

MAKE Decisions Medical Information Science for Decision Support



Assoc. Prof. Dr. Andreas HOLZINGER (Medical University Graz)



<https://hci-kdd.org/mini-course-make-decisions-practice>

Day 1 > Part 4 > 19.09.2018

DSS: from Expert Systems to explainable Artificial Intelligence

Keywords



- Artificial intelligence
- Case based reasoning
- Computational methods in cancer detection
- Cybernetic approaches for diagnostics
- Decision support models
- Decision support system (DSS)
- Explainable AI
- Fuzzy sets
- MYCIN – Expert System
- Reasoning under uncertainty
- Radiotherapy planning



Day 1 - Hot Ideas

01 Information Sciences meets
Life Sciences

02 Data, Information
and Knowledge

03 Decision Making and
Decision Support

04 DSS: from Expert Systems
to explainable AI

Day 2 - Cool Practice

05 Methods of Explainable-AI

Groupwork: Planning of a 500
bed Hospital - Bringing AI into
the workflows

Plenary: Presenting of the
developed concepts

Advance Organizer (1/2)



- **Case-based reasoning (CBR)** = process of solving new problems based on the solutions of similar past problems;
- **Certainty factor model (CF)** = a method for managing uncertainty in rule-based systems;
- **CLARION** = Connectionist Learning with Adaptive Rule Induction ON-line (CLARION) is a cognitive architecture that incorporates the distinction between implicit and explicit processes and focuses on capturing the interaction between these two types of processes. By focusing on this distinction, CLARION has been used to simulate several tasks in cognitive psychology and social psychology. CLARION has also been used to implement intelligent systems in artificial intelligence applications.
- **Clinical decision support (CDS)** = process for enhancing health-related decisions and actions with pertinent, organized clinical knowledge and patient information to improve health delivery;
- **Clinical Decision Support System (CDSS)** = expert system that provides support to certain reasoning tasks, in the context of a clinical decision;
- **Collective Intelligence** = shared group (symbolic) intelligence, emerging from cooperation/competition of many individuals, e.g. for consensus decision making;
- **Crowdsourcing** = a combination of "crowd" and "outsourcing" coined by Jeff Howe (2006), and describes a distributed problem-solving model; example for crowdsourcing is a public software beta-test;
- **Decision Making** = central cognitive process in every medical activity, resulting in the selection of a final choice of action out of several alternatives;
- **Decision Support System (DSS)** = is an IS including knowledge based systems to interactively support decision-making activities, i.e. making data useful;

- **DXplain** = a DSS from the Harvard Medical School, to assist making a diagnosis (clinical consultation), and also as an instructional instrument (education); provides a description of diseases, etiology, pathology, prognosis and up to 10 references for each disease;
- **Etiology** = in medicine (many) factors coming together to cause an illness (see causality)
- **Explainable AI** = Explainability = upcoming fundamental topic within recent AI; answering e.g. **why** a decision has been made
- **Expert-System** = emulates the decision making processes of a human expert to solve complex problems;
- **GAMUTS** in Radiology = Computer-Supported list of common/uncommon differential diagnoses;
- **ILIAD** = medical expert system, developed by the University of Utah, used as a teaching and testing tool for medical students in problem solving. Fields include Pediatrics, Internal Medicine, Oncology, Infectious Diseases, Gynecology, Pulmonology etc.
- **Interpretability** = there is no formal technical definition yet, but it is considered as a prerequisite for trust
- **MYCIN** = one of the early medical expert systems (Shortliffe (1970), Stanford) to identify bacteria causing severe infections, such as bacteremia and meningitis, and to recommend antibiotics, with the dosage adjusted for patient's body weight;
- **Reasoning** = cognitive (thought) processes involved in making medical decisions (clinical reasoning, medical problem solving, diagnostic reasoning);
- **Transparency** = opposite of opacity of black-box approaches, and connotes the ability to understand how a model works (that does not mean that it should always be understood, but that – in the case of necessity – it can be re-enacted)

Agenda

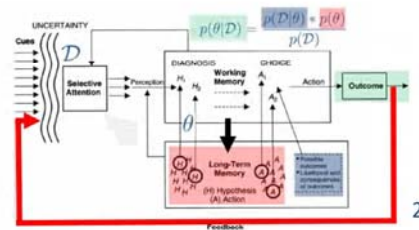
- **00 Reflection – follow-up from last lecture**
- **01 Decision Support Systems (DSS)**
- **02 Computers help making better decisions?**
- **03 History of DSS = History of AI**
- **04 Example: Towards Personalized Medicine**
- **05 Example: Case Based Reasoning (CBR)**
- **06 Towards Explainable AI**

- ... can apply your knowledge gained in the previous lectures to example systems of decision support;
- ... have an overview about the core principles and architecture of decision support systems;
- ... are familiar with the certainty factors as e.g. used in MYCIN;
- ... are aware of some design principles of DSS;
- ... have seen similarities between DSS and KDD on the example of computational methods in cancer detection;
- ... have seen basics of CBR systems;

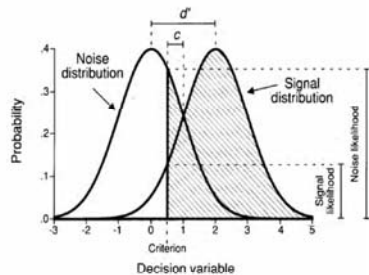




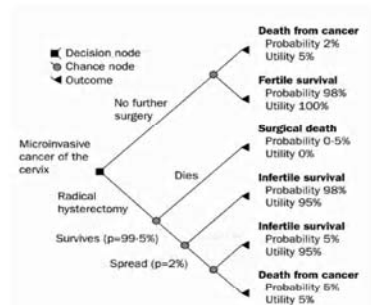
1



2



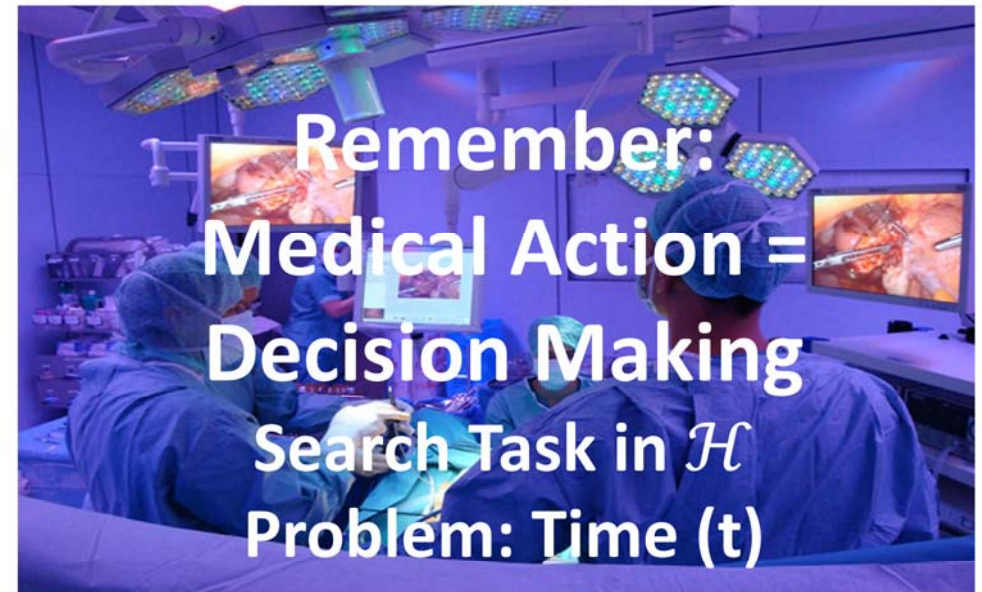
3



4

- Remember: Medicine is an complex application domain – dealing most of the time with **probable information!**
- Some challenges include:
 - (a) defining hospital system architectures in terms of generic tasks such as diagnosis, therapy planning and monitoring to be executed for (b) medical reasoning in (a);
 - (c) patient information management with (d) minimum uncertainty.
- Other challenges include: (e) knowledge acquisition and encoding, (f) human-computer interface and interaction; and (g) system integration into existing clinical legacy and proprietary environments, e.g. the enterprise hospital information system; to mention only a few.

01 Decision Support Systems





Health Informatics

13

Andreas Holzinger



Source: Cisco (2008).
Cisco Health Presence
Trial at Aberdeen Royal
Infirmary in Scotland

Health Informatics

14

Andreas Holzinger

The Medical Domain and Decision Making

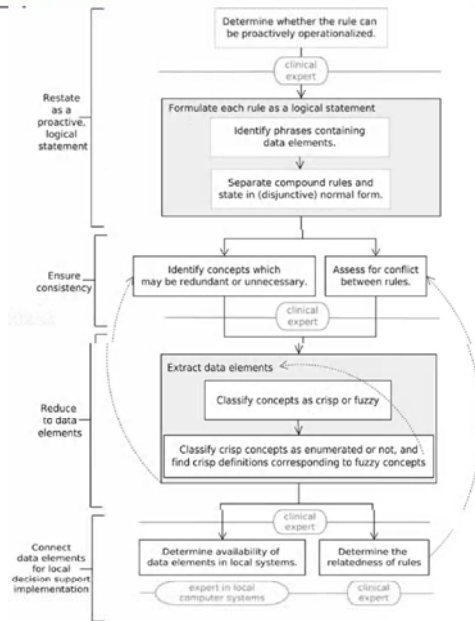
- 400 BC Hippocrates (460-370 BC), father of western medicine:
 - A medical record should accurately reflect the course of a disease
 - A medical record should indicate the probable cause of a disease
- **1890** William Osler (1849-1919), father of modern western medicine
 - **Medicine is a science of uncertainty and an art of probabilistic decision making**
- Today
 - Prediction models are based on data features, patient health status is modelled as high-dimensional feature vectors ...

