

The Bayesian Web

Adding Reasoning with Uncertainty to the Semantic Web

Ken Baclawski

Northeastern University

The Semantic Web

- The Semantic Web is an extension of the current web in which information is given well defined meaning... (Berners-Lee, Hendler & Lassila)
- The Semantic Web is based on formal logic for which one can only assert facts that are unambiguously certain.
- Unfortunately, there are many sources of uncertainty, such as measurements, unmodeled variables, and subjectivity.

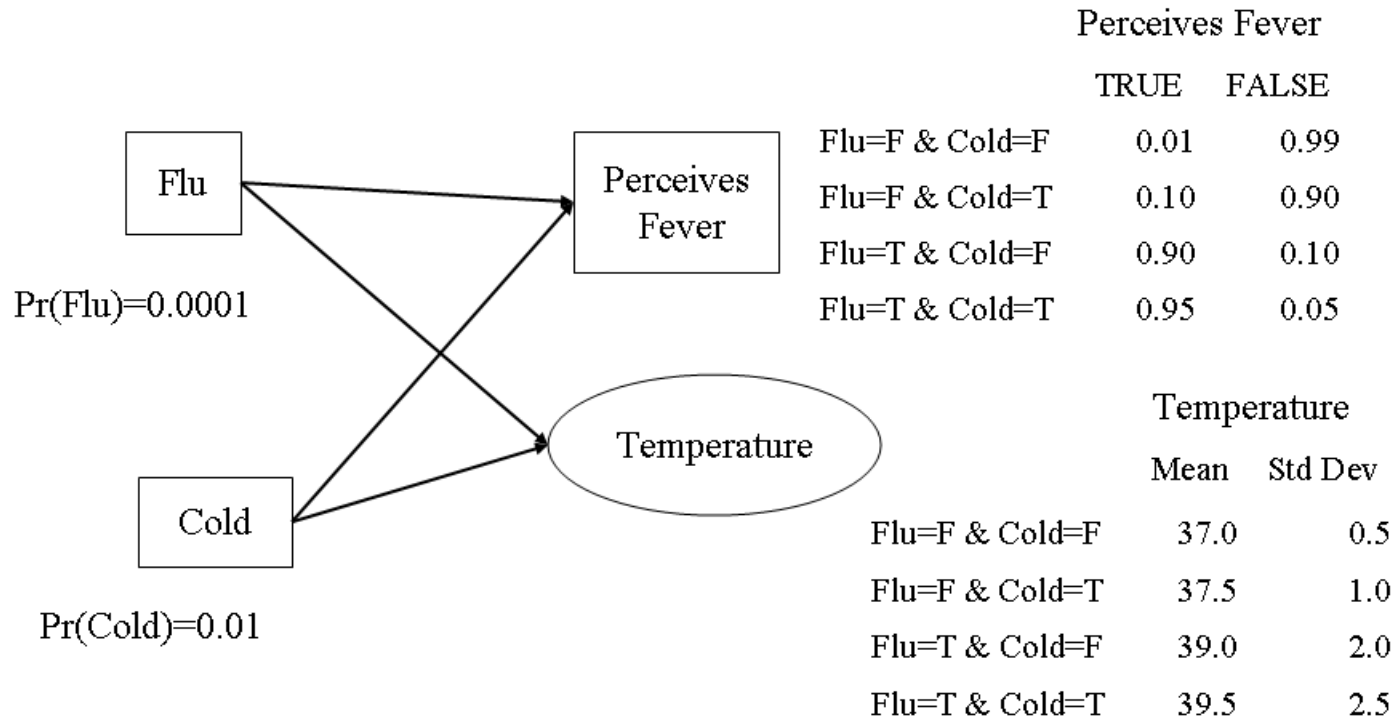
Adding Uncertainty

- The challenge is to develop a full-featured stochastic reasoning infrastructure, comparable to the logical reasoning infrastructure of the Semantic Web.
- The *Bayesian Web* is a proposal to add reasoning about uncertainty to the Semantic Web.
- The basis for the Bayesian Web is the concept of a Bayesian network (BN).

The Bayesian Network Formalism

- A BN is a graphical mechanism for specifying joint probability distributions (JPDs).
- The nodes of a BN are random variables.
- The edges of a BN represent stochastic dependencies.
- The graph of a BN must not have any directed cycles.
- Each node of a BN has an associated CPD.
- The JPD is the product of the CPDs.

