Medical Information Science for Decision Support

EXECUTIVE ACADEMY

Assoc. Prof. Dr. Andreas HOLZINGER (Med. Uni Graz)

Day 1 –Part 3 -17.4.2018

Decision Making and Decision Support

Overview

Day 1 - Fundamentals

01 Information Sciences meets Life Sciences

02 Data, Information and Knowledge

03 Decision Making and Decision Support

04 From Expert Systems to Explainable AI

Keywords

- Decision
- Cognition
- Intelligence
- Expected Utility Theory
- Probabilistic Inference
- Probabilistic Decision Theory
- Signal Detection Theory
- ROC curve
- Learning and Inference
- Naïve Bayes Classifier

Advance Organizer (1/2)

- **Argmax/argmin** = set of points for which f(x) attains the function's largest/smallest value.
- **Brute Force** = systematically computing all possible candidates for a solution and checking whether each candidate satisfies the problem's statement;
- **Cognition** = mental processes of gaining knowledge, comprehension, including thinking, attention, remembering, language understanding, decision making and problem-solving;
- **Cognitive Science** = interdisciplinary study of human information processing, including perception, language, memory, reasoning, and emotion;
- **Confounding Variable** = an unforeseen, unwanted variable that jeopardizes reliability and validity of a study outcome.
- **Correlation coefficient** = measures the relationship between pairs of interval variables in a sample, from r = -1.00 to 0 (no correlation) to r = +1.00
- **Decision Making** = a central cognitive process in every medical activity, resulting in the selection of a final choice of action out of alternatives; according to Shortliffe (2011) DM is still the key topic in medical informatics;
- **Diagnosis** = classification of a patient's condition into separate and distinct categories that allow medical decisions about treatment and prognostic;
- **Differential Diagnosis (DDx)** = a systematic method to identify the presence of an entity where multiple alternatives are possible, and the process of elimination, or interpretation of the probabilities of conditions to negligible levels;
- **Evidence-based medicine (EBM)** = aiming at the best available evidence gained from the scientific method to clinical decision making. It seeks to assess the strength of evidence of the risks and benefits of treatments (including lack of treatment) and diagnostic tests. Evidence quality can range from meta-analyses and systematic reviews of double-blind, placebo-controlled clinical trials at the top end, down to conventional wisdom at the bottom; NOTE: Evidence (English) is NOT Evidenz (Deutsch)!
- **Expected Utility Theory (EUT)** = states that the decision maker selects between risky or uncertain prospects by comparing their expected utility values.

Advance Organizer (2/2)

- **External Validity** = the extent to which the results of a study are generalizable or transferable;
- **Hypothetico-Deductive Model (HDM)** = formulating a hypothesis in a form that could conceivably be falsified by a test on observable data, e.g. a test which shows results contrary to the prediction of the hypothesis is the falsification, a test that could but is not contrary to the hypothesis corroborates the theory – then you need to compare the explanatory value of competing hypotheses by testing how strong they are supported by their predictions;
- **Internal Validity** = the rigor with which a study was conducted (e.g., the design, the care taken to conduct measurements, and decisions concerning what was and was not measured);
- **PDCA** = Plan-Do-Check-Act, The so called PDCA-cycle or Deming-wheel can be used to coordinate a systematic and continuous improvement. Every improvement starts with a goal and with a plan on how to achieve that goal, followed by action, measurement and comparison of the gained output.
- **Perception** = sensory experience of the world, involving the recognition of environmental stimuli and actions in response to these stimuli;
- **Qualitative Research** = empirical research exploring relationships using textual, rather than quantitative data, e.g. case study, observation, ethnography; Results are not considered generalizable, but sometimes at least transferable.
- **Quantitative Research** = empirical research exploring relationships using numeric data, e.g. surveys, quasi-experiments, experiments. Results should be generalized, although it is not always possible.
- **Reasoning** = cognitive (thought) processes involved in making medical decisions (clinical reasoning, medical problem solving, diagnostic reasoning, behind every action;
- **Receiver-operating characteristic (ROC)** = in signal detection theory this is a graphical plot of the sensitivity, or true positive rate, vs. false positive rate (1 – specificity or 1 – true negative rate), for a binary classifier system as its discrimination threshold is varied;
- **Symbolic reasoning** = logical deduction
- **Triage** = process of judging the priority of patients' treatments based on the severity of their condition;

Glossary

- CES = Central Executive System
- DDx = Differential Diagnosis
- DM = Decision Making
- DSS = Decision Support System
- EBM = Evidence-based medicine
- fMRI = functional Magnetic Resonance Image
- HDM = Hypothetico-Deductive Model
- IOM = Institute of Medicine
- LTS = Long Term Storage
- ME = Medical Error
- PDCA = Plan-Do-Check-Act
- QM = Quality Management
- ROC = Receiver Operating Characteristic
- RST = Rough Set Theory
- STS = Short Term Storage
- USTS = Ultra Short Term Storage (Sensory Register)

Learning Goals: At the end of this lecture you …

- … are familiar with some principles and elements of human information processing;
- … can discriminate between perception, cognition, thinking, reasoning & problem solving;
- … have got insight into some basics of human decision making processes;
- … got an overview of the Hypothetico-Deductive Method HDM versus PCDA Deming approach;
- … have acquired some basics on modeling patient health and differential diagnosis

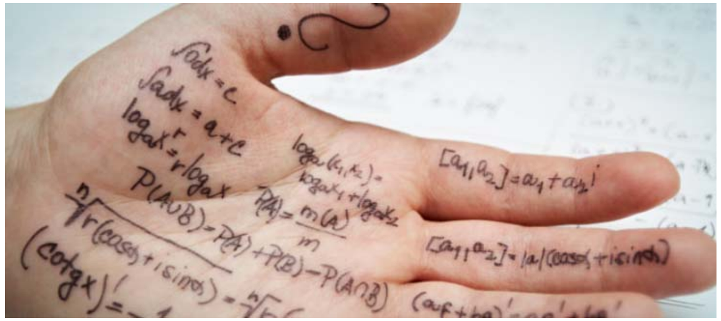
Agenda for today

- **00 Reflection – follow-up from last lecture**
- **01 Medical Action = Decision Making**
- **02 Cognition**
- **03 Human vs. Computer**
- **04 Human Information Processing**
- **05 Probabilistic Decision Theory**
- **06 Example: Naïve Bayes Classifier**

00 Reflection

Reflection from last lecture

- The Quiz-Slide will be shown during the course



Key Challenges

- Medicine is an extremely complex application domain – dealing most of the time with uncertainties -> **probable information!**
- Key: Structure learning and prediction in large-scale biomedical networks with **probabilistic graphical models**
- Causality and Probabilistic Inference:
- Uncertainties are present at all levels in health related systems
- Data sets are noisy, mislabeled, atypical, dirty, wrong, etc. etc.
- Even with data of high quality from different real-world sources requires **processing uncertain information to make viable decisions**.
- In the increasingly complicated settings of modern science, model structure or **causal relationships may not be known a-priori** [1].
- Approximating probabilistic inference in Bayesian belief networks is NP-hard [2] -> here we need the "human-in-the-loop" [3]

[1] Sun, X., Janzing, D. & Schölkopf, B. Causal Inference by Choosing Graphs with Most Plausible Markov Kernels. ISAIM, 2006.
[2] Dagum, P. & Luby, M. 1993. Approximating probabilistic inference in Bayesian belief networks is NP-hard. Artificial intelligence, 60, (1), 141-153.
[3] Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? Springer Brain Informatics (BRIN), 3, 1-13, doi:10.1007/s40708-016-0042-6.

01 Cognition

Cognitive Science vs. Computer Science

- **Cognitive Science $\rightarrow$ human intelligence**
  - Study the principles of ***human learning*** to understand biological intelligence
- **Human-Computer Interaction $\rightarrow$ the bridge**
  - Interacting with algorithms that learn shall enhance user friendliness and let concentrate on problem solving - Opening the “black-box” to a “glass-box”
- **Computer Science $\rightarrow$ computational intelligence**
  - Study the principles of ***machine learning*** to understand artificial intelligence

What is this: Cognitive Science?

