-nets model a collection of CP-nets, they seem suitable for multi-agent settings
- Given a collection of CP-nets modelling the preferences of several agents
  - *Can we aggregate them into a single PCP-net?*
  - *Can we use probabilities to model conflicting preferences?*

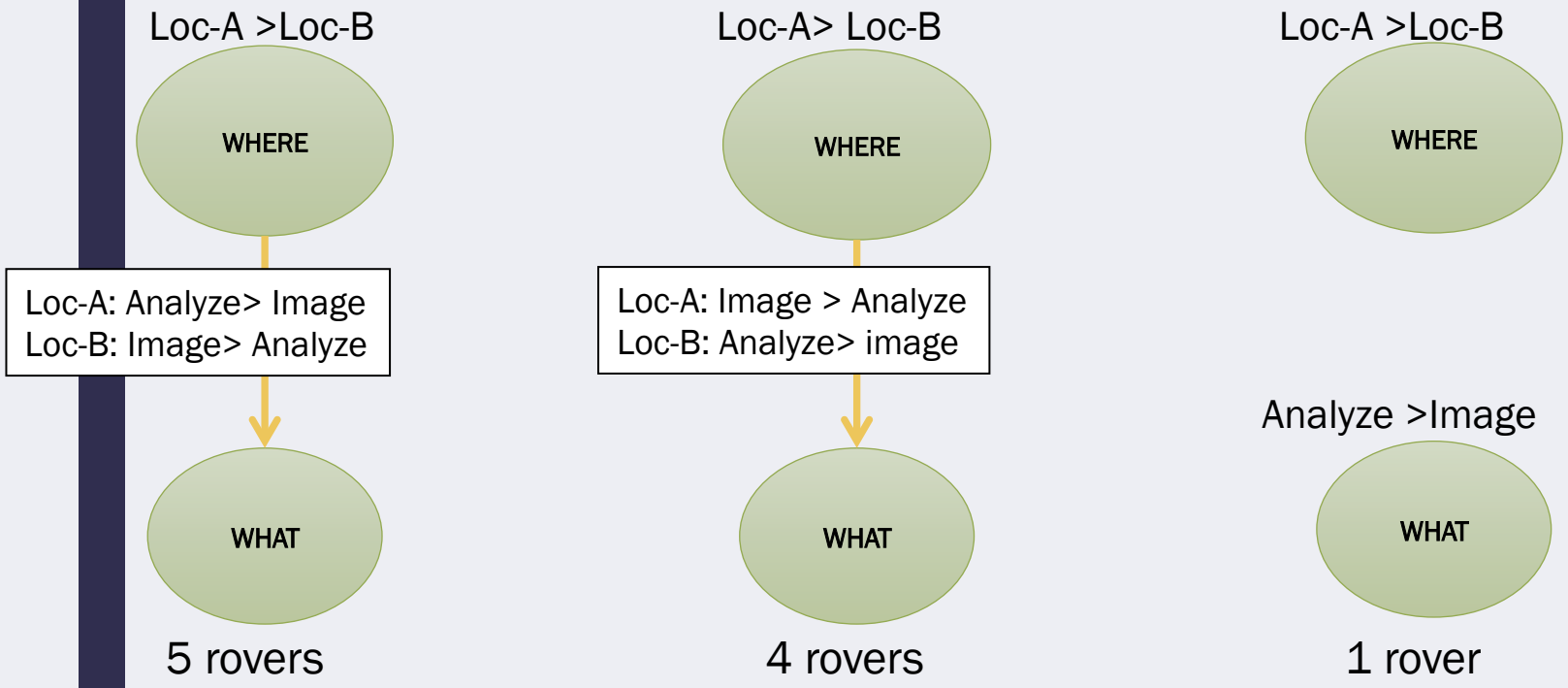
Multi-agent setting

- A collection of CP-nets
- Each agent expresses his preferences with a CP net
- All CP nets have the same features and the same domains
- Possibly different CP-tables
- Compatibility condition:
  - *There is an ordering of the features such that all dependency arcs go in the same direction*
  - *Ordering O s.t., in every CP-net, arc* (X_i,X_j) *iff* $X_i <_O X_J$

