An Independence of Causal Interactions Model for Opposing Influences

Paul P. Maaskant[1] & Marek J. Druzdzel[2,3]

1 Department of Media and Knowledge Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands
2 Faculty of Computer Science, Bialystok Technical University, Wiejska 45A, 15-351 Bialystok, Poland
3 Decision Systems Laboratory, School of Information Sciences and Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA 15260, USA

The problem

- **A major problems with Bayesian Networks is the exponential growth of conditional probability tables (CPTs) in the number of parents**

5 parents

Node2	State0															
Node3	State0								State1							
Node4	State0				State1				State0				State1			
Node5	State0		State1		State0		State1		State0		State1		State0		State1	
Node6	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1
State0	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
State1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Node2	State1															
Node3	State0								State1							
Node4	State0				State1				State0				State1			
Node5	State0		State1		State0		State1		State0		State1		State0		State1	
Node6	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1	State0	State1
State0	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
State1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

- **It is not uncommon for a node to have 10+ parents**
- **This is a serious practical problem for representation, learning, and elicitation**

Two popular classes of solutions

- **Independence of Causal Influences (ICI) gates**

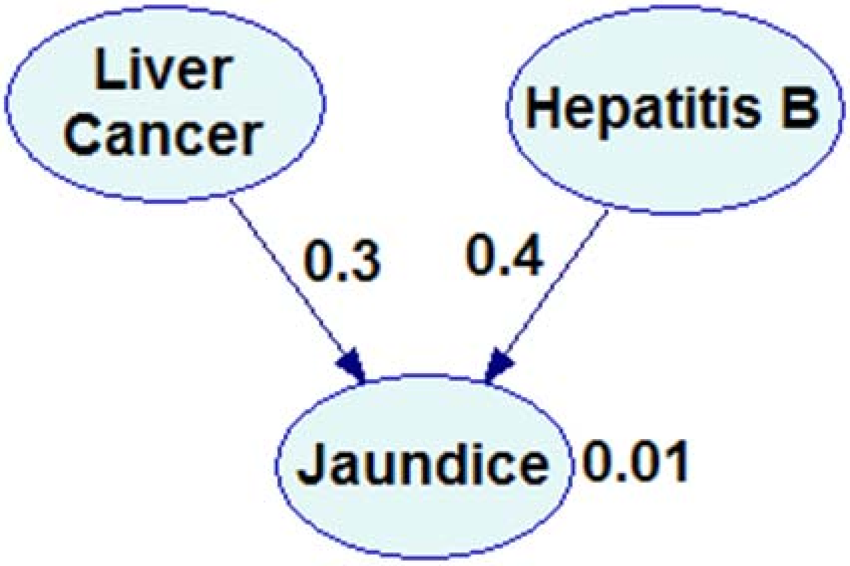
  Assume a model that defines interactions between the parents (causes) to determine the probability over the effect variable (e.g., noisy-OR, noisy-AND)

- **Context Specific Independence (CSI)**

  Suitable when a CPT contains symmetries – independences in some contexts (parents' instantiations)

Independence of Causal Influences: Canonical gates

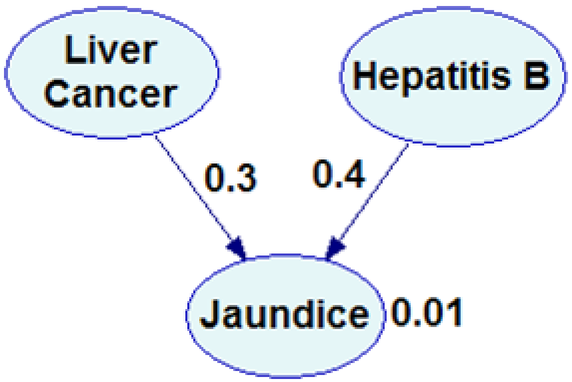
- Reduce the number of parameters from exponential to polynomial in the number of parents
- The most popular canonical gate is Noisy-OR
- In binary case, one numerical parameter q_i per parent plus one "leak" q_0
- The parameters q_i have very clear meaning, due to "amechanistic property"



Independence of Causal Influences: Canonical gates

The parameters q_i can be used to derive the complete CPT

Parent	Liver Cancer	Hepatitis B	LEAK
State	Cancer	Present	
Jaundice	0.3	0.4	0.01
NoJaundice	0.7	0.6	0.99



Liver Cancer	Cancer		NoCancer	
Hepatitis B	Present	Absent	Present	Absent
Jaundice	0.5842	0.307	0.406	0.01
NoJaundice	0.4158	0.693	0.594	0.99

Weakness of Noisy-OR/MAX

- Noisy-OR/MAX and Noisy AND/MIN gates model only positive influences
- Existing proposals to fix this stop short of offering a sound and intuitive combination of positive and negative influences

Related work

Approach 1: "Solving everything"

