Mini Course

Fundamentals of Medical AI

Part 03

From Decision Making to Decision Support

Andreas Holzinger

Human-Centered AI Lab (Holzinger Group)

Institute for Medical Informatics/Statistics, Medical University Graz, Austria

and

Explainable AI-Lab, Alberta Machine Intelligence Institute, Edmonton, Canada

Overview

Primer on Probability & Information	
Part 1 Theory	Part 2 Practice
01 Introduction to Medical AI and Machine Learning for Health	05 Methods of Explainable AI
02 Data, Information and Knowledge	06 Social, Ethical and Legal Aspects of Medical AI
03 Human Decision Making and AI Decision Support	07 Project: Bringing AI into medical workflows
04 Causal Reasoning and Interpretable AI	08 Presentation of the developed concepts
Written Exam	

Agenda

- **00 Reflection – follow up from last lecture**
- **01 Medical Action = Decision Making**
- **02 Can AI help doctors to make better decisions ?**
- **03 Human Information Processing**
- **04 Probabilistic Decision Theory**
- **05 Example: P4 Medicine**
- **06 Example: Case Based Reasoning**
- **Conclusion**

00 Reflection

Reflection from last lecture

1

INTERNATIONAL

2

$$C_1 \sqcup \ldots \sqcup C_n$$

3

World Health Organization

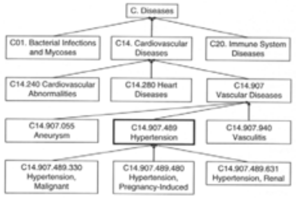
ICD 11

4

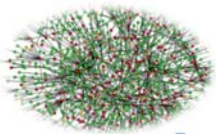
A
24184005|Finding of increased blood pressure (finding) →
38936003|Abnormal blood pressure (finding) AND
roleGroup SOME
(363714003|Interprets (attribute) SOME
75367002|Blood pressure (observable entity))

B
12763006|Finding of decreased blood pressure (finding)→
392570002|Blood pressure finding (finding) AND
roleGroup SOME
(363714003|Interprets (attribute) SOME
75367002|Blood pressure (observable entity))

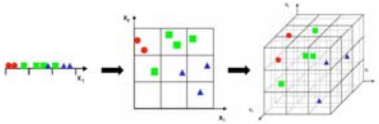
5



6



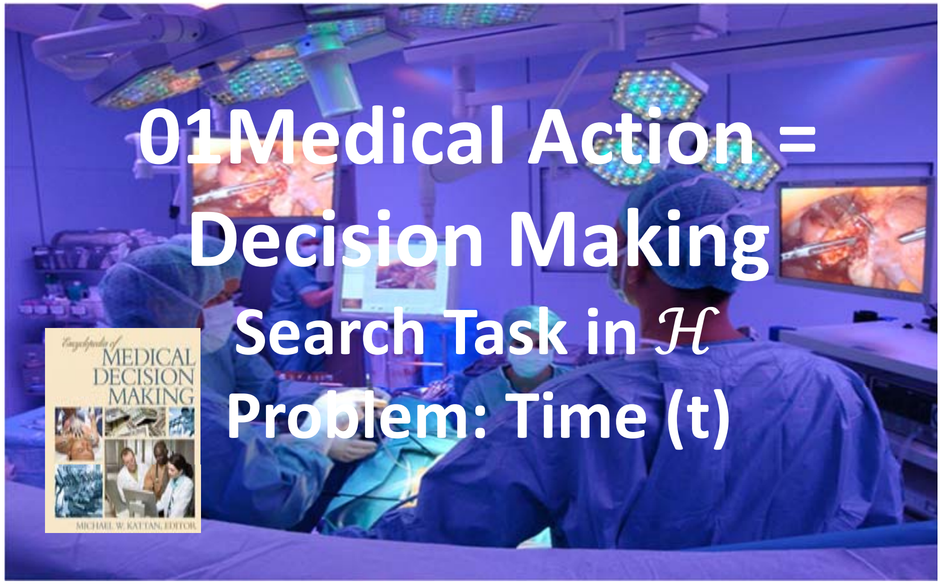
7



8

Solutions in the Appendix

What is medical decision making ?



Michael W. Kattan (ed.) (2009). Encyclopedia of medical decision making, London: Sage.

What is the main limitation in decision making ?

