Lecture

Foundations of Artificial Intelligence

Part 7 – Uncertainty

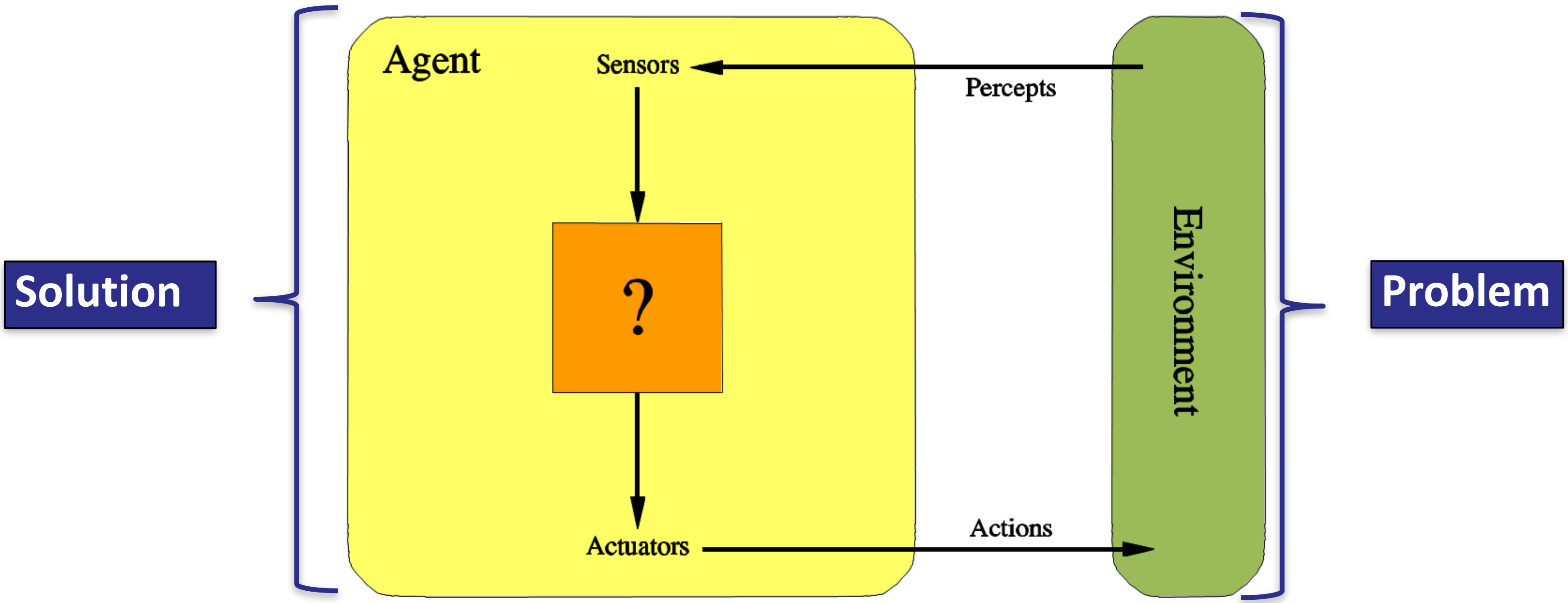
Dr. Mohsen Mesgar

Universität Duisburg-Essen

Recall

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken



Recall …

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

- Two major optimization problems
  - **Optimal path finding:** start and goal states are clearly defined
  - **Optimal goal state finding:** start state is given but the goal state is unknown. Instead we **have a set of constraints that should be satisfied**

Recall …

Environment variables

- **Observability**: fully vs partial
- **Certainty**: deterministic vs strategic vs stochastic
- **Temporal succession**: episodic vs sequential
- **Continuity**: static vs dynamic
- **Scale**: discrete vs continuous
- **Population**: single-agent vs multi-agent

