Algorithms for context prediction in Ubiquitous Systems

Introduction to probability theory

Stephan Sigg

Institute of Distributed and Ubiquitous Systems
Technische Universität Braunschweig

November 11, 2008

Overview and Structure

- Introduction to context aware computing
- Basics of probability theory
- Algorithms
  - Simple prediction approaches: ONISI and IPAM
  - Markov prediction approaches
  - The State predictor
  - Alignment prediction
  - Prediction with self organising maps
  - Stochastic prediction approaches: ARMA and Kalman filter
  - Alternative prediction approahces
    - Dempster shafer
    - Evolutionary algorithms
    - Neural networks
    - Simulated annealing

Overview and Structure

- Introduction to context aware computing
- **Basics of probability theory**
- Algorithms
  - Simple prediction approaches: ONISI and IPAM
  - Markov prediction approaches
  - The State predictor
  - Alignment prediction
  - Prediction with self organising maps
  - Stochastic prediction approaches: ARMA and Kalman filter
  - Alternative prediction approaohces
    - Dempster shafer
    - Evolutionary algorithms
    - Neural networks
    - Simulated annealing

Outline

Basics of probability theory

1. Introduction
2. Notation
3. Calculation with probabilities

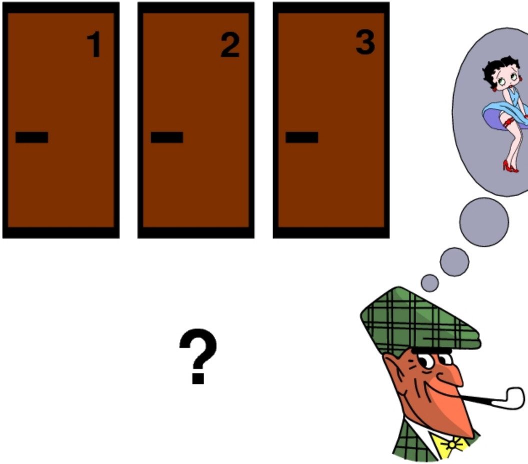
Probability in everyday life

We are confronted with Probability constantly:

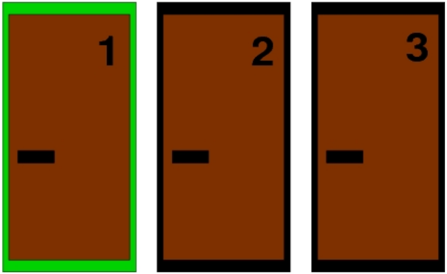
- Weather forecasts
- Quiz shows
- ...

Example

The treasure behind the doors



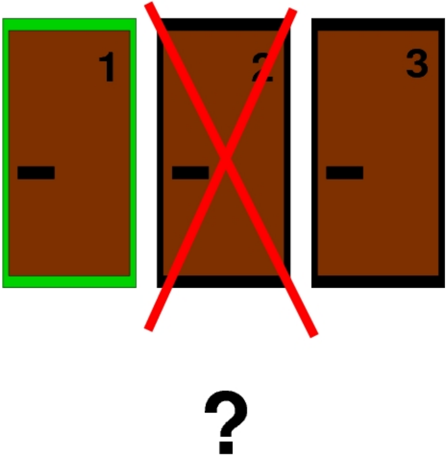
Example
The treasure behind the doors



?

Example

The treasure behind the doors



Example

The treasure behind the doors

- What shall the candidate do?
  - Alter his decision?
  - Retain his decision?
  - Does it make a difference?

Example

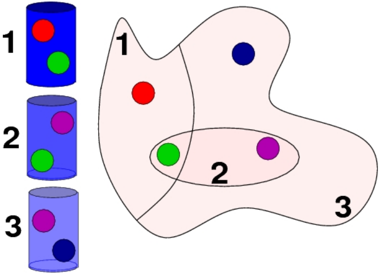
The treasure behind the doors

- What shall the candidate do?
  - Alter his decision?
  - Retain his decision?
  - Does it make a difference?
- We will consider the solution to this Problem in some minutes

Outline

Notation

Experiments, Events and sample points



- The results of experiments or observations are called events.
- Events are sets of sample points.
- The sample space is the set of all posible events.

Example

Sample spaces

- Three distinct balls (a,b,c) are to be placed in three distinct bins.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	abc			ab	ab	c		c		ac	ac	b	
2		abc		c		ab	ab		c	b		ac	ac
3			abc		c		c	ab	ab		b		b

14	15	16	17	18	19	20	21	22	23	24	25	26	27
b		bc	bc	a		a		a	a	b	b	c	c
	b	a		bc	bc		a	b	c	a	c	a	b
ac	ac		a		a	bc	bc	c	b	c	a	b	a

Example

Sample spaces

- Suppose that the three balles are not distinguishable.

