

Where are we?

Ahead of Schedule

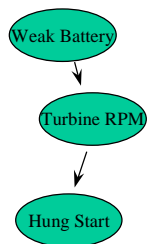
Next week: Approximation Algorithms,
Read 9.1 - 9.8 in CGH
Read Boyen & Koller, Kozlov & Koller, Pradhan & Dagum
Skim Dagum and Luby

Today

Brief GeNIe Demo
Summary of Model Construction Techniques
Hugin Conflict Measure
(Well... it kind of fits here.)
Summary of Application Areas (top level)

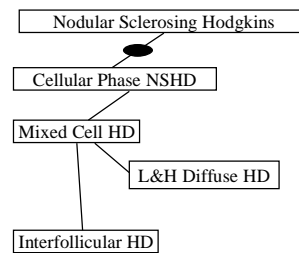
Assessing Structure

Causal Input Lists



Causal independence list
implies correct dependency structure
Easiest and most compact probability
assessments
Counter:
Tomorrow's Weather --> Weather Report

Similarity Nets

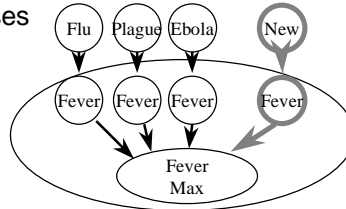


Best features derived from most similar
faults
Sim Net -->
Context Specific Independence
Partitions

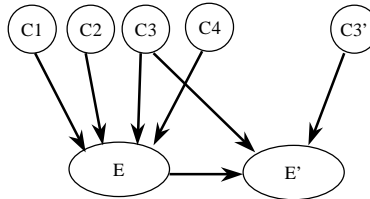
Probability Assessment

Causal Independence Dist'ns

Incremental addition of causes
Linear-time assessment



Causal Independence Assumption



Clarity and Assessment

Clarity

IMHO: Most important factor in model construction
Clarity required for good probability assessment

Probability Assessment

Priors only affect the initial phases of question asking and are quickly washed out.

Likelihoods sensitive to qualitative behavior (order of magnitude).

Noisy-Or:

Link probabilities matter the most.

Classes of Graphical Model Applications

Prediction/Dynamic Probabilistic Networks

Predict effects based on causes.

Influence Diagrams

Select the decisions that have the best outcomes.

Dynamic Probabilistic Networks

Infer state of system based on noisy observations

“Data-Mining” (Model selection)

Infer probabilities based on observed data.

Infer structure based on observed data.

Diagnosis

Infer causes based on observed symptoms (effects)

Hugin Conflict

The problem:

How do you tell when your model no longer works?

Example: failure of single fault assumption.

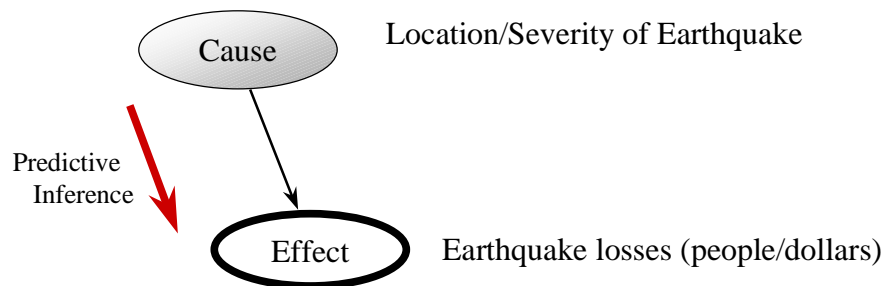
One solution:

Hugin conflict measure

$$\text{conf}(E) = \log \frac{P\{e_1\} \cdots P\{e_n\}}{P\{e_1, \dots, e_n\}}$$

Predictive models

Infer effects based on causes



Predictive Models

Probabilities do matter

LA County Hospital Expansion

Constructed by the National Health Foundation
using Analytica (Lumina Decision Systems)

Summary:

Predict req'd hospital beds based on demographics, immigration patterns, aging, and disease prevalence.