Bayesian Network Specification



CPDs:

1. Perceives Fever given Flu and/or Cold.
2. Temperature given Flu and/or Cold.
3. Probability of Flu (unconditional).
4. Probability of Cold (unconditional).

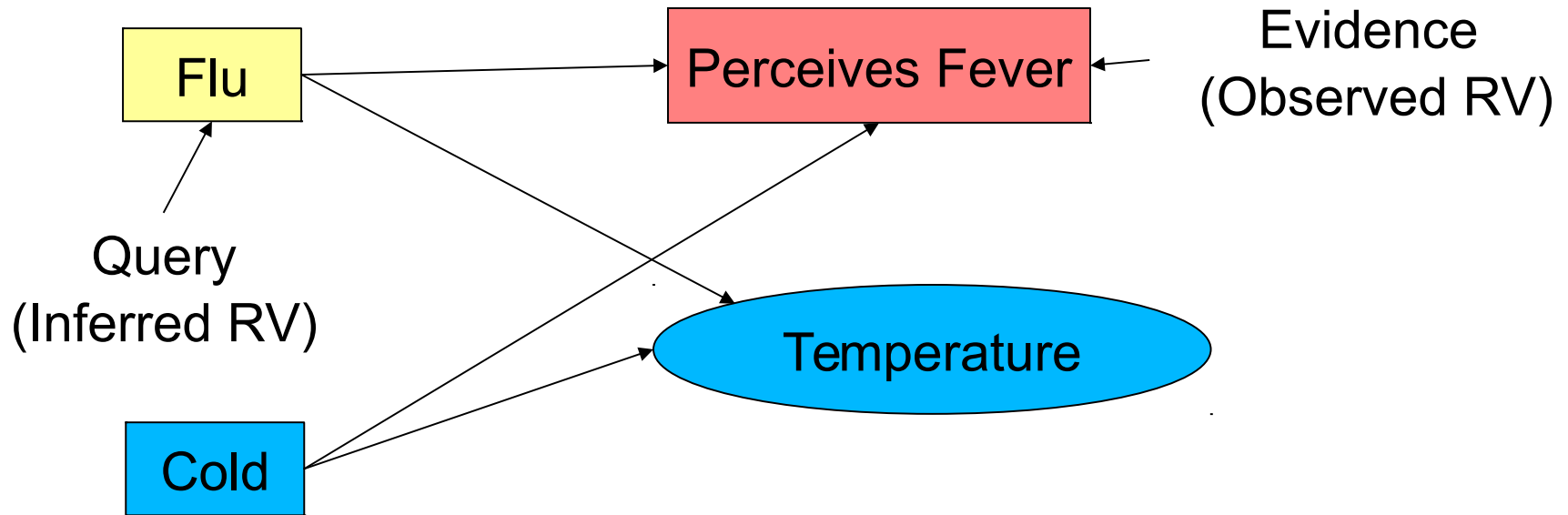
Discrete RV

Continuous RV

Stochastic Inference

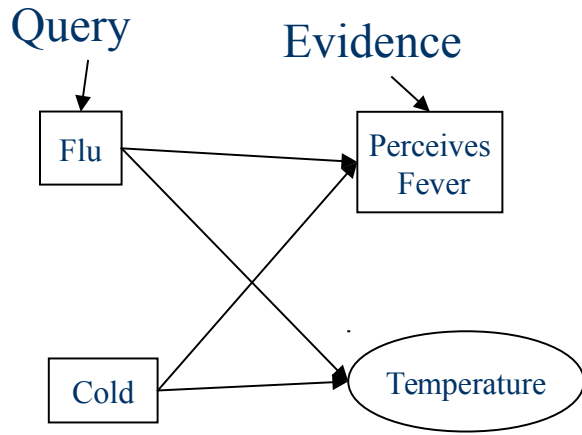
- Stochastic inference.
 - The main use of BNs.
 - Analogous to the process of logical inference and querying performed by rule engines.
 - Based on Bayes' law.
- Evidence
 - Can be either hard observations with no uncertainty or uncertain observations specified by a probability distribution.
 - Can be given for any nodes, and any nodes can be queried.
- Nodes can be continuous random variables, but inference in this case is more complicated.
- BNs can be augmented with other kinds of nodes, and used for making decisions based on stochastic inference.