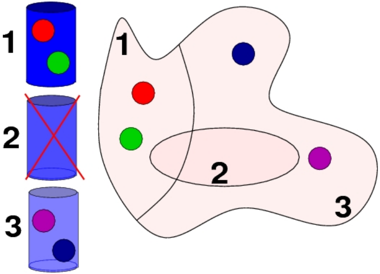
Event / Bin	1	2	3	4	5	6	7	8	9	10
1	***			**	**	*		*		*
2		***		*		**	**		*	*
3			***		*		*	**	**	*

Example
Sample spaces

- Indistinguishable sample spaces and indistinguishable bins

Event / Bin	1	2	3
1	***	**	*
2		*	*
3			*

Impossible events



Impossible event

With $\chi = \{\}$ we denote the fact that event χ contains no sample points. It is impossible to observe event χ as an outcome of the experiment.

Probability of events

Probability of events

Given a sample space Π and an event $\chi \in \Pi$, the occurence probabiltiy $P(\chi)$ of event χ is the sum probability of all sample points from χ:

$$P(\chi) = \sum_{x \in \chi} P(x). \tag{1}$$

Statistical independence

Independence

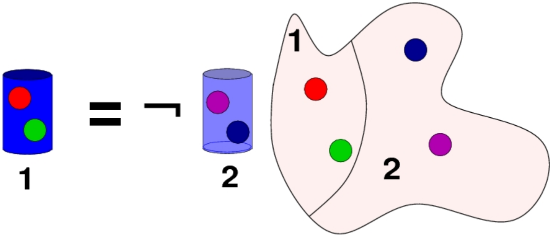
A collection of events χ_i that form the sample space Π is independent if for all subsets $S \subseteq \Pi$

$$P\left(\bigcap_{\chi_i \in S} \chi_i\right) = \prod_{\chi_i \in S} P(\chi_i). \tag{2}$$

- Statistical independence is required for many useful results in probability theory.
- Be careful to apply such results not in cases where independence between sample points is not provided.

Outline

Negation of events

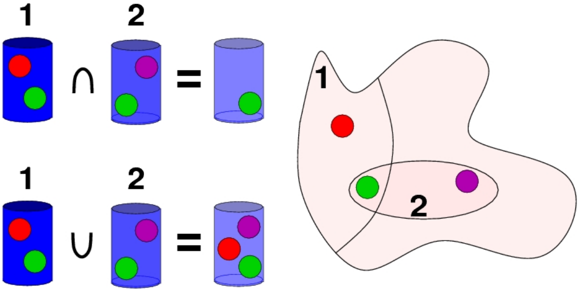


For every event χ there is an event $\neg\chi$ that is defined as 'χ does not occur'.

Negation of events

The event consisting of all sample points x with $x \notin \chi$ is the complementary event (or negation) of χ and is denoted by $\neg\chi$.

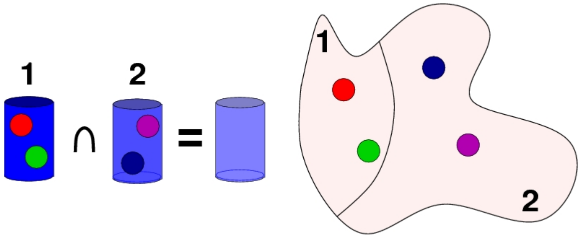
Subsumming events



$$\chi_1 \cap \chi_2 = \{x | x \in \chi_1 \land x \in \chi_2\} \tag{3}$$

$$\chi_1 \cup \chi_2 = \{x | x \in \chi_1 \lor x \in \chi_2\} \tag{4}$$

Mutual exclusive events



Mutual exclusive events

When the events χ_1 and χ_2 have no sample point x in common, the event $\chi_1 \cap \chi_2$ is impossible: $\chi_1 \cap \chi_2 = \{\}$.
The events χ_1 and χ_2 are mutually exclusive.

