

Uncertainty in Ontology Mapping: *A Bayesian Perspective*

Yun Peng, Zhongli Ding, Rong Pan

**Department of Computer Science and
Electrical engineering**

University of Maryland Baltimore County

ypeng@umbc.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

| | | | | | |
|--|------------------------------------|-------------------------------------|----------------------------|---|---------------------------------|
| 1. REPORT DATE AUG 2004 | | 2. REPORT TYPE | | 3. DATES COVERED 00-00-2004 to 00-00-2004 | |
| 4. TITLE AND SUBTITLE Uncertainty in Ontology Mapping: A Bayesian Perspective | | | | 5a. CONTRACT NUMBER | |
| | | | | 5b. GRANT NUMBER | |
| | | | | 5c. PROGRAM ELEMENT NUMBER | |
| 6. AUTHOR(S) | | | | 5d. PROJECT NUMBER | |
| | | | | 5e. TASK NUMBER | |
| | | | | 5f. WORK UNIT NUMBER | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Maryland Baltimore County, Department of Computer Science and Electrical engineering, 1000 Hilltop Circle, Baltimore, MD, 21250 | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | | | 10. SPONSOR/MONITOR'S ACRONYM(S) | |
| | | | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) | |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited | | | | | |
| 13. SUPPLEMENTARY NOTES Proceedings of the 2004 Performance Metrics for Intelligent Systems Workshop (PerMIS '04), Gaithersburg, MD on August 24-26 2004 | | | | | |
| 14. ABSTRACT | | | | | |
| 15. SUBJECT TERMS | | | | | |
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT unclassified | b. ABSTRACT unclassified | c. THIS PAGE unclassified | | | |

Outline

- Motivations
 - Uncertainty in ontology representation, reasoning and mapping
 - Why Bayesian networks (BN)
- Overview of the approach
- Translating OWL ontology to BN
 - Representing probabilistic information in ontology
 - Structural translation
 - Constructing conditional probability tables (CPT)
- Ontology mapping
 - Formalizing the notion of “mapping”
 - Mapping reduction
 - Mapping as evidential reasoning
- Conclusions

Motivations

- Uncertainty in ontology engineering
 - In representing/modeling the domain
 - Besides A *subclassOf* B , also A is a *small* subset of B
 - Besides A *hasProperty* P , also *most* objects with P are in A
 - A and B overlap, but none is a subclass of the other
 - In reasoning
 - How close a description D is to its most specific subsumer and most general subsumee?
 - Noisy data: leads to over generalization in subsumptions
 - Uncertain input: the object is *very likely* an instance of class A

Motivations

- In mapping concepts from one ontology to another
 - Similarity between concepts in two ontologies often cannot be adequately represented by logical relations
 - Overlap rather than inclusion
 - Mappings are hardly 1-to-1
 - If A in onto1 is similar to B in onto2, A would also be similar to the sub and super classes of B (with different degree of similarity)
- Uncertainty becomes more prevalent in web environment
 - One ontology may import other ontologies
 - Competing ontologies for the same or overlapped domain

Bayesian Networks

- Why Bayesian networks (BN)
 - Existing approaches
 - Logic based approaches are inadequate
 - Others often based on heuristic rules
 - Uncertainty is resolved during mapping, and not considered in subsequent reasoning
 - Loss of information
 - BN is a graphic model of dependencies among variables:
 - Structural similarity with OWL graph
 - BN semantics is compatible with that of OWL
 - Rich set of efficient algorithms for reasoning and learning