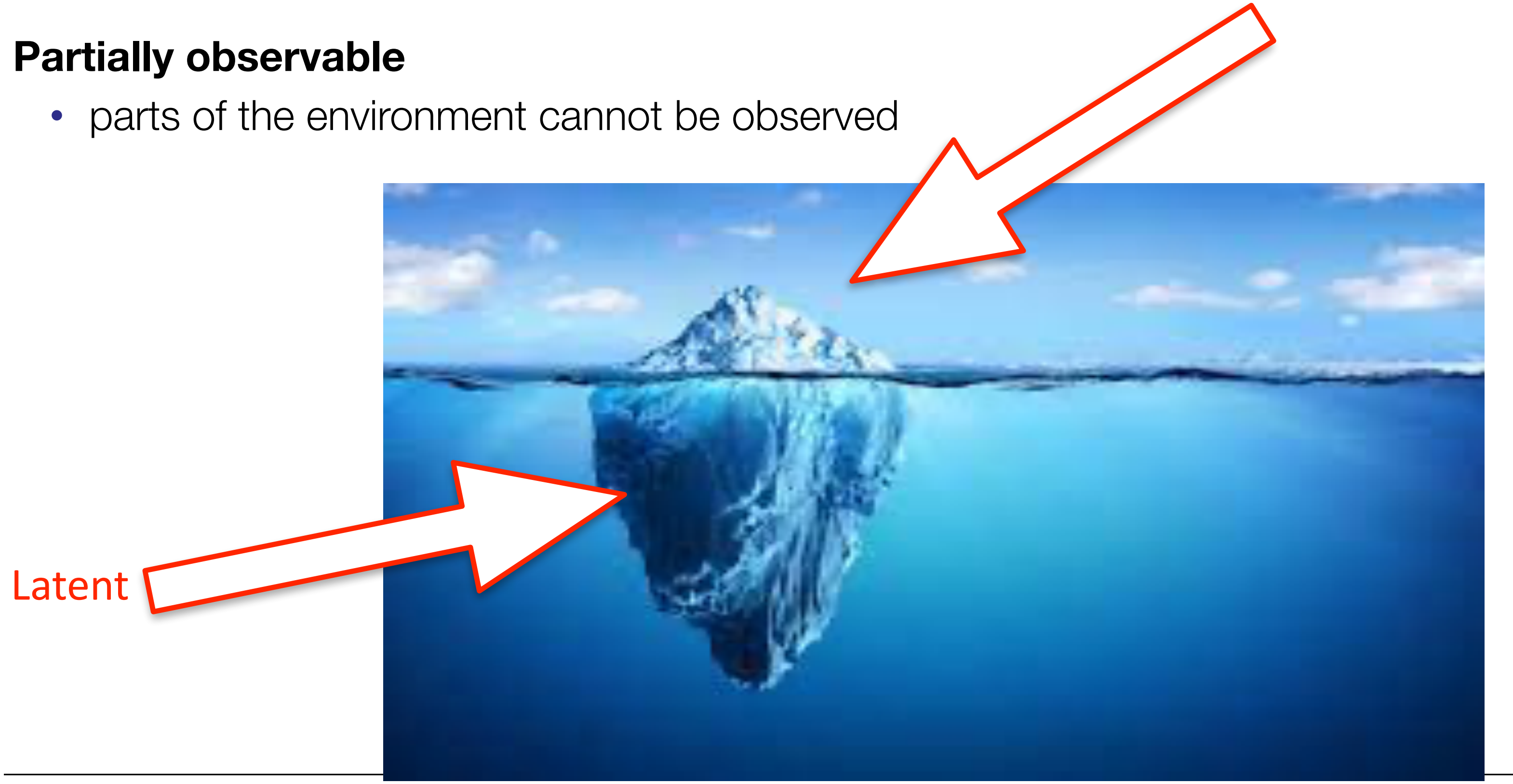
Observability

Fully observable

- the complete state of the environment can be observed (relevant parts)
- no need to keep track of internal states

Partially observable

- parts of the environment cannot be observed



https://mewiseminar.wordpress.com/2014/01/11/zdf-mit-dem-zweiten-sieht-man-besser/

Certainty

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Deterministic

- next environment state is determined only by current state and executed action

Strategic

- only the opponents' actions cannot be foreseen

Stochastic

- next environment state is uncertain (e.g., throwing dices).

https://commons.wikimedia.org/wiki/File:Two_red_dice_01.svg

Any other open questions?

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Motivation

In previous lectures, we developed agent functions for environments that are **fully observable and deterministic**

- states true or false (maybe unknown) and
- Given state and an action, next state is always predictable

Unfortunately, the real world is not like that

- the whole truth about the world is not accessible
- the next state in many environments is not predictable **certainly**

→ agents must deal with uncertainty in many environments

Certainty

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Deterministic

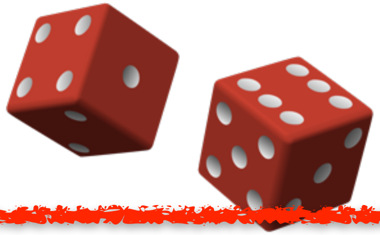
- next environment state is determined only by current state and executed action

Strategic

- only the opponents' actions cannot be foreseen

Stochastic

- next environment state is **uncertain** (e.g., throwing dices).



In this lecture you learn about …

- **Uncertainty**
  - What is it?
  - Why is it happening?
  - How to deal with it?
- **Probability**
  - Prior probability
  - Posterior probability
  - Joint probability
  - Bayes’ rule