- Srinivas 1993: "feeding lines model" embodying a world of possible functions
- Heckerman & Breese 1994: decomposable ICI models
- Lucas 2005: 2^{2^n} possible n-ary Boolean functions
- Xiang & Jia 2007: AND and negation (capable of representing any Boolean function)

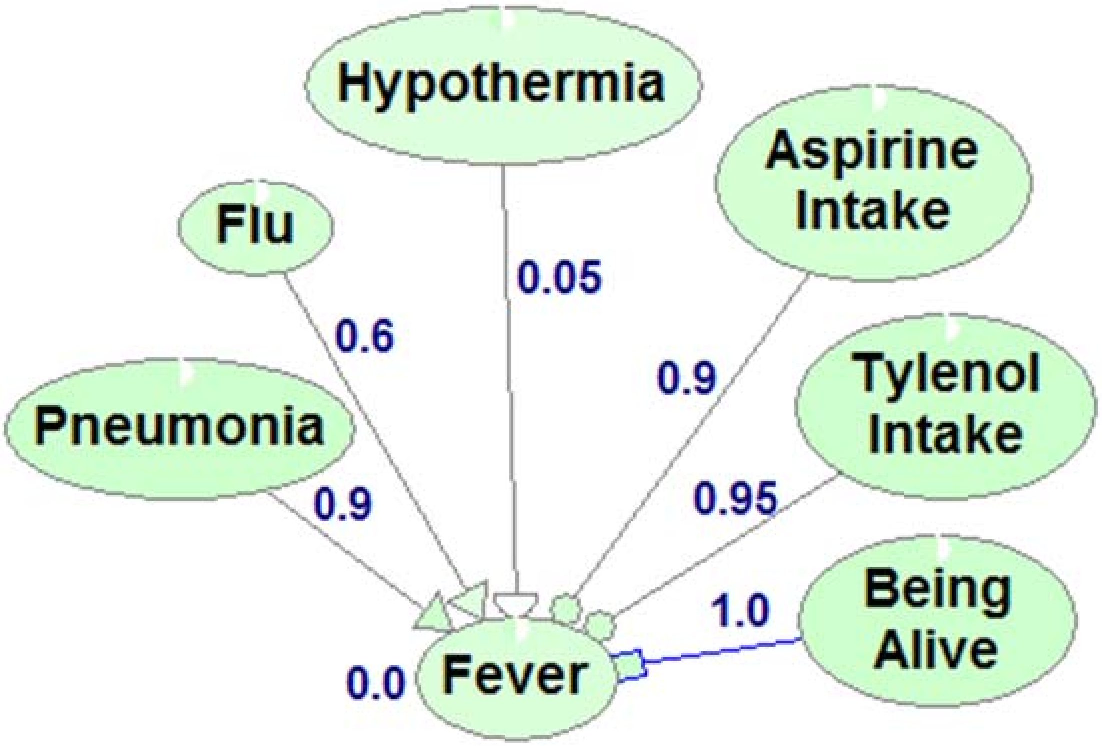
Related work

Approach 2: Attempts to address the problem

- Pearl 1988: introduced "global inhibitors" (a single AND gate at the output of an OR gate)
- Chang et al. 1994: CAST model, combining positive and negative influences (based on Noisy-OR, suffers from unclear probabilistic semantic)
- Lemmer & Gossink 2004: "recursive Noisy-OR," treats positive and negative case separately

This proposal

- A simple, clearly defined gate that can express combinations of positive and negative influences
- The gate refers to simple and clear causal concepts: The starting point is the human side
- Uses OR, AND, and negation (we can do a lot with these ☺)
- Resembles one of the De Morgan's canonical forms (a conjunction of disjuncts)