Example

10 Rovers must decide

- Where to go: Location A or Location B
- What to do: Analyze a rock or Take and image



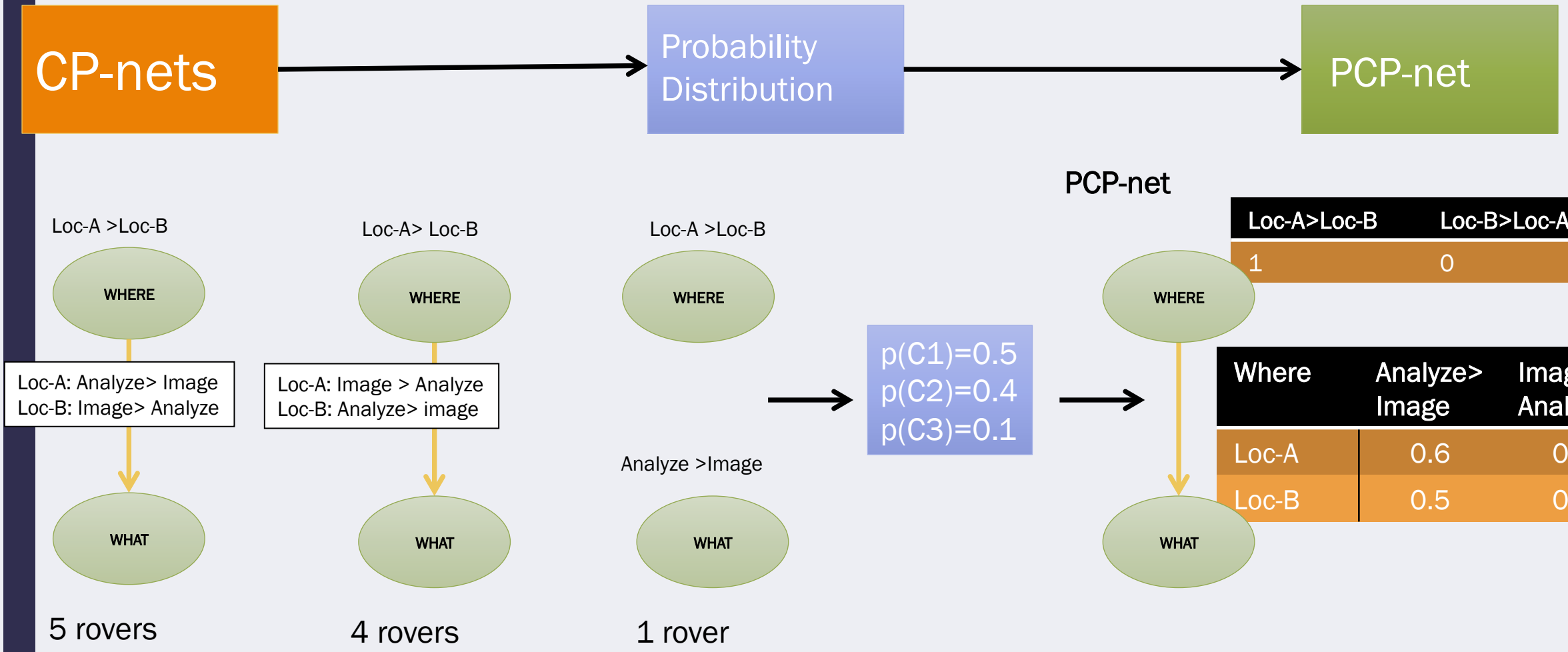
Proportion Aggregation Method

- From the CP-nets to a PCP-net which induces a probability distribution over CP-nets which approximates the given one
- PCP-net dependency graph = union of the dependency graphs of all the CP-nets
- PCP-tables defined using the probability of the CP-nets in the profile