Combining probabilities

- To compute the probability $P(\chi_1 \cup \chi_2)$ that either χ_1 or χ_2 or both occur we add the occurence probabilities

$$P(\chi_1 \cup \chi_2) \leq P(\chi_1) + P(\chi_2) \tag{5}$$

Combining probabilities

- To compute the probability $P(\chi_1 \cup \chi_2)$ that either χ_1 or χ_2 or both occur we add the occurence probabilities

$$P(\chi_1 \cup \chi_2) \leq P(\chi_1) + P(\chi_2) \tag{5}$$

- The '$\leq$'-relation is correct since sample points might be contained in both events:

$$P(\chi_1 \cup \chi_2) = P(\chi_1) + P(\chi_2) - P(\chi_1 \cap \chi_2). \tag{6}$$

Example
Coin tosses

Question

What is the probability that in two toin cosses either head occurs first or tail occurs second ?

Example
Coin tosses

Events	coin tosses	probability
head - head		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
head - tail		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
tail - head		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
tail - tail		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Example
Coin tosses

Events	coin tosses	probability	sum probability
head - head		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$	
head - tail		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$	
tail - tail		$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$	$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

Conditional probability

Conditional probability

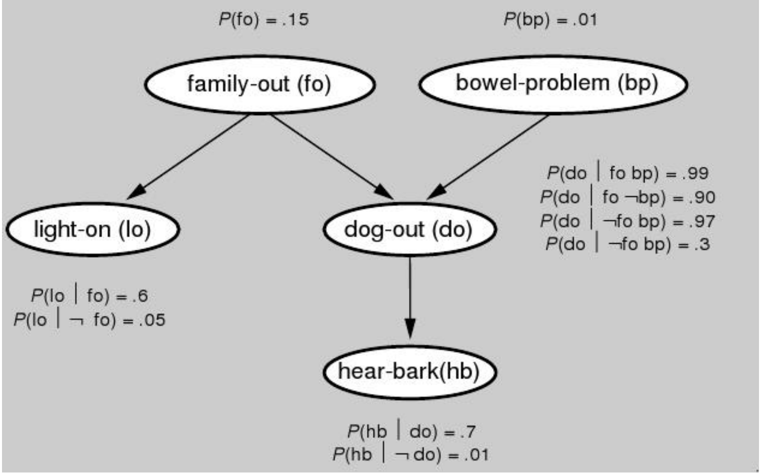
The conditional probability of two events χ_1 and χ_2 with $P(\chi_2) > 0$ is denoted by $P(\chi_1|\chi_2)$ and is calculated by

$$\frac{P(\chi_1 \cap \chi_2)}{P(\chi_2)} \tag{7}$$

$P(\chi_1|\chi_2)$ describes the probability that event χ_2 occurs in the presence of event χ_2.

Example

Conditional probability



Bayes Rule

With rewriting and some simple algebra we obtain the bayes rule:

Bayes Rule

$$P(\chi_1|\chi_2) = \frac{P(\chi_2|\chi_1) \cdot P(\chi_1)}{\sum_i P(\chi_2|\chi_i) \cdot P(\chi_i)}. \quad (8)$$

- This equation is useful in many statistical applications.
- With Bayes rule we can calculate $P(\chi_1|\chi_2)$ provided that we know $P(\chi_2|\chi_1)$ and $P(\chi_1)$.

Expectation

Expectation

The expectation of an event χ is defined as

$$E[\chi] = \sum_{x \in \mathbb{R}} x \cdot P(\chi = x) \tag{9}$$

Example
Expectation

Example

Consider the event χ of throwing a dice. The Sample space is given by $S_\chi = \{1, 2, 3, 4, 5, 6\}$.

What is the expectation of this event?

Example

Expectation

Example

Consider the event χ of throwing a dice. The Sample space is given by $S_\chi = \{1, 2, 3, 4, 5, 6\}$.

What is the expectation of this event?

- The expectation of this event is

$$E[\chi] = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = 3.5 \tag{10}$$

Calculation with expectations

Linearity of expectation

For any two random variables χ_1 and χ_2,

$$E[\chi_1 + \chi_2] = E[\chi_1] + E[\chi_2]. \tag{11}$$

Multiplying expectations

For an independent random variables χ_1 and χ_2,

$$E[\chi_1 \cdot \chi_2] = E[\chi_1] \cdot E[\chi_2]. \tag{12}$$

Variance

Variance

The variance of a random variable χ is defined as

$$var[\chi] = E[(\chi - E[\chi])^2]. \tag{13}$$

Calculation with variance

Add variances

For any independent random variables χ_1 and χ_2

$$var[\chi_1 + \chi_2] = var[\chi_1] + var[\chi_2]. \tag{14}$$

Multiplying variances

For any random variable χ and any $c \in \mathbb{R}$,

$$var[c\chi] = c^2 var[\chi]. \tag{15}$$

Example

The treasure behind the doors