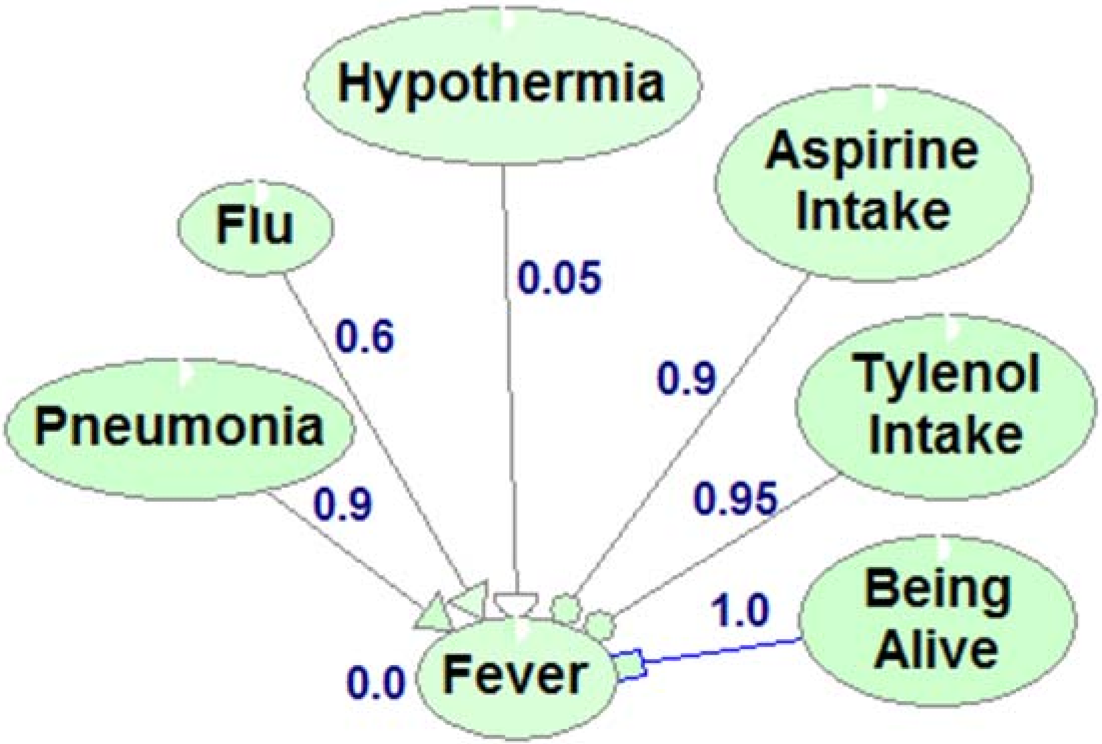


Four types of causal influences

- **Cause**: (Flu, Pneumonia) A positive influence on the child (increases the probability of the child)
- **Barrier**: (Hypothermia) A negative influence on the child (decreases the probability of the child)
- **Requirement**:

  (Being alive) Is required for the child to happen
- **Inhibitor**:

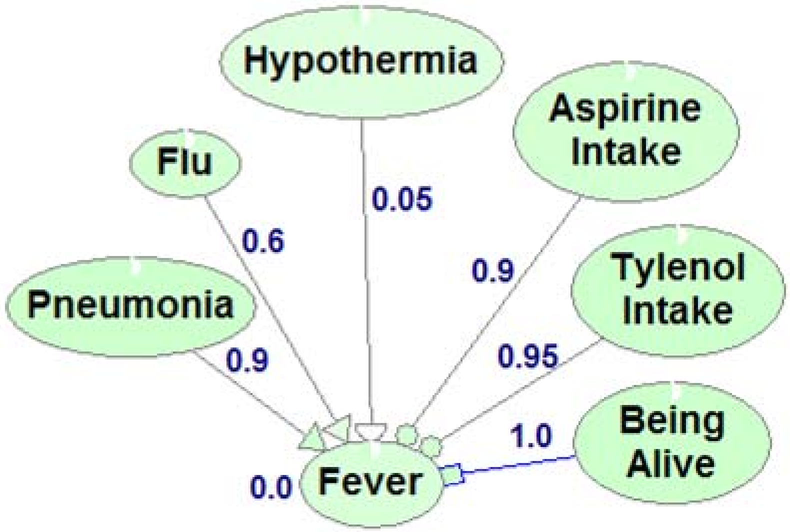
  (Aspirine and Tylenol intake) When present, it prevents the child from happening



Four types of causal influences

Here is how they combine formally in logic

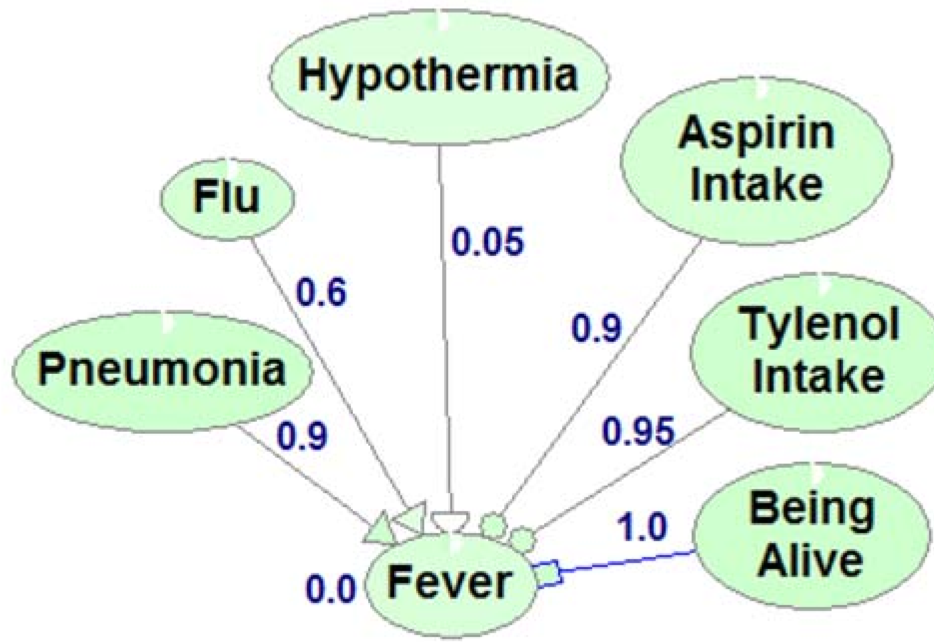
$$F=(C_1|C_2|\ldots|C_i|{\sim}B_1|{\sim}B_2|\ldots|{\sim}B_j)$$
$$\& \ R_1 \ \& \ R_2 \ \& \ \ldots \ \& \ R_k$$
$$\& \ {\sim}I_1 \ \& \ {\sim}I_2 \ \& \ \ldots \ \& \ {\sim}I_l$$