- “By 1960 it was clear that something interdisciplinary was happening. At Harvard we called it cognitive studies, at Carnegie-Mellon they called it information-processing psychology, and at La Jolla they called it cognitive science. “

George A. Miller (1920-2012), Harvard University, well known for:

The **magical number seven**, plus or minus two: Some limits on our capacity for processing information.
GA Miller - Psychological review, 1956 - psycnet.apa.org
Abstract A variety of researches are examined from the standpoint of information theory. It is shown that the unaided observer is severely limited in terms of the amount of information he can receive, process, and remember. However, it is shown that by the use of various ...
Cited by 27734 Related articles All 74 versions Import into EndNote

Why fitting Cognitive Science with Machine Learning?

- CS aims to reverse engineer **human intelligence**;
- ML provides powerful sources of insight into ***how*** **machine intelligence** is possible.
- CS therefore raises challenges for, and draws inspiration from ML;
- ML could inspire **new directions** by novel insights about the human mind

Some definitions (very incomplete)

- Intelligence
  - Hundreds of controversial definitions – very hard to define;
  - For us: ability to solve problems, to make decisions and to acquire and apply knowledge and skills.
- Learning
  - Different definitions – relatively hard to define
  - basically acquisition of knowledge through prior experience
- Problem Solving
  - Process of finding solutions to complex issues
- Reasoning
  - ability of our mind to think and understand things
- Decision Making
  - Process of “de-ciding” (“ent-scheiden”) between alternative options
- Sense Making
  - Process of giving meaning to experience
- Causality
  - Relationship between cause and effect

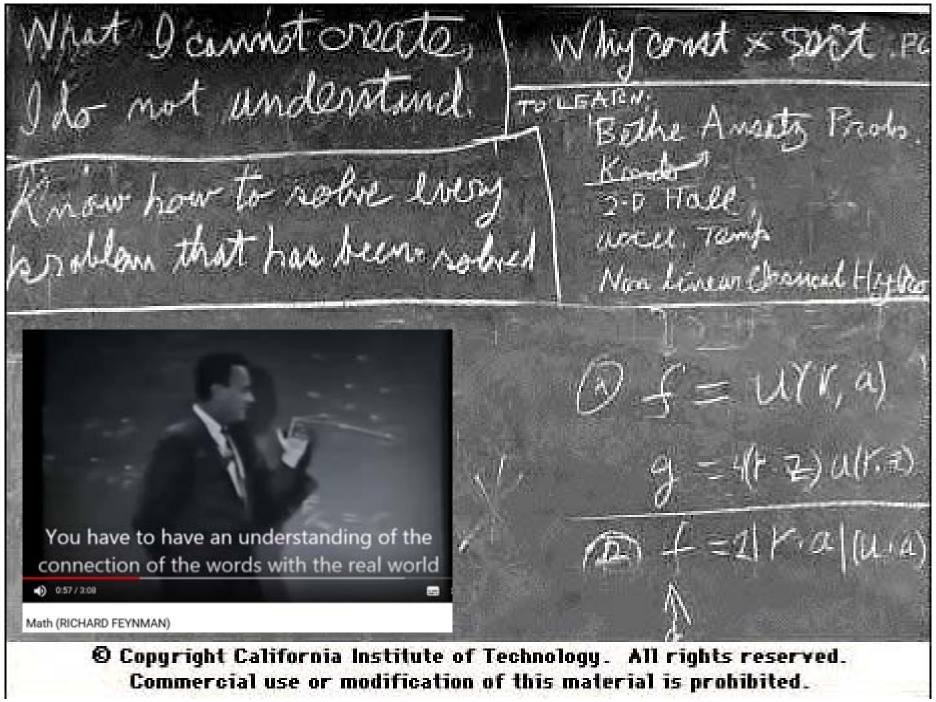
What is learning?

Learning: mathematical and computational principles allowing one to learn from examples in order to acquire knowledge

- **adaptive behavior** change caused by experience
- to act successfully in a complex environment

Kupfermann, I. (1991). Learning and memory. Principles of neural science, 997-1008.

Examples adaptive behavior to complex environments



Typical Questions of Cognitive Science Research

- How do we process information?
- How do we learn and generalize?
- How do we solve problems?
- How do we reason and make decisions?
- How do we make predictions?
- How do we behave in new situations?

Intelligence

Motto of Google Deepmind: “Solving Intelligence …”

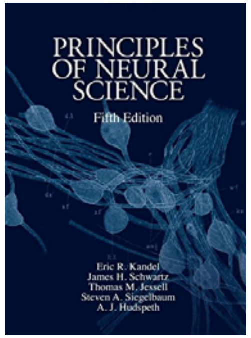
“Solve intelligence – then solve everything else”

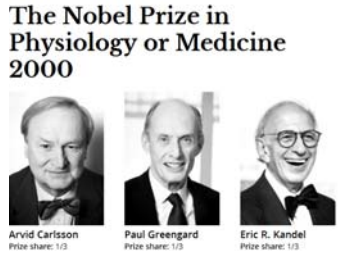
Demis Hassabis, 22 May 2015

The Royal Society,
Future Directions of Machine Learning Part 2

https://youtu.be/XAbLn66iHcQ?t=1h28m54s

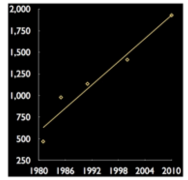
Do we understand how our brain works?





This book doubled in Volume every decade …

Number of Pages ->



Editions ->

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A. & Hudspeth, A. 2012. Principles of neural science, 5th Edition (1760 pages), New York: McGraw-Hill.

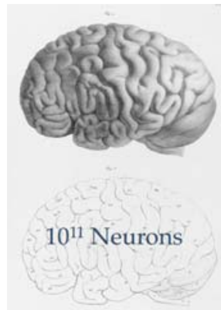
- Facts ≠ Knowledge, Descriptions ≠ Insight
- **Our goal should be the opposite: To make this book shorter!**

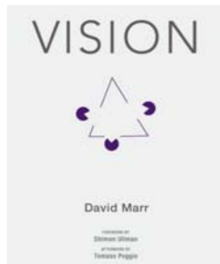
CS vs ML did NOT harmonize in the past

- Cognitive Science had its focus on specific experimental paradigms because it was embedded deeply in Psychology and Linguistics; and aimed to be cognitively/neutrally plausible …
- ML had its focus on standard learning problems and tried to optimize in the range of 1 % because it was embedded in Computer Engineering; and aimed to have working systems whether mimicking the human brain or not …

David Marr (1945 – 1980) Neuroscientist

- Cerebellum: big memory to support motor learning
- Neocortex: big memory flexibly learns statistical structure from input patterns
- Hippocampus: big memory encoding memory traces via Hebbian learning
- Example Vision: process of discovering properties (what, where) of things in the real-world from 3D-images (on 2D)
- Vision = information processing task + rich internal representation
- Understanding of vision requires multiple levels of analysis: computational – algorithmic and physical (hardware)



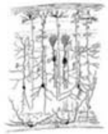


Marr, D. 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, New York, Henry Holt.

Marr: Three Levels of Information Processing Systems

- **Computation**
  - "What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?"

- **Representation and algorithm**
  - "What is the representation for the input and output, and the algorithm for the transformation?"



- **Implementation**
  - "How can the representation and algorithm be realized physically?"

Marr, D. 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, New York, Henry Holt.

Cognitive Science versus Machine Learning

- Human learning
- Categorization
- Causal learning
- Function learning
- Representations
- Language
- Experiment design

- Machine learning
- Density estimation
- Graphical models
- Regression
- Nonparametric Bayes
- Probabilistic grammars
- Inference algorithms

"People who are interested in machine learning should be cognitive scientists and vice versa" Joshua Tenenbaum, MIT

http://web.mit.edu/cocosci/josh.html

Is the human brain a inference engine ?

- Learning concepts from examples (babies!)
- Causal inference and reasoning
- Predicting everyday events
- Even little children solve complex problems unconsciously, effortlessly, and … successfully!
- **Compare your best Machine Learning algorithm with a seven year old child!**

Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Griffiths, T. L. Connecting human and machine learning via probabilistic models of cognition. Interspeech 2009, 2009 Brighton (UK). ISCA, 9-12. available online via: https://cocosci.berkeley.edu/tom/papers/probmods.pdf

03 Human versus Computer

People are awesome …

See Youtube: “people are awesome” … hundreds of examples

Problem Solving: Humans vs. Computers

When is the human *) better?

