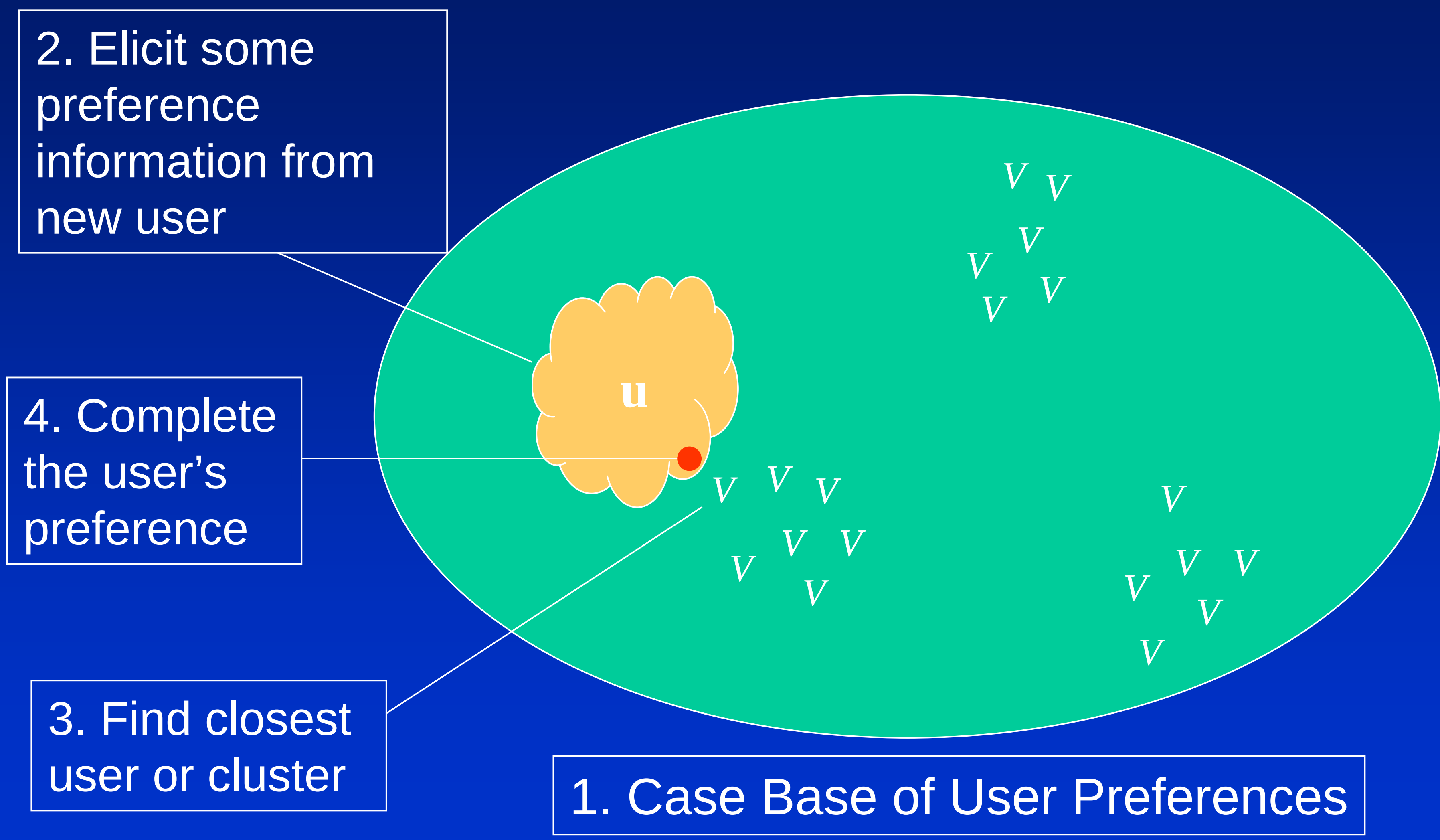
Using Prior Knowledge to Reduce the Cost of Elicitation