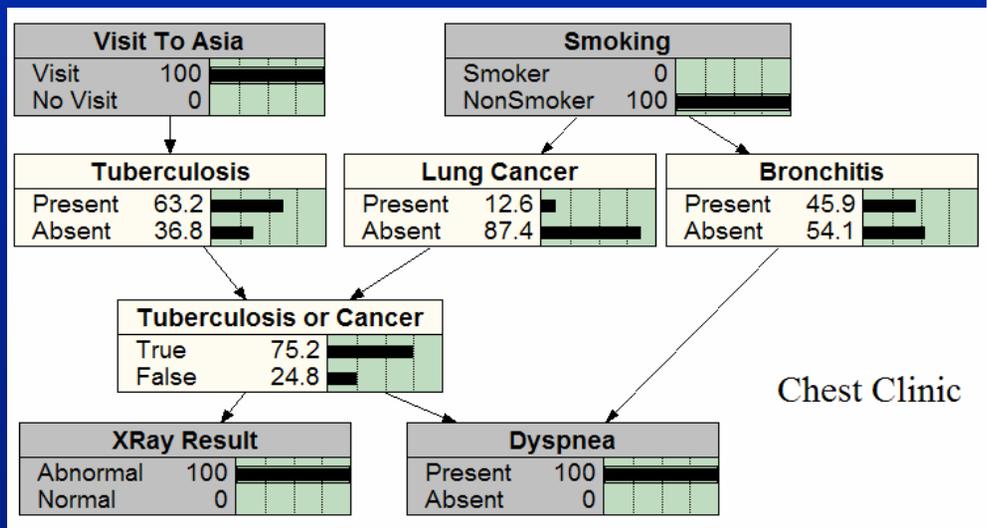
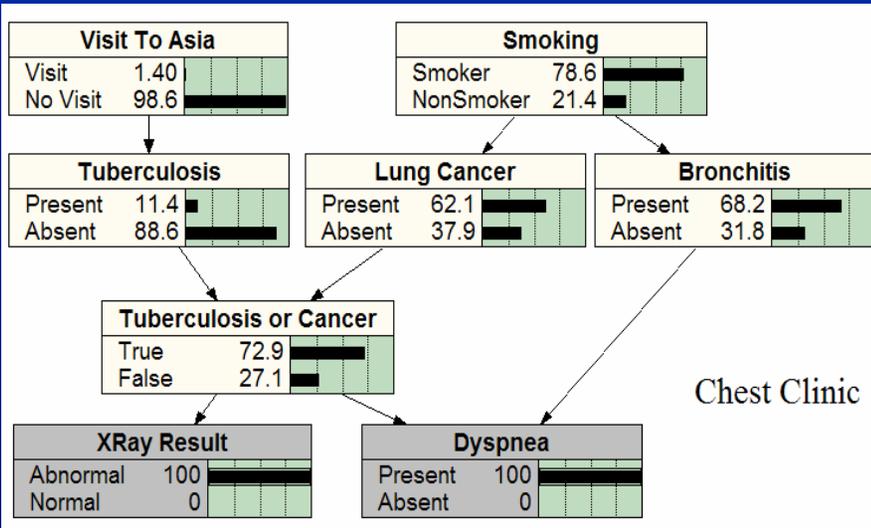
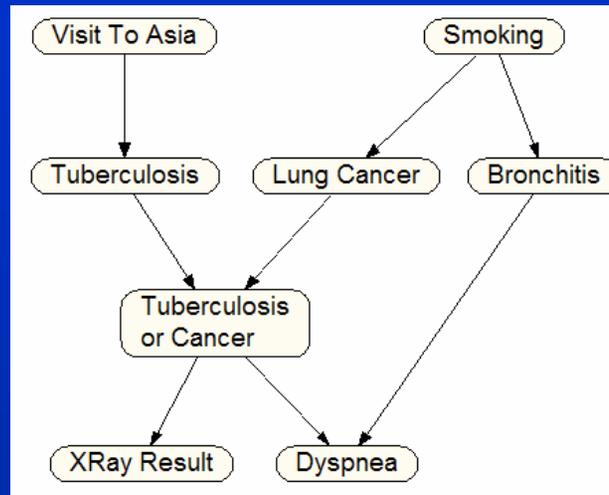
Bayesian Networks

- Directed acyclic graph (DAG)
 - Nodes: (discrete) random variables
 - Arcs: causal/influential relations
 - A variable is independent of all other non-descendent variables, given its parents
- Conditional prob. tables (CPT)
 - To each node: $P(x_i | \pi_i)$ where π_i is the parent set of x_i
- Chain rule:
 - $P(x_1, \dots, x_n) = \prod_i P(x_i | \pi_i)$
 - Joint probability as product of CPT