*) human intelligence/natural intelligence/human mind/human brain/ learning

- **Natural Language Translation/Curation**

  Computers cannot understand the context of sentences [3]
- **Unstructured problem solving**

  Without a pre-set of rules, a machine has trouble solving the problem, because it lacks the creativity required for it [1]
- **NP-hard Problems**

  Processing times are often exponential and makes it almost impossible to use machines for it, but human make heuristic decisions which are often not perfect but sufficiently good [4]

When is the computer **) better?

**) Computational intelligence, Artificial Intelligence/soft computing/ML

- **High-dimensional data processing**

  Humans are very good at dimensions less or equal than 3, but computers can process data in arbitrarily high dimensions
- **Rule-Based environments**

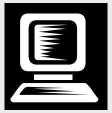
  Difficulties for humans in rule-based environments often come from not recognizing the correct goal in order to select the correct procedure or set of rules [2]
- **Image optimization**

  Machine can look at each pixel and apply changes without human personal biases, and with more speed [1]

[1] Kipp, M. 2006. Creativity Meets Automation: Combining Nonverbal Action Authoring with Rules and Machine Learning. In: LNCS 4133, pp. 230-242, doi:10.1007/11821830_19.
[2] Cummings, M. M. 2014. Man versus Machine or Man + Machine? IEEE Intelligent Systems, 29, (5), 62-69, doi:10.1109/MIS.2014.87.
[3 Pizlo, Z., Joshi, A. & Graham, S. M. 1994. Problem Solving in Human Beings and Computers. Purdue TR 94-075.
[4] Griffiths, T. L. Connecting human and machine learning via probabilistic models of cognition. Interspeech, 2009, ISCA, 9-12..

Comparison: Human vs. Computer

Superiority

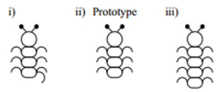


Human	Computer
sensitiveness for stimuli (visual, auditory, tactile, olfactory)	Precise Counting and Measuring of physical entities
Ability for inductive Reasoning and complex Problem Solving	Deductive Operations, formal Logic, Application of Rules
Creating of networked knowledge and storage for a live-long time	Storage of huge amounts of data which are not necessarily connected
Flexibility in decisions, even in totally new situations	Reliable reaction to unambiguous input signals
Discovering of ambiguous signals even when distorted	Reliable performance over long periods without tiredness

Holzinger, A. 2000. *Basiswissen Multimedia 2: Lernen. Kognitive Grundlagen multimedialer Informationssysteme,* Würzburg, Vogel.

Human Cognitive capacities of Inference and Prediction

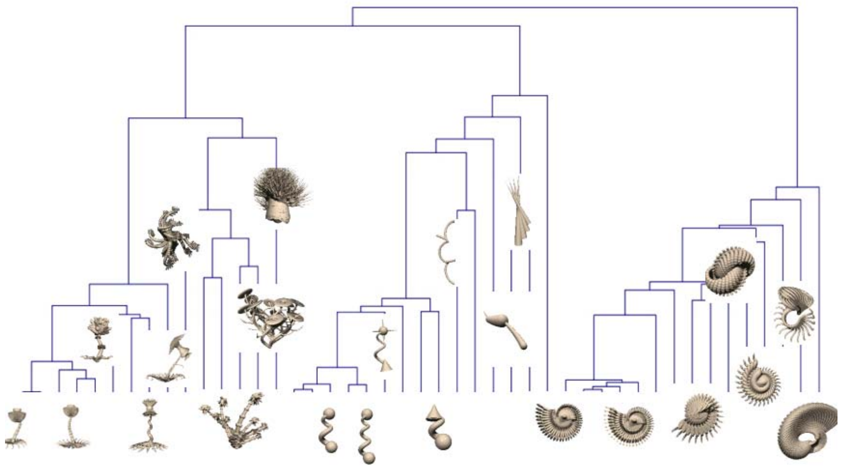
- Similarity [1]



- Representativeness and evidential support
- Causal judgment [2]
- Coincidences and causal discovery
- Clinical diagnostic inference [3]
- Predicting the future

[1] Kemp, C., Bernstein, A. & Tenenbaum, J. B. A generative theory of similarity. Proceedings of the 27th Annual Conference of the Cognitive Science Society, 2005. 1132-1137.
[2] Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J. & Blum, B. 2003. Inferring causal networks from observations and interventions. Cognitive science, 27, (3), 453-489.
[3] Krynski, T. R. & Tenenbaum, J. B. 2007. The role of causality in judgment under uncertainty. Journal of Experimental Psychology: General, 136, (3), 430.

Humans learn from very few examples …



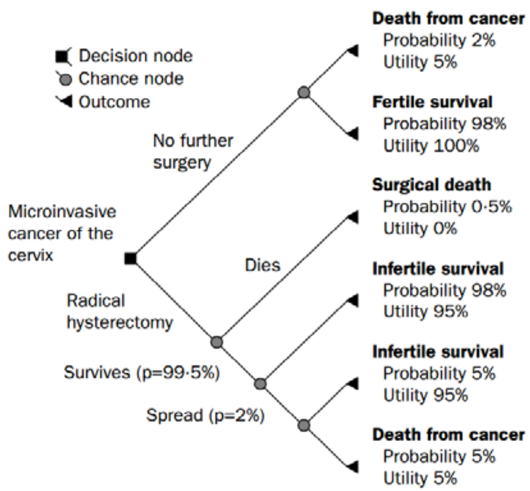
Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Important: Statistics meet Knowledge

- 1. How does abstract knowledge guide learning and inference from sparse data?
  - (Approximate) Bayesian inference in probabilistic models.
- 2. What are the forms and contents of that knowledge?
  - Probabilities defined over a range of structured representations: graphs, grammars, predicate logic, schemas… programs.
- 3. How is that knowledge itself acquired?
  - Hierarchical Bayesian models, with inference at multiple levels of abstraction ("learning to learn"). Learning as (hierarchical Bayesian) program induction.
- **Central Question:**
  How does our mind get so much out of so little?

Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Decision trees are coming from Clinical Practice



Physician treating a patient approx. 480 B.C. Beazley (1963), Attic Red-figured Vase-Painters, 813, 96. Department of Greek, Etruscan and Roman Antiquities, Sully, 1st floor, Campana Gallery, room 43 Louvre, Paris

Elwyn, G., Edwards, A., Eccles, M. & Rovner, D. 2001. Decision analysis in patient care. The Lancet, 358, (9281), 571-574.

Example: Discovery of causal relationships from data …

15

Hans Holbein d.J., 1533, The Ambassadors, London: National Gallery

Lopez-Paz, D., Muandet, K., Schölkopf, B. & Tolstikhin, I. 2015. Towards a learning theory of cause-effect inference. Proceedings of the 32nd International Conference on Machine Learning, JMLR, Lille, France.

https://www.youtube.com/watch?v=9KiVNIUMmCc

Humans can understand the context

15b

- **"How do humans generalize from so few examples?"**
  - Learning relevant representations
  - Disentangling the explanatory factors
  - Finding the shared underlying explanatory factors, in particular between $P(x)$ and $P(Y|X)$, with a causal link between $Y \rightarrow X$

Bengio, Y., Courville, A. & Vincent, P. 2013. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35, (8), 1798-1828, doi:10.1109/TPAMI.2013.50.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Expected Utility Theory $E\ (U|d)$

For a single decision variable an agent can select $D = d$ for any $d \in dom(D)$.
The expected utility of decision $D = d$ is

http://www.eoht.info/page/Oskar+Morgenstern

$$E(U \mid d) = \sum_{x_1,\ldots,x_n} P(x_1,\ldots,x_n \mid d) U(x_1,\ldots,x_n,d)$$

An optimal single decision is the decision $D = dmax$ whose expected utility is maximal:

$$d_{\max} = \arg \max_{d \in \mathrm{dom}(D)} E(U \mid d)$$

Von Neumann, J. & Morgenstern, O. 1947. Theory of games and economic behavior, Princeton university press.