Peter Haddawy
Asian Institute of Technology

I. Case-Based Preference Elicitation



Distance on Preference Orders

- Need a distance measure that
  - applies to partial preferences under certainty and uncertainty
  - has nice theoretical properties
  - is computationally tractable.
- Proposed measure: *probabilistic distance*, defined as *the probability of conflict.*
- On the space of completely specified preference orders, defined as:

$$\delta(\prec_1, \prec_2) = \Pr(\prec_1 \text{ and } \prec_2 \text{ disagree wrt } a, b)$$

$$(\text{For example, } a \prec_1 b \text{ and } b \prec_2 a)$$

Probabilistic Distance

- The probabilistic distance on complete preference orders is a *metric*.

- Question: How to define it on partial preference orders?
- Observation: a partial order can be viewed as a set of complete orders consistent with it (linear extensions).

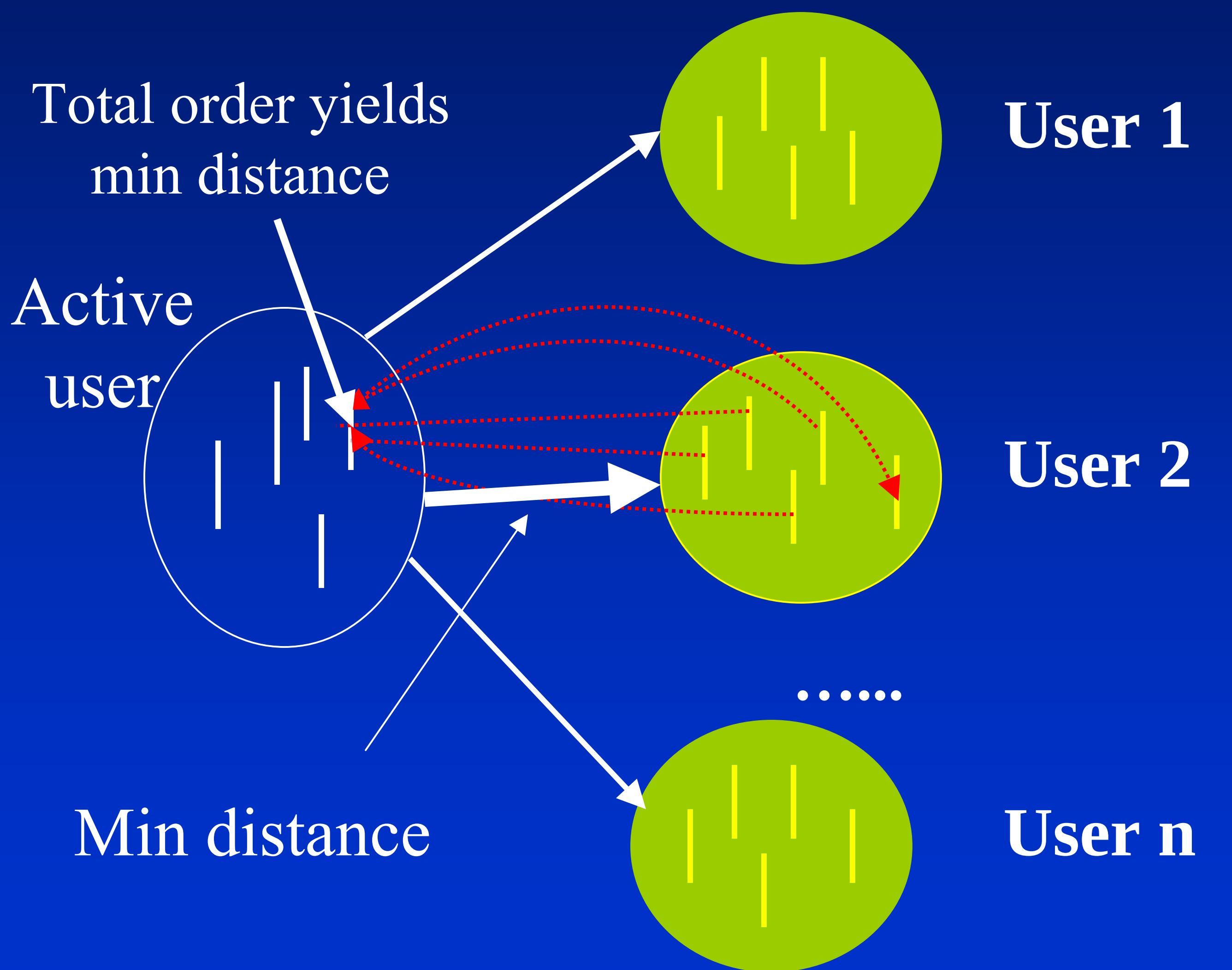
Probabilistic Distance

- The probabilistic distance on partial preference orders is defined as:

$$\delta(\prec_1, \prec_2) = \Pr(\prec_{v1} \text{ and } \prec_{v2} \text{ disagree wrt } a, b)$$
$$\text{where } \prec_{v1} \text{ is consistent with } \prec_1$$
$$\text{and } \prec_{v2} \text{ is consistent with } \prec_2$$

- Computing this distance is not straightforward: the number of extensions can be exponential.
- Use MCMC technique to sample the space of linear extensions
- Can approximate distance in polynomial time

Computing Probabilistic Distance



II. Preference Elicitation via Theory Refinement

- Would like to be able to use assumptions to guide but not constrain the elicitation process

- Theory refinement
  - Make assumptions that at least approximately apply to a large segment of users
    - Direct flights
    - Faster CPU
    - Better gas mileage
  - Correct for inaccuracies in assumptions when we encounter an individual to whom they do not apply

KBANN

- Domain theory:
  - acyclic set of propositional Horn-clauses
- Encoded in a back-propagation neural network.
- Network trained with examples that may not be consistent with the domain theory
- Domain theory supplemented and corrected

Choosing a Flight

- Find a flight from Bangkok to Boston
- Three attributes: time (T), cost(C), and whether or not a layover is involved (L).
- Assume
  - Attributes are mutually preferentially independent
  - People prefer shorter flights, cheaper flights, and flights without layovers

Encoding the Assumptions

- Dominance:

$$C_1 < C_2 \ \&\ T_1 \leq T_2 \ \&\ L_1 \leq L_2 \ \Rightarrow\ F_1 \succ F_2$$

$$C_1 \leq C_2 \ \&\ T_1 < T_2 \ \&\ L_1 \leq L_2 \ \Rightarrow\ F_1 \succ F_2$$

$$C_1 \leq C_2 \ \&\ T_1 \leq T_2 \ \&\ L_1 < L_2 \ \Rightarrow\ F_1 \succ F_2$$

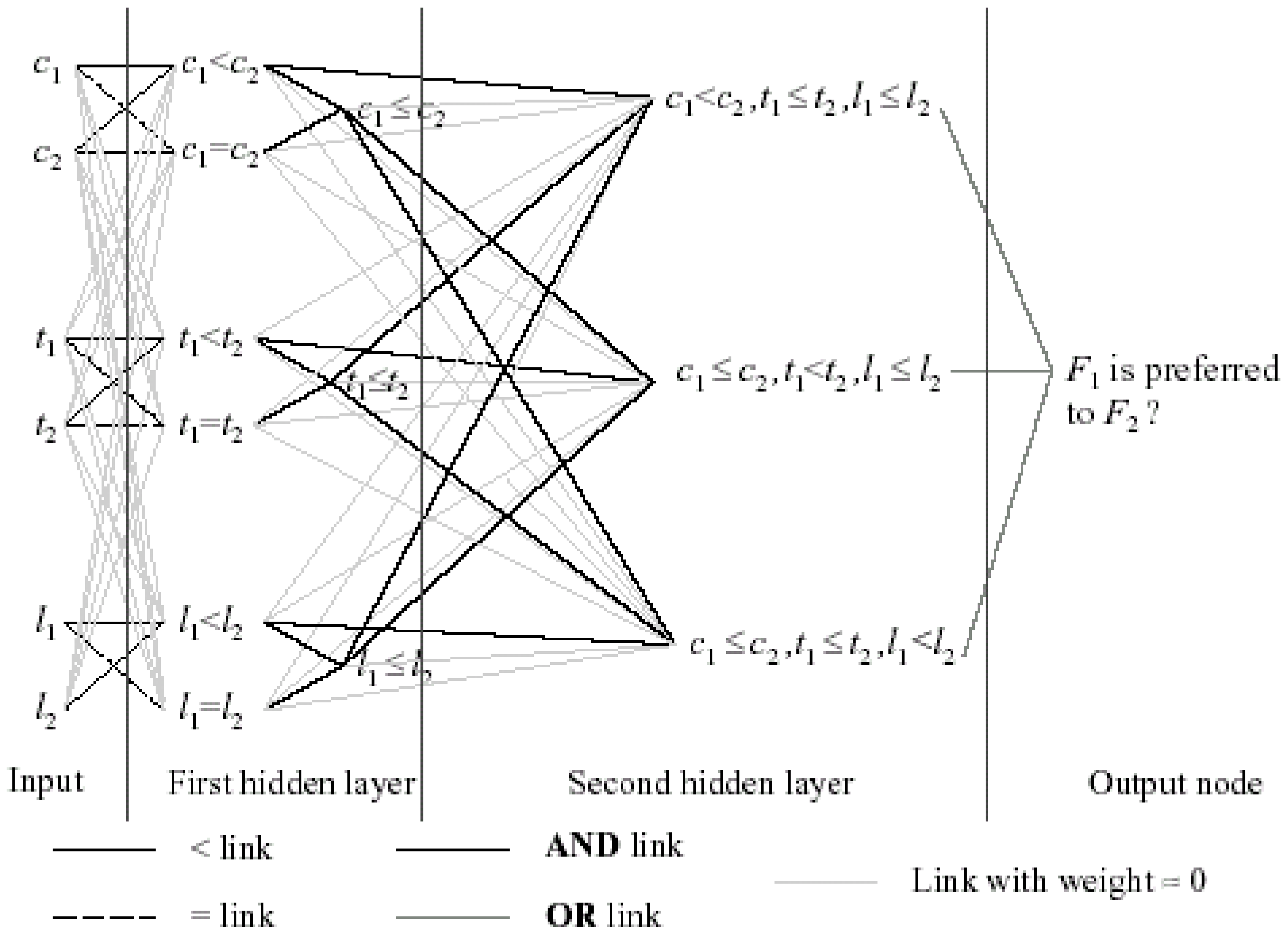
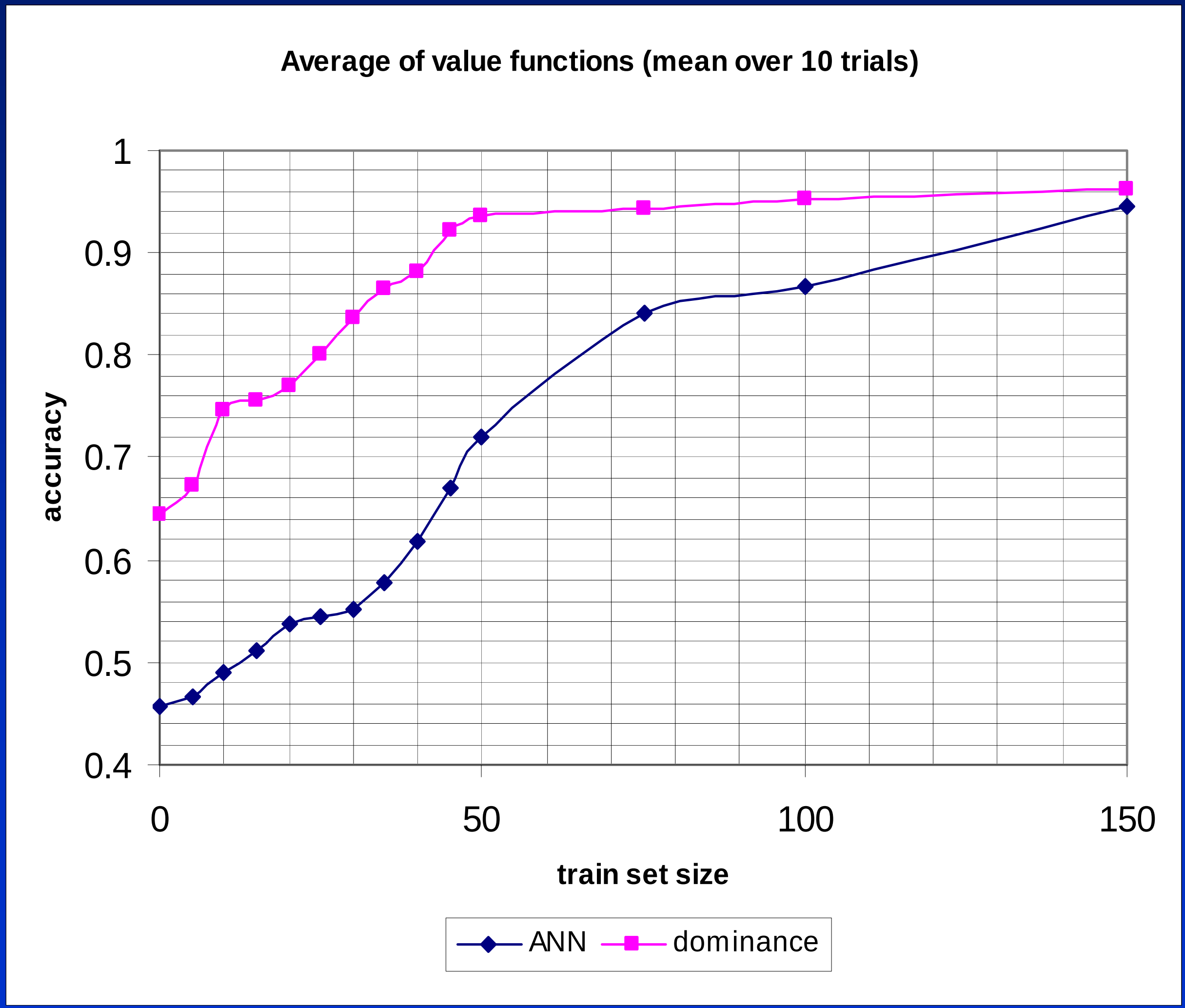


