Bayesian Networks

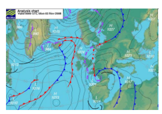
Pascal Wiggers

5-2-2009

TUDelft

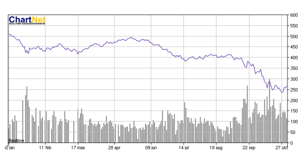
- I have to get to "Landgoed Ehzerwold" by public transport
  - I want to be on time
  - I want to leave home as late as possible
- How should I (*or some decision support system*) plan this trip?

Life is uncertain

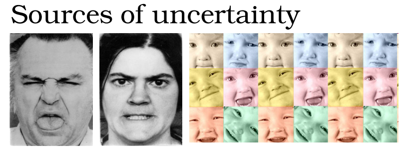


Often, complete information is impossible
Incomplete information leads to uncertainty

- Throwing a dice
- The weather
- Economy
- Diseases
- Global warming
- DNA
- Traffic (jams)
- ...

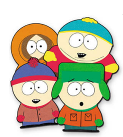


Sources of uncertainty



- Laziness - physical randomness
  - Throwing a coin
- Practical ignorance - hard or impossible to measure
  - Sliding scale (emotions)
- Theoretical ignorance
  - Taxreduction is good for the economy.
  - Increasing the taxes is good for the economy

Uncertainty



- Correlations
  - Television makes children aggressive
- Stereotypes
  - Not all vegetarians vote for the green party
  - Not all people that play golf vote for the conservative party
- Rules of thumb
  - If it rains today, it will rain tomorrow
- Exceptions
  - All birds can fly
  - Pinguins are birds but can not fly

Uncertainty

- Decision support
  - How can I best plan my trip to the Hotel?
- Diagnosis
  - What illness do symptoms indicate?
- Classification
  - Which emotion does someone have?
  - Which task is someone performing?
- Monitoring
  - Are people in a public place getting aggressive?
- Prediction
  - Based on X-ray results, MRI and blood tests, will the patient survive the proposed operation?
  - What will the next word be?
    - "Do you want sugar in your…"

"One sees, from this Essay, that the theory of probabilities is basically just common sense reduced to calculus; it makes one appreciate with exactness that which accurate minds feel with a sort of instinct, often without being able to account for it."

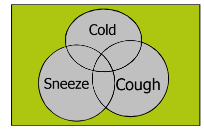
Pierre-Simon Laplace (1749–1827)

De Finetti (1931): If Agent 1 expresses a set of degrees of belief that violate the axioms of probability theory then there is a combination of bets for Agent 2 that **guarantees** that Agent 1 will lose money **every** time.

Probability theory

P(Cold, Sneeze, Cough)

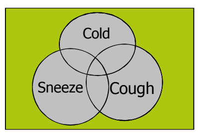
Cold	Sneeze	Cough	
false	false	false	0.300
false	true	false	0.100
false	false	true	0.100
false	true	true	0.050
true	false	false	0.001
true	true	false	0.112
true	false	true	0.112
true	true	true	0.225



Probability theory

P(Cold, Sneeze, Cough)

Cold	Sneeze	Cough	
false	false	false	0.300
false	true	false	0.100
false	false	true	0.100
false	true	true	0.050
true	false	false	**0.001**
true	true	false	**0.112**
true	false	true	**0.112**
true	true	true	**0.225**

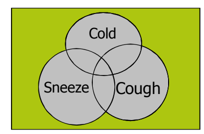


P(Cold = true) = 0.001 + 0.112 + 0.112 + 0.225 = 0.45

Probability theory

P(Cold, Sneeze, Cough)

Cold	Sneeze	Cough	
false	false	false	0.300
false	true	false	0.100
false	**false**	**true**	**0.100**
false	true	true	0.050
true	false	false	0.001
true	true	false	0.112
true	**false**	**true**	**0.112**
true	true	true	0.225



P(Sneeze = false and Cough = true) = 0.100 + 0.112 = 0.212

Marginalization

Throw a dice and a coin

- What is the probability that 6 comes up on the coin?
- P(6) = P(6, heads) + P(6, tails) = 1/6 * ½ + 1/6 * ½ = 1/6
- In general:

$$P(A) = \sum_{B} P(AB)$$

Marginalisation

$$P(Cold, Sneeze) = \sum_{Cough} P(Cold, Sneeze, Cough)$$

Cold	Sneeze	Cough	
false	false	false	0.3
false	true	false	0.1
false	false	true	0.1
false	true	true	0.05
true	false	false	0.001
true	true	false	0.112
true	false	true	0.112
true	true	true	0.225