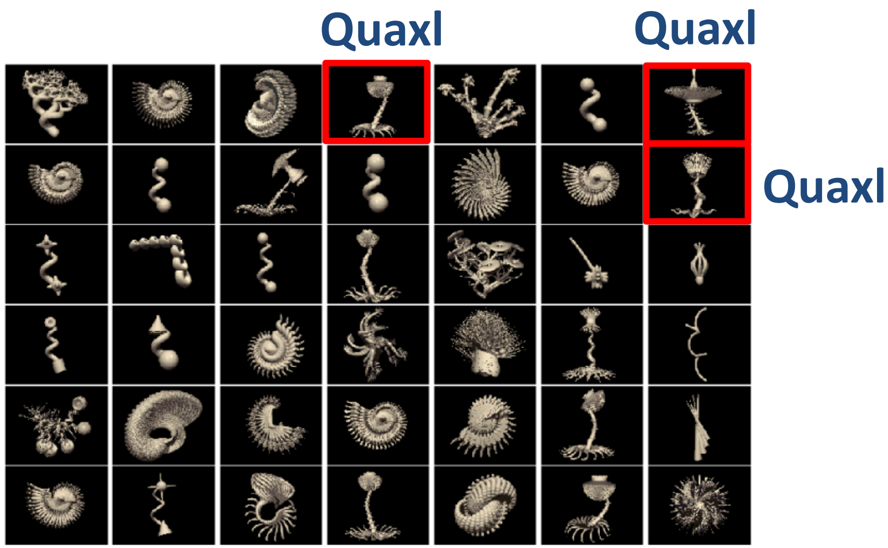
04 Human Information Processing

04 Human Information Processing

How does our mind get so much out of so little ...

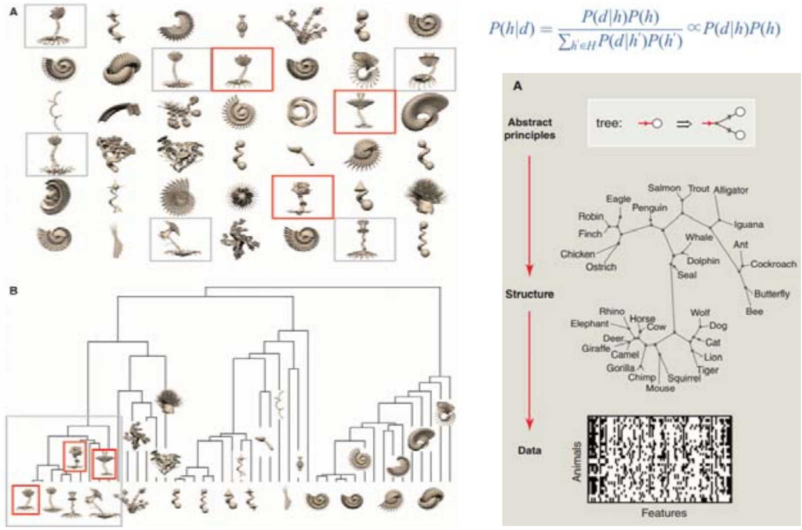
Salakhutdinov, R., Tenenbaum, J. & Torralba, A. 2012. One-shot learning with a hierarchical nonparametric Bayesian model. Journal of Machine Learning Research, 27, 195-207.

Learning words for objects – concepts from examples



Salakhutdinov, R., Tenenbaum, J. & Torralba, A. 2012. One-shot learning with a hierarchical nonparametric Bayesian model. Journal of Machine Learning Research, 27, 195-207.

How do we understand our world …



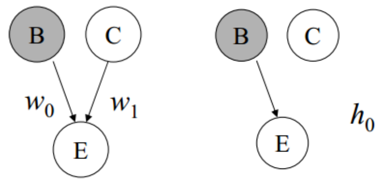
Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285.

One of the unsolved problems in human concept learning

- which is highly relevant for ML research, concerns the factors that determine the subjective difficulty of concepts:
- Why are some concepts psychologically extremely simple and easy to learn,
- while others seem to be extremely difficult, complex, or even incoherent?
- These questions have been studied since the 1960s but are still unanswered …

Feldman, J. 2000. Minimization of Boolean complexity in human concept learning. Nature, 407, (6804), 630-633, doi:10.1038/35036586.

A few certainties

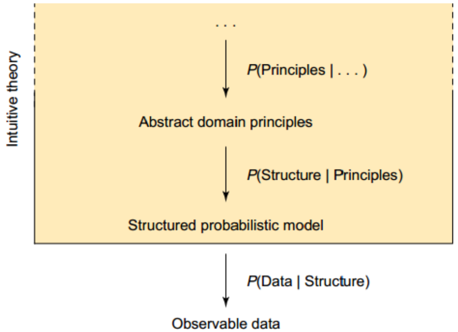


- Cognition as probabilistic inference
  - Visual perception, language acquisition, motor learning, associative learning, memory, attention, categorization, reasoning, causal inference, decision making, theory of mind
- Learning concepts from examples
- Learning and applying intuitive theories (balancing complexity vs. fit)

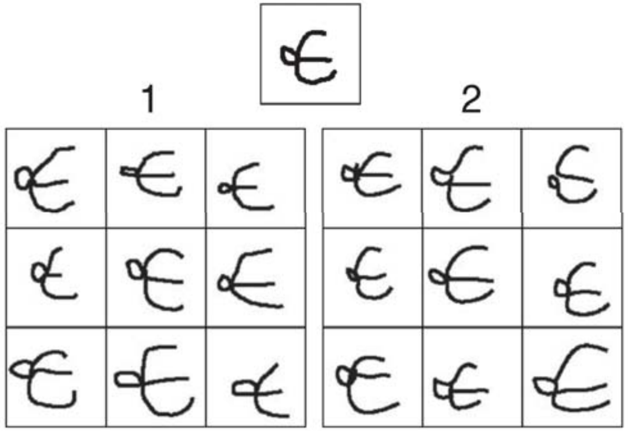
Modeling basic cognitive capacities as intuitive Bayes

- Similarity
- Representativeness and evidential support
- Causal judgement
- Coincidences and causal discovery
- Diagnostic inference
- Predicting the future

Tenenbaum, J. B., Griffiths, T. L. & Kemp, C. 2006. Theory-based Bayesian models of inductive learning and reasoning. Trends in cognitive sciences, 10, (7), 309-318.



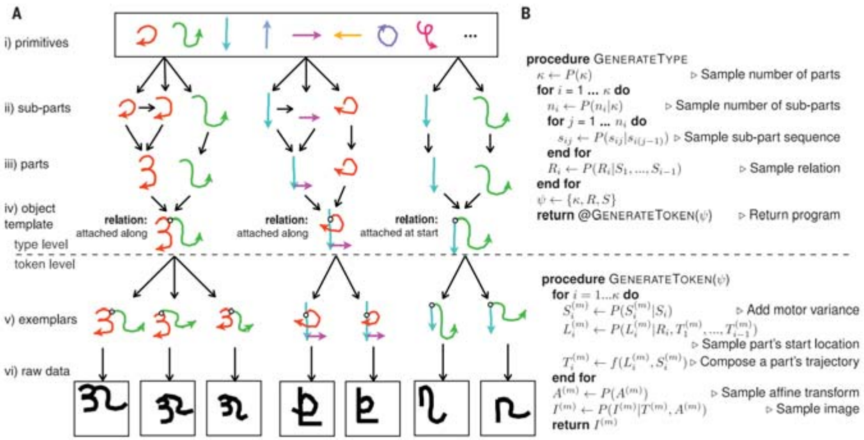
Drawn by Human or Machine Learning Algorithm?



Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. 2015. Human-level concept learning through probabilistic program induction. Science, 350, (6266), 1332-1338, doi:10.1126/science.aab3050.