Given feature *X* and assignment *u* to its parents:

$$P(x \succ \bar{x} \mid u) = \sum_{C_i : x \succ \bar{x} \mid u} P(C_i) \qquad P(\bar{x} \succ x \mid u) = 1 - P(x \succ \bar{x} \mid u)$$

Proportion Aggregation Method: example



Least Square Aggregation Method

- PCP-net dependency graph = union of the dependency graphs of all the CP-nets
- PCP-tables defined by minimizing the mean square error between the probability distribution induced by the PCP-net and the one in the profile

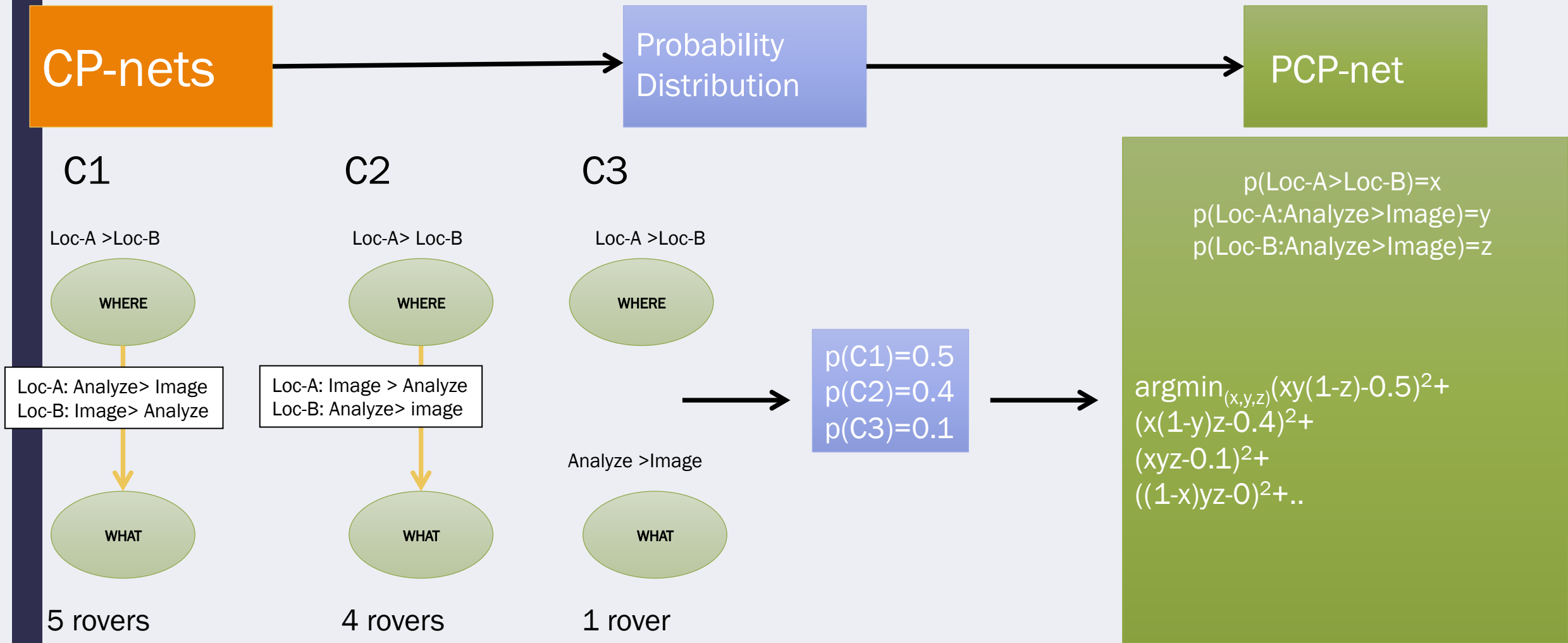
$$\operatorname{argmin}_{q \in [0,1]^r} \sum_C (f_C(q) - p(C))^2$$