Example questions to experts

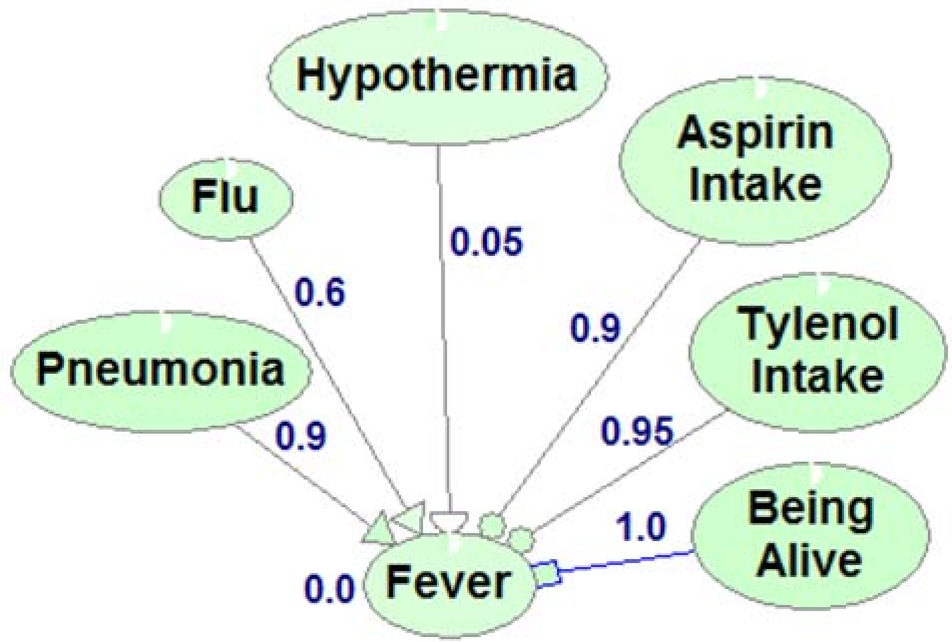
(Please note that there is a natural discrepancy between what one has to say formally and what sounds clear to a human. Each of the questions listed below can be adjusted to the needs of particular context, i.e., their elements can be rephrased or omitted if they do not make sense.)

Leak: ***"What is the probability of fever if (pneumonia, flue, and hypothermia are all absent, the patient takes both aspirin and Tylenol and) the patient is dead?"***



Example questions to experts

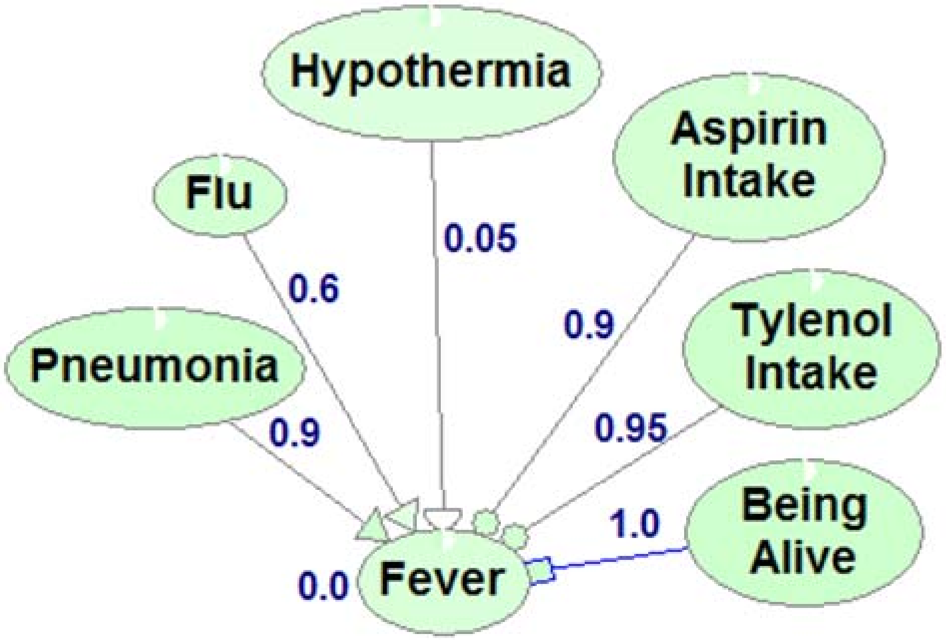
- **Cause**: *"What is the probability of fever in (an alive) patient who has pneumonia, but neither flu nor hypothermia, and does not take any drugs?*
- **Barrier**: *"What is the probability of fever in (an alive) patient with hypothermia who has both flu and pneumonia but takes neither aspirin nor Tylenol?*