Human-Level concept learning – probabilistic induction

A Bayesian program learning (BPL) framework, capable of learning a large class of visual concepts from just a single example and generalizing in ways that are mostly indistinguishable from people

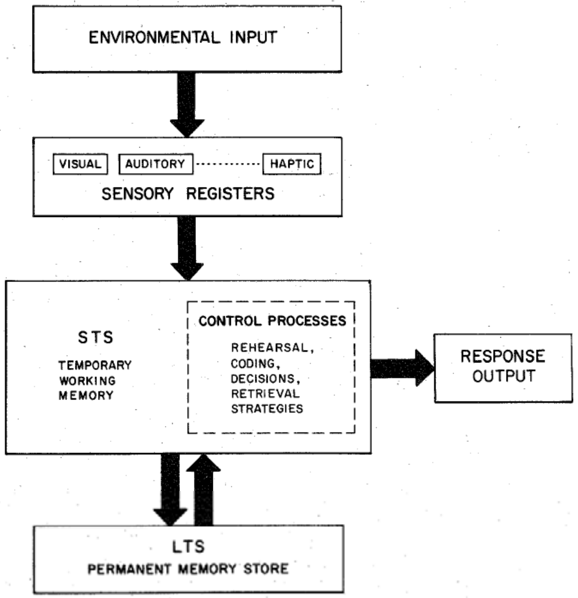


Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. 2015. Human-level concept learning through probabilistic program induction. Science, 350, (6266), 1332-1338, doi:10.1126/science.aab3050.

How does our mind get so much out of so little?

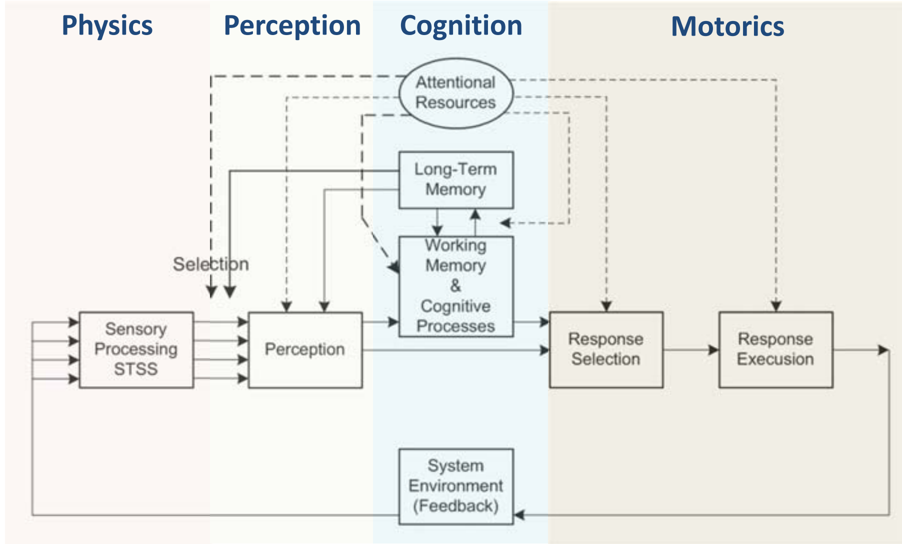
Tenenbaum, J.B., Kemp, C., Griffiths, T.L. & Goodman, N.D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Human Information Processing Model (A&S)



Atkinson, R. C. & Shiffrin, R. M. (1971) *The control processes of short-term memory (Technical Report 173, April 19, 1971). Stanford, Institute for Mathematical Studies in the Social Sciences, Stanford University.*

General Model of Human Information Processing



Wickens, C., Lee, J., Liu, Y. & Gordon-Becker, S. (2004) *Introduction to Human Factors Engineering: Second Edition. Upper Saddle River (NJ), Prentice-Hall.*

Learning and Inference

04

d … data

h … hypotheses

$\mathcal{H}$ … $\{H_1, H_2, \ldots, H_n\}$ $\forall\, h, d$ …

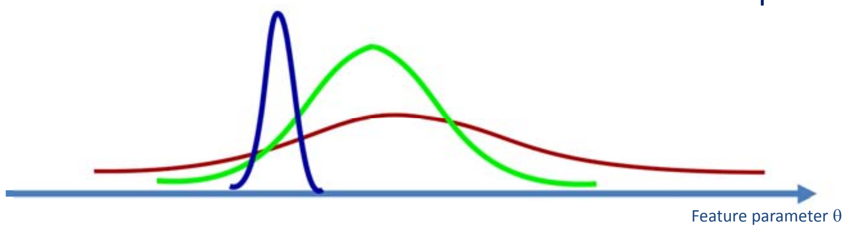
Likelihood

Prior Probability

$$p(h|d) = \frac{p(d|h) * p(h)}{\sum_{h \in H} p(d|h')\, p(h')}$$

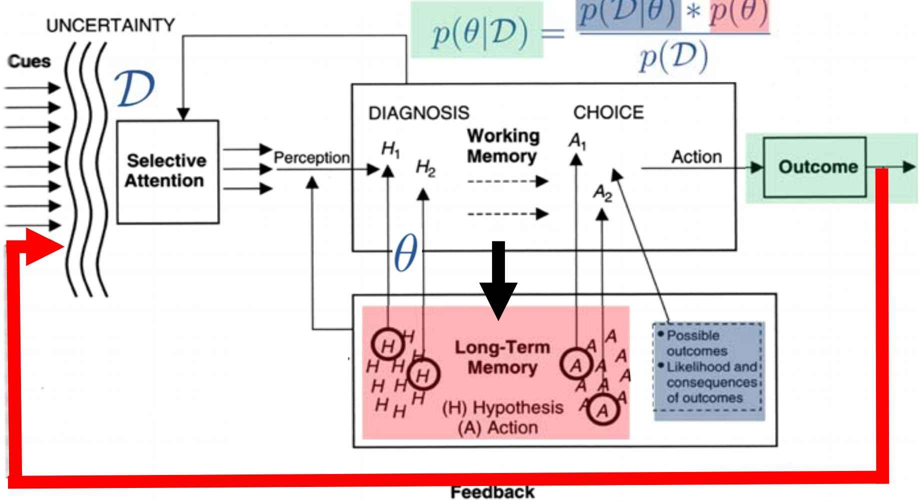
Posterior Probability

Problem in $\mathbb{R}^n \rightarrow$ complex



Connection to Cognitive Science: Decision Making

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta) * p(\theta)}{p(\mathcal{D})}$$



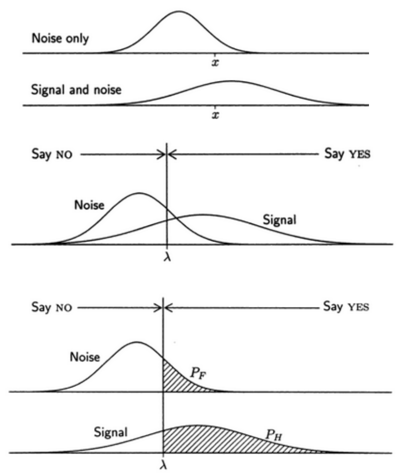
Wickens, C. D. (1984) *Engineering psychology and human performance. Columbus (OH), Charles Merrill, modified by Holzinger, A.*

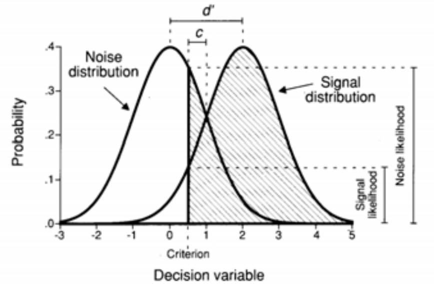
05 Probabilistic Decision Theory

"It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge"
Pierre Simon de Laplace, 1812

Signal Detection Theory (SDT)

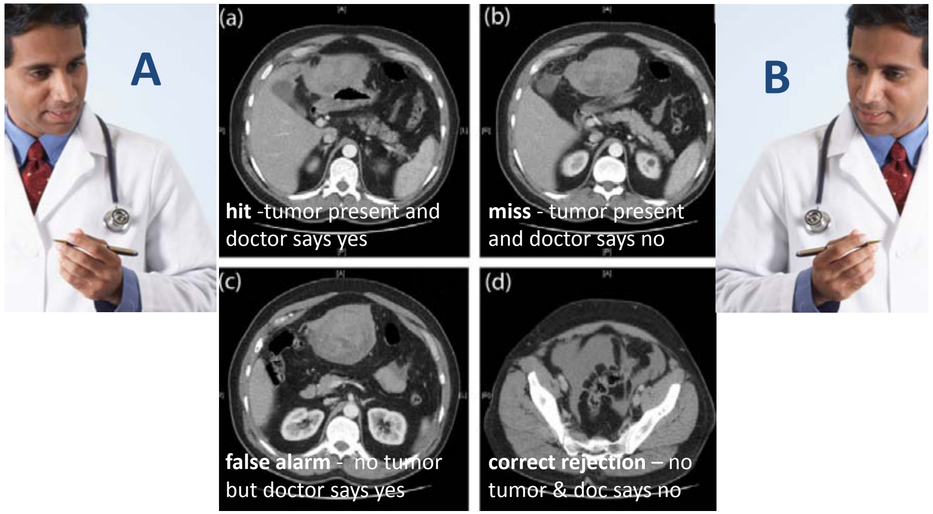
Image source: Staffordshire University Computing Futures Museum http://www.fcet.staffs.ac.uk/jdw1/sucfm/malvern.htm





Stanislaw, H. & Todorov, N. 1999. Calculation of signal detection theory measures. Behavior research methods, instruments, & computers, 31, (1), 137-149.