Q is the vector of probabilities in the PCP-net, one for each row of the PCP-tables

Probability of CP-net C Expressed in terms of q

Probability of CP-net C In the given set

Least square method: example



Complexity of aggregation methods

- Proportion aggregation
  - *May be exponential even if in each CP-net the number of parents of a feature is bounded by a constant k*
  - *Polynomial if the maximum number of parents in the PCP-net is bounded as well by constant k'>k*
- Least square aggregation
  - *Requires considering all possible CP-nets induced by a PCP-net, which can be exponential*
  - *Can be approximated by just considering the CP-nets that have non-zero probabilities in the formula*

Which collective optimal outcome?

	Proportional aggregation	Least square Aggregation
Most probable optimal outcome	PR_O	LS_O
Optimal outcome of the most probable induced CP-net	PR_I	LS_I

Properties

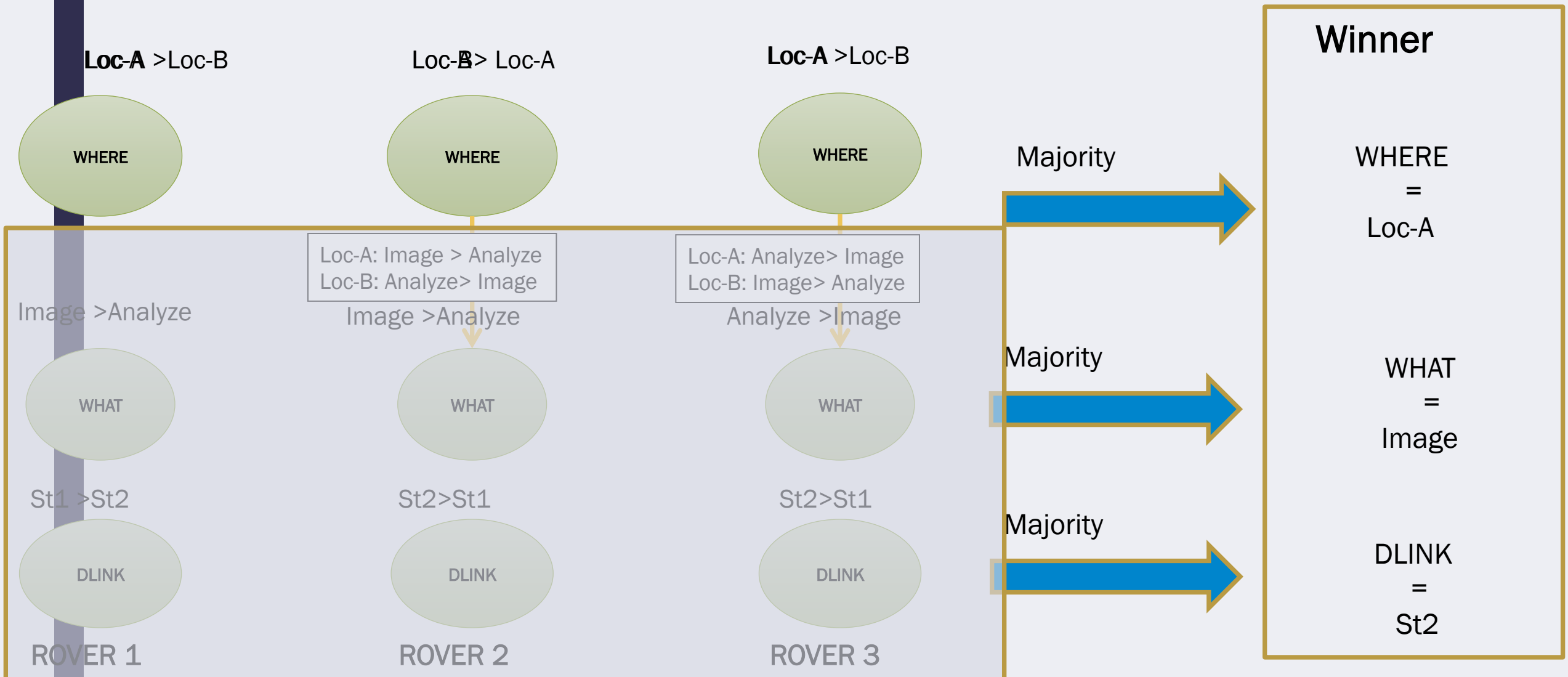
- Given any profile of CP-nets, PR_I produces the same result as sequential majority voting