Today

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

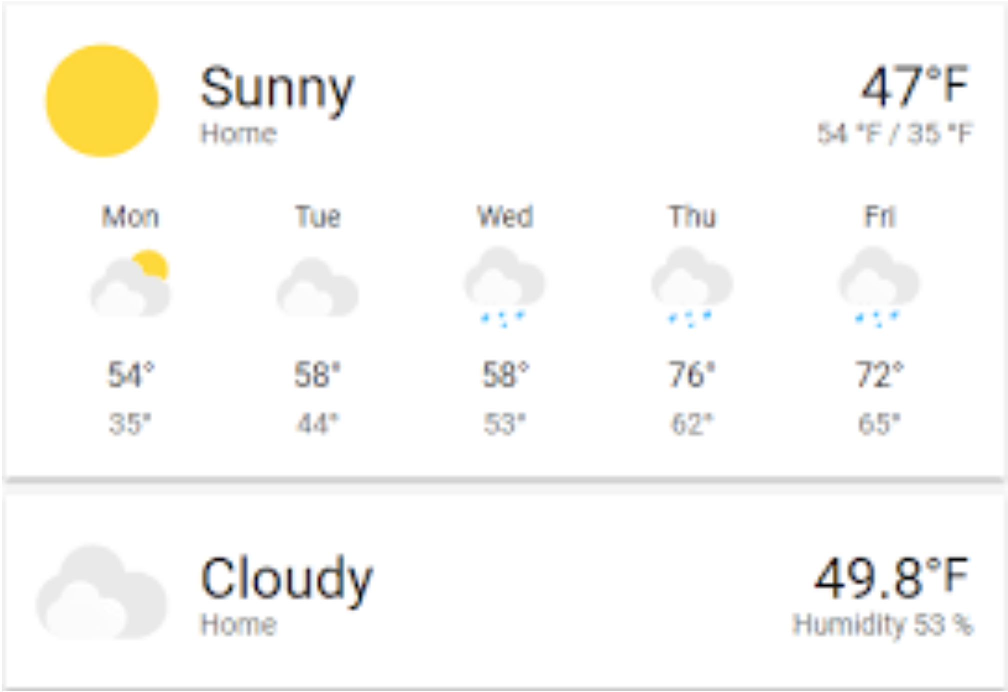
Uncertainty

Uncertainty

- The action in the backgammon game is to throw the dices
- The state of the board depends on the output of the action
  - Although we know all possibilities for the output of the action
  - We cannot be certain, which possible numbers show up on dices

Uncertainty

- Weather forecasting might be certain for today
- It's uncertain for tomorrow and next days



Uncertainty

- Many different actions for getting to the airport:
- Action A_t = leave for the airport t minutes before departure
- typical questions:
  - Will action A_t get me to the airport in time?
  - Which action is the best choice?

Uncertainty

- Refers to situations involving imperfect or unknown information
- Happens because of the lack of knowledge
- Uncertainty arises in partially observable or stochastic environments, as well as due to ignorance, indolence, or both.

Problems with uncertainty

Risks involved in plan "A_{90} will get me to the airport in time"

- **partial observability** (road state, other drivers' plans, etc.)
- **noisy sensors** (traffic reports may be wrong)
- **uncertainty in action outcomes** (flat tire, accident, etc.)
- **immense complexity** of modeling and predicting traffic

An **ideal rational plan**

A_{90} will get me to the airport in time as long as my car doesn't break down, I don't run out of gas, no accident, the bridge doesn't fall down, …

- impossible to model **all things that can go wrong**

A **more cautious plan**

- A_{1440} will get me to the airport in time
- will (almost) certainly succeed, but clearly suboptimal
  → we have to pay for a night in a hotel

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Dealing with Uncertainty

Probabilities

Probabilities are one way of handling uncertainty

- e.g. A_{90} will get me to the airport in time with probability 0.5

The probability summarizes effects that are due to

- **Theoretical ignorance:** our theory might be incomplete
  - e.g., we cannot completely model the weather
- **Practical ignorance:** even if we know all theoretical odds, we might not be aware of the actual current situation.
  - e.g. “Today starts a big festival just next to the airport.”

Probabilities and beliefs

Probabilities are related to one's beliefs

- A probability p attached to a statement means that I believe that the statement will be true in $p * 100\%$ of the cases
  - Example: I assume a 0.1 probability of traffic jam on the A40
    → there might be jam (10% of cases), but usually there is none
- It does not mean that it is true with $p\%$
  - Example: the traffic on the A40 is jammed with a degree of 10%

Probabilities and beliefs

Probabilities are related to one's beliefs

- A probability p attached to a statement means that I believe that the statement will be true in $p * 100\%$ of the cases
  - Example: I assume a 0.1 probability of traffic jam on the A40
    → there might be jam (10% of cases), but usually there is none
- It does not mean that it is true with $p\%$
  - Example: the traffic on the A40 is jammed with a degree of 10%

Probability theory is about **degree of belief**.

- other techniques (e.g., fuzzy logic) deal with **degree of truth**

Probability and knowledge

Probability of an event depends on how much knowledge about other events

P(A_{25} at airport in time | no reported accidents) = 0.06

- in 6% of the days I get to the airport in 25 minutes if no accidents are reported

P(A_{25} at airport in time | no reported accidents, 5 a.m.) = 0.15

- chances are higher at 5 in the morning

Making decisions under uncertainty

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Suppose I believe the following

- P(A_{25} at airport on time |current knowledge) = 0.04
- P(A_{90} at airport on time | current knowledge) = 0.70
- P(A_{120} at airport on time | current knowledge) = 0.95
- P(A_{1440} at airport on time | current knowledge) = 0.9999

Which action should I choose?