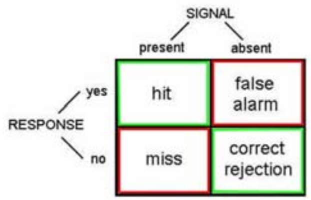
Signal Detection Theory on the MDM process

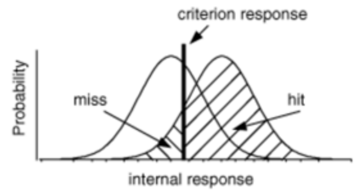


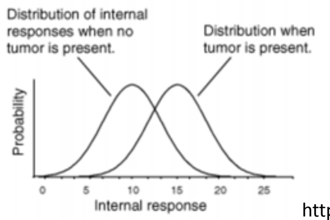
Two doctors, with equally good training, looking at the same CT scan, will have the same information … but they may have a different bias/criteria!

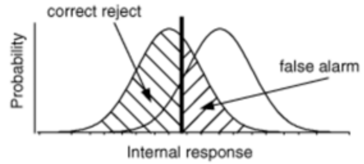
Decision Making Process - Signal Detection

Remember: Two doctors, with equally good training, looking at the same CT scan data, will have the same information … but they may gain different knowledge due to *bias/criteria*.





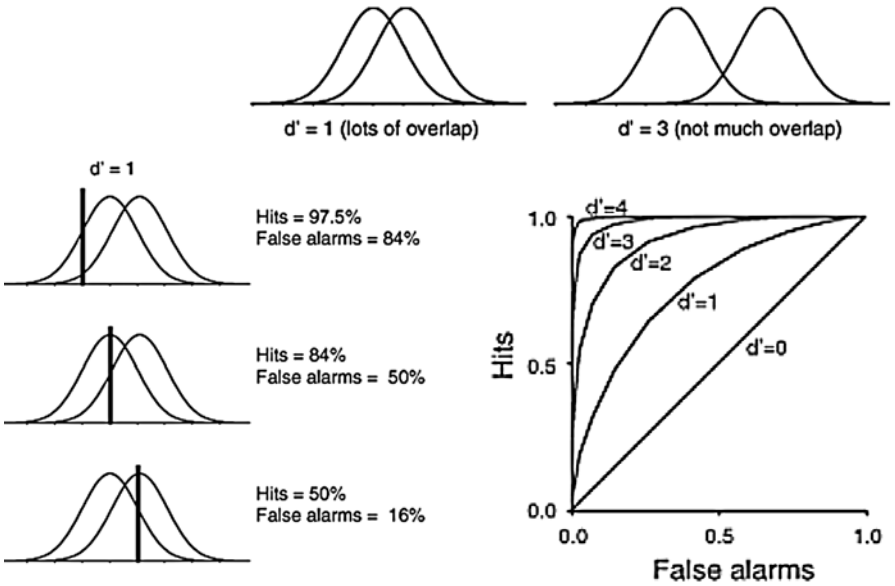




http://www-psych.stanford.edu/~lera/psych115s/notes/signal

For an example see: Braga & Oliveira (2003) Diagnostic analysis based on ROC curves: theory and applications in medicine. *Int. Journal of Health Care Quality Assurance, 16, 4, 191-198.*

Receiver Operating Characteristics (ROC curve)



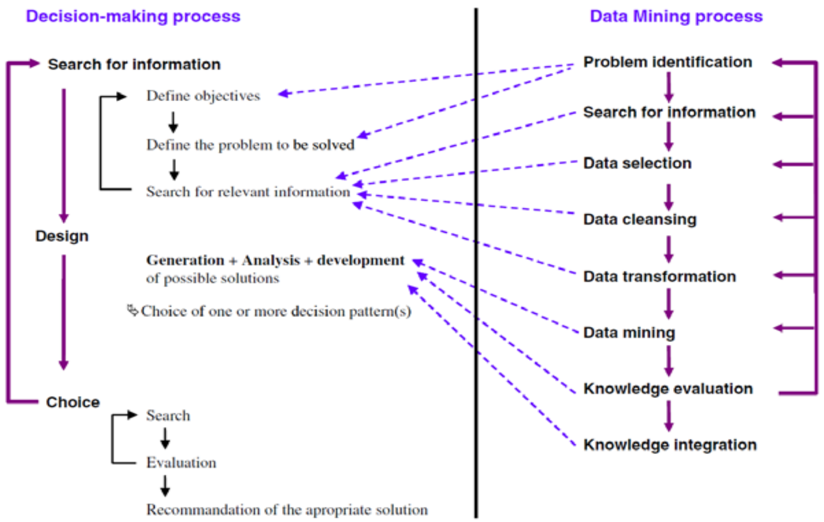
http://gru.stanford.edu/doku.php/tutorials/sdt

Slide 7-23 Information Acquisition and criteria - bias

- **Information acquisition:** in the CT data, e.g. healthy lungs have a characteristic shape; the presence of a tumor might distort that shape (= anomaly).
- Tumors have different image characteristics: brighter or darker, different texture, etc.
- With proper training a doctor learns what kinds of things to look for, so with more practice/training they will be able to acquire more (and more reliable) information.
- Running another test (e.g., MRI) can be used to acquire more (relevant!) information.
- The effect of information is to increase the likelihood of getting either a hit or a correct rejection, while reducing the likelihood of an outcome in the two error boxes (slide 33).
- **Criterion:** Additionally to relying on technology/testing, the medical profession allows doctors to use their own judgment.
- Different doctors may feel that the different types of errors are not equal.
- For example, a doctor may feel that missing an opportunity for early diagnosis may mean the difference between life and death.
- A false alarm, on the other hand, may result only in a routine biopsy operation. They may chose to err toward ``yes'' (tumor present) decisions.
- Other doctors, however, may feel that unnecessary surgeries (even routine ones) are very bad (expensive, stress, etc.).
- They may chose to be more conservative and say ``no'' (no turmor) more often. They will miss more tumors, but they will be doing their part to reduce unnecessary surgeries. And they may feel that a tumor, if there really is one, will be picked up at the next check-up.