Example questions to experts

- **Requirement**: *"What is the probability of no fever in a dead patient (who has both flu and pneumonia but no hypothermia and does not use any drugs)?"*

- **Inhibitor**: *"What is the probability of no fever in (an alive) patient with both flu and pneumonia but no hypothermia if the patient takes aspirin but no Tylenol?"*



From DeMorgan to CPT ☺

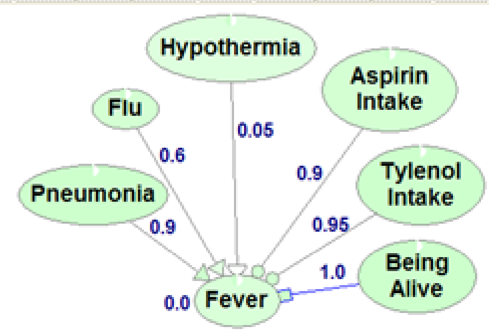
Pneumonia	True																							
Tylenol Intake	True																							
Flu	True								False								True							
Aspirin Intake	True				False				True				False				True				False			
Being Alive	True		False		True		False		True		False		True		False		True		False		True		False	
Hypothermia	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False
True	0.0098	0.00981	0	0	0.098	0.0981	0	0	0.0095	0.009525	0	0	0.095	0.09525	0	0	0.098	0.0981	0	0	0.98	0.981	0	0
False	0.9902	0.99019	1	1	0.902	0.9019	1	1	0.9905	0.990475	1	1	0.905	0.90475	1	1	0.902	0.9019	1	1	0.02	0.019	1	1

…

False									True																
	False								True								False								
	True				False				True				False				True				False				
lse	True		False		True		False		True		False		True		False		True		False		True		False		True
False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	Fa
0	0.095	0.09525	0	0	0.95	0.9525	0	0	0.006	0.0062	0	0	0.06	0.062	0	0	0	0.0005	0	0	0	0.005	0	0	0.06	0.0
1	0.905	0.90475	1	1	0.05	0.0475	1	1	0.994	0.9938	1	1	0.94	0.938	1	1	1	0.9995	1	1	1	0.995	1	1	0.94	0.9

…

False																													
True														False															
True						False								True								False							
e		False				True				False				True				False				True				False			
False		True		False		True		False		True		False		True		False		True		False		True		False		True		False	
True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False	True	False
0	0	0.06	0.062	0	0	0	0.0005	0	0	0	0.005	0	0	0.06	0.062	0	0	0.6	0.62	0	0	0	0.005	0	0	0	0.05	0	0
1	1	0.94	0.938	1	1	1	0.9995	1	1	1	0.995	1	1	0.94	0.938	1	1	0.4	0.38	1	1	1	0.995	1	1	1	0.95	1	1



Empirical validation

Experimental design:

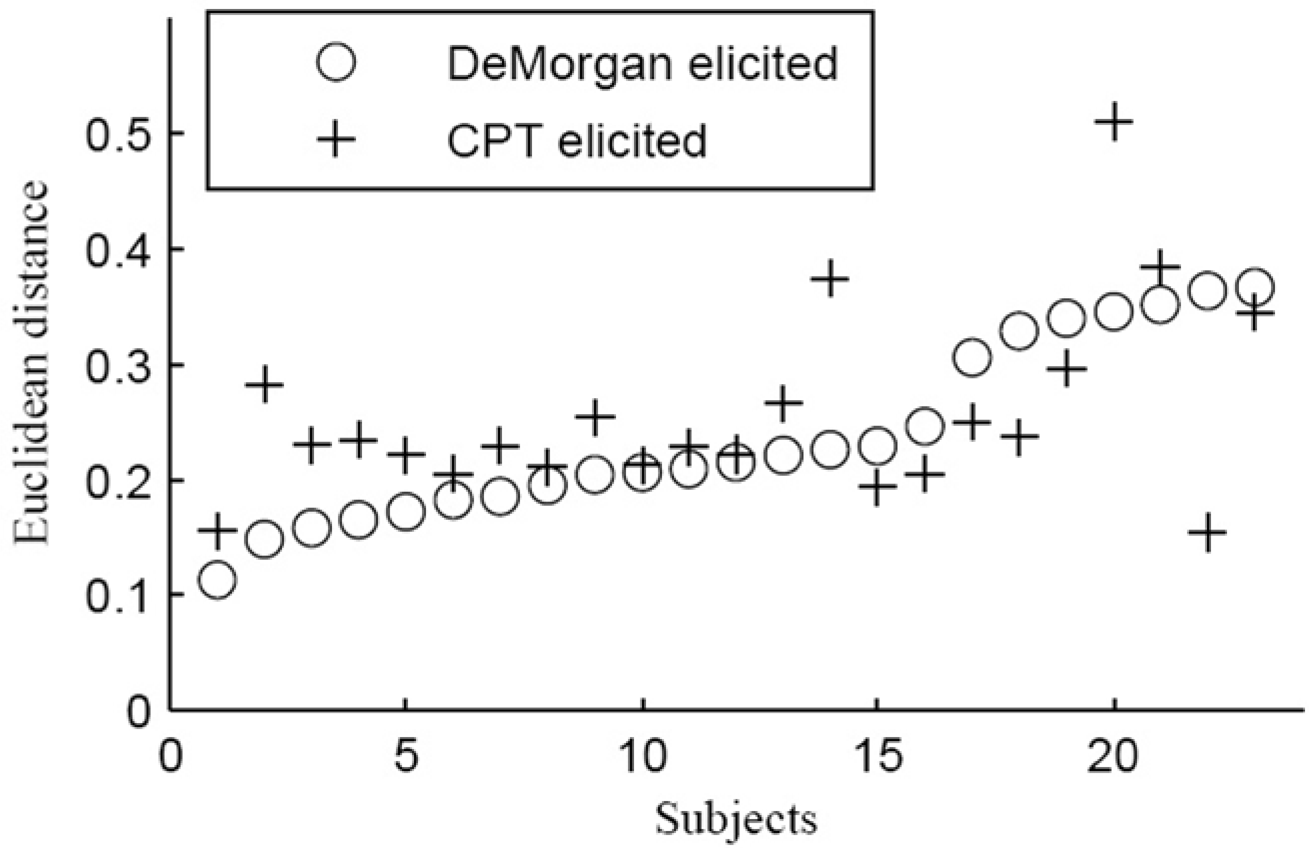
- Elicitation of conditional probability distribution in a gate with four parents
- Methodology proposed by Wang et al. (2002): Have subjects play a computer game (and learn a fictional domain in the process), treat the observed probabilities as a gold standard
- Within-subject design, cross-over study, 25 subjects, students in the DA&DSS class ☺

Results:

- A DeMorgan gate can be elicited in shorter time than a CPT at practically no loss in accuracy
- Results not statistically significant at $\alpha=0.05$ with the current sample size and model size
- Expect stronger effects for larger models