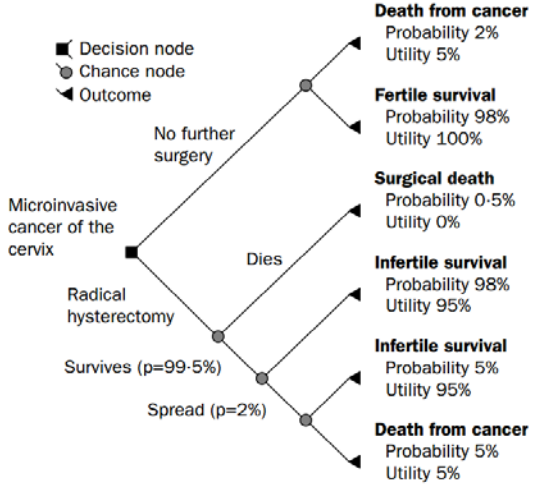
Decision Making is central in any (medical) work !

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

What are the roots of medical 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 – 1890 (!)**
  - **this is exactly what current AI/machine learning is tackling -** prediction models are based on data features, patient health status is modelled as high-dimensional feature vectors …

Where are decision trees coming from ?

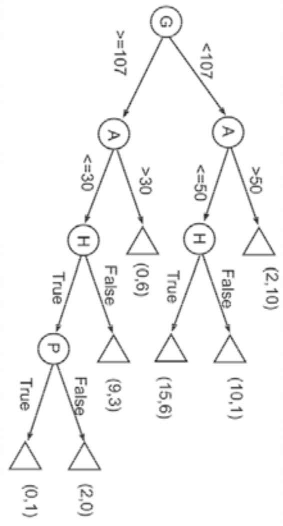


Physician treating a patient approx. 480 B.C.
Beazley (1963), Attic Red-figured Vase-Painters, 813, 96.
Department of Greek, Etruscan and Roman Antiquities, Sully, 1st floor, Campana Gallery, room 43 Louvre, Paris

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

Decision forests > decision tree > "interpretable AI"

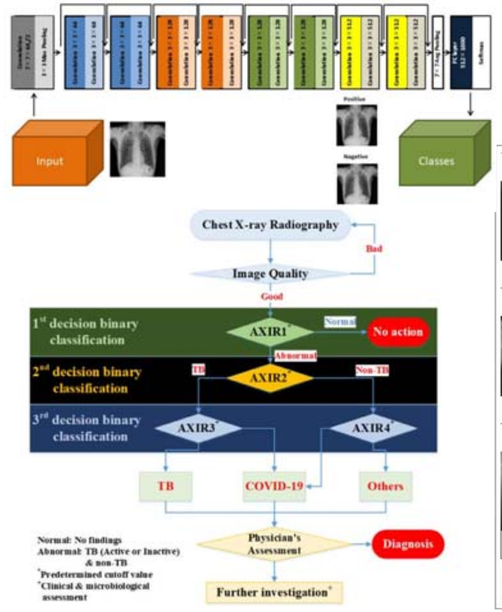
Omer Sagi & Lior Rokach (2020). Explainable decision forest: Transforming a decision forest into an interpretable tree. Information Fusion, 61, 124-138, doi:10.1016/j.inffus.2020.03.013.




```
Algorithm 1: Generate a decision tree using an existing decision forest
Input: T (A set of n decision trees {t_0, t_1, ..., t_n}), D_pruning (pruning
       dataset), forest_min_size (minimum forest size), L (Maximum
       size of conjunctions at each iteration)
Output: DT (a decision tree)
T' = {} - Pruned decision forest, AUC_T'(D_pruning) = 0
while AUC_T'(D_pruning) was improved or |T'| < forest_min_size do
    Add to T' the decision tree t from T that maximizes
     AUC_T'(D_pruning)
end
for each decision tree t_i in T' do
    if i = 0 then
        CS_0 =conjunction set of t_0
        (CS_i = {(c_i1, ŷ_ci1), (c_i2, ŷ_ci2)...(c_iJ, ŷ_ciJ)} where c_ij is a
        conjunction of rules and ŷ_cj is a vector of classes probabilities
    end
    else
        CS_i = {};
        for leaf c_j in t_i do
            for conjunction c_k in CS_i-1 do
                if c_j does not contradict c_k then
                    add (c_j ∧ c_k, ŷ_cj + ŷ_ck) to CS_i
                end
            end
        end
        order CS_i by P(c_ij)
        CS_i = top L conjunctions in CS_i
    end
end
Define DT as a single node that contains CS_|T'|
for each node at DT that wasn't splitted or defined as a leaf do
    for each splitting candidate rule r_ij in {r_11, r_12...r_iL} do
        Find Information Gain IG from split(CS_node, r_ij)
    end
    define r' as the attribute with the highest information gain
    if IG(r') = 0 then
        define current node as a leaf
    end
    else
        split current node by r'
    end
end
```