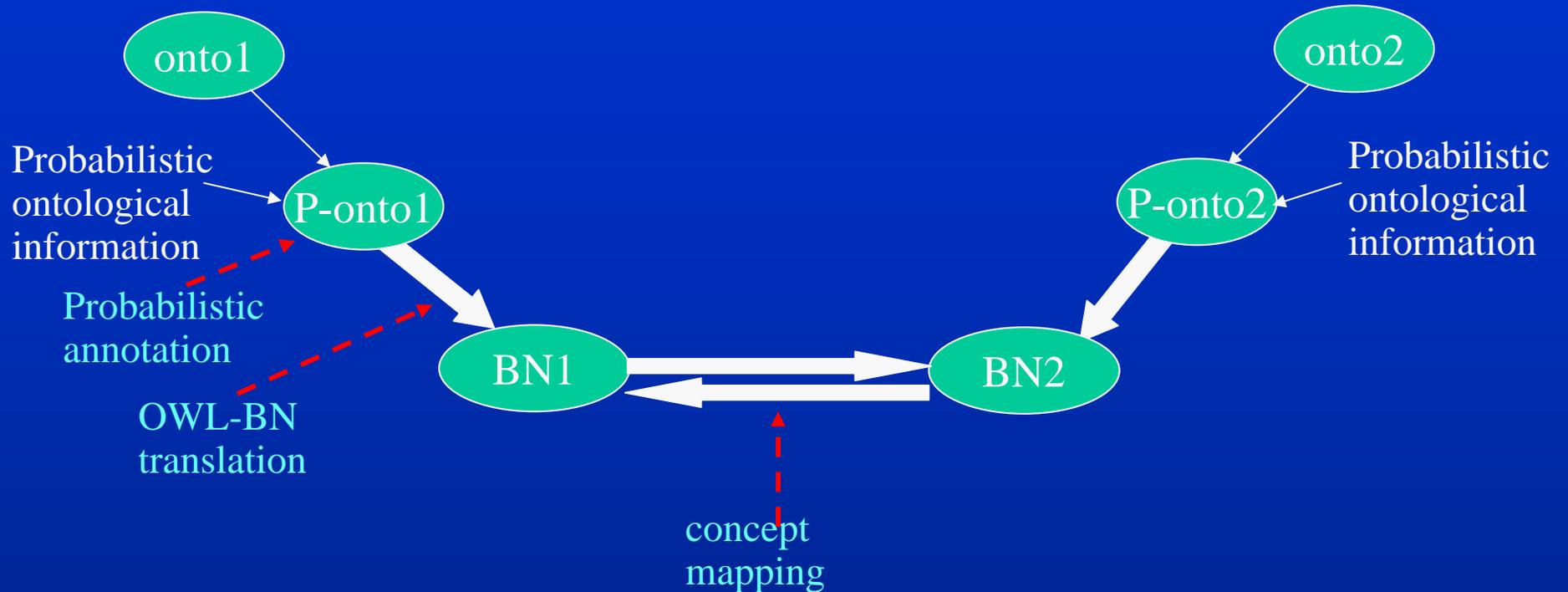
Bayesian Networks

| Visit | No_Visit |
|-------|----------|
| 1.000 | 99.000 |

| Tuberculosis | Cancer | Tb... |
|--------------|---------|-------|
| Present | Present | True |
| Present | Absent | True |
| Absent | Present | True |
| Absent | Absent | False |



Overview of The Approach



- OWL-BN translation
 - By a set of translation rules and procedures
 - Maintain OWL semantics
 - Ontology reasoning by probabilistic inference in BN

- Ontology mapping
 - A parsimonious set of links
 - Capture similarity between concepts by joint distribution
 - Mapping as evidential reasoning

OWL-BN Translation

- Encoding probabilities in OWL ontologies
 - Not supported by current OWL
 - Define new classes for prior and conditional probabilities
- Structural translation: a set of rules
 - Class hierarchy: set theoretic approach
 - Logical relations (equivalence, disjoint, union, intersection...)
 - Properties
- Constructing CPT for each node:
 - Iterative Proportional Fitting Procedure (IPFP)
- Translated BN will preserve
 - Semantics of the original ontology
 - Encoded probability distributions among relevant variables

Encoding Probabilities

- Allow user to specify prior and conditional Probabilities.
 - Two new OWL classes: "PriorProbObj" and "CondProbObj"
 - A probability is defined as an instance of one of these classes.
- $P(A)$: e.g., $P(\text{Animal}) = 0.5$

```
<prob:PriorProbObj rdf:ID="P(Animal)">  
  <prob:hasVariable><rdf:value>&ont;Animal</rdf:value></prob:hasVariable>  
  <prob:hasProbValue>0.5</prob:hasProbValue>  
</prob:PriorProbObj>
```