We make a decision based on our preferences.

- How bad would it be to miss the flight?
- How bad would it be to wait for an hour at the airport?

Utility theory is used to represent and infer preferences.

Decision theory = probability theory + utility theory

Today

Probability

Terminology

- **Experiment (or trail): i**s any procedure that produces a definite outcome that cannot be predicted with certainty. An experiment is said to be *random* if it has more than **one possible** outcome,
- **Outcome: a possible** result of an experiment
- **Sample space (S):** the set of all possible outcomes of that experiment.
- **Event:** a property that we are studying in an experiment
- **Event space:** A subset from the sample space with a common property.

- **Probability Theory:** We are interested in estimating the probability of occurring an event = observing an outcome that belongs to the event space

Flipping a coin

- **Experiment (or trail):** flipping a coin (https://g.co/kgs/dmoetH)
- **Outcome:** one possible output is “H”
- **Sample space:** S = {“H”,”T”}
- **Event (A):** How likely is to see a head?
- **Event space:** **A** = {“H”}
- **Event (B):** How likely is to see a tail?
- **Event space:** B = {“T”}

Rolling a dice

- **Experiment (or trail):** rolling a dice (https://www.calculator.net/dice-roller.html)
- **Outcome:** one possible output is "1"
- **Sample space:** S = {"1", "2", "3", "4", "5", "6"}
- **Event A:** How likely is it to see an even number?
- **Event space:** A = {"2", "4", "6"}
- **Event B:** How likely is it to see a number greater than 3?
- **Event space:** B = {"4", "5", "6"}

Identifying gender and job of a person

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- **Experiment (or trail):** randomly select a person and ask about her job and/or gender
- **Outcome:** one possible output is "Female/Teacher"
- **Sample space for job:** S1 = {Teacher, Healer}
- **Sample Space for gender:** S2 = {Male, Female}
- **Event A:** How likely is it to see a female worker?
- Event space A1 = {Female}
- **Event B:** How likely is it to see a teacher?
- Event space B = {Teacher}

Probability of an event

Given a sample space for an experiment, we are interested in the probability of a specific event

There are different ways to estimate the probability of an event

- **Count-based**
- Neural models
- …

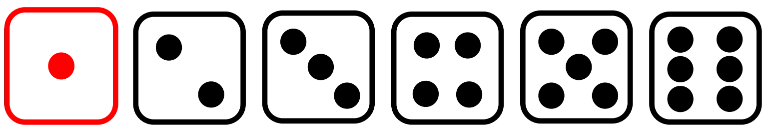
Prior probability

Prior probability: probability of an event in a sample space without any additional knowledge

$$P(A) = \frac{|\text{Event Space}|}{|\text{Sample Space}|}$$

Example

- Probability of rolling "1" with a single 6-sided fair dice?



$$P\left(\frac{\text{“good“ events}}{\text{all events}}\right) = \frac{1}{6}$$

Prior probability

Prior probability: probability of an event without any additional knowledge

- p(female) = 18.3 million/44 million = 0.42
- p(teacher) = 673225/44 million = 0.015
- p(healer) = 45000/44 million = 0.001

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225		
holistic healer	45000*		

Boy-or-girl problem

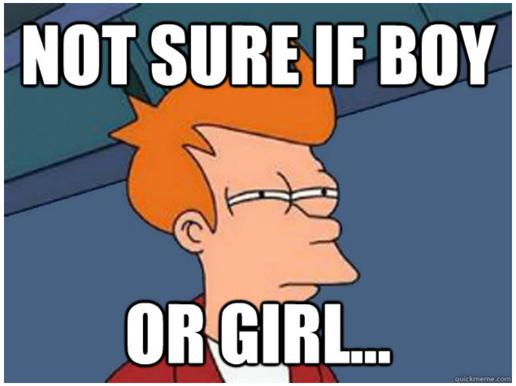
Experiment: Mr. Jones has two children.