Decision trees state of the art in 2020 for Covid-19

Seung Hoon Yoo, Hui Geng, Tin Lok Chiu, Siu Ki Yu, Dae Chul Cho, Jin Heo, Min Sung Choi, Il Hyun Choi, Cong Cung Van & Nguen Viet Nhung (2020). Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging. Frontiers in medicine, 7, (427), 1-8, doi:10.3389/fmed.2020.00427.



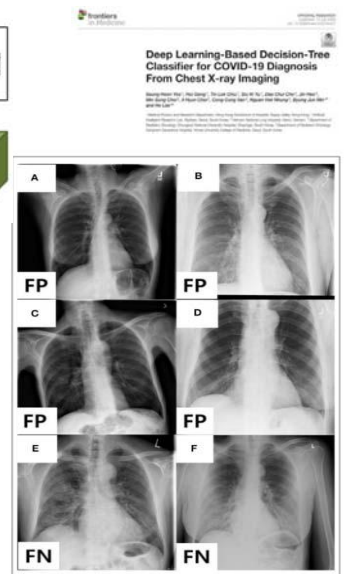


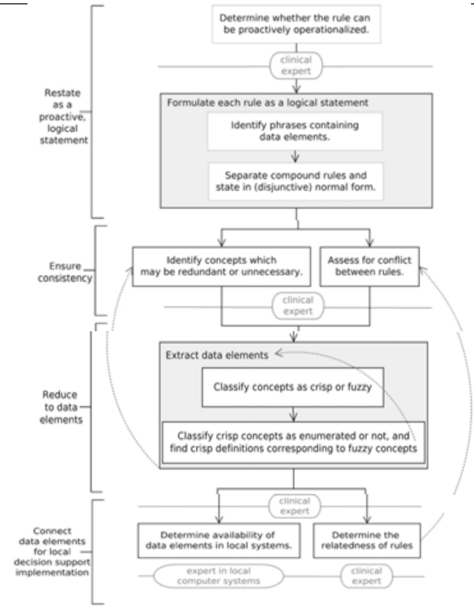
FIGURE 3 | Images showing four FPs predicted as COVID-19 and two FNs predicted as non-TB lung disease in AXIR4 results: infiltration (A–C), pleural thickening (D), and COVID-19 (E, F) symptoms.

How to ensure quality measures ?

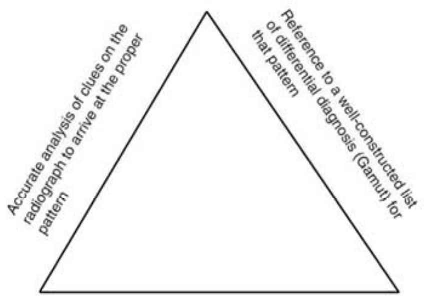
- Clinical guidelines are **systematically** developed documents to assist doctors and patient decisions about appropriate care;
- In order to build DS, based on a guideline, it is **formalized** (transformed from natural language to a logical algorithm), and
- **implemented** (using the algorithm to 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.*



Classical Example: Triangulation to find diagnoses



Correlation of radiographic findings and Gamut with patients' clinical and lab findings to arrive at the most likely diagnosis

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

Gamut F-137

PHRENIC NERVE PARALYSIS OR DYSFUNCTION

COMMON

1. Iatrogenic (eg, surgical injury; chest tube; therapeutic avulsion or injection; subclavian vein puncture)
2. Infection (eg, tuberculosis; fungus disease; abscess)
3. Neoplastic invasion or compression (esp. carcinoma of lung)

UNCOMMON

1. Aneurysm$_g$, aortic or other
2. Birth trauma (Erb's palsy)
3. Herpes zoster
4. Neuritis, peripheral (eg, diabetic neuropathy)
5. Neurologic disease$_g$ (eg, hemiplegia; encephalitis; polio; Guillain-Barré S.)
6. Pneumonia
7. Trauma