- $P(A|B)$: e.g., $P(\text{Male}|\text{Animal}) = 0.48$

```
<prob:CondProbObjT rdf:ID="P(Male|Animal)">  
  <prob:hasCondition><rdf:value>&ont;Animal</rdf:value></prob:hasCondition>  
  <prob:hasVariable><rdf:value>&ont;Male</rdf:value></prob:hasVariable>  
  <prob:hasProbValue>0.5</prob:hasProbValue>  
</prob:CondProbObjT>
```

Structural Translation

- Set theoretic approach
 - Each OWL class is considered a **set** of objects/instances
 - Each class is defined as a node in BN
 - An arc in BN goes from a superset to a subset
 - Consistent with OWL semantics

```
<owl:Class rdf:ID="Human">  
  <rdfs:subclassOf rdf:resource="#Animal">  
  <rdfs:subclassOf rdf:resource="#Biped">  
</owl:Class>
```

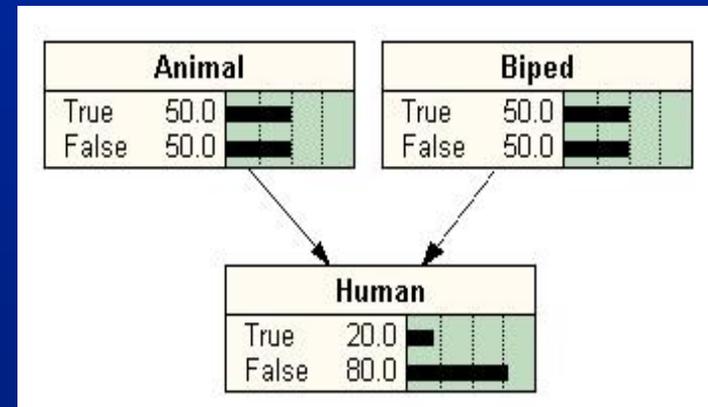
RDF Triples:

(Human rdf:type owl:Class)

(Human rdfs:subClassOf Animal)

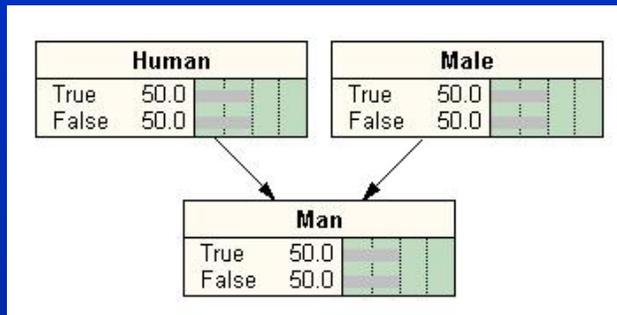
(Human rdfs:subClassOf Biped)

Translated to BN



Structural Translation

- Logical relations
 - Some can be encoded by CPT (e.g.. $\text{Man} = \text{Human} \cap \text{Male}$)



| Human | Male | True | False |
|-------|-------|--------|--------|
| True | True | 100.00 | 0.000 |
| True | False | 0.000 | 100.00 |
| False | True | 0.000 | 100.00 |
| False | False | 0.000 | 100.00 |