Cold	Sneeze	
false	false	0.3+0.1 =0.4
false	true	0.1+0.05 =0.15
true	false	0.001+0.112 =0.113
true	true	0.112+0.225 =0.337

Conditional distributions

$$P(Cold|Sneeze) = \frac{P(Cold, Sneeze)}{P(Sneeze)}$$

Cold	Sneeze	
false	false	0.4
false	true	0.15
true	false	0.113
true	true	0.337

Sneeze	
false	0.513
true	0.487

Cold	Sneeze	
false	false	0.4 / 0.513 = 0.78
true	false	0.113 / 0.513 = 0.22
false	true	0.15 / 0.487 =0.31
true	true	0.337 / 0.487 = 0.69

Probabilistic reasoning

The probability of one or more random variables given some evidence.

$$P(X_Q|X_E) = \frac{P(X_Q, X_E)}{P(X_E)} = \frac{\sum_{X_H} P(X_Q, X_H, X_E)}{\sum_{X_H, X_Q} P(X_Q, X_H, X_E)}$$

Full joint distribution

Cold	Sneeze	Cough	Allergy	
false	false	false	false	?
false	false	false	true	?
false	true	false	false	?
false	true	false	true	?
false	false	true	false	?
false	false	true	true	?
false	true	true	false	?
false	true	true	true	?
true	false	false	false	?
…				

The product rule

P(Sneeze, Cough|Cold) = P(Cold, Sneeze, Cough) / P(Cold)

P(Cold, Sneeze, Cough) = P(Sneeze, Cough|Cold) P(Cold)

= P(Sneeze|~~Cough~~, Cold) P(Cough|Cold) P(Cold)

Are Sneezing and Coughing dependent?

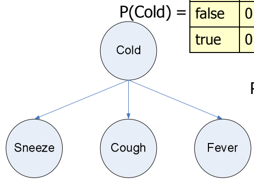
Yes!

Unless… We know the reason behind the dependency
The common cause

Bayesian networks

P(Cold) =

Cold	
false	0.95
true	0.05



P(Sneeze|Cold)=

Sneeze	Cold	
false	false	0.7
true	false	0.3
false	true	0.05
true	true	0.95

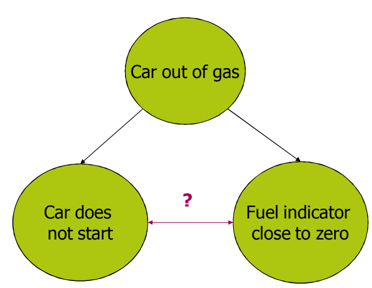
P(Cough|Cold)=

Cough	Cold	
false	false	0.7
true	false	0.3
false	true	0.03
true	true	0.97

P(Fever|Cold)=

Fever	Cold	
false	false	0.95
true	false	0.05
false	true	0.3
false	true	0.7

Independence



Bayes' rule

From the productrule it follows that:

$$p(H|E)p(E) = p(E|H)p(H)$$

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}$$

Bayes' rule tells us how our belief in some event H should change in the light of new evidence E

Bayes rule

Why would it be any easier to obtain/use P(E|H) than to use P(H|E) directly?

P(Cold|Sneeze) Diagnostic: from symptom to cause

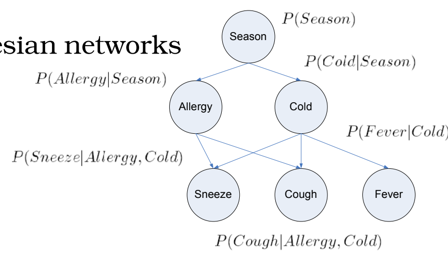
There is a sudden epidemic of cold P(Cold) goes up

How does P(Cold|Sneeze) change?

P(Sneeze|Cold) Causal: not affected by the epidemic, it models the way colds work

Diagnostic knowledge is often more fragile than causal knowledge

Bayesian networks



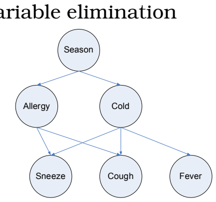
The network represents the joint probability distribution:

$P(Sneeze, Cough, Fever, Cold, Allergy, Season)$

With 2+4+4+8+8+4=30 entries instead of 2^6=64 entries.