Digression: Clinical Guidelines as DSS & Quality Measure

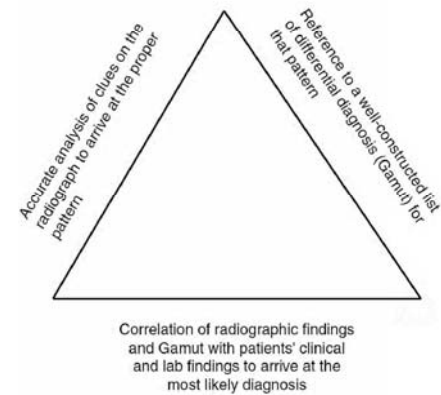
- Clinical guidelines are **systematically** developed documents to assist doctors and patient decisions about appropriate care;
- In order to build DS, based on a guideline, it is **formalized** (transformed from natural language to a logical algorithm), and
- **implemented** (using the algorithm to program a DSS);
- To increase the quality of care, they must be linked to a process of care, for example:
 - "80% of diabetic patients should have an HbA1c below 7.0" could be linked to processes such as:
 - "All diabetic patients should have an annual HbA1c test" and
 - "Patients with values over 7.0 should be rechecked within 2 months."
- **Condition-action rules** specify one or a few conditions which are linked to a specific action, in contrast to narrative guidelines which describe a series of branching or iterative decisions unfolding over time.
- Narrative guidelines and clinical rules are two ends of a continuum of clinical care standards.

Example: Clinical Guidelines

Medlock, S., Opondo, D., Eslami, S., Askari, M., Wierenga, P., de Rooij, S. E. & Abu-Hanna, A. (2011) LERM (Logical Elements Rule Method): A method for assessing and formalizing clinical rules for decision support. *International Journal of Medical Informatics*, 80, 4, 286-295.



Example: Triangulation to find diagnoses



Reeder, M. M. & Felson, B. 2003. *Reeder and Felson's gamuts in radiology: comprehensive lists of roentgen differential diagnosis*, New York, Springer Verlag.

Example - Gamuts in Radiology

REEDER AND FELSON'S GAMUTS IN RADIOLOGY

GAMUT G-25 EROSIVE GASTRITIS*

COMMON

1. Acute gastritis (eg. alcohol abuse)
2. Crohn's disease [1] [2]
3. Drugs (eg. aspirin [3] NSAID [4] steroids)
4. Helicobacter pylori infection [5]
5. Idiopathic
6. [Normal areas gastricae [6]]
7. Peptic ulcer hyperacidity

UNCOMMON

1. Corrosive gastritis [7]
2. Cryptosporidium antritis
3. [Lymphoma]
4. Opportunistic infection (eg. candidiasis [moniliasis] [8] herpes simplex; cytomegalovirus)
5. Postoperative gastritis
6. Radiation therapy
7. Zollinger-Ellison 3. [9] multiple endocrine neoplasia (MEN) 3.

* Superficial erosions or aphthoid ulcerations seen especially with double contrast technique.

[1] This condition does not actually cause the gamuted imaging finding, but can produce imaging changes that simulate it.

Reeder, M. M. & Felson, B. (2003) *Reeder and Felson's gamuts in radiology: comprehensive lists of roentgen differential diagnosis*. New York, Springer

<http://gamuts.isradiology.org/Gamuts.htm>

Example: Radiology Gamuts Ontology

Gamuts

Search Gamuts

Follow @RadiologyGamuts

Embolus

May Be Caused by

Common

- Atheromatous plaque with mural thrombus
- Atheromatous ulcer with mural thrombus
- Bacterial endocarditis
- Catheterization
- Endarterectomy
- Iatrogenic injury
- Paradoxical embolus
- Septic embolus
- Thrombophlebitis
- Venous thrombosis

Uncommon

- Arterial aneurysm
- Atrial fibrillation with left atrial thrombus
- Chagas myocardiopathy with intracardiac thrombus
- Left atrial myxoma
- Myocardial infarction with left ventricular thrombus
- Tumor embolus

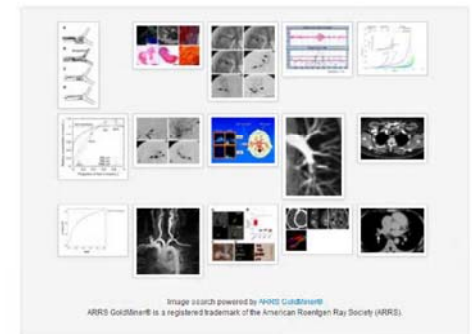
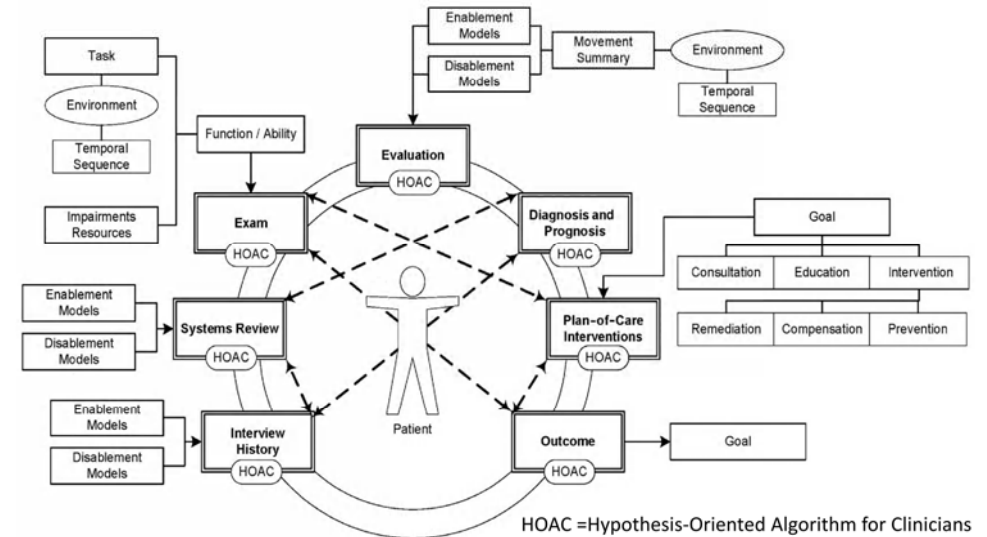


Image search powered by [RISORT GoldServer®](#)
RISORT GoldServer® is a registered trademark of the American Roentgen Ray Society (ARRS).



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



Schenkman, M., Deutsch, J. E. & Gill-Body, K. M. (2006) An Integrated Framework for Decision Making in Neurologic Physical Therapist Practice. *Physical Therapy*, 86, 12, 1681-1702.

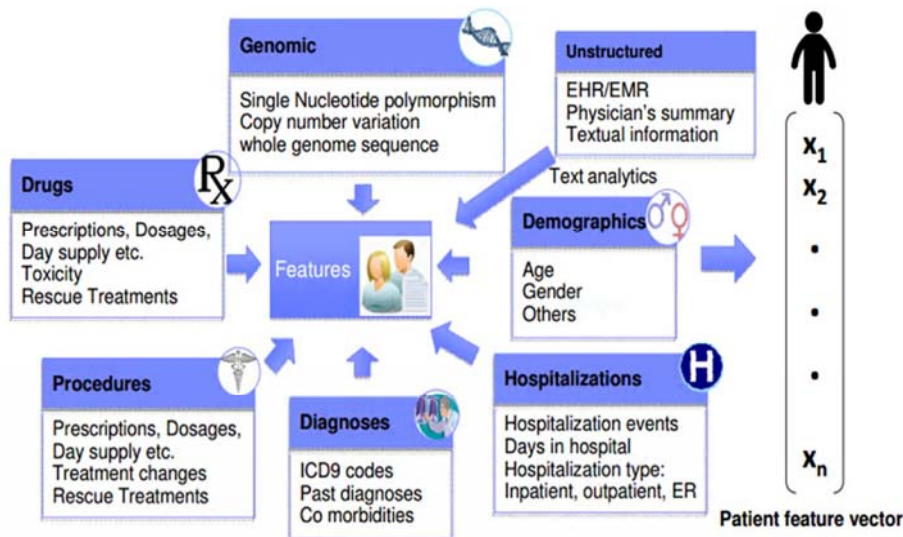
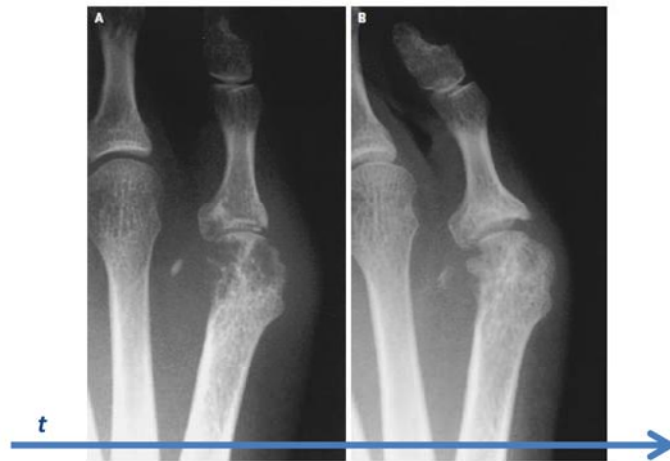


Image credit to Michal Rosen-Zvi

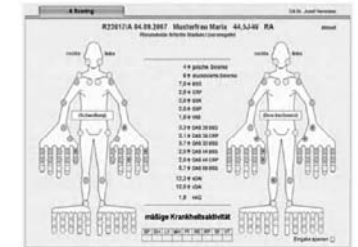


Chao, J., Parker, B. A. & Zvaifler, N. J. (2009) Accelerated Cutaneous Nodulosis Associated with Aromatase Inhibitor Therapy in a Patient with Rheumatoid Arthritis. *The Journal of Rheumatology*, 36, 5, 1087-1088.

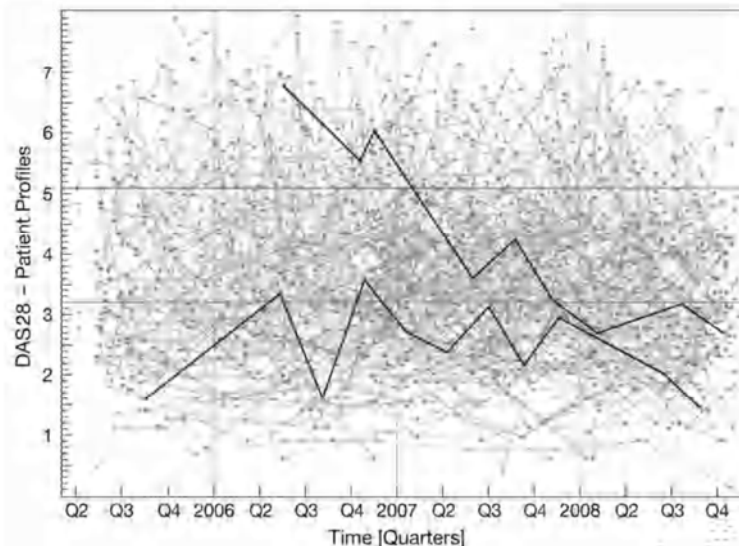


Ikari, K. & Momohara, S. (2005) Bone Changes in Rheumatoid Arthritis. *New England Journal of Medicine*, 353, 15, e13.