- Others can be realized by adding control nodes

$\text{Man} \subset \text{Human}$

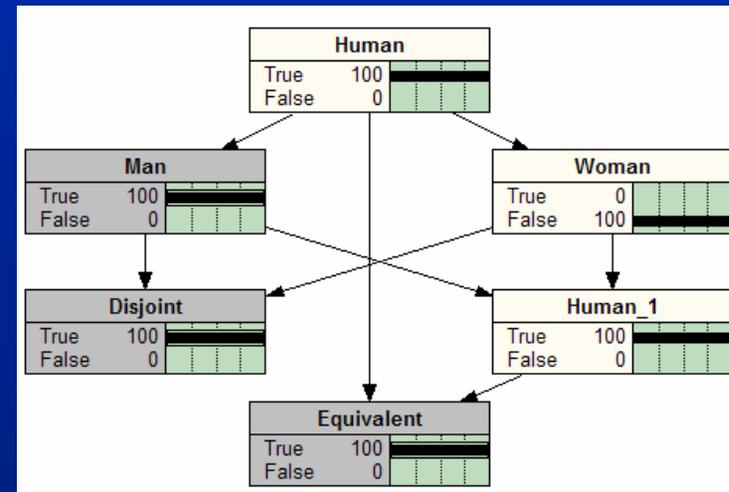
$\text{Woman} \subset \text{Human}$

$\text{Human} = \text{Man} \cup \text{Woman}$

$\text{Man} \cap \text{Woman} = \emptyset$

auxiliary node: Human_1

Control nodes: Disjoint, Equivalent



Constructing CPT

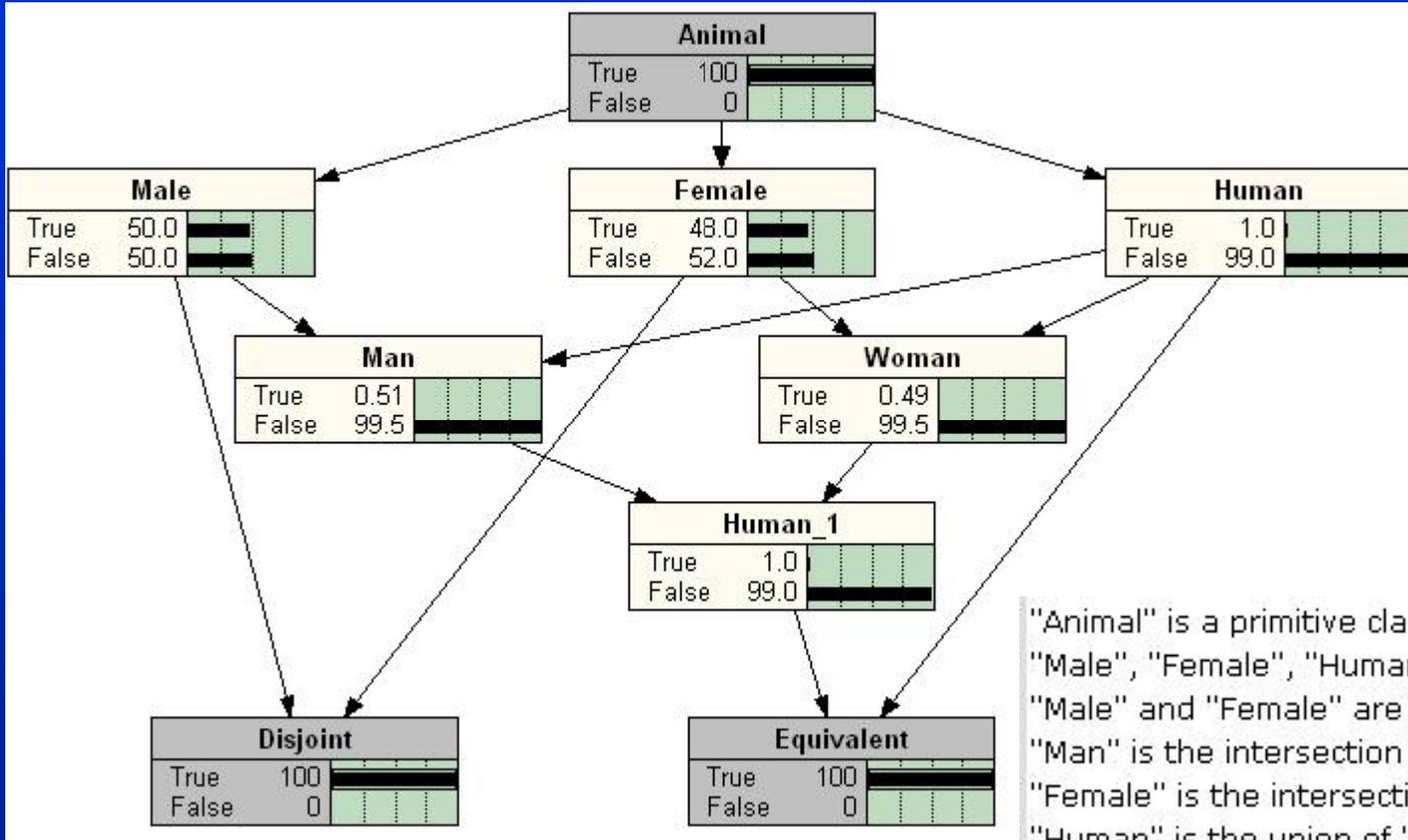
- Imported Probability information is not in the form of CPT
- Assign initial CPT to the translated structure by some default rules
- Iteratively modify CPT to fit imported probabilities while setting control nodes to *true*.
 - IPFP (Iterative Proportional Fitting Procedure)