Bayesian networks – Variable elimination

$$P(X_Q|X_E) = \alpha \sum_{X_H} P(X_Q, X_H, X_E)$$



$$P(C|Sn = truc)$$

$$= \alpha \sum_{Se,A,Cou,F} P(Se, A, C, Sn, Cou, F)$$

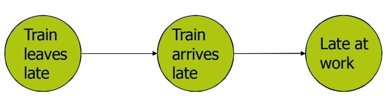
$$= \alpha \sum_{Se,A,Cou,F} P(Se)P(A|Se)P(C|Se)P(Sn|A,C)P(Cou|A,C)P(F|C)$$

$$= \alpha \sum_{A} \left[P(Sn|A,C) \sum_{Cou} (P(Cou|A,C)) \sum_{Se} (P(Se)P(A|Se)P(C|Se)) \sum_{F} (P(F|C)) \right]$$

Bayesian networks

IF Train leaves late THEN Train arrives late

IF Train arrives late THEN Late at work



Leaves late	
false	0.8
true	0.2

Arrives late	Leaves late	
false	false	0.9
true	false	0.1
false	true	0.4
true	true	0.6

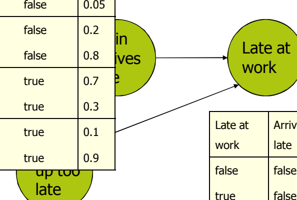
Late at work	Arrives late	
false	false	0.95
true	false	0.05
false	true	0.2
true	true	0.8

…twerken

IF … …s late

IF … …k

IF … …rk



Late at work	Train arrives late	Wake up too late	
false	false	false	0.95
true	false	false	0.05
false	true	false	0.2
true	true	false	0.8
false	false	true	0.7
true	false	true	0.3
false	true	true	0.1
true	true	true	0.9

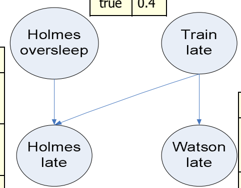
Late at work	Arrives late	
false	false	0.95
true	false	0.05
false	true	0.2
true	true	0.8

Bayesian networks - use

`Explaining away'

O	
false	0.6
true	0.4

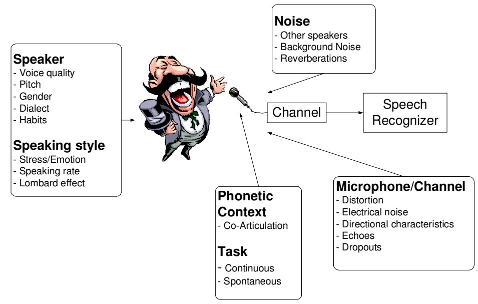
Train	
false	0.9
true	0.1



H	O	Train	
false	false	false	0.9
true	false	false	0.1
false	true	false	0.5
true	true	false	0.5
false	false	true	0.4
true	false	true	0.6
false	true	true	0.2
true	true	true	0.8

W	Train	
false	false	0.9
true	false	0.1
false	true	0.2
true	true	0.8

Why is speech recognition difficult?



What to do, what to do?

- Many degrees of freedom
- But as the numbers get smaller things work well

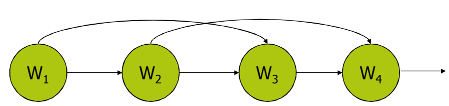
- Relatively simple models
- Much unused knowledge

- Use the context to reduce complexity of the task
- Add sufficient structure

Context

- 10 year old girl vs. elderly man
  - gender, dialect, age, educational level, professional background
- Conversation vs. presentation vs. debate
  - Speaking style, choice of words
- Topic of conversation
  - Vocabulary

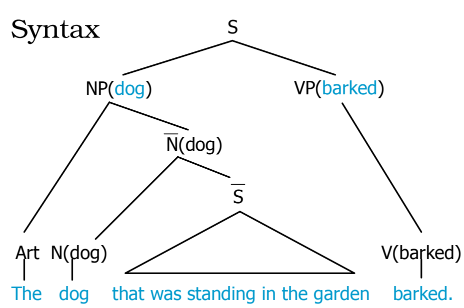
Sentence structure



The dog barked.

The dog that was standing in the garden barked.

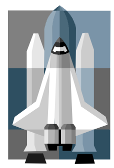
Syntax



Semantics

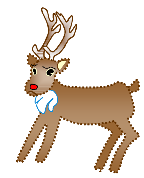
- The Astronauts stepped into the space shuttle.
- The Astronauts stepped into the great puddle.

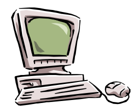
- In Africa one can meet a lion face to face.
- In Africa one can meet alien face to face.



Pragmatics

- Rudolf the red nose reindeer.
- Rudolf the Red knows rain, dear.
- Rudolf the Red Nose reigned here.