- The four results can all be different
  - *There exist cases P where*
    - $PR_O(P) \neq PR_I(P)$, $LS_O(P) \neq LS_I(P)$
    - $\{PR_O(P), PR_I(P)\}$ disjoint from $\{LS_O(P), LS_I(P)\}$

Sequential Majority: example

3 Rovers must decide:

- Where to go: Location A or Location B
- What to do: Analyze a rock or Take and image
- Which station to downlink the data to: Station 1 or Station 2



Desirable properties

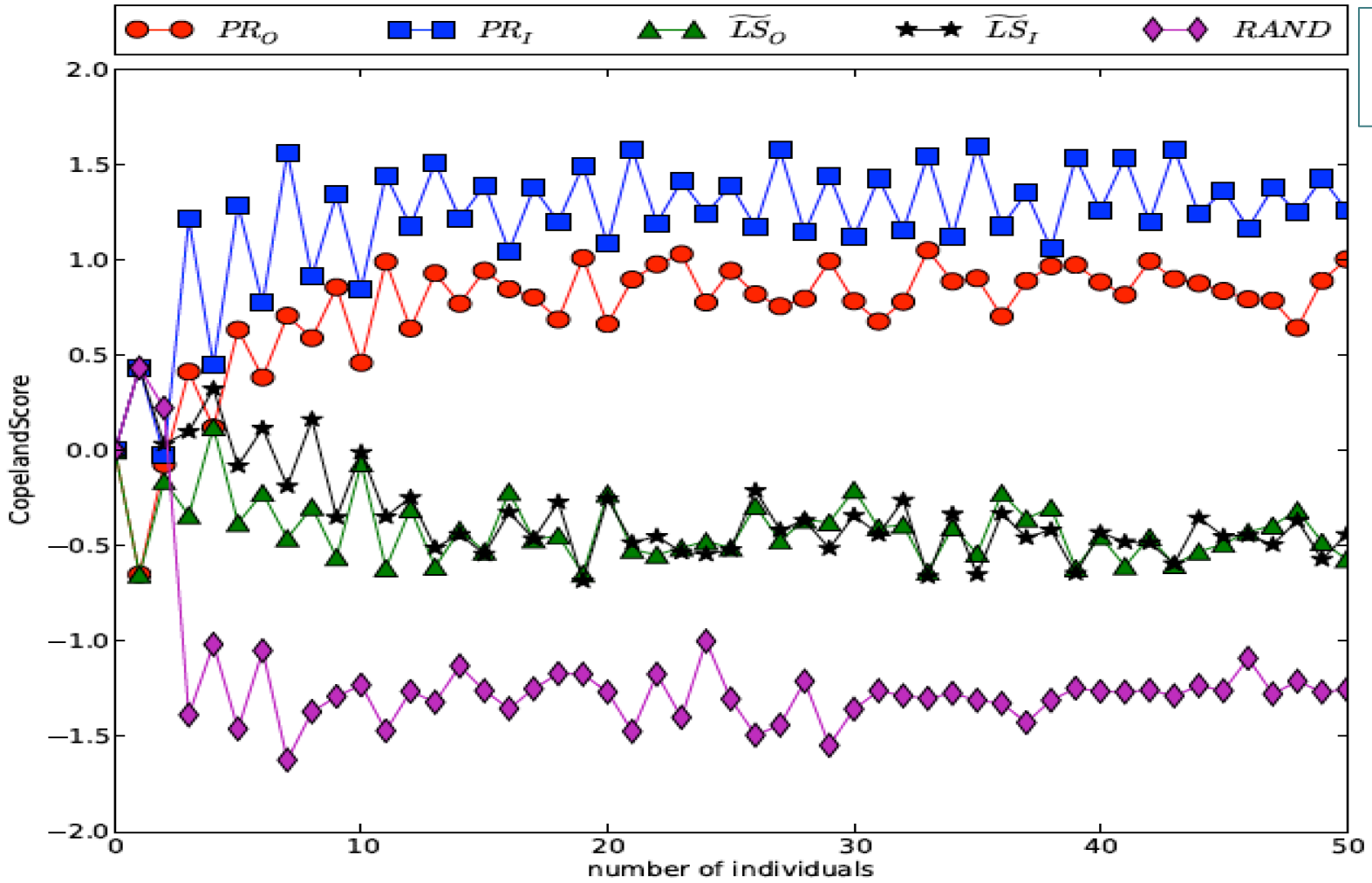
	Anonymity	Neutrality	Homogeneity	Opt-Monoton.	Consistency	Participation	Consensus
PR_O	Y	Y	Y	Y	Y	Y	Y
PR_I	Y	Y	Y	Y	Y	Y	Y
LS_O	Y	Y	Y	?	?	?	?
LS_I	Y	Y	Y	?	?	?	?