To find $Q(x)$ that fit $Q(E_1), \dots, Q(E_k)$ to the given $P(x)$

- $Q_0(x) = P(x)$; then repeat $Q_i(x) = Q_{i-1}(x) Q(E_j) / Q_{i-1}(E_j)$ until converging
- $Q_\infty(x)$ is an I-projection of $P(x)$ on $Q(E_1), \dots, Q(E_k)$ (minimizing Kullback-Leibler distance to P)

- Modified IPFP for BN

Example



"Animal" is a primitive class

"Male", "Female", "Human" are subclasses of "Animal"

"Male" and "Female" are disjoint with each other

"Man" is the intersection of "Male" and "Human"

"Female" is the intersection of "Female" and "Human"

"Human" is the union of "Man" and "Woman"

Probability information:

$$P(\text{Animal}) = 0.5$$

$$P(\text{Male}|\text{Animal}) = 0.5$$

$$P(\text{Female}|\text{Animal}) = 0.48$$

$$P(\text{Human}|\text{Animal}) = 0.01$$

Ontology Mapping

- Formalize the notion of *mapping*
- Mapping involving multiple concepts
- Reasoning under ontology mapping
- Assumption: ontologies have been translated to BN

Formalize The Notion of Mapping

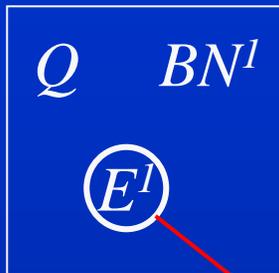
- Simplest case: Map concept E^1 in Onto^1 to E^2 in Onto^2
 - How similar between E^1 and E^2
 - How to impose belief (distribution) of E^1 to Onto^2
- Cannot do it by simple Bayesian conditioning
$$P(x|E^1) = \sum_{E^2} P(x|E^2)P(E^2|E^1) \text{similarity}(E^1, E^2)$$
 - Onto^1 and Onto^2 have different probability space (Q and P)
 - $Q(E^1) \neq P(E^1)$
 - New distribution, given E^1 in Onto^1 : $P^*(x) \neq \sum P(x|E^1)P(E^1)$
 - $\text{similarity}(E^1, E^2)$ also needs to be formalized

Formalize The Notion of Mapping

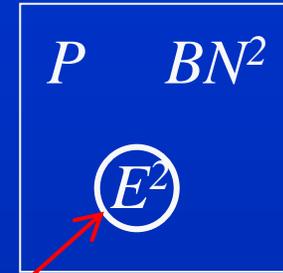
- Jeffrey's rule
 - Conditioning cross prob. spaces
 - $P^*(x) = \sum P(x/E^1)Q(E^1)$
 - P^* is an I-projection of $P(x)$ on $Q(E^1)$ (minimizing Kullback-Leibler distance to P)
 - Update P to P^* by applying $Q(E^1)$ as soft evidence in BN
- $\text{similarity}(E^1, E^2)$
 - Represented as joint prob. $R(E^1, E^2)$ in another space R
 - Can be obtained by learning or from user
- Define