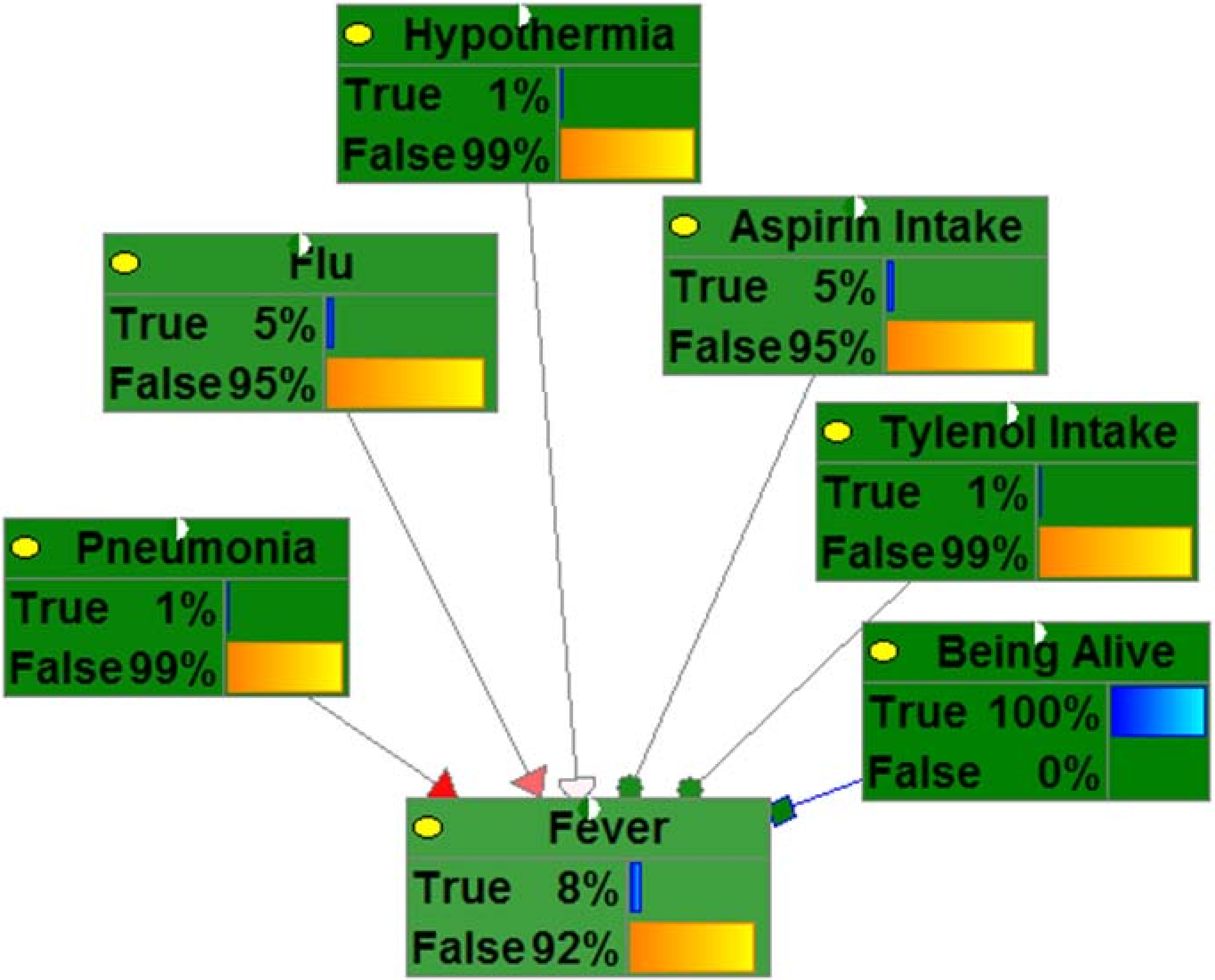
Raw data

•



Euclidean distance from the true distribution for each of the 23 subjects (sorted from smallest to largest)

Embedding the DeMorgan gate in a practical modeling system



Implementation (in GeNIe and SMILE☺)

A developer's environment for graphical decision models (http://genie.sis.pitt.edu/).

Support for model building: ImaGeNIe

Diagnosis: Diagnosis

Qualitative interface: QGeNIe

Learning and discovery module: SMiner

Model developer module: GeNIe.

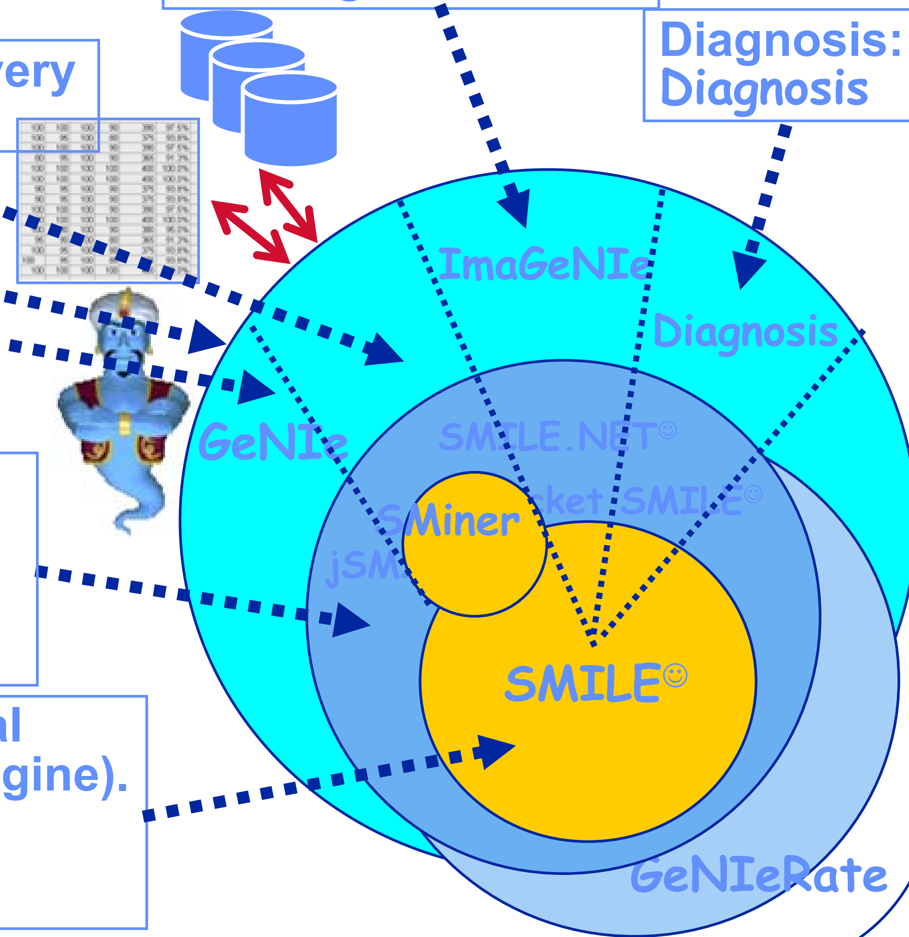
Implemented in Visual C++ in Windows environment.