- 50+ Patients per day ~ 5000 data points per day ...
- Aggregated with specific scores (Disease Activity Score, DAS)
- Current patient status is related to previous data
- = convolution over time
- ⇒ **time-series data**



Simonik, K. M., Holzinger, A., Bloice, M. & Hermann, J. (2011). *Optimizing Long-Term Treatment of Rheumatoid Arthritis with Systematic Documentation*. *Pervasive Health - 5th International Conference on Pervasive Computing Technologies for Healthcare, Dublin, IEEE*, 550-554.



Simonik, K. M., Holzinger, A., Bloice, M. & Hermann, J. (2011). *Optimizing Long-Term Treatment of Rheumatoid Arthritis with Systematic Documentation*. *Pervasive Health - 5th International Conference on Pervasive Computing Technologies for Healthcare, Dublin, IEEE*, 550-554.

02 Can Computers help doctors to make better decisions?





<http://biomedicalcomputationreview.org/content/clinical-decision-support-providing-quality-healthcare-help-computer>



Two types of decisions (Diagnosis vs. Therapy)

- **Type 1 Decisions:** related to the **diagnosis**, i.e. computers are used to assist in diagnosing a disease on the basis of the individual patient data. Questions include:
 - What is the probability that this patient has a myocardial infarction on the basis of given data (patient history, ECG, ...)?
 - What is the probability that this patient has acute appendices, given the signs and symptoms concerning abdominal pain?
- **Type 2 Decisions:** related to **therapy**, i.e. computers are used to select the best therapy on the basis of clinical evidence, e.g.:
 - What is the best therapy for patients of age x and risks y , if an obstruction of more than z % is seen in the left coronary artery?
 - What amount of insulin should be prescribed for a patient during the next 5 days, given the blood sugar levels and the amount of insulin taken during the recent weeks?

Bemmel, J. H. V. & Musen, M. A. 1997. *Handbook of Medical Informatics*, Heidelberg, Springer.

Example: Knee Surgery of a Soccer Player



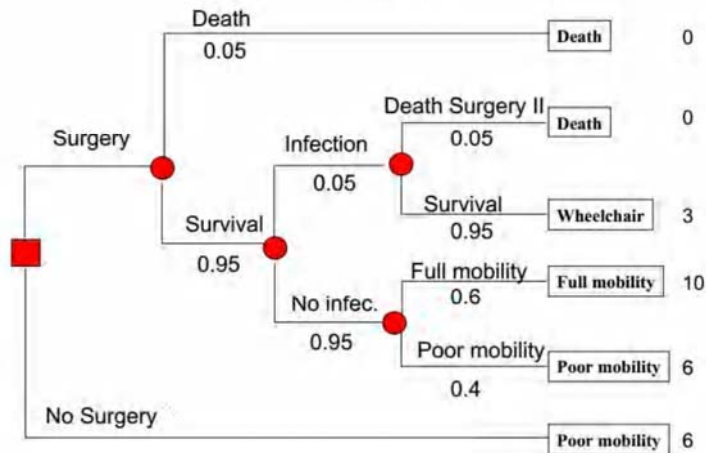
- Example of a Decision Problem
- Soccer player considering knee surgery
- Uncertainties:
 - Success: recovering full mobility
 - Risks: infection in surgery (if so, needs another surgery and may loose more mobility)
- Survival chances of surgery

Harvard-MIT Division of Health Sciences and Technology

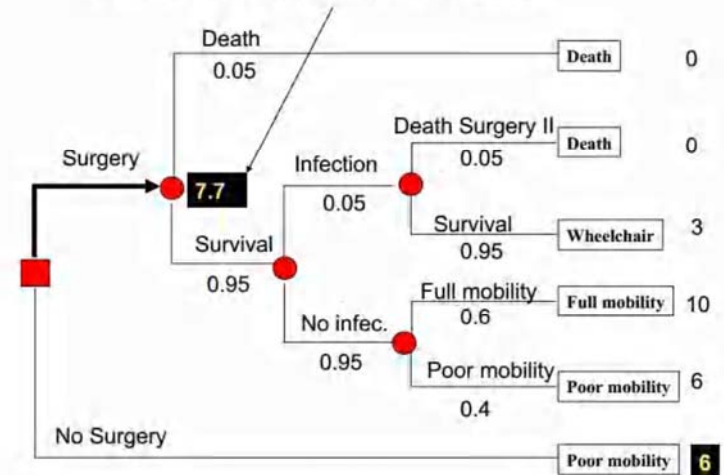
HST.951J: Medical Decision Support, Fall 2005

Instructors: Professor Lucila Ohno-Machado and Professor Staal Vinterbo

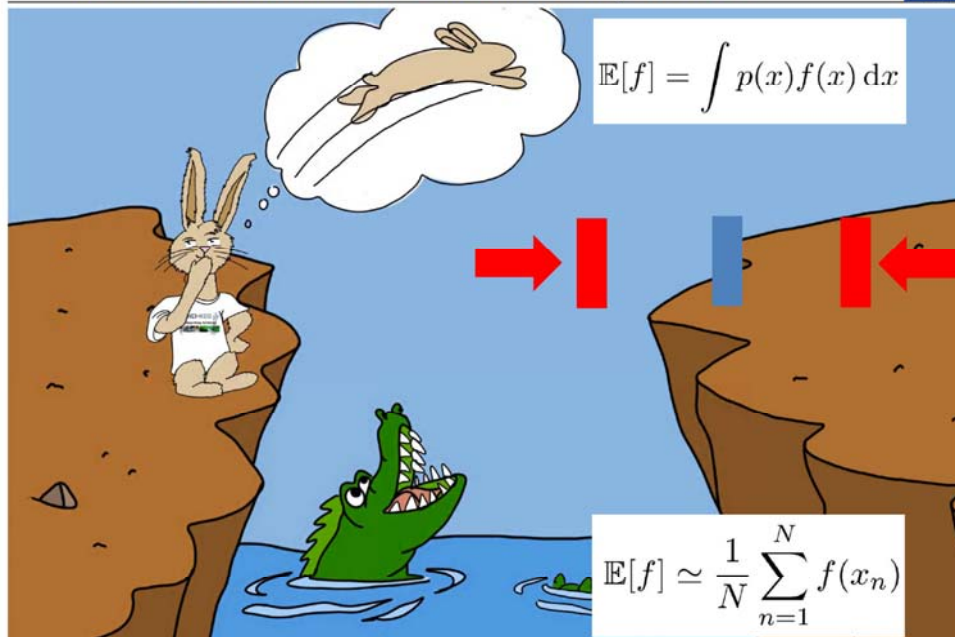
Knee Surgery



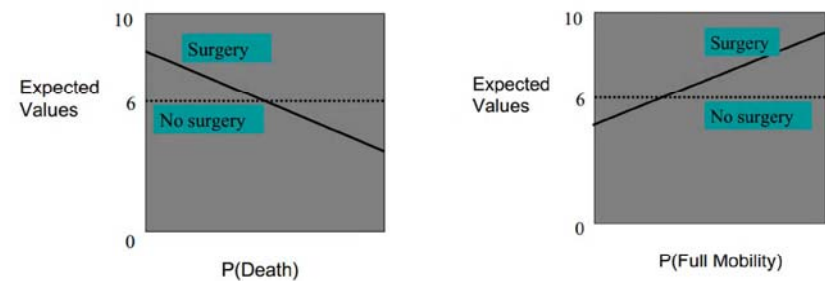
Expected Value of Surgery

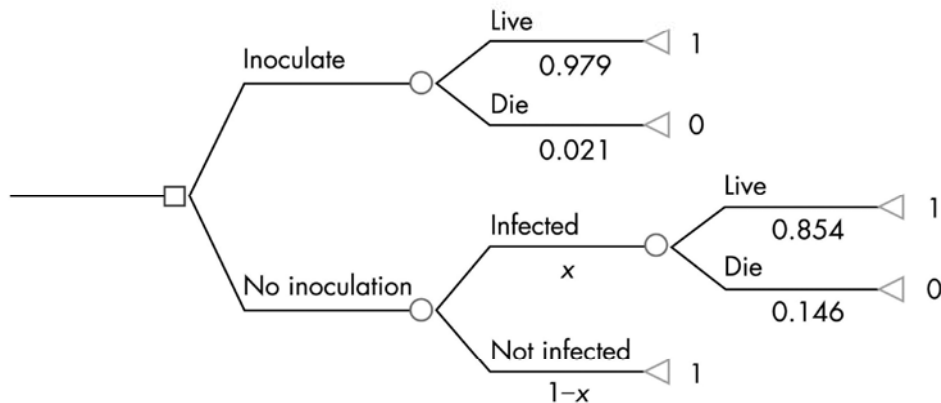


Estimate Confidence Interval: Uncertainty matters !



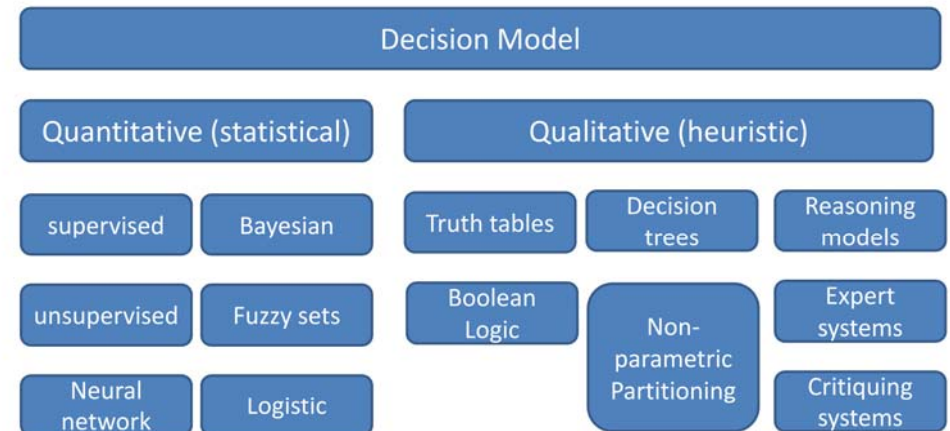
Effect of probabilities in the decision





Ferrando, A., Pagano, E., Scaglione, L., Petrinco, M., Gregori, D. & Ciccone, G. (2009) A decision-tree model to estimate the impact on cost-effectiveness of a venous thromboembolism prophylaxis guideline. *Quality and Safety in Health Care*, 18, 4, 309-313.