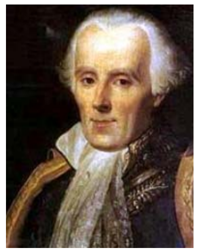
Mohamed, A. et al. (2010) Traumatic rupture of a gastrointestinal stromal tumour with intraperitoneal bleeding and haematoma formation. *BMJ Case Reports, 2010.*

Decision Making Process vs. Data Mining process



Ayed, B. M., Ltifi, H., Kolski, C. & Alimi, A. (2010) A user-centered approach for the design & implementation of KDD-based DSS: A case study in the healthcare domain. *Decision Support Systems, 50, 64-78.*

Repetition Bayes Foundations



What is the simplest mathematical operation for us?

$$p(x) = \sum_x (p(x,y))$$

How do we call repeated adding?

$$p(x,y) = p(y|x) * p(y)$$

Laplace (1773) showed that we can write:

$$p(x,y) * p(y) = p(y|x) * p(x)$$

Now we introduce a third, more complicated operation:

$$\frac{p(x,y) * p(y)}{p(y)} = \frac{p(y|x) * p(x)}{p(y)}$$

We can reduce this fraction by $p(y)$ and we receive what is called Bayes rule:

$$p(x,y) = \frac{p(y|x) * p(x)}{p(y)} \qquad p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

Always Remember: Learning and Inference

04

d … data

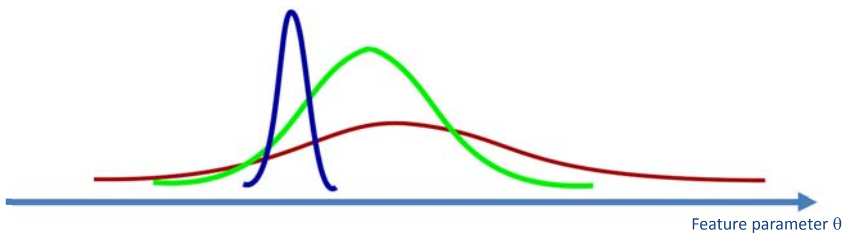
h … hypotheses

$\mathcal{H}$ …$\{H_1, H_2, \ldots, H_n\}$ $\quad \forall\, h, d$ …

Likelihood · Prior Probability

$$p(h|d) = \frac{p(d|h) * p(h)}{\sum_{h \in H} p(d|h')\, p(h')}$$

Posterior Probability · Evidence $p(d)$ = marginal likelihood



Bayes Law of Total Probability = data modelling

d … data; h … hypothesis

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)}$$

$P(h)$: prior belief (probability of hypothesis h before seeing any data)

$P(d\,|\,h)$: likelihood (probability of the data if the hypothesis h is true)

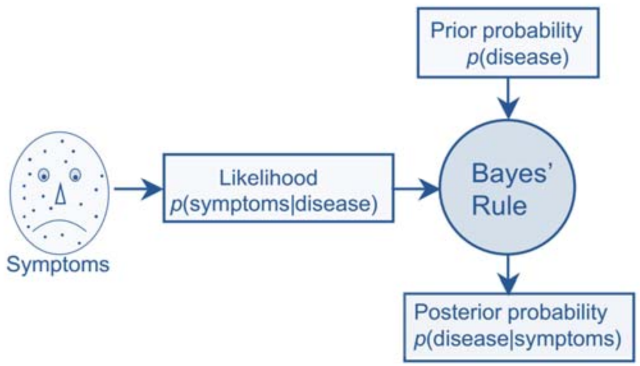
$P(d) = \sum_h P(d\,|\,h)P(h)$: data evidence (marginal probability of the data)

$P(h\,|\,d)$: posterior (probability of hypothesis h after having seen the data d)

$$posterior = \frac{likelihood * prior}{evidence}$$

- **evidence = marginal likelihood = "normalization"**
- **Remember: The inverse probability allows to infer unknowns, learn from data and make predictions … machine learning!**

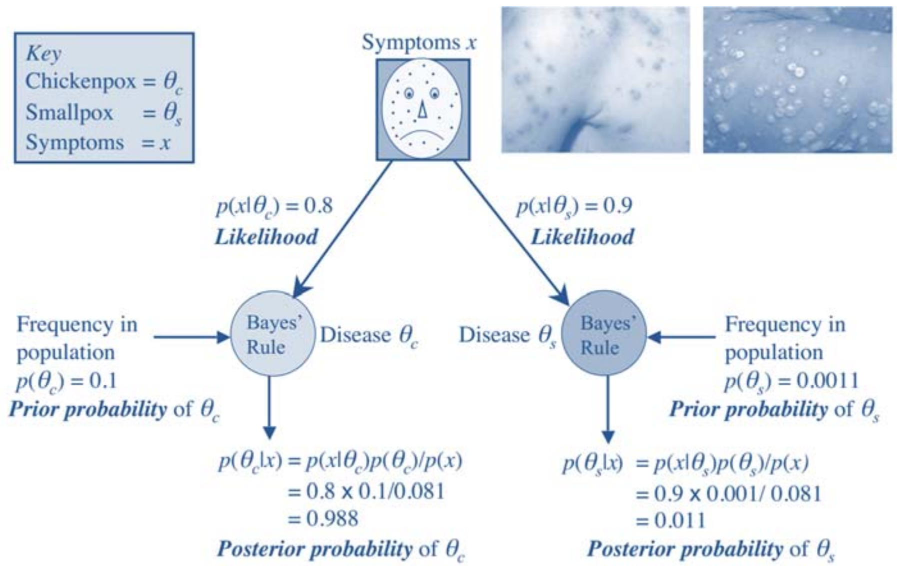
Bayes Rule for Medical Diagnosis



$$p(\text{disease}|\text{symptoms}) = \frac{p(\text{symptoms}|\text{disease})\,p(\text{disease})}{p(\text{symptoms})}$$

Stone, J. V. 2013. Bayes' rule: a tutorial introduction to Bayesian analysis. Sebtel Press.

Bayesian Inference



Practical Example: Diagnoses

- Your MD has bad news and good news for you.
- Bad news first: You are tested positive for a serious disease, and the test is 99% accurate (T)
- Good news: It is a rare disease, striking 1 in 10,000 (D)
- **How worried would you now be?**

$$posterior\ p(x) = \frac{likelyhood * prior\ p(x)}{evidence} \qquad p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

$$p(T = 1|D = 1) = p(d|h) = 0{,}99 \ \ and$$
$$p(D = 1) = p(h) = 0{,}0001$$

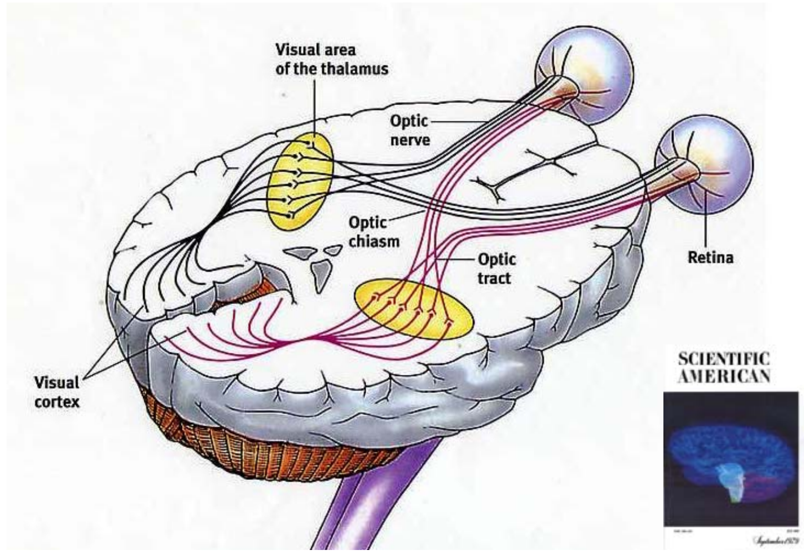
$$p(D = 1 \,|\, T = 1) = \frac{(0{,}99)*(0{,}0001)}{(1-0{,}99)*(1-0{,}0001)+0{,}99*0{,}0001} =$$

= 0,0098

Thank you!