Reference

1. Prasad S, Athreya BH: Transient paralysis of the phrenic nerve associated with head injury. JAMA 1976;236:2532–2533

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
3. Drugs (eg, aspirin ; NSAID ; steroids)
4. *Helicobacter pylori* infection
5. Idiopathic
6. [Normal areae gastricae]
7. Peptic ulcer; hyperacidity

UNCOMMON

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

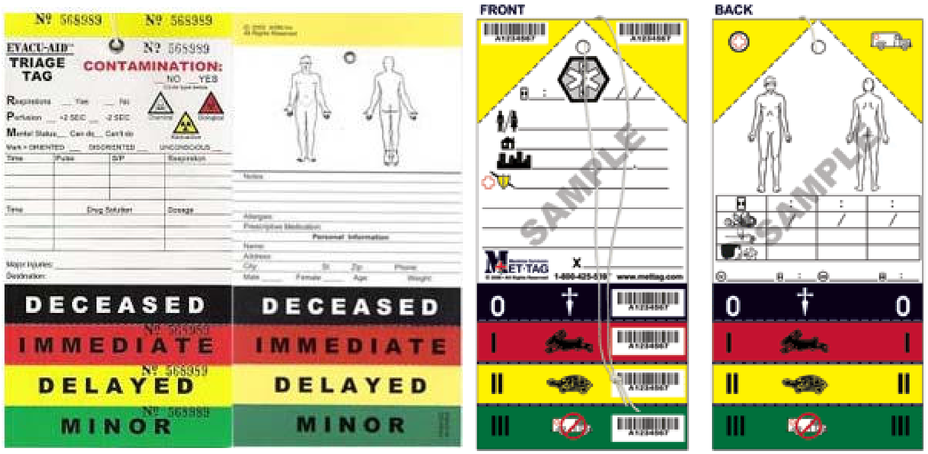
* Superficial erosions or aphthoid ulcerations seen especially with double contrast technique.

[ ] 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://rfs.acr.org/gamuts/data/G-25.htm

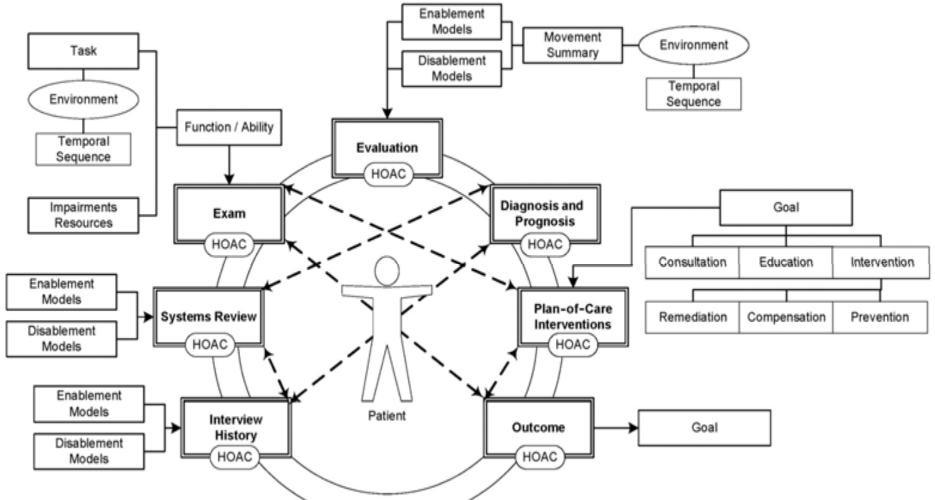
What is the simplest decision support process ?