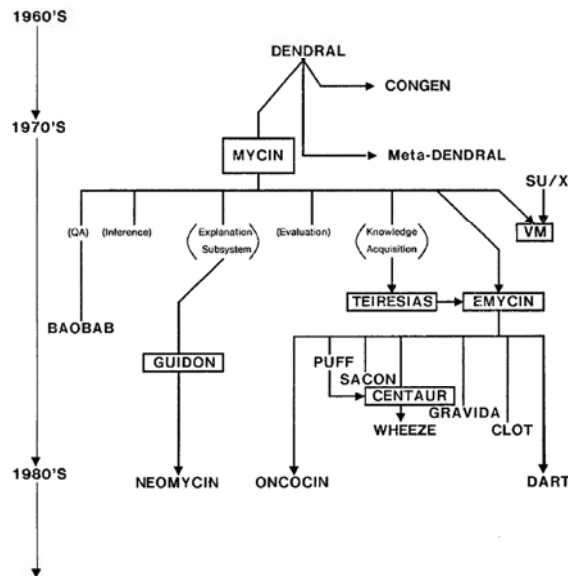
03 History of DSS = History of AI



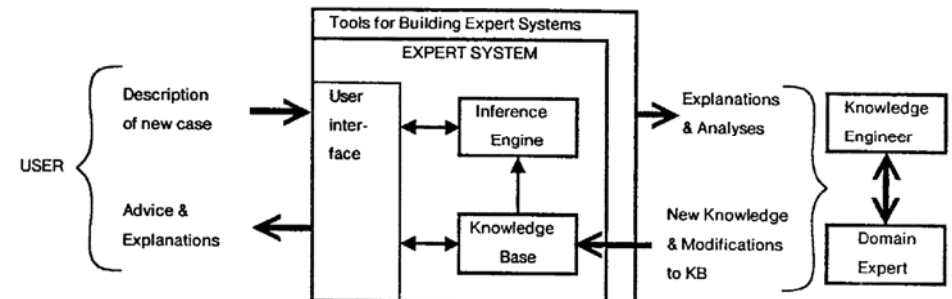
Extended by A. Holzinger after: Bammel, J. H. v. & Musen, M. A. (1997) *Handbook of Medical Informatics*. Heidelberg, Springer.

A ultrashort history of Early AI

- **1943** McCulloch, W.S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5, (4), 115-133, doi:10.1007/BF02459570.
- **1950** Turing, A.M. Computing machinery and intelligence. *Mind*, 59, (236), 433-460.
- **1959** Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3, (3), 210-229, doi:10.1147/rd.33.0210.
- **1975** Shortliffe, E.H. & Buchanan, B.G. 1975. A model of inexact reasoning in medicine. *Mathematical biosciences*, 23, (3-4), 351-379, doi:10.1016/0025-5564(75)90047-4.

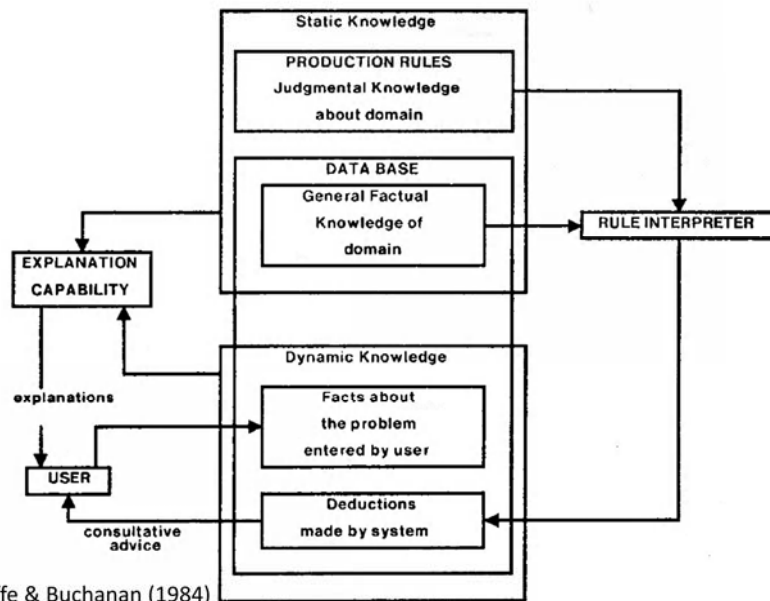


Shortliffe, E. H. & Buchanan, B. G. (1984) *Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project*. Addison-Wesley.



Shortliffe, T. & Davis, R. (1975) Some considerations for the implementation of knowledge-based expert systems *ACM SIGART Bulletin*, 55, 9-12.

Static Knowledge versus dynamic knowledge



Shortliffe & Buchanan (1984)

Dealing with uncertainty in the real world

- The information available to humans is often imperfect – imprecise - uncertain.
- This is especially in the medical domain the case.
- An **human agent** can cope with deficiencies.
- Classical logic permits only **exact reasoning**:
- IF A is true THEN A is non-false and
IF B is false THEN B is non-true
- Most real-world problems do not provide this exact information, mostly it is inexact, incomplete, uncertain and/or **un-measurable!**

Harcourt Fenton Mudd: Now listen, Spock, you may be a wonderful science officer but, believe me, you couldn't sell fake patents to your mother!

Spock: I fail to understand why I should care to induce my mother to purchase falsified patents.



- MYCIN is a rule-based Expert System, which is used for therapy planning for patients with bacterial infections
- Goal oriented strategy (“Rückwärtsverkettung”)
- To every rule and every entry a certainty factor (CF) is assigned, which is between 0 und 1
- Two measures are derived:
 - MB: measure of belief
 - MD: measure of disbelief
- Certainty factor – CF of an element is calculated by:

$$CF[h] = MB[h] - MD[h]$$
- CF is positive, if more evidence is given for a hypothesis, otherwise CF is negative
- $CF[h] = +1 \rightarrow h$ is 100 % true
- $CF[h] = -1 \rightarrow h$ is 100% false

Original Example from MYCIN

h_1 = The identity of ORGANISM-1 is streptococcus

h_2 = PATIENT-1 is febrile

h_3 = The name of PATIENT-1 is John Jones

$CF[h_1, E] = .8$: There is strongly suggestive evidence (.8) that the identity of ORGANISM-1 is streptococcus

$CF[h_2, E] = -.3$: There is weakly suggestive evidence (.3) that PATIENT-1 is not febrile

$CF[h_3, E] = +1$: It is definite (1) that the name of PATIENT-1 is John Jones

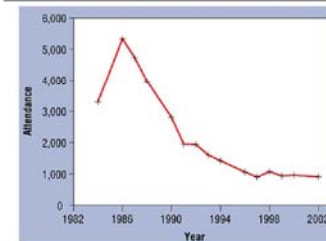
Shortliffe, E. H. & Buchanan, B. G. (1984) *Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project*. Addison-Wesley.

MYCIN was *no* success in the clinical practice

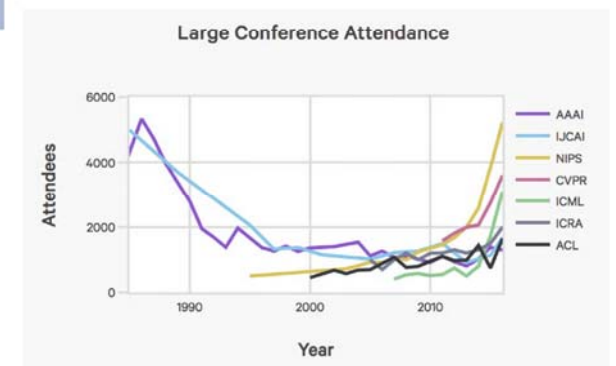




<https://blogs.dxc.technology/2017/04/25/are-we-heading-toward-an-ai-winter/>



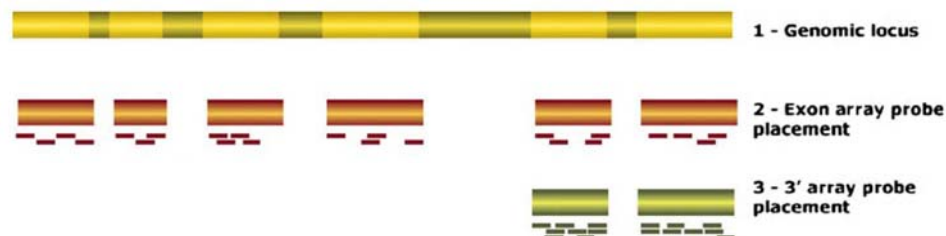
<https://www.computer.org/csli/mags/ex/2003/03/x3018.html>



<https://medium.com/machine-learning-in-practice/nips-accepted-papers-stats-26f124843aa0>

04 Towards P4-Medicine

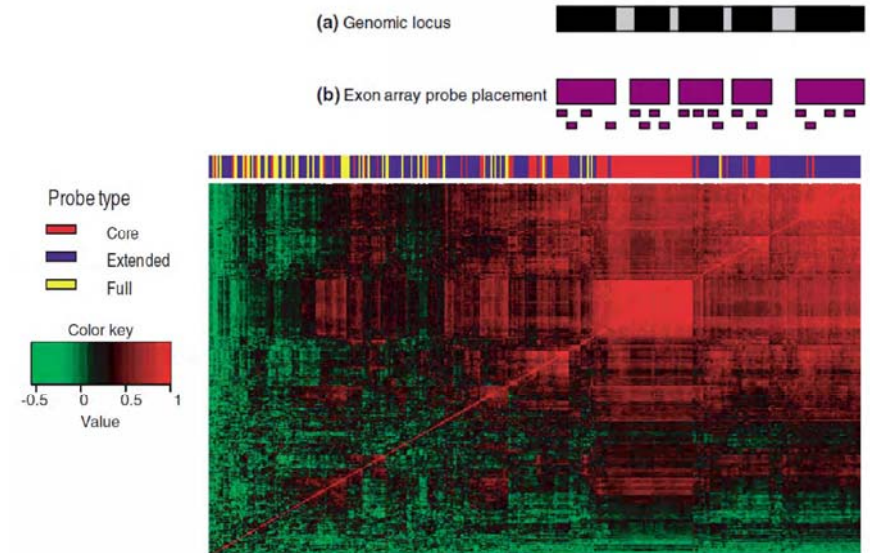
Slide 8-23 Computational leukemia cancer detection 1/6



Exon array structure. Probe design of exon arrays. (1) Exon—intron structure of a gene. Gray boxes represent introns, rest represent exons. Introns are not drawn to scale. (2) Probe design of exon arrays. Four probes target each putative exon. (3) Probe design of 30expression arrays. Probe target the 30end of mRNA sequence.