Estimated savings: \$ 300 M

References:

Morgan and Henrion, *Uncertainty*

"Analytica": Monte Carlo integration of
general probabilistic models.



Contact: Max Henrion, henrion@lumina.com

Influence Diagrams

Industry:

Used for plant expansion decisions, marketing decisions, medical decision making.

Tool used by a large number of management consulting companies

Strategic Decisions Group (SDG)

References:

Influence Diagrams:

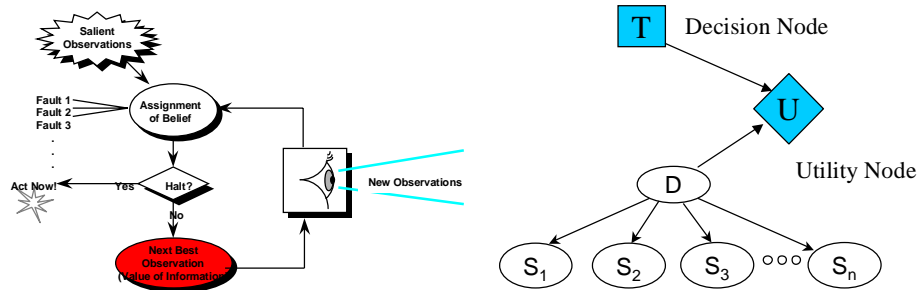
Howard + Matheson, *The Principles and Applications of Decision Analysis*

Evaluation:

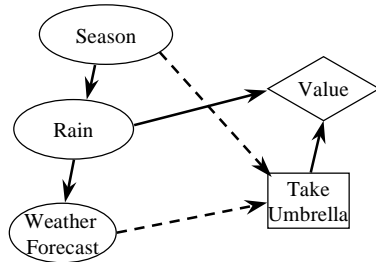
Shachter + Peot, UAI-94

Jensen + Jensen, UAI-96

Influence Diagrams and Pathfinder



Generic Influence Diagram

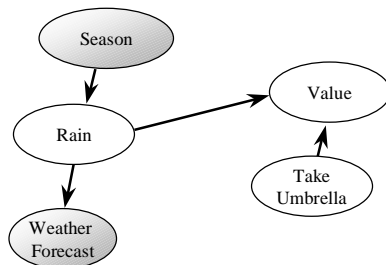


Infer optimal *Policy*

$$D = f(S, F)$$

Policy Example:

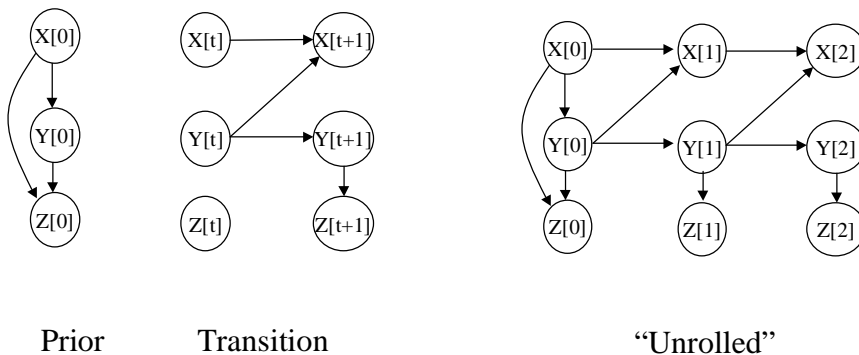
Take Umbrella if the Forecast says "Rain"
and it is the Rainy Season.
Leave the Umbrella if the Forecast says "Sun"
or it is the Dry Season.



$$D^* = \underset{D}{\operatorname{argmax}} V(D, E)$$

Dynamic Probabilistic Network

Model of state evolution with time.