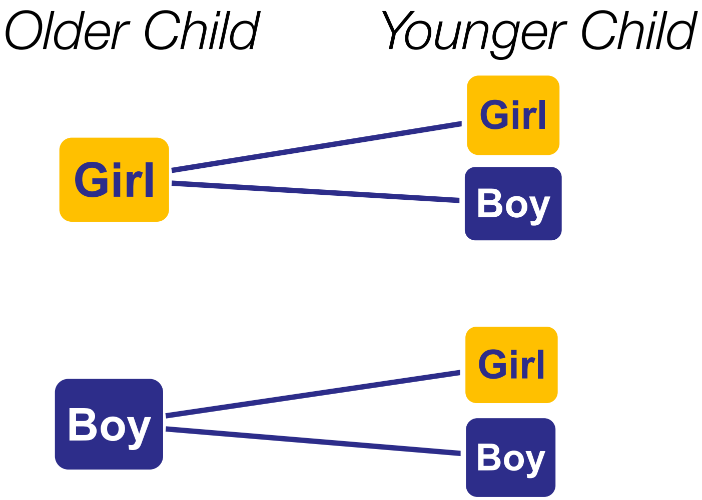
Outcome: {B, G}

Sample Space: $\{BG, BB, GB, GG\}$

Event Space: {GG}

$$p = \frac{\text{event space}}{\text{sample space}} = \frac{1}{4}$$





Boy-or-girl problem

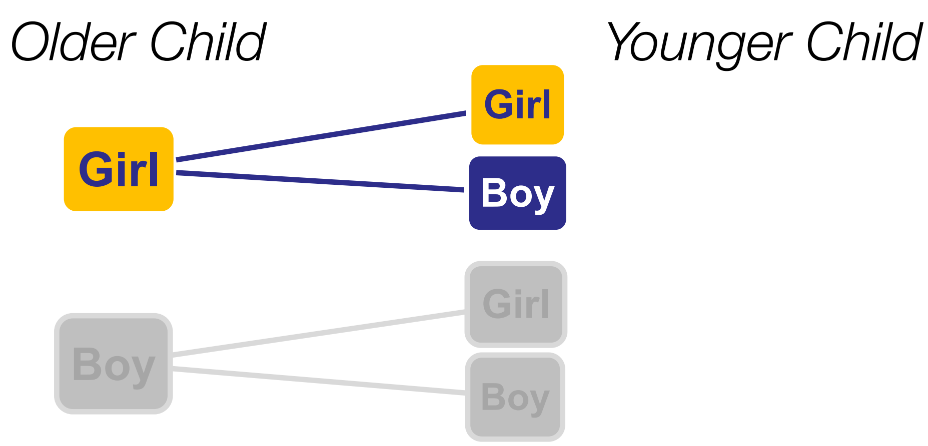
Experiment: Mr. Jones has two children. The older child is a girl.

Event: What is the probability that both children are girls?

Sample Space: $\{GB, GG\}$

True Events: $\{GG\}$

$$p = \frac{\textbf{event space}}{\textbf{sample space}} = \frac{1}{2}$$



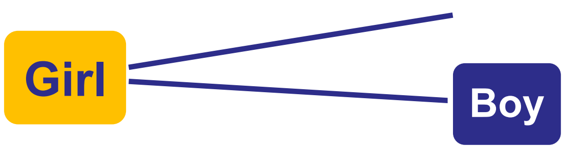
Boy-or-girl problem

Experiment: Mr. Jones has two children. At least one of them is a girl.

Event: What is the probability that both children are girls?

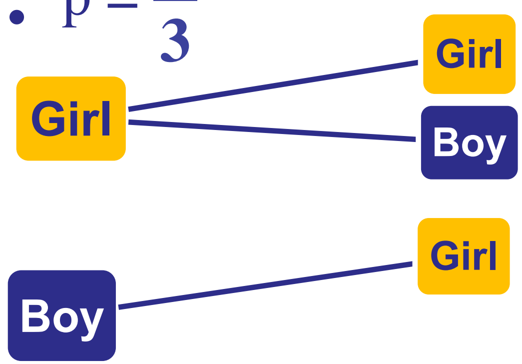
First way of reasoning

- Sample Space: $\{GB, GG\}$
- Event Space: {GG}
- $p = \frac{1}{2}$



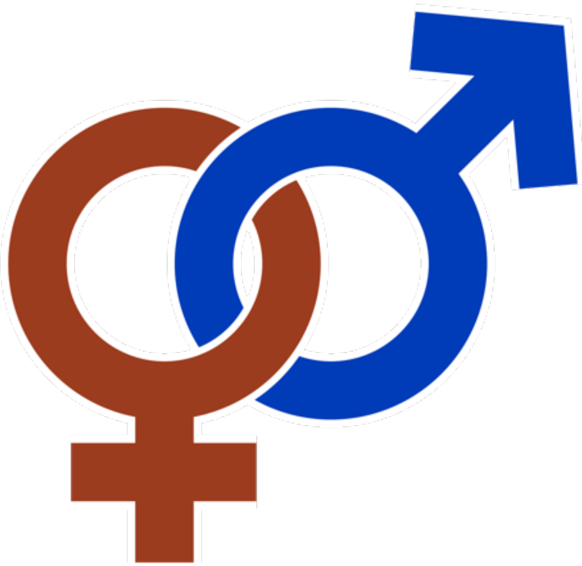
Second way of reasoning

- Sample Space: $\{GB, GG, BG\}$
- Event Space: **{GG}**
- $p = \frac{1}{3}$



Paradox

- If two ways of reasoning lead to different answers for the same question, it is often called a paradox.
- It has been shown that the answers for the boy-or-girl paradox vary when the wording of the question is modified slightly.