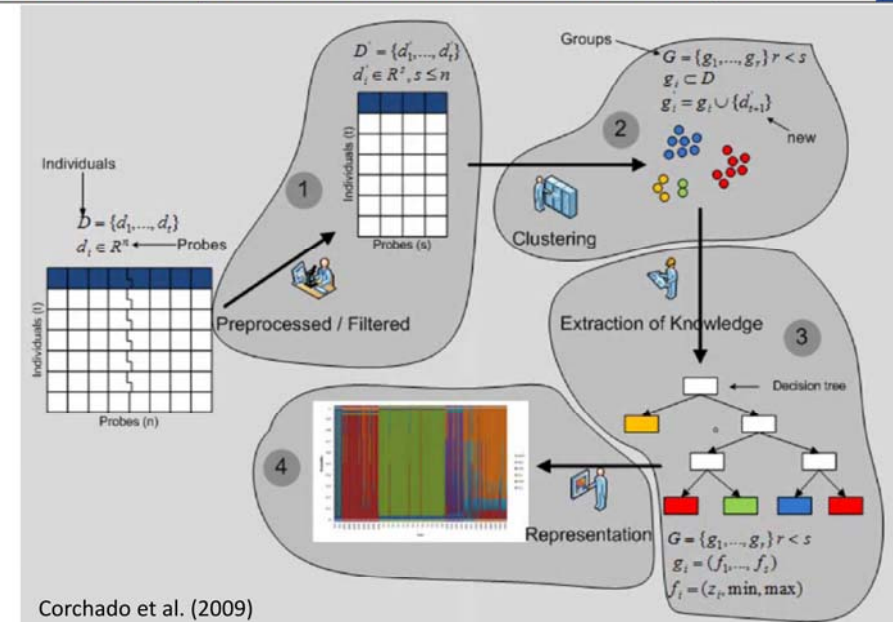
Corchado, J. M., De Paz, J. F., Rodriguez, S. & Bajo, J. (2009) Model of experts for decision support in the diagnosis of leukemia patients. *Artificial Intelligence in Medicine*, 46, 3, 179-200.

Slide 8-22 Example: Exon Arrays



Kapur, K., Xing, Y., Ouyang, Z. & Wong, W. (2007) Exon arrays provide accurate assessments of gene expression. *Genome Biology*, 8, 5, R82.

Slide 8-24 Computational leukemia cancer detection 2/6



Corchado et al. (2009)

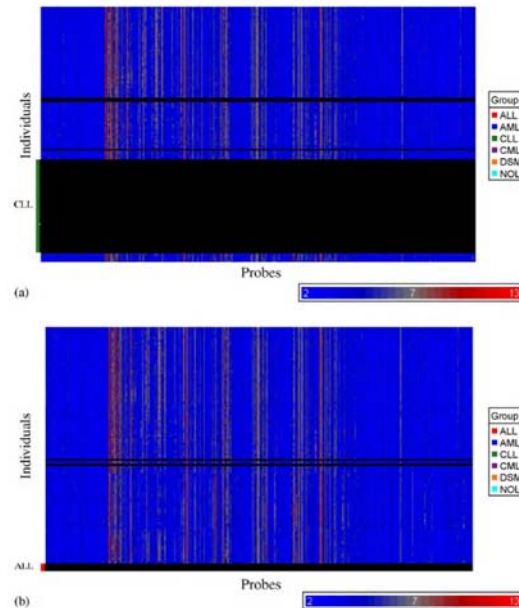
Slide 8-25 Computational leukemia cancer detection 3/6



A = acute, C = chronic,
L = lymphocytic, M = myeloid

- **ALL** = cancer of the blood AND bone marrow caused by an abnormal proliferation of lymphocytes.
- **AML** = cancer in the bone marrow characterized by the proliferation of myeloblasts, red blood cells or abnormal platelets.
- **CLL** = cancer characterized by a proliferation of lymphocytes in the bone marrow.
- **CML** = caused by a proliferation of white blood cells in the bone marrow.
- **MDS** (Myelodysplastic Syndromes) = a group of diseases of the blood and bone marrow in which the bone marrow does not produce a sufficient amount of healthy cells.
- **NOL** (Normal) = No leukemias

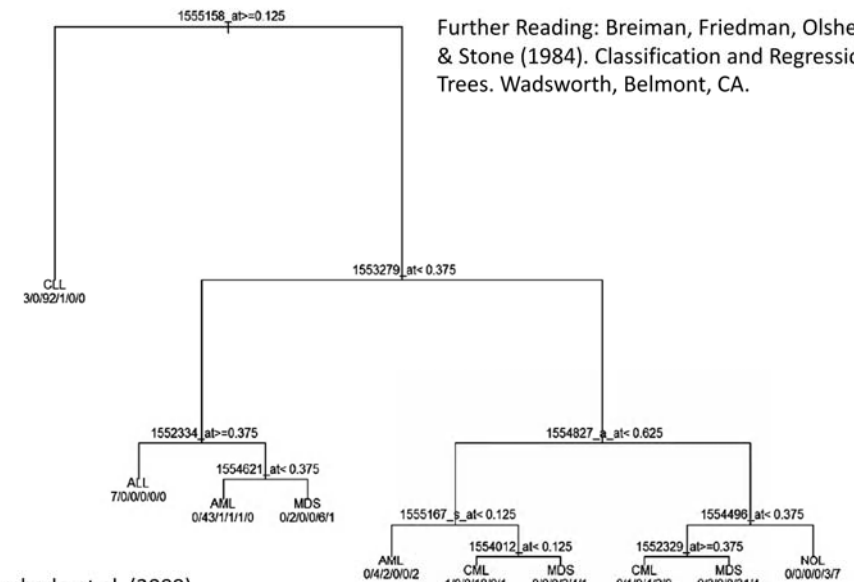
Corchado et al. (2009)



8-26 Computational leukemia cancer detection 4/6



Further Reading: Breiman, Friedman, Olshen, & Stone (1984). Classification and Regression Trees. Wadsworth, Belmont, CA.

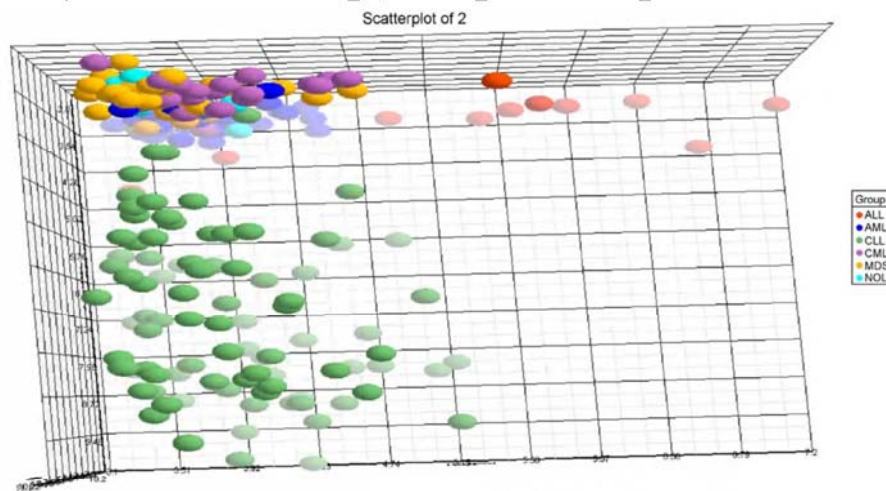


Corchado et al. (2009)

8-27 Computational leukemia cancer detection 5/6



Classification CLL—ALL. Representation of the probes of the decision tree which classify the CLL and ALL to 155158_at, 1553279_at and 1552334_at



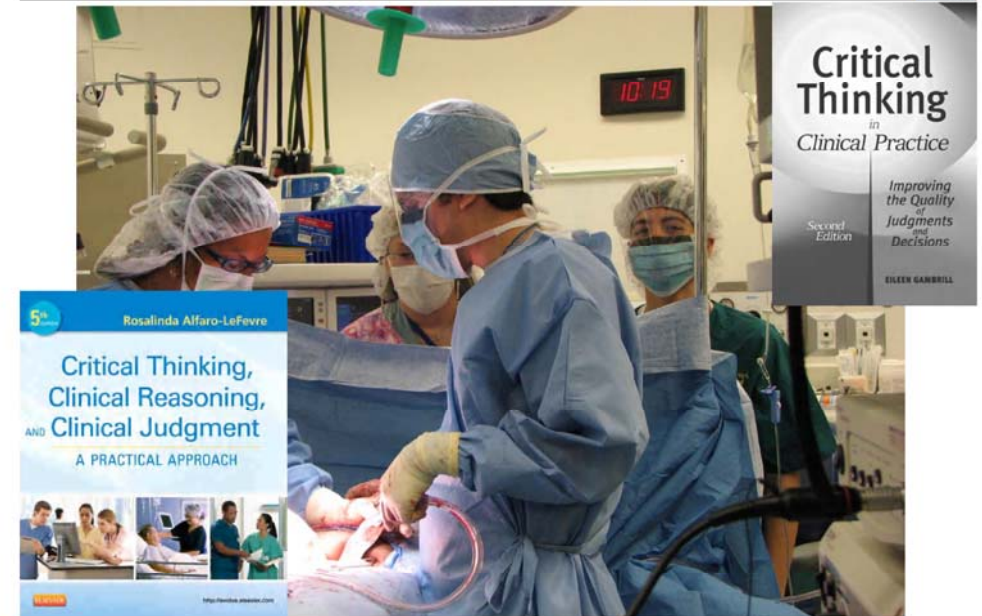
Corchado et al. (2009)