Figure 1. KBANN architecture for modeling user preferences in flights.

Empirical Results
(for preferences consistent with assumptions)

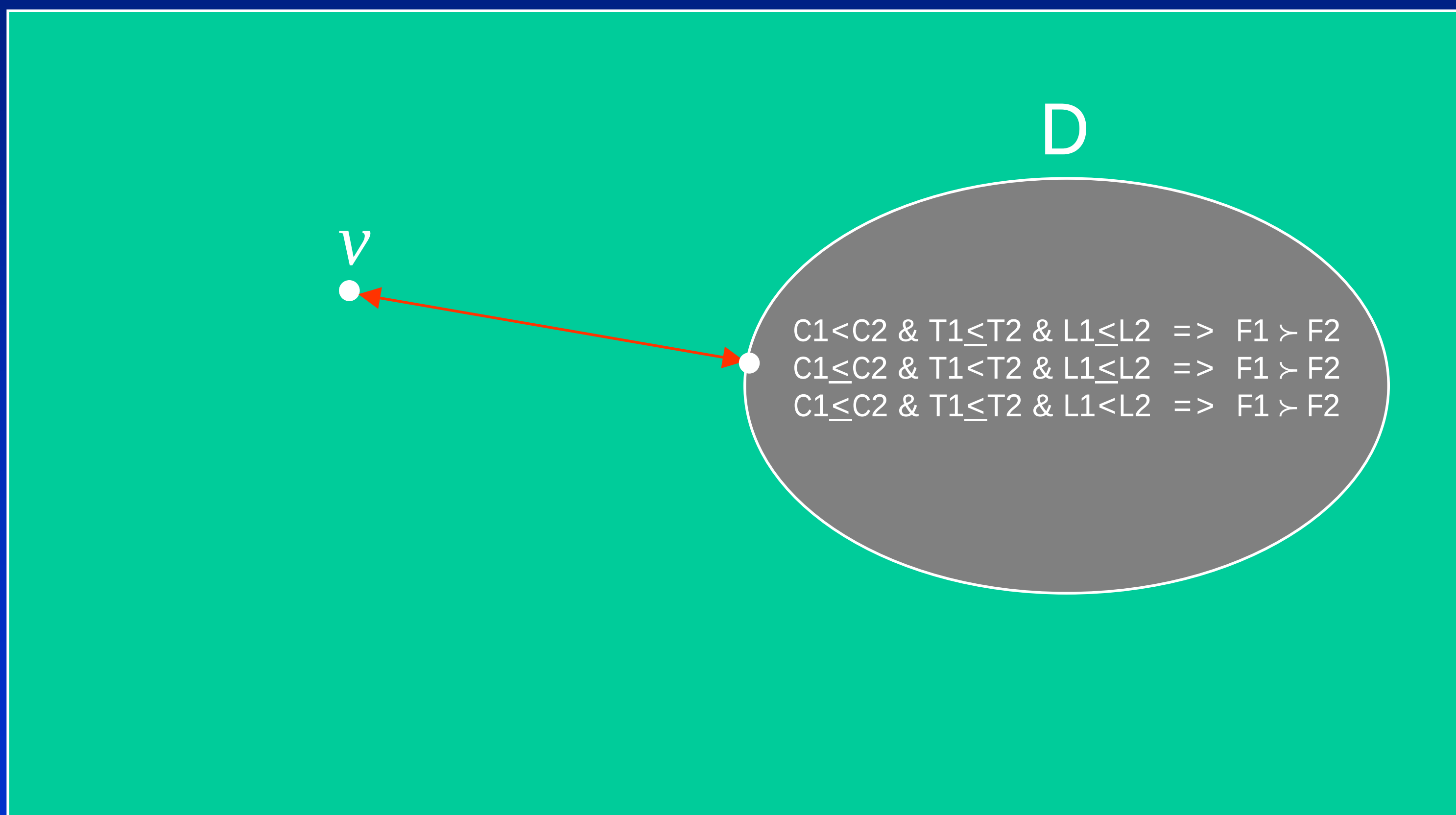


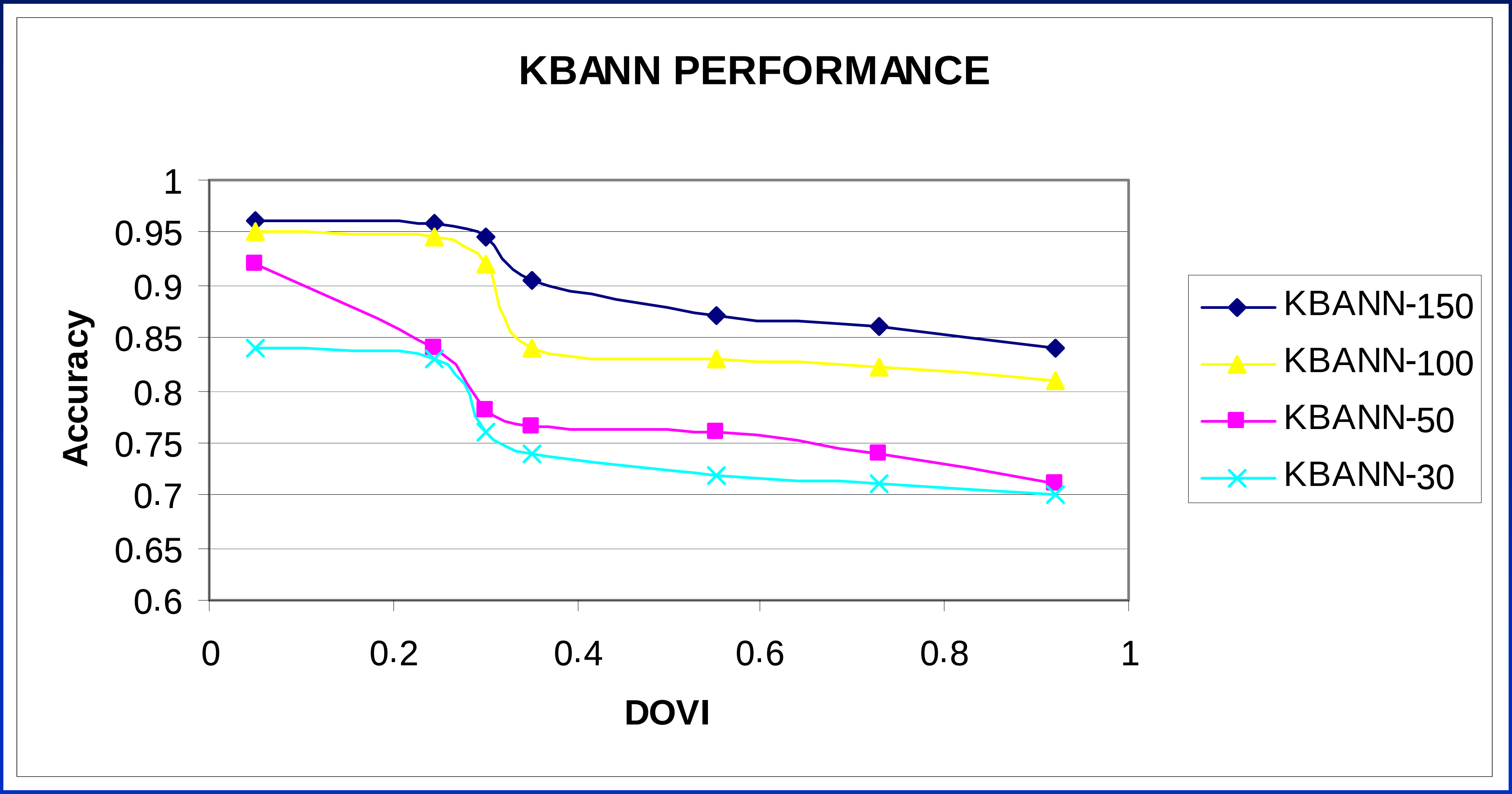
Robustness Analysis

- What about when preferences are not consistent with the assumptions?
- Evaluate by examining the performance of the network for value functions that violate the assumptions to various degrees.
- Need a definition of the degree to which a preference structure (v) violates a domain theory (D).
  - Use Probabilistic Distance

Degree of Violation

DOV(v,D) = Distance between v and the member of D closest to v.





KBANN PERFORMANCE
Accuracy
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
0.2
0.4
0.6
0.8
1
DOVI
KBANN-150
KBANN-100
KBANN-50
KBANN-30

Uncertainty

- Network input: standard-gamble questions
- Assumptions
  - Dominance
  - Attitude toward risk
- Results similar to those for certainty

Open Questions

- Field in practical applications
- Compare with other theory refinement techniques
- Stronger/weaker theories

- Learning of preferences in the context of negotiation
- Dynamics of preference and the implications for elicitation