Wrappers: SMILE.NET☺, jSMILE☺, Pocket SMILE☺

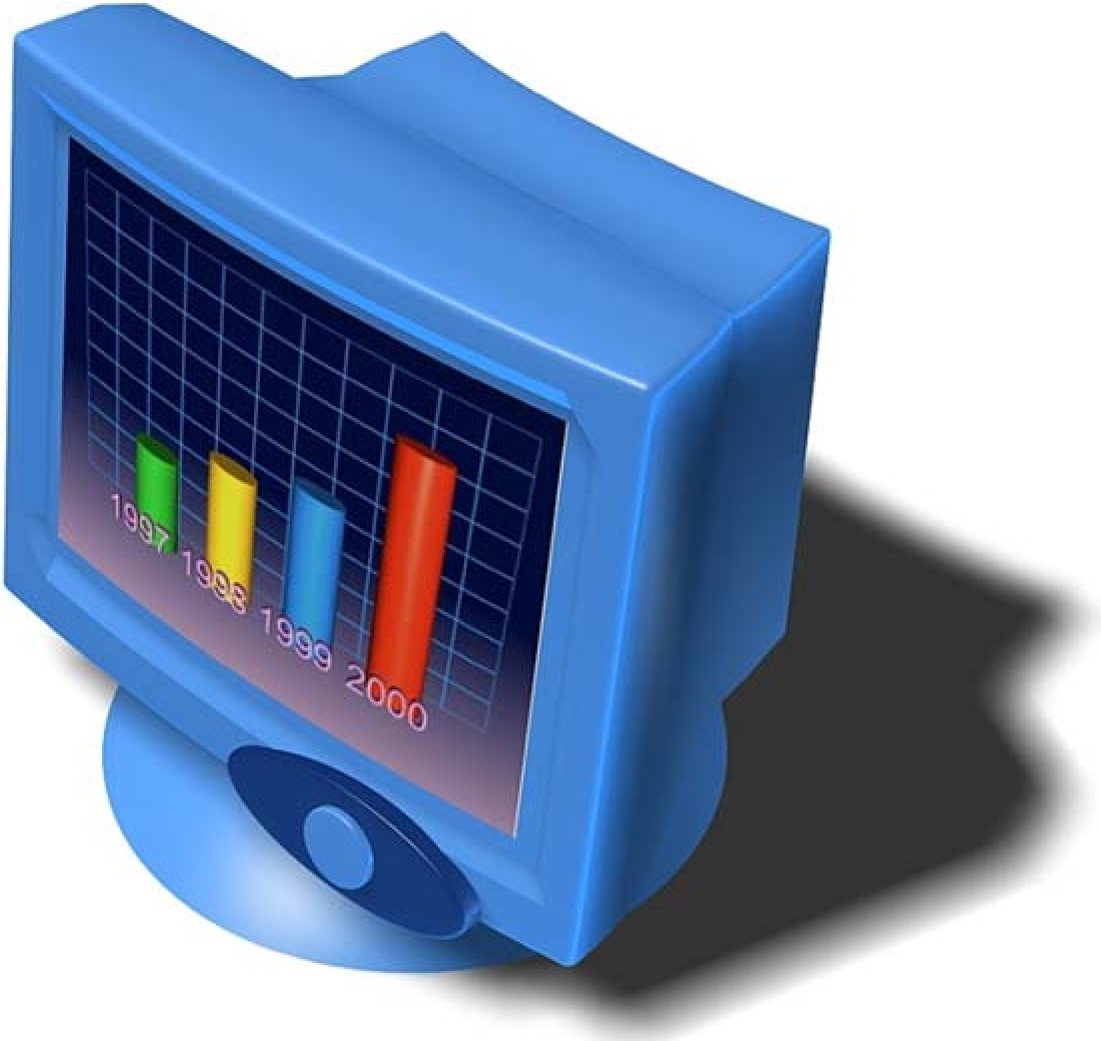
Allow SMILE☺ to be accessed from applications other than C++compiler

Reasoning engine: SMILE☺ (Structural Modeling, Inference, and Learning Engine).

A platform independent library of C++ classes for graphical models.



Demonstration



Concluding remarks

- There are significant advantages that stem from canonical gates in probability elicitation, learning, and computation
- The DeMorgan gate offers simple semantics and is able to express any logical function
- It seems to be fairly intuitive for humans
- We believe that it offers a simple and powerful tool for model building (rapid prototyping)
- We are working on extending the DeMorgan gate to multiple outcomes (along the lines of MAX and MIN gates)