Prior

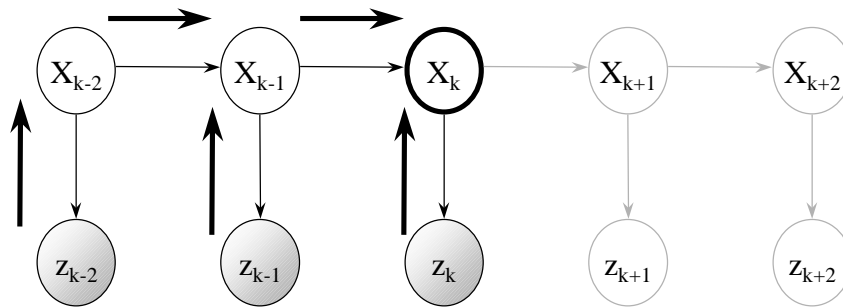
Transition

"Unrolled"

Filtering/Smoothing

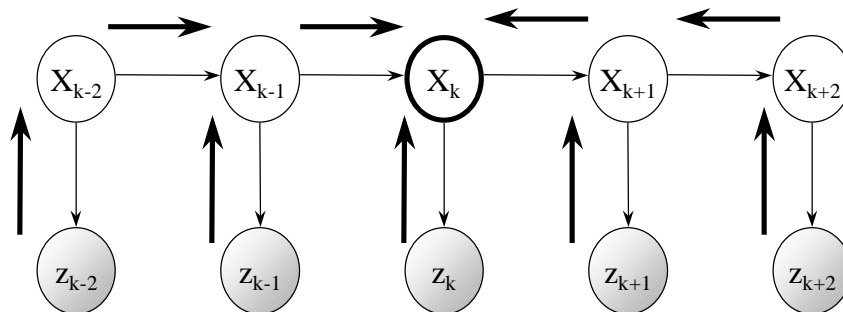
Kalman filter

Used in aviation and space for navigation and target tracking.
Estimate position/velocity based on noisy sensor data and kinematics.