Computational leukemia cancer detection 6/6

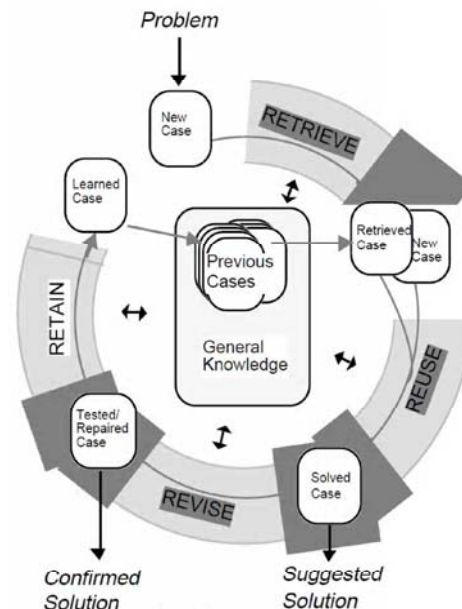


- The model of Corchado et al. (2009) combines:
 - 1) methods to **reduce the dimensionality** of the original data set;
 - 2) pre-processing and data filtering techniques;
 - 3) a clustering method to classify patients; and
 - 4) extraction of knowledge techniques
- The system reflects how human experts work in a lab, but
 - 1) **reduces the time** for making predictions;
 - 2) **reduces the rate of human error**; and
 - 3) **works with high-dimensional data** from exon arrays

05 Example: Case Based Reasoning (CBR)

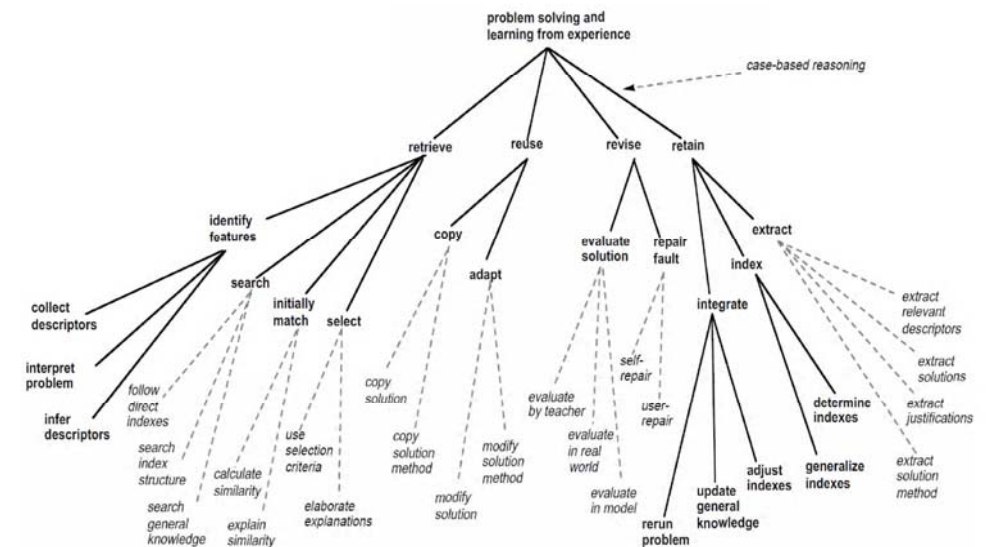


Slide 8-30 Case Based Reasoning (CBR) Basic principle



Aamodt, A. & Plaza, E. (1994) Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7, 1, 39-59.

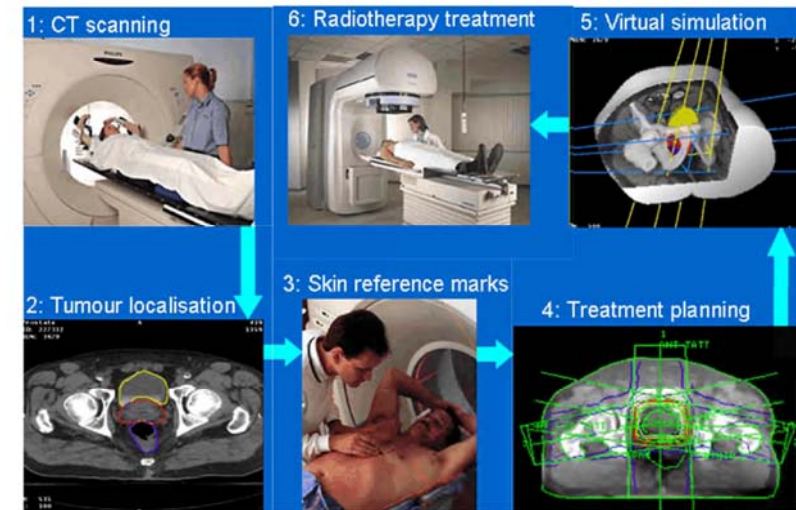
Slide 8-31 The task-method decomposition of CBR



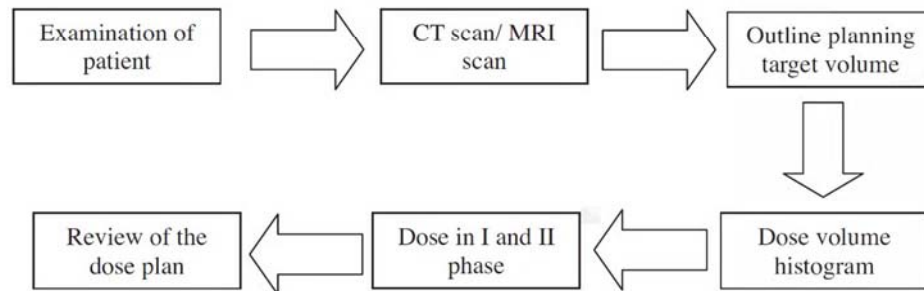
Aamodt & Plaza (1994)



Source: <http://www.teachingmedicalphysics.org.uk>



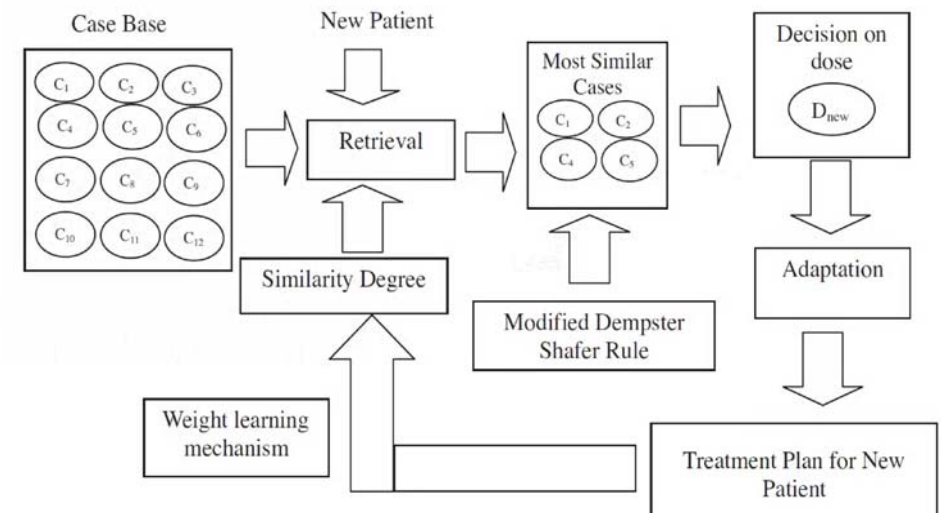
Source: Imaging Performance Assessment of CT Scanners Group, <http://www.impactscan.org>



Measures:

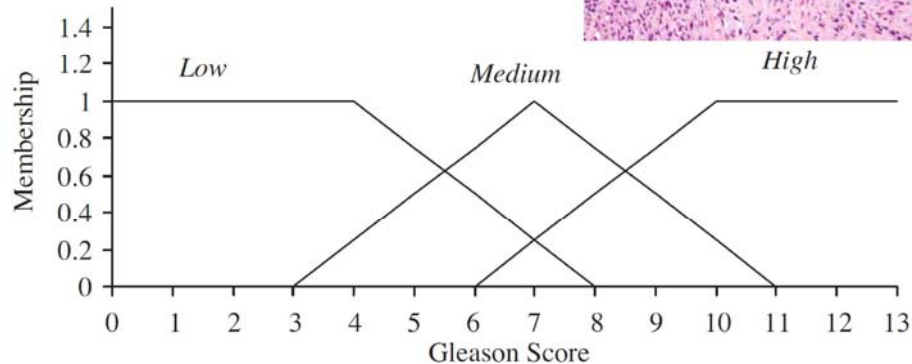
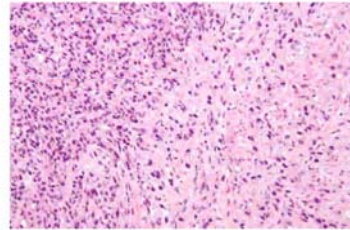
- 1) Clinical Stage = a labelling system
- 2) Gleason Score = grade of prostate cancer = integer between 1 to 10; and
- 3) Prostate Specific Antigen (PSA) value between 1 to 40
- 4) Dose Volume Histogram (DVH) = pot. risk to the rectum (66, 50, 25, 10 %)

Petrovic, S., Mishra, N. & Sundar, S. (2011) A novel case based reasoning approach to radiotherapy planning. *Expert Systems With Applications*, 38, 9, 10759-10769.



Petrovic, S., Mishra, N. & Sundar, S. (2011) A novel case based reasoning approach to radiotherapy planning. *Expert Systems With Applications*, 38, 9, 10759-10769.

Gleason score evaluates the grade of prostate cancer. Values: integer within the range



Petrovic, S., Mishra, N. & Sundar, S. (2011) A novel case based reasoning approach to radiotherapy planning. *Expert Systems With Applications*, 38, 9, 10759-10769.