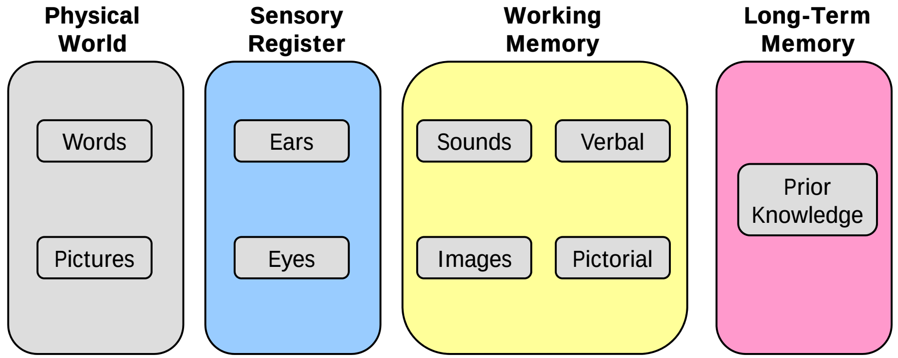
Appendix

Slide 7-7 Example: Visual Information Processing



Source: Department of Neuroscience, The Mount Sinai School of Medicine (2004)

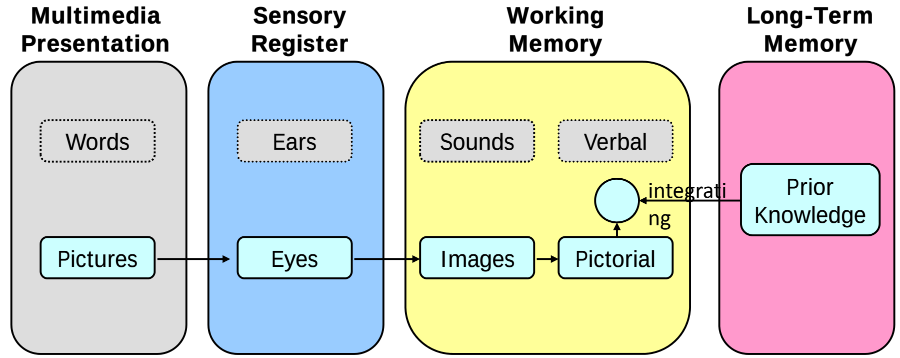
Slide 7-8 Schematic Information Processing Chain



cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

Slide 7-9 Information processing of images/pictures

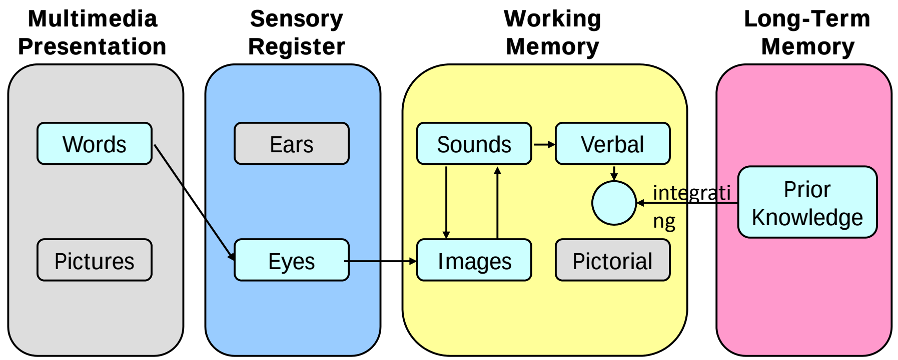
a) Processing of visual information (PICTURES)



cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

Slide 7-10 Information processing of words/pictures

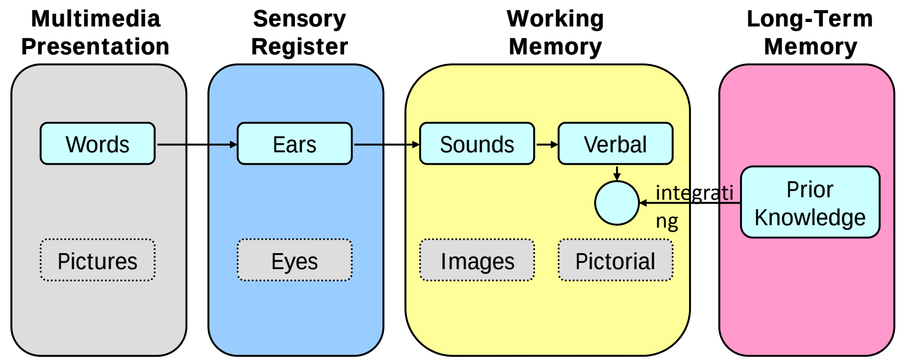
b) Processing of visual information (PRINTED WORDS)



cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

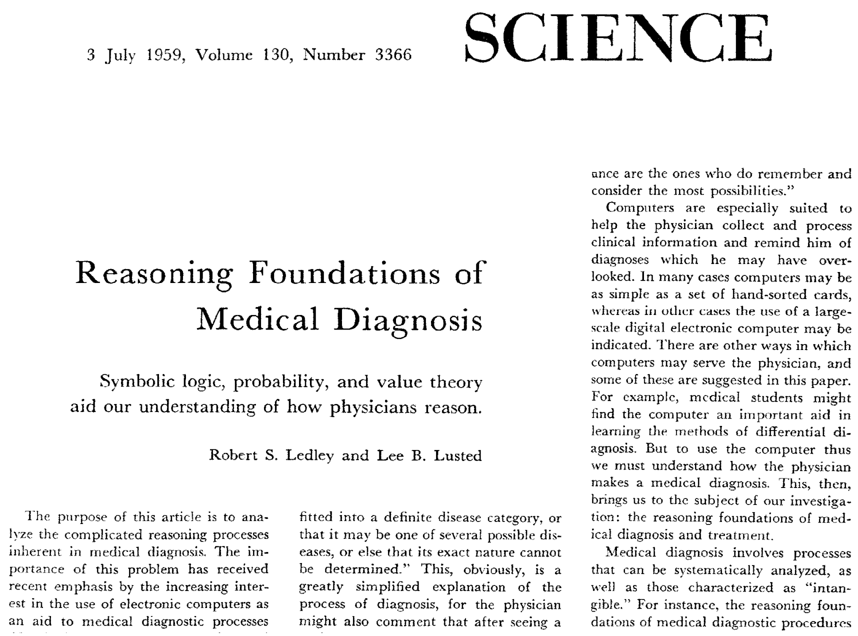
Slide 7-11 Information processing of words/sounds

c) Processing of audio information (SPOKEN WORDS)



cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

Slide 7-3: Reasoning Foundations of Medical Diagnosis



3 July 1959, Volume 130, Number 3366 SCIENCE

Reasoning Foundations of Medical Diagnosis

Symbolic logic, probability, and value theory aid our understanding of how physicians reason.

Robert S. Ledley and Lee B. Lusted

The purpose of this article is to analyze the complicated reasoning processes inherent in medical diagnosis. The importance of this problem has received recent emphasis by the increasing interest in the use of electronic computers as an aid to medical diagnostic processes

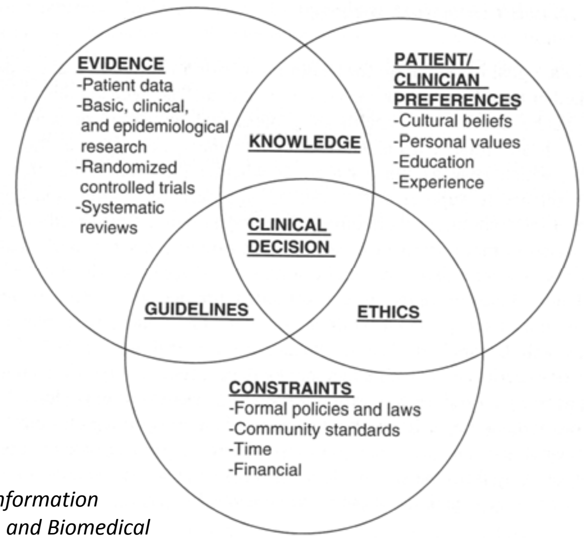
fitted into a definite disease category, or that it may be one of several possible diseases, or else that its exact nature cannot be determined." This, obviously, is a greatly simplified explanation of the process of diagnosis, for the physician might also comment that after seeing a

ance are the ones who do remember and consider the most possibilities."

Computers are especially suited to help the physician collect and process clinical information and remind him of diagnoses which he may have overlooked. In many cases computers may be as simple as a set of hand-sorted cards, whereas in other cases the use of a large-scale digital electronic computer may be indicated. There are other ways in which computers may serve the physician, and some of these are suggested in this paper. For example, medical students might find the computer an important aid in learning the methods of differential diagnosis. But to use the computer thus we must understand how the physician makes a medical diagnosis. This, then, brings us to the subject of our investigation: the reasoning foundations of medical diagnosis and treatment.

Medical diagnosis involves processes that can be systematically analyzed, as well as those characterized as "intangible." For instance, the reasoning foundations of medical diagnostic procedures

Slide 7-4 Decision Making is central in Medicine!



Hersh, W. (2010) *Information Retrieval: A Health and Biomedical Perspective. New York, Springer.*

Example for Decision Support

Iserson, K. V. & Moskop, J. C. 2007. Triage in Medicine, Part I: Concept, History, and Types. Annals of Emergency Medicine, 49, (3), 275-281.