It does not help much to choose one of the four method!

Experimental Analysis

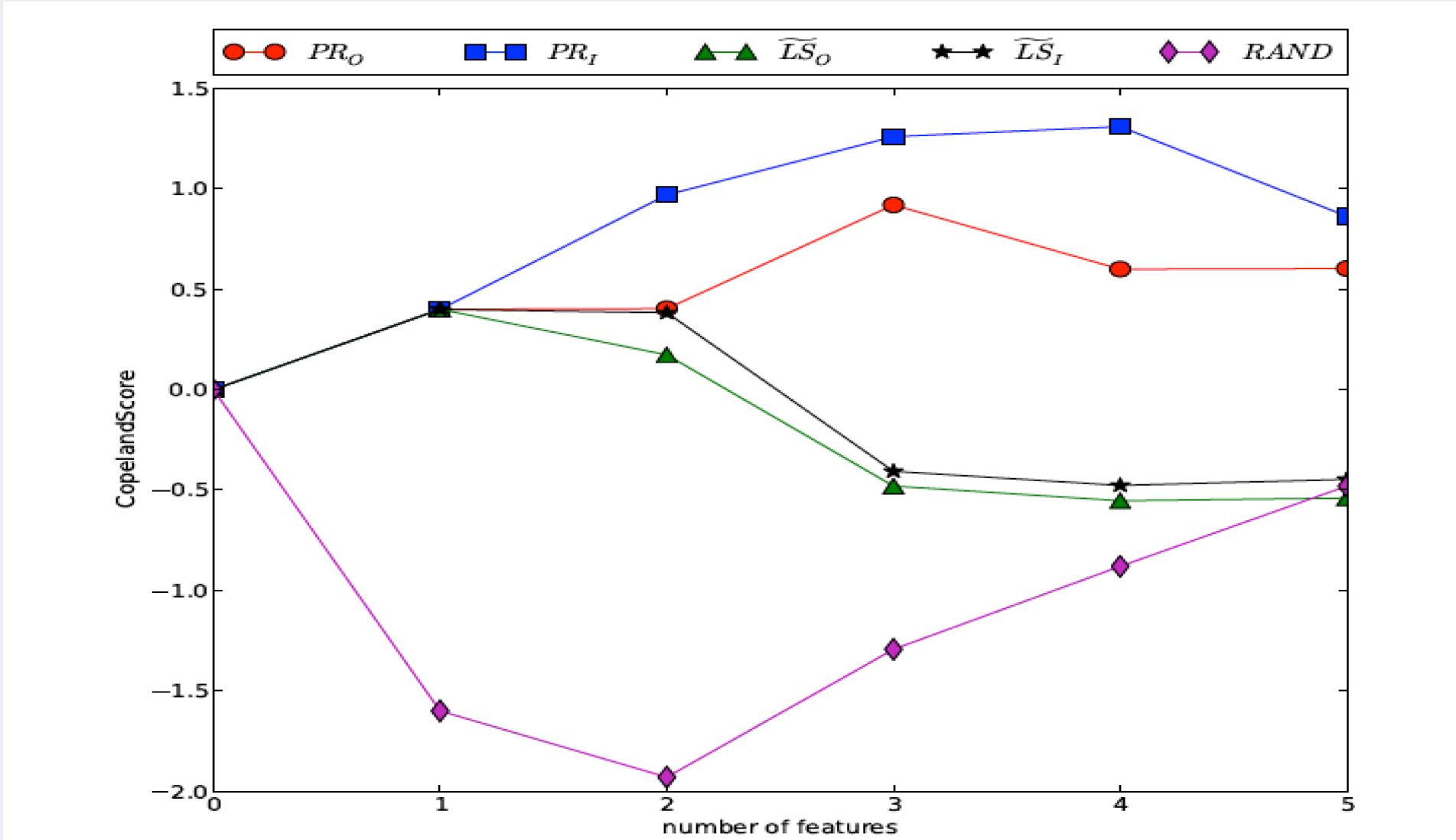
- Compute a score for each result
- Counting how many CP-nets prefer it to the others
  - *In their preference ordering*
- Averaging over all cases considered