06 Towards Explainable AI

Petrovic et al. (2011)

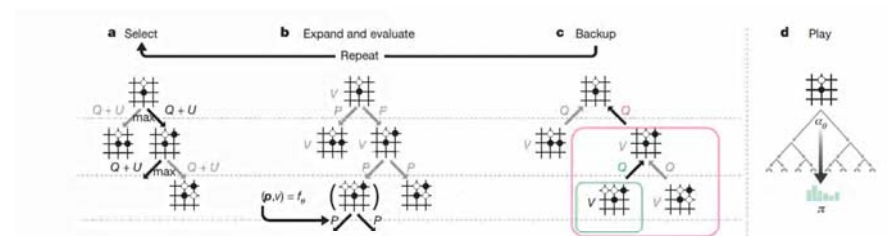
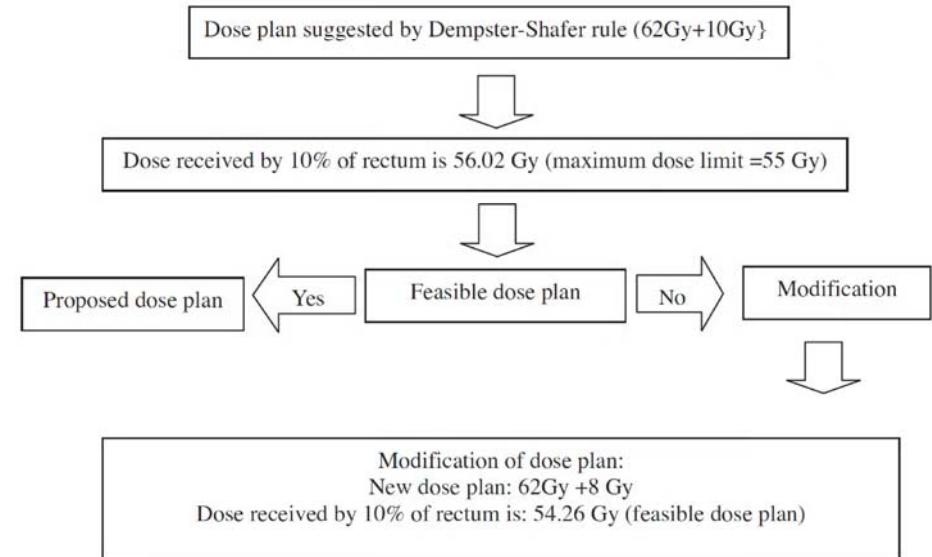


Figure 2 | MCTS in AlphaGo Zero. a. Each simulation traverses the tree by selecting the edge with maximum action value Q , plus an upper confidence bound U that depends on a stored prior probability P and visit count N for that edge (which is incremented once traversed). b. The leaf node is expanded and the associated position s is evaluated by the neural network ($P(s, \cdot), V(s) = f_\theta(s)$); the vector of P values are stored in

the outgoing edges from s . c. Action value Q is updated to track the mean of all evaluations V in the subtree below that action. d. Once the search is complete, search probabilities π are returned, proportional to $N^{1/\tau}$, where N is the visit count of each move from the root state and τ is a parameter controlling temperature.

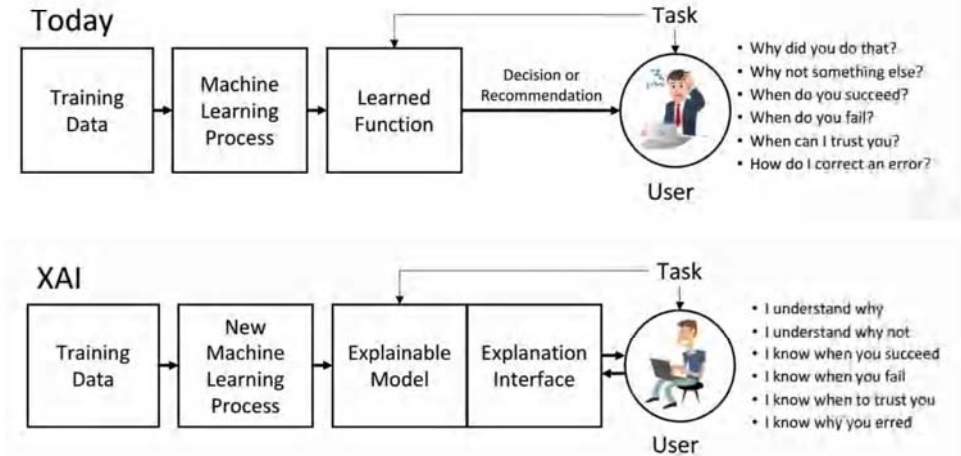
19 OCTOBER 2017 | VOL 550 | NATURE | 355

$$(p, v) = f_\theta(s) \text{ and } l = (z - v)^2 - \pi^T \log p + c \|\theta\|^2$$

David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George Van Den Driessche, Thore Graepel & Demis Hassabis 2017. Mastering the game of go without human knowledge. *Nature*, 550, (7676), 354-359, doi:doi:10.1038/nature24270.

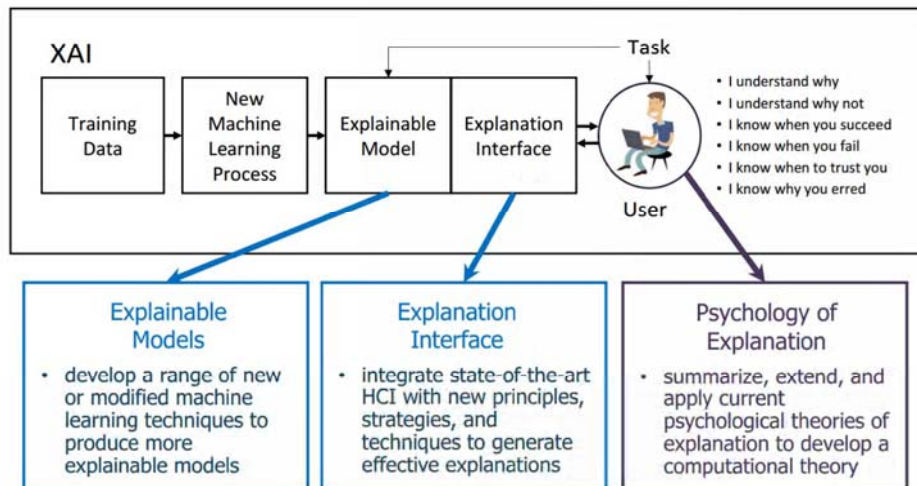


David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis 2016. Mastering the game of Go with deep neural networks and tree search. Nature, 529, (7587), 484-489, doi:10.1038/nature16961.



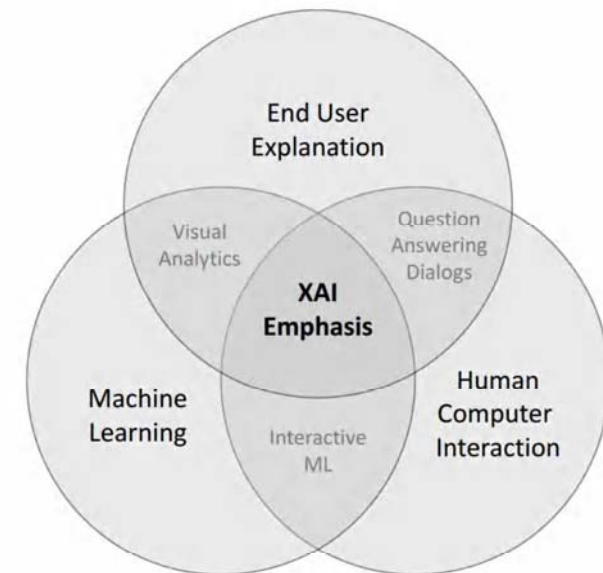
David Gunning 2016. Explainable artificial intelligence (XAI): Technical Report Defense Advanced Research Projects Agency DARPA-BAA-16-53, Arlington, USA, DARPA (free for public distribution)

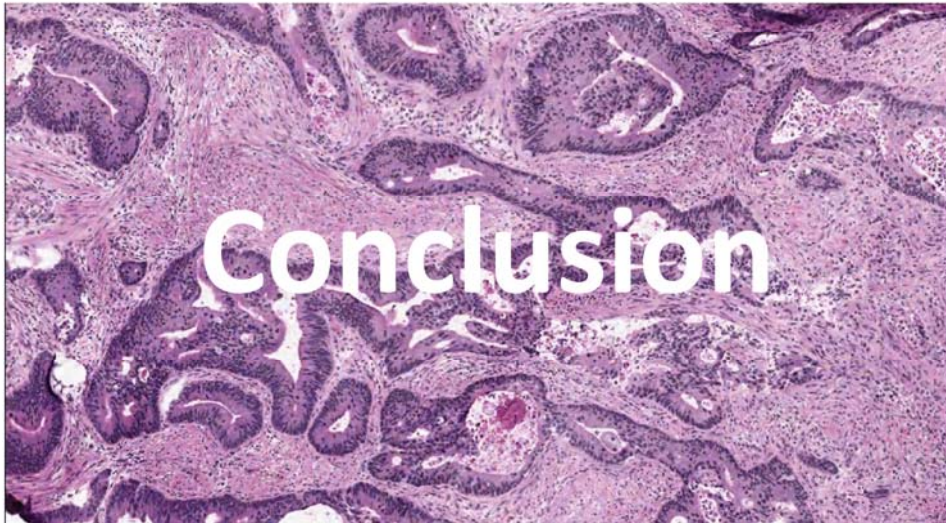
Explainable Models – explainable Interfaces



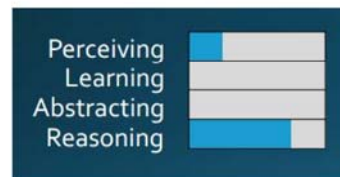
David Gunning 2016. Explainable artificial intelligence (XAI): Technical Report Defense Advanced Research Projects Agency DARPA-BAA-16-53, Arlington, USA, DARPA (free for public distribution)

DARPA – interactive Machine Learning is important





The first wave of AI (1943-1975): Handcrafted Knowledge



- Engineers create a set of logical rules to represent knowledge (Rule based Expert Systems)
- Advantage: works well in narrowly defined problems of well-defined domains
- Disadvantage: No adaptive learning behaviour and poor handling of $p(x)$

Image credit to John Launchbury

- Computational approaches can find in R^n what no human is able to see
- However, still there are many hard problems where a human expert in R^2 can understand the **context** and bring in experience, expertise, knowledge, intuition, ...
- Black box approaches can not explain **WHY** a decision has been made ...