Today

Posterior Probability

Posterior probability

How likely is an event after incorporating knowledge that *B* is true.

- P(teacher | femaleWorker): how likely is it that Sarah is a teacher?
- P(tumor | headache): I have a headache. Do I have a tumor?
- P(pregnant | noPeriod): A girl's period is three days overdue. Is she pregnant?
- P(passExam | noSubmissions): I haven‘t done any exercises. Will I pass?

Posterior probability

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Conditional probability: probability of an event A given that another event B has occurred.

P (A | B)

- *“the conditional probability of A given B”*
- *“the probability of A under the condition B”*

P(teacher | femaleWorker): if we take a female worker from our sample, how likely is it that she is a teacher?

P(teacher, femaleWorker): if we take a worker from our sample, how likely is it that the worker is female and a teacher.

Posterior probability

$$P(A \mid B) = \frac{P(A,B)}{P(B)}$$

$$P(femaleworker \mid teacher) = \frac{P(femaleworker, teacher)}{P(teacher)}$$

$$P(femaleworker \mid teacher) = \frac{\frac{491454}{44million} = 0.011}{0.015} = 0.73$$

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225	181 771	491 454
holistic healer	45000*	N/A	N/A

Joint Probability

Joint probability

Joint probability: probability of two events occurring together

- $P(teacher, femaleworker)$

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225		
holistic healer	45000*		

Joint probability

- Multiplication rule for joint probabilities of dependent variables:

$$P(A,B) = P(A) * P(B|A)$$

$$P(\text{femaleWorker}, \text{teacher}) = P(\text{femaleWorker}) * P(\text{teacher}|\text{femaleWorker})$$

$$= \mathbf{0.42} * \frac{\mathbf{491454}}{\mathbf{18.3\ million}} = 0.011$$

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225	181 771	491 454
holistic healer	45000*	N/A	N/A

Joint probability for independent events

$$P(A, B) = P(A) \times P(B)$$

P(teacher, femaleWorker) = P (teacher) $*$ P (femaleWorker)

- P(teacher, femaleWorker) = 0.015 * 0.42 = 0.0063

→ this would lead to an estimate of 277 200 female teachers

Not independent!

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225	181 771	491 454
holistic healer	45000*	N/A	N/A

Bayes' rule

- Estimating the probability of an event, based on prior knowledge of conditions that might be related to the event

$$P\left(A \middle| B\right) = \frac{P(B|A) * P(A)}{P(B)}$$

Today

Bayes' Rule

Conditional probability

- $P(\text{teacher} \mid \text{femaleWorker}) = \dfrac{P(\text{femaleWorker} \mid \text{teacher}) * P(\text{teacher})}{P(\text{femaleWorker})}$

- $P(\text{teacher} \mid \text{femaleWorker}) = \dfrac{\frac{491454}{673225} * 0.015}{0.42} = 0.026$

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

	total	male	female
working population (our sample)	44 million	25.7 million	18.3 million
teacher	673 225	181 771	491 454
holistic healer	45000*	N/A	N/A

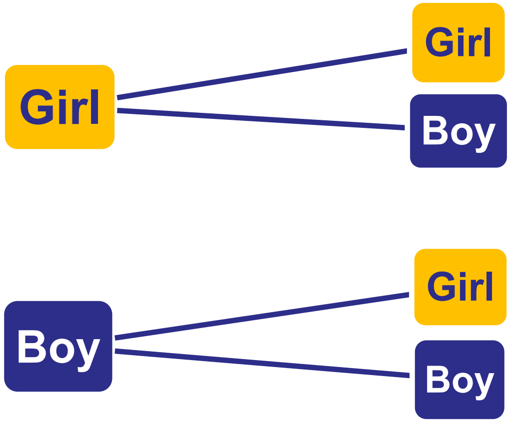
Back to girls and boys

Sample Space: $\{BG, BB, GB, GG\}$

- $P(GG) = \frac{1}{4} = 0.25$
- $P(BB) = \frac{1}{4} = 0.25$
- $P(B) = \frac{3}{4} = 0.75$
- $P(G) = \frac{3}{4} = 0.75$

- $P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$
- $P(GG|G) = \frac{P(G|GG) * P(GG)}{P(G)}$
- $P(GG|G) = \frac{1 * 0.25}{0.75} = \frac{1}{3}$

Given one child is a girl, how likely is it that the other child is a girl, too?



Today

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Knowledge reduces uncertainty

Monty Hall Problem

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Scenario: You're on a game show, and you are supposed to choose one of three doors: Behind one door is a car; behind the other two are goats. You pick a door, say number 1, and the host, who knows what's behind each door, opens another door, say number 2, which reveals a goat.

He then asks you: *"Do you want to switch your choice to door No. 3?"*
What do you do??