Bayesian Network Inference

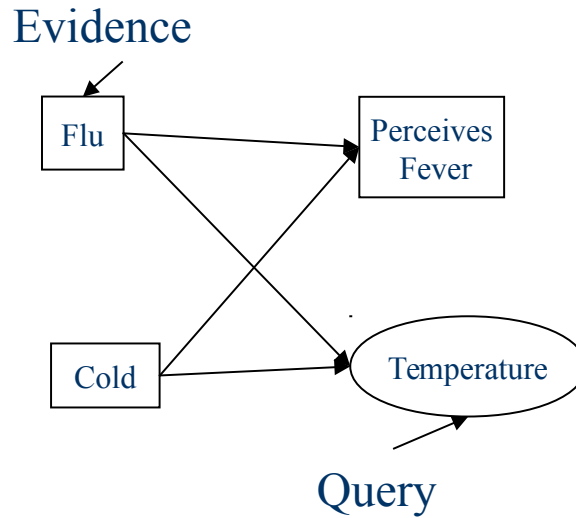


Inference is performed by observing some RVs (evidence) and computing the distribution of the RVs of interest (query). The evidence can be a value or a probability distribution. The BN combines the evidence probability distributions even when there are probabilistic dependencies.

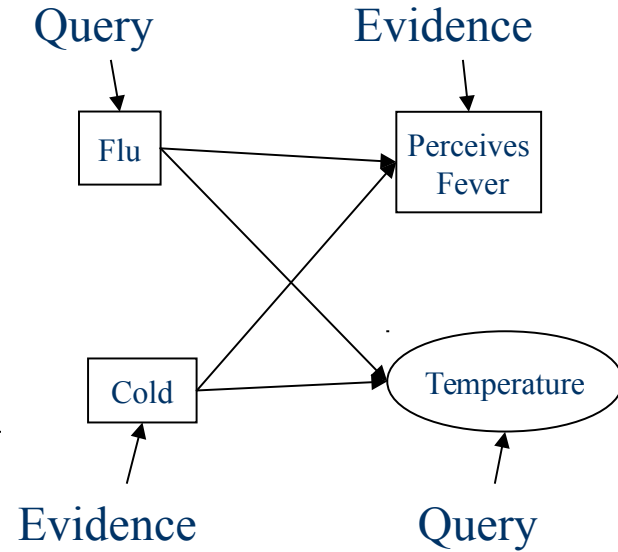
Bayesian Network Inference



Diagnostic Inference



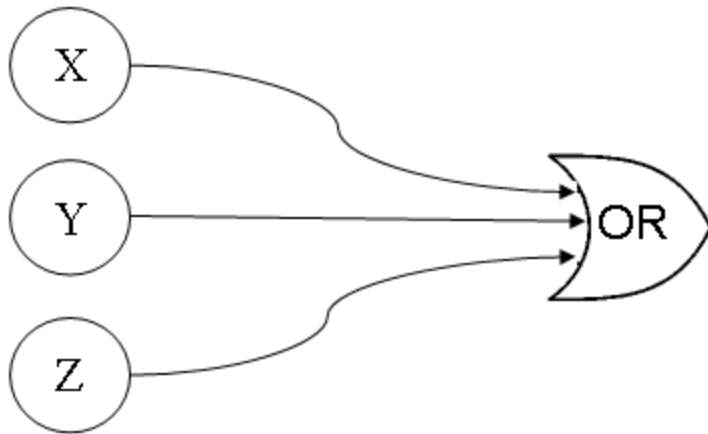
Causal Inference



Mixed Inference

BN Design Patterns

- One methodology for designing BNs is to use design patterns or idioms.
- Many BN design patterns have been identified, but most are only informally specified.