The second wave of AI (1975 –): Statistical Learning



- Engineers create learning models for specific tasks and train them with “big data” (e.g. Deep Learning)
- Advantage: works well for standard classification tasks and has prediction capabilities
- Disadvantage: No contextual capabilities and minimal reasoning abilities

Image credit to John Launchbury

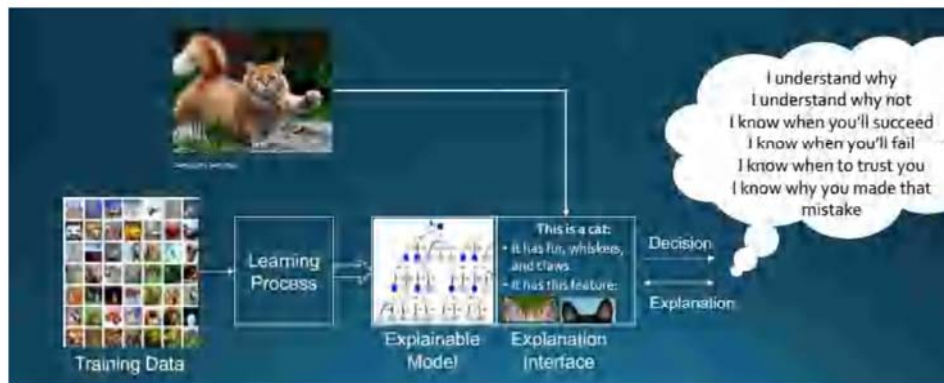
Perceiving
Learning
Abstracting
Reasoning



- A contextual model can perceive, learn and understand and abstract and reason
- Advantage: can use transfer learning for adaptation on unknown unknowns
- Disadvantage: Superintelligence ...

Image credit to John Launchbury

Contextual adaptations can explain decisions



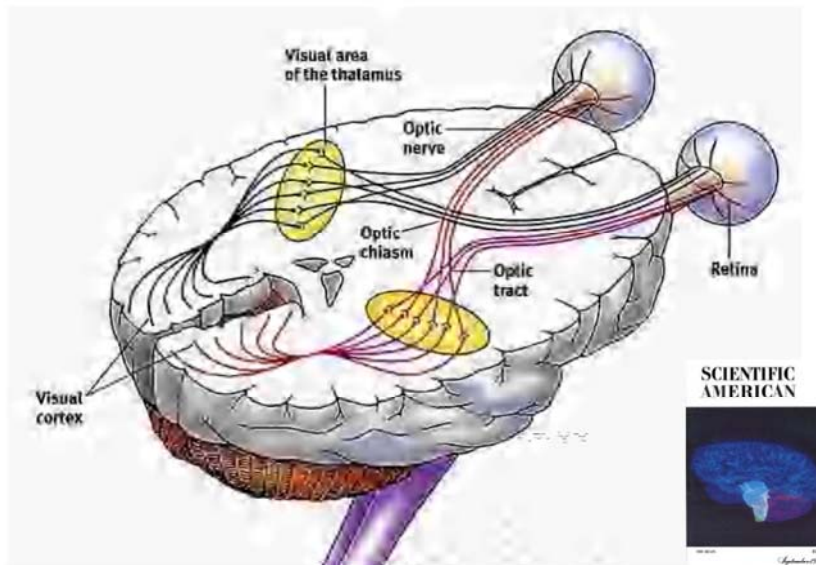
- Myth 1a: Superintelligence by 2100 is inevitable!
- Myth 1b: Superintelligence by 2100 is impossible!
- **Fact: We simply don't know it!**
- Myth 2: Robots are our main concern
Fact: Cyberthreats are the main concern: it needs no body – only an Internet connection
- Myth 3: AI can never control us humans
Fact: Intelligence is an enabler for control: We control tigers by being smarter ...



Thank you!

Appendix

Slide 7-7 Example: Visual Information Processing

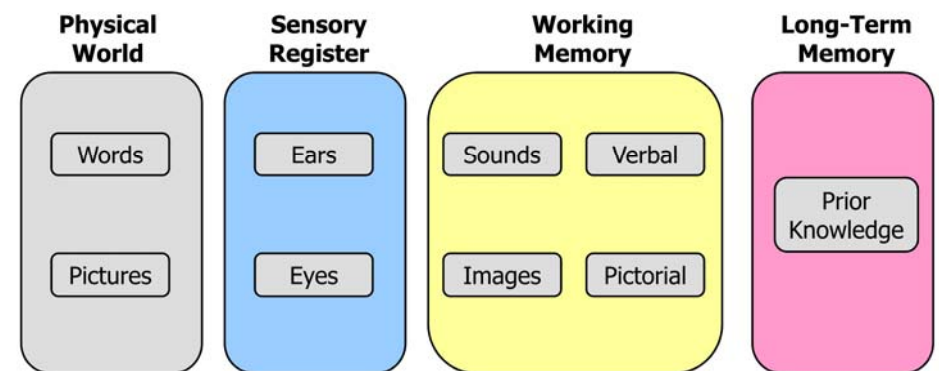


Source: Department of Neuroscience, The Mount Sinai School of Medicine (2004)

Answers to the Reflection Questions

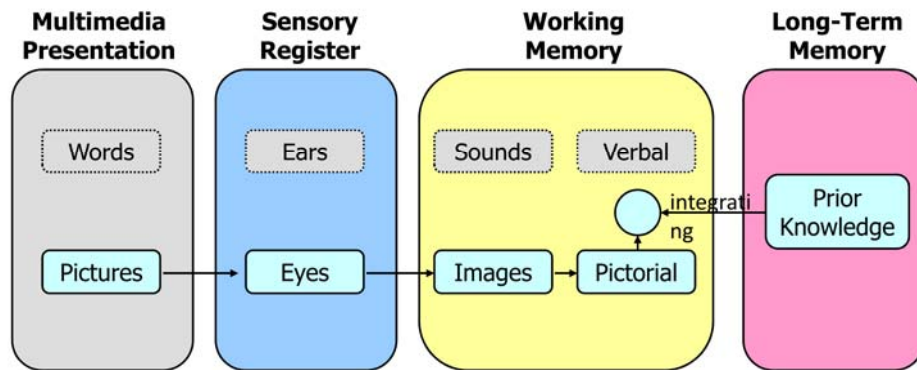
- 1) Humans are good in pattern recognition of dimensions less than three; Natural language and NP-hard problems; computers are better in high-dim, rule-based environments etc.
- 2) How we make decisions – constant hypotheses generation and the most likely fitting is selected
- 3) ROC-Curve - Signal-Noise detection theory: describes when we say yes/no
- 4) Decision analysis was founded on the work of von Neumann and Morgenstern (1947) who described a model for human decision making known as the theory of expected utility.

Slide 7-8 Schematic Information Processing Chain



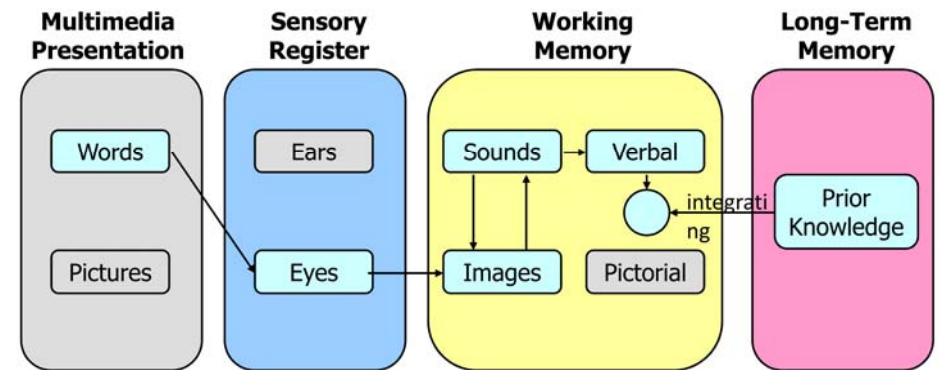
cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

a) Processing of visual information (PICTURES)



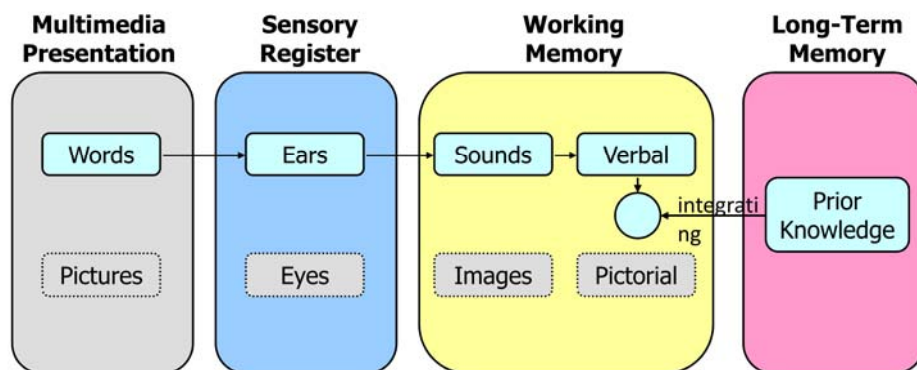
cf. with Paivio (1973), Mayer & Moreno (1998), Holzinger (2000), Schnotz & Bannert (2002)

b) Processing of visual information (PRINTED WORDS)



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

c) Processing of audio information (SPOKEN WORDS)



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

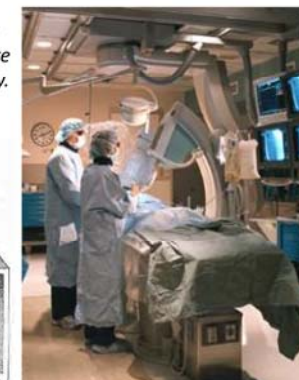


Stanford Heuristic Programming Project
Memo HPP-78-1
Computer Science Department
Report No. STAN-CS-78-649

February 1978

E. Feigenbaum, J. Lederberg, B. Buchanan, E. Shortliffe

Rheingold, H. (1985) *Tools for thought: the history and future of mind-expanding technology*. New York, Simon & Schuster.



DENDRAL AND META-DENDRAL:
THEIR APPLICATIONS DIMENSION
by
Bruce G. Buchanan and Edward A. Feigenbaum

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY



Buchanan, B. G. & Feigenbaum, E. A. (1978) DENDRAL and META-DENDRAL: their applications domain. *Artificial Intelligence*, 11, 1978, 5-24.

- The Quiz-Slide will be shown during the course