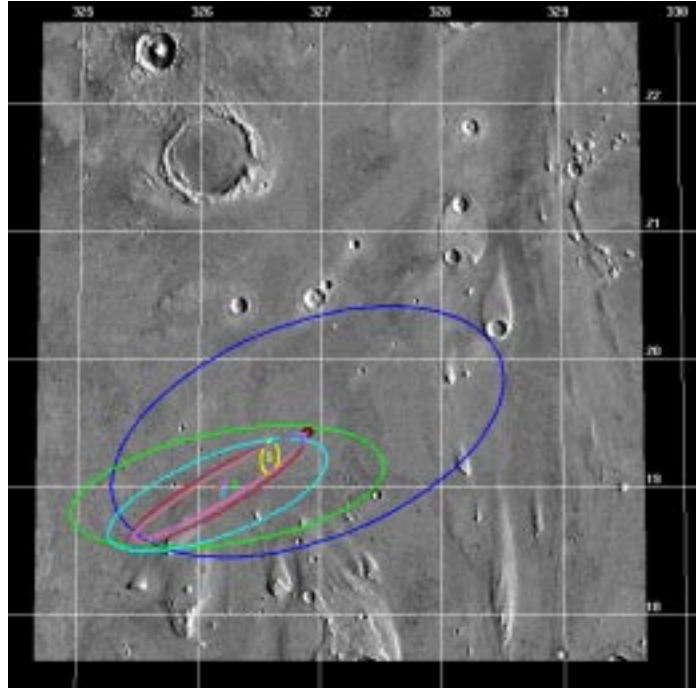


Filtering/Smoothing

“Rauch-Tung-Streibel” smoother



Kalman Filters:
Mars Pathfinder

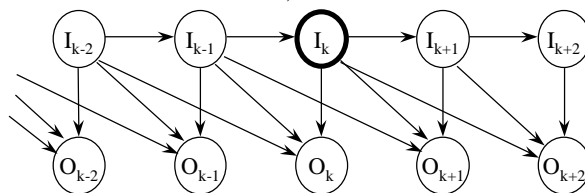


Convolutional Codes

Used in satellite data transmission, cellular phones

Increases S/N ratio, allowing lower power transmission

$$\bar{\mathbf{O}}(t) = \sum_{i=0, n} \bar{\mathbf{A}}(t) \mathbf{I}(t-i)$$



Optimal decoder is Viterbi decoder

DPN with normal measurements and discrete transitions

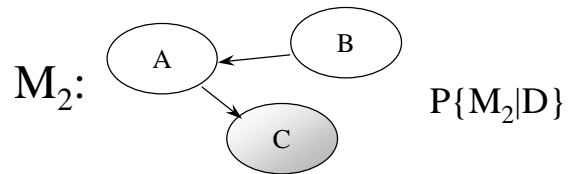
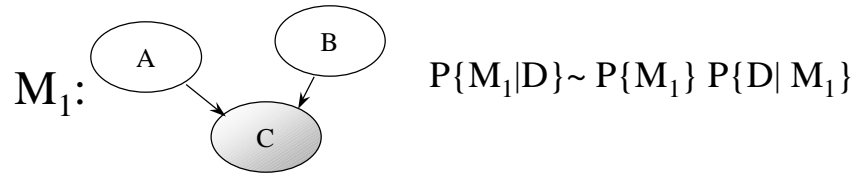
Viterbi path (the answer) is the maximum probability instantiation for DPN.

Also used for speech recognition.

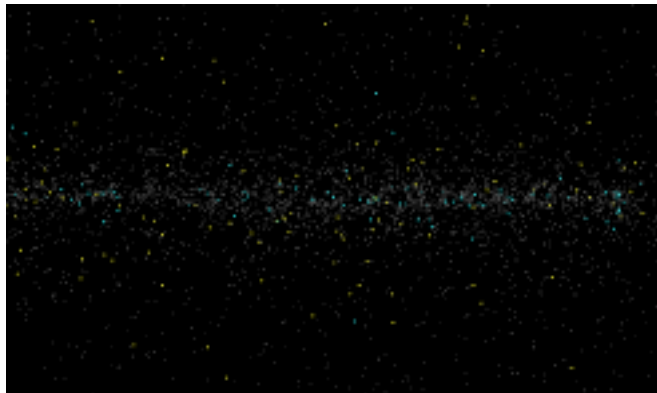
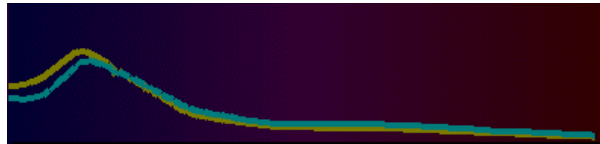
Hot topic

Turbo codes: Used for deep space communication (Galileo)

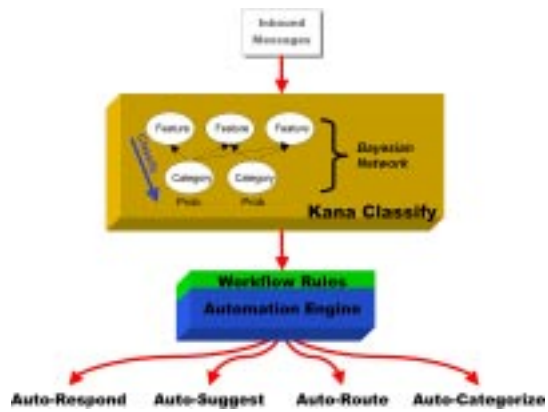
Inferring Structure



Autoclass: IRAS Star Catalog



Kana Mail Sorting



Online Troubleshooters

Application Description

Web-based diagnosis of faults for >70 applications

Web Address: <http://www.microsoft.com/support/tshooters.asp>

Type

Multiple fault with utility model (*troubleshooting* [Breese+Heckerman, UAI 96])

Size

40-50 nodes/network

4 to 5 main symptoms. 5-10 causes per symptom.