$$\Pr(W=\text{true}|X=\text{true}, Y=\text{true}, Z=\text{true}) = 1 - q_x q_y q_z$$

$$\Pr(W=\text{true}|X=\text{true}, Y=\text{false}, Z=\text{true}) = 1 - q_x q_z$$

$$\Pr(W=\text{true}|X=\text{true}, Y=\text{true}, Z=\text{false}) = 1 - q_x q_y$$

$$\vdots$$

$$\Pr(W=\text{true}|X=\text{false}, Y=\text{false}, Z=\text{false}) = 0$$

The noisy OR-gate design pattern

Bayesian Web facilities

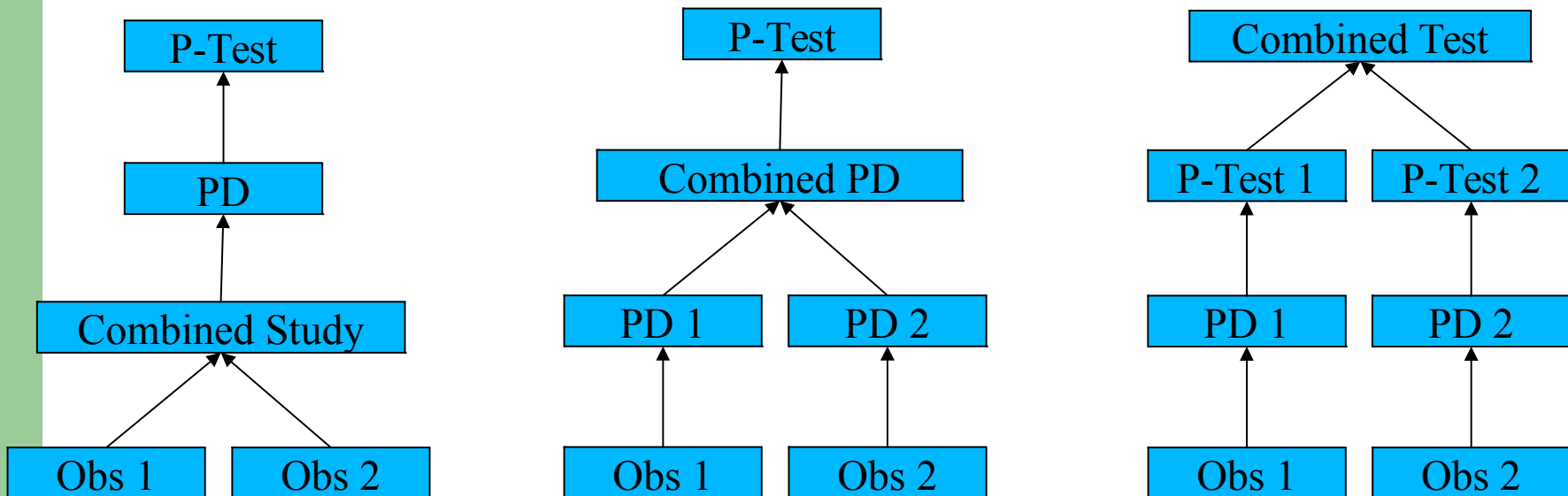
- A common interchange format for stochastic models and statistical test results
- Allow specification of the context of a model or result
- Open hierarchy of probability distribution types
- Component based construction of stochastic models
- Stochastic inference engines

Bayesian Web Capabilities

- Use a BN developed by another group as easily as navigating from one Web page to another.
- Perform stochastic inference using information from one source and a BN from another.
- Combine BNs from the same or different sources.
- Reconcile and validate BNs.

Meta-Analysis

Meta-analysis is the process of combining information from disparate sources. Information can be combined at many levels.



The various levels where information can be combined has been standardized by the Joint Defense Laboratories (JDL) model. The whole process is called *data fusion*.

Level	Name	Process	Estimation	Product	Medicine
0	Signal Assessment	Identify features	Detection	Signal State	Observation
1	Object Assessment	Identify entities	Attributive State	Entity State	Symptom
2	Situation Assessment	Relationships among entities	Relation	Situation State	Diagnosis
3	Impact Assessment	Evaluation	Game Theory	Situation Utility	Prognosis

Situation Awareness

- Situation awareness (SAW) is “knowing what is going on around oneself.”
 - More precisely, SAW is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future (Endsley & Garland).
- SAW occurs at level 2 of the JDL model.

Research & Development Challenge

- To develop semantic data fusion tools for biomedicine that support researchers and clinicians in the task of situation awareness (diagnosis) and impact assessment (prognosis).
- Some examples of applications of such a tool include:
 - Tracking epidemics
 - Monitoring the patient during surgery
 - Meta-analysis services for researchers
 - Assessing the health of populations by region or recognized group