Copeland Score, varying number of CP-nets



RAND: random result, baseline

Copeland Score, varying number of features



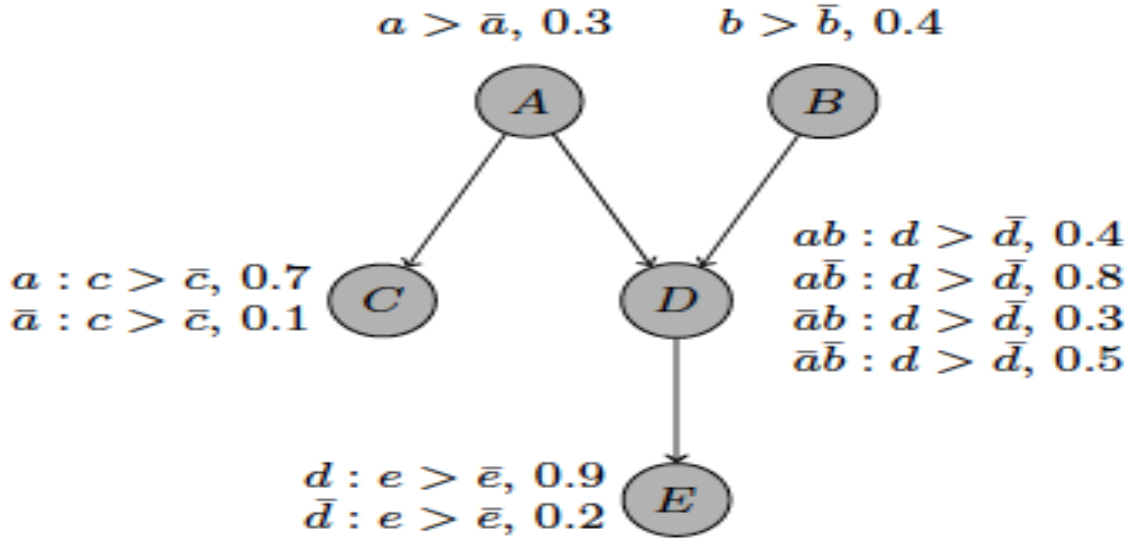
Which collective optimal outcome?