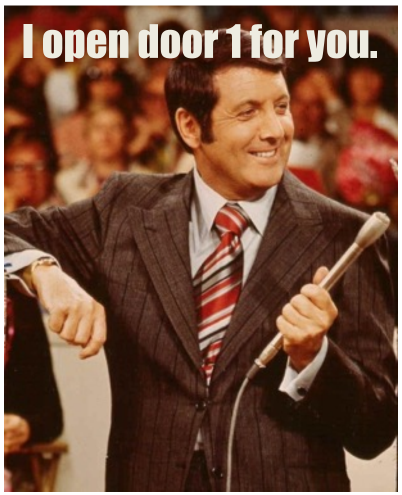
Monty Hall Problem

Monty Hall Problem

Monty Hall Problem

P() $= \frac{1}{3}$ P() $= \frac{1}{3}$ P() $= \frac{1}{3}$

Monty Hall Problem



What would you do?

stay? or switch?

Monty Hall Solution

Switching is always better than sticking to your primary selected door.

Why?

.

Monty Hall Solution

P () = $\frac{1}{3}$

P () = $\frac{2}{3}$

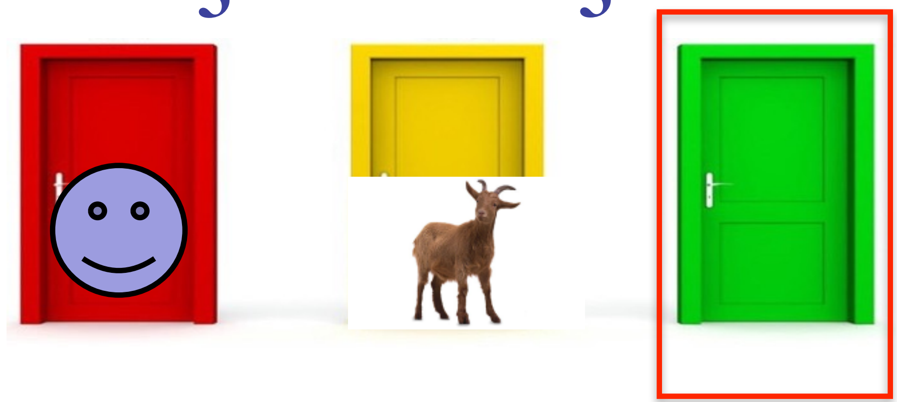
Monty Hall Solution

P () $= \dfrac{1}{3}$

P () $= \dfrac{2}{3}$

Monty Hall Solution

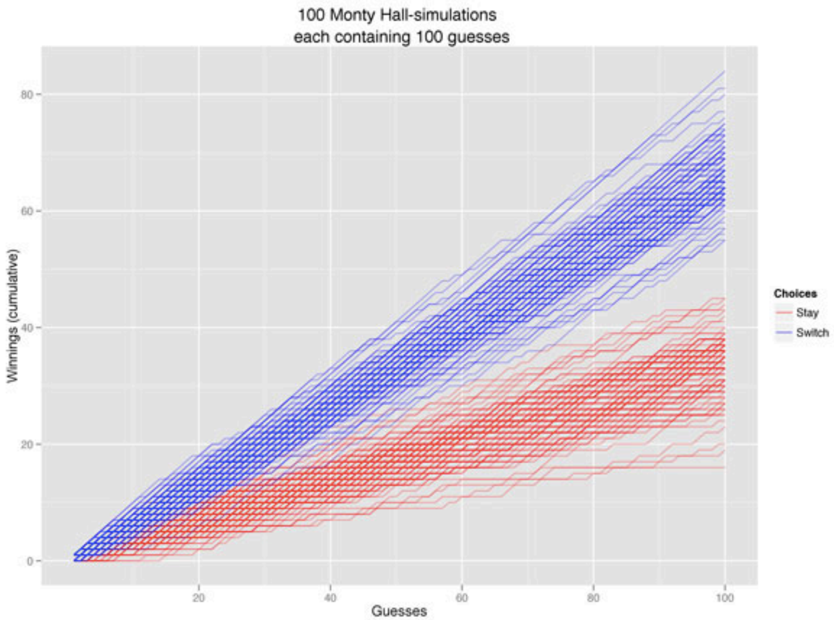
P () = $\frac{1}{3}$ P () = $\frac{0}{3}$ P () = $\frac{2}{3}$