Contact: Jack Breese, breese@microsoft.com

Online Troubleshooters



Microsoft/KI OnParenting



Microsoft OnParenting

Find by Symptom feature for a multimedia children's health encyclopedia

Computes the relevance of articles in the database, based on symptoms

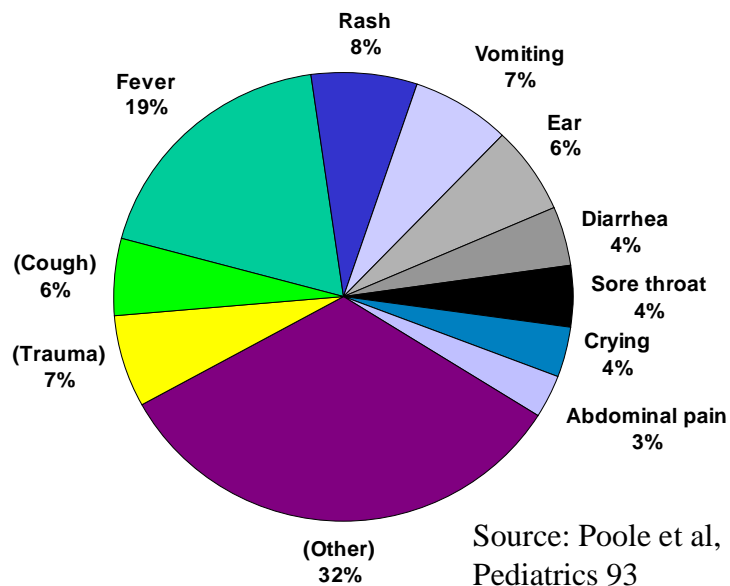
Launched in Feb 1995 on MSN as MS Pregnancy and Childcare

Re-introduced in April 1997 on MSN as MS OnParenting

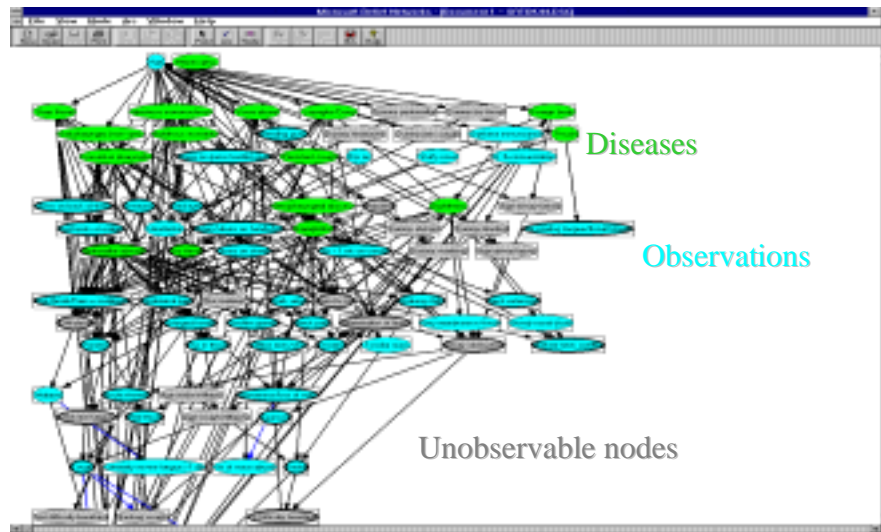
<http://onparenting.msn.com> (Moved again)

Contact: Michael Shwe, shwe@kana.com

Top 10 Pediatric Complaints



Chief Complaint: Sore Throat



TTC Internetworking Detective

Application Description

Isolates H/W and configuration errors in T1/Frame Relay networks.

Software add-on for the **FIREBERD 500** Internetwork Analyzer

Type

Multiple Fault

Contact: Mark Peot, peot@kic.com



Ricoh



Ricoh Corp:
Office Equipment



In use at customer help desk centers in the US and Japan.

Diagnosis of fax machines and other office equipment.

Single-fault with observation/
observation dependencies.