$$\text{map}(E^1, E^2) = \langle E^1, E^2, BN^1, BN^2, R(E^1, E^2) \rangle$$

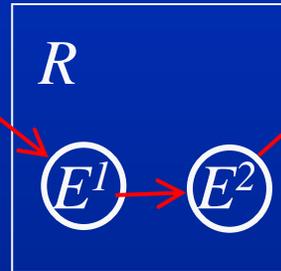
Reasoning With $\text{map}(E^1, E^2)$



Applying $Q(E^1)$ as soft evidence to update R to R^* by Jeffrey's rule

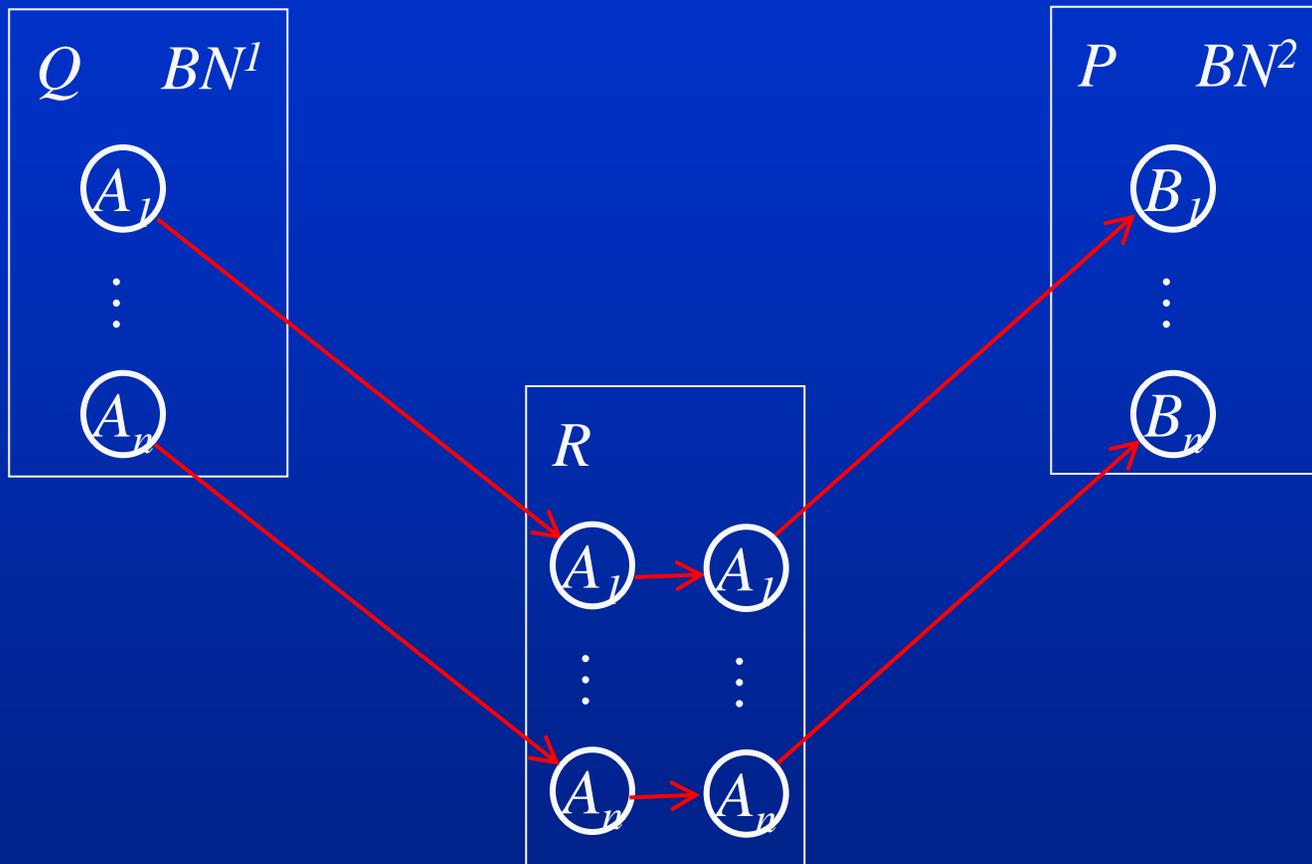


Applying $R^*(E^2)$ as soft evidence to update P to P^* by Jeffrey's rule



$$\begin{aligned} &\text{Using similarity}(E^1, E^2): \\ &R^*(E^2) \\ &= R^*(E^1, E^2)/R^*(E^1) \end{aligned}$$

Reasoning With Multiple $\text{map}(E^1, E^2)$



Multiple pair-wise mappings: $\text{map}(A_k, B_k)$:
Realizing Jeffrey's rule by IPFP

Mapping Reduction

- Multiple mappings
 - One node in BN1 can map to all nodes in BN2
 - Most mappings with little similarity
 - Which of them can be removed without affecting the overall
- Similarity measure:
 - *Jaccard-coefficient*: $\text{sim}(E^1, E^2) = P(E^1 \cap E^2) / R(E^1 \cup E^2)$
 - A generalization of subsumption
 - Remove those mappings with very small sim value
- Question: can we further remove other mappings
 - Utilizing knowledge in BN

Conclusions

- Summary
 - A principled approach to uncertainty in ontology representation, reasoning and mapping
- Current focuses:
 - OWL-BN translation: properties
 - Ontology mapping: mapping reduction
- Prototyping and experiments
- Issues
 - Complexity
 - How to get these probabilities