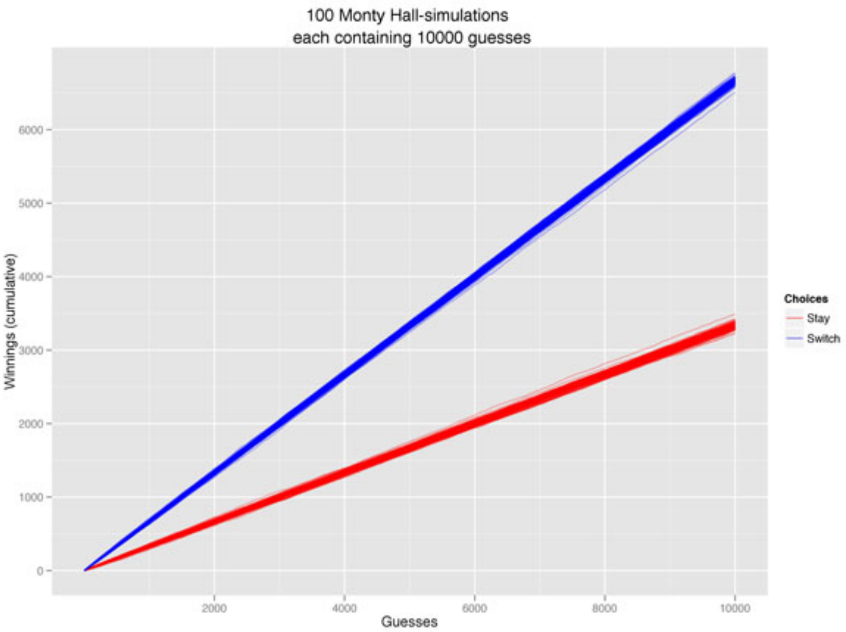


Monty Hall Solution

Simulation

Simulation





http://rpsychologist.com/monty-hall-simulation

Summary

- **Uncertainty**
  - What is it?
  - Why is it happening?
  - How to deal with it?
- **Probability**
  - Prior probability
  - Posterior probability
  - Joint probability
  - Bayes' rule

Readings

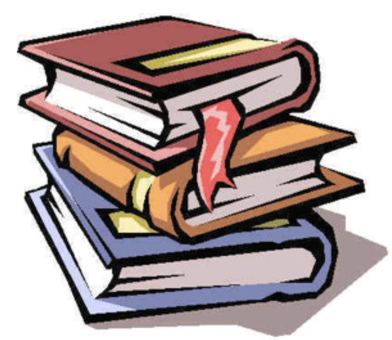
- Christopher D. Manning & Hinrich Schütze (1999): *Foundations of Statistical Natural Language Processing*

Mandatory

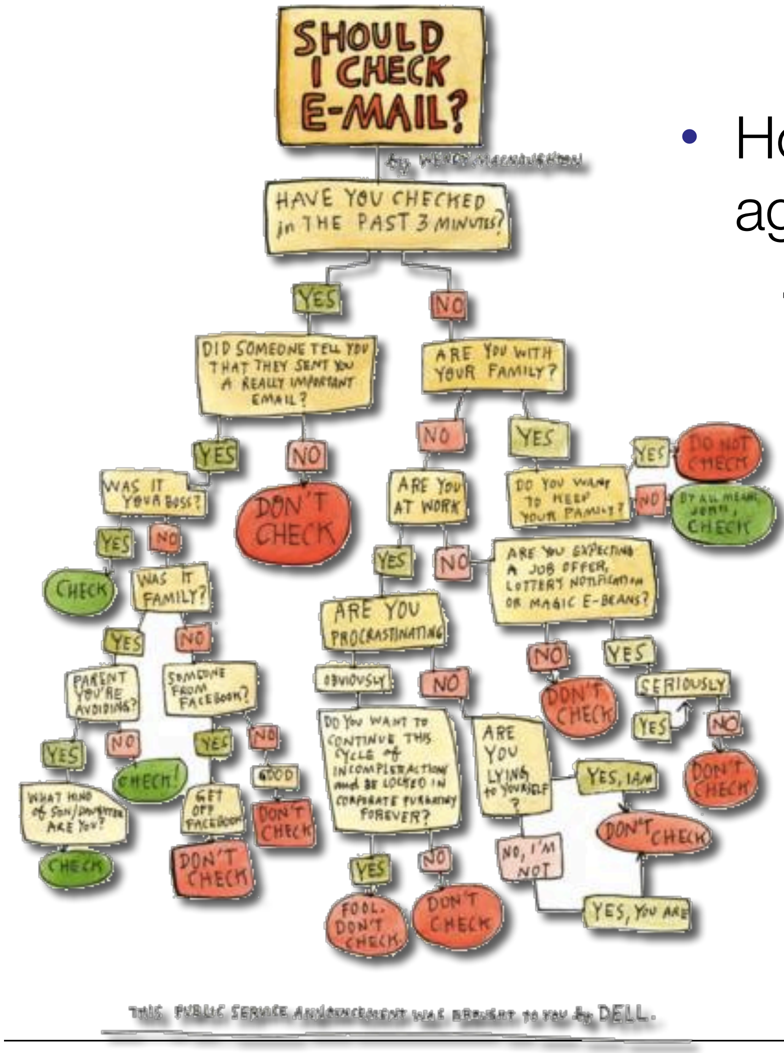
- Chapter 2: *Mathematical Foundations*,
  - 2.1 *Elementary Probability Theory*, p.40-50.

Optional

- Rest of chapter 2.



Next lecture



- How do we use probabilities to let an agent act in an uncertain environment?
  - → Machine learning basics

Thank You