- PR_l (optimal outcome of the most probable CP-net) is the best
  - *Sequential majority* ➔ *polynomial to compute*

Dominance queries in PCP-nets

- Given O1 and O2, return the probability that O1 is preferred to O2
  - *Def.: the sum of the probabilities of the induced CP-nets where O1 is preferred to O2*
- Computing it is NP-hard
  - *Upper and a lower bound in polynomial time on polytrees*
    - Lower bound: $O(n \times 2^k)$, k max number of parents, n features
    - Upper bound: O(n)

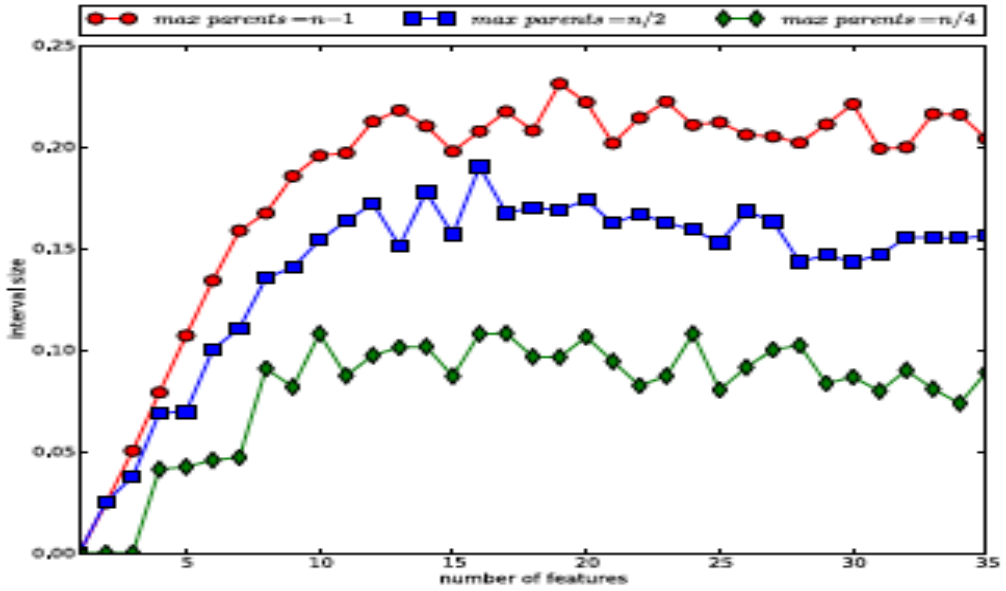
Example



- $o = (ab\bar{c}de)$
- $o' = (a\bar{b}cd\bar{e})$
- $\mathbb{P}(o > o') = ?$

Upper bound: $0.12\ (= 0.4 \cdot 0.3)$
Lower bound: $0.108\ (= 0.4 \cdot 0.3 \cdot 0.9)$

Usually very small interval



- Maximal interval size: 0.24
- Mean interval size: 0.15

- number of features $n \in [0, 35]$
- 30 PCP-net each iteration
- 30 outcomes couple for each PCP-net
- parents $\leq n/2$

How to learn a (P)CP-net?

- From examples of dominance pairs (O1 > O2)
- Active or passive learning
- Conversational recommender systems
- Special cases
  - *Ex. Lexicographic preferences*
- How to extract dominance pairs from text/blogs/tweets/ news/likes/scores/… ?

Summary

- (P)CP-nets are useful to
  - *Model the preferences of a single individual*
    - Probabilities model noise or uncertainty
  - *Represent accurately the collective preferences of a collection of CP-nets*
    - Probabilities model conflict resolution
- Optimality and dominance are computationally difficult in general, but easy under some reasonable restrictions

Future work

- Techniques to extract dominance pairs
- Eliciting methods for both CP-nets and PCP-nets
- Approaches to learn a PCP-net from dominance pairs and conditional statements
- Constrained (P)CP-nets
  - *Usually some outcomes are not available*