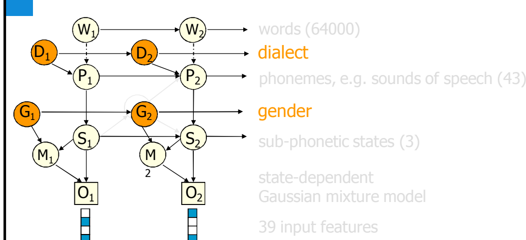


- This new display can recognize speech.
- This nudist play can wreck a nice beach.

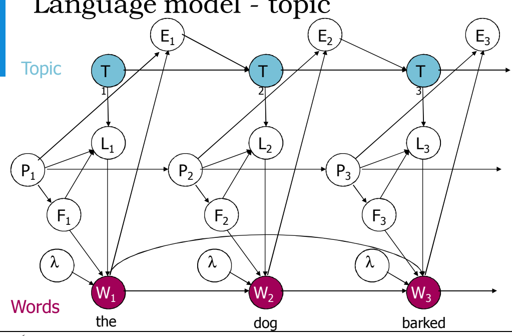
Nothing is impossible… it's just that not everything is equally probable

- A tasty brick
- I cannot eat my car
- Colorless green ideas sleep furiously
- The <uhm> dog that stood…was standing in the garden, well it <uh> barked.

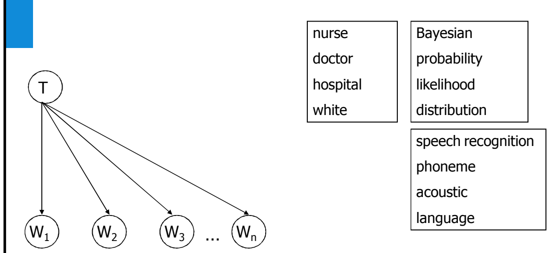
Dynamic Bayesian networks (DBNs)



Language model - topic



Language model – topic



0.25		0.25		0.25		0.25	
F1	0.025	speech	0.05	politics	0.04	coffee	0.07
Ferrari	0.05	recognition	0.05	politician	0.02	sugar	0.03
engine	0.01	sounds	0.02	speech	0.005	tea	0.05
crash	0.001	words	0.02	spoken	0.001	spoken	0.0001
…		spoken	0.005	debate	0.01	politics	0.0002
		debate	0.0001	…		…	
		…					

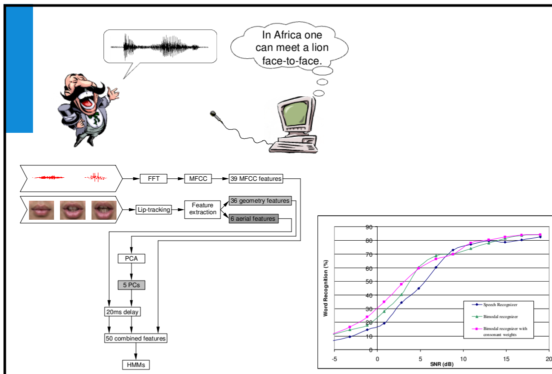
Speech

0.22 ↓		0.3 ↑		0.26 ↑		0.22 ↓	
F1	0.025	speech	0.05	politics	0.04	coffee	0.07
Ferrari	0.05	recognition	0.05	politician	0.02	sugar	0.03
engine	0.01	sounds	0.02	speech	0.005	tea	0.05
crash	0.001	words	0.02	spoken	0.001	spoken	0.0001
…		spoken	0.005	debate	0.01	politics	0.0002
		debate	0.0001	…		…	
		…					

Speech recognition

0.21 ↓		0.33 ↑		0.25 ↓		0.21 ↓	
F1	0.025	speech	0.05	politics	0.04	coffee	0.07
Ferrari	0.05	recognition	0.05	politician	0.02	sugar	0.03
engine	0.01	sounds	0.02	speech	0.005	tea	0.05
crash	0.001	words	0.02	spoken	0.001	spoken	0.0001
…		spoken	0.005	debate	0.01	politics	0.0002
		debate	0.0001	…		…	
		…					

Speech recognition finds the words



What is the probability that the sun will still exist tomorrow?

- Unknown, since there hasn't been any experiment that tested the existence of the sun *tomorrow*
- 1, because the sun did exist on all previous days
- 1-e, where e is the percentage of stars in the universe that explodes per day
- Depends on the age, temperature, size and type of the sun
- (d + 1)/(d + 2), where d is the number of days that the sun existed (Laplace)

Questions? → Go to lunch