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

Image Source: http://store.gomed-tech.com

Example Clinical DSS: Hypothesis-Oriented Algorithm



HOAC =Hypothesis-Oriented Algorithm for Clinicians

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.*

Example Prediction Models > Feature Generation

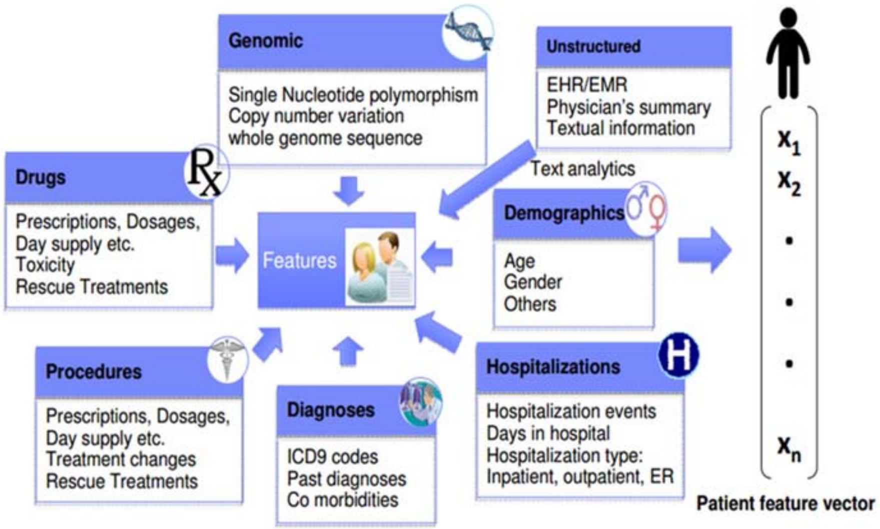


Image credit to Michal Rosen-Zvi

02 Can AI help doctors to make better decisions?

What are the two types of decisions ?

- **Type 1 Decisions:** <u>related to the **diagnosis**</u>, 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:** <u>related to **therapy**</u>, 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?

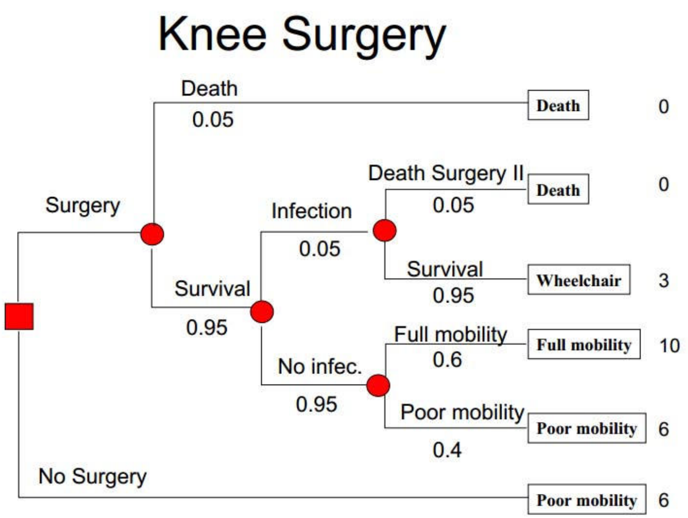
Bemmel, J. H. V. & Musen, M. A. 1997. *Handbook of Medical Informatics, Heidelberg, Springer.*

Example: Knee Surgery of a Soccer Player

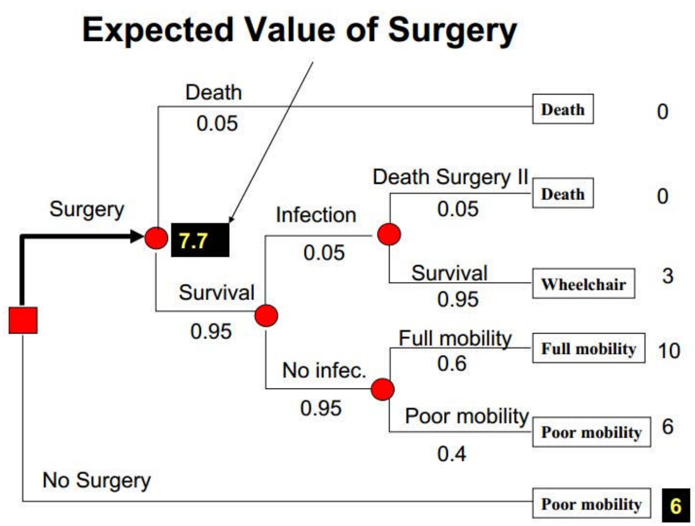
- Example of a Decision Problem
- Soccer player considering knee surgery
- Uncertainties:
- Success: recovering full mobility
- Risks: infection in surgery (if so, needs another surgery and may loose more mobility)
- Survival chances of surgery

Harvard-MIT Division of Health Sciences and Technology
HST.951J: Medical Decision Support, Fall 2005
Instructors: Professor Lucila Ohno-Machado and Professor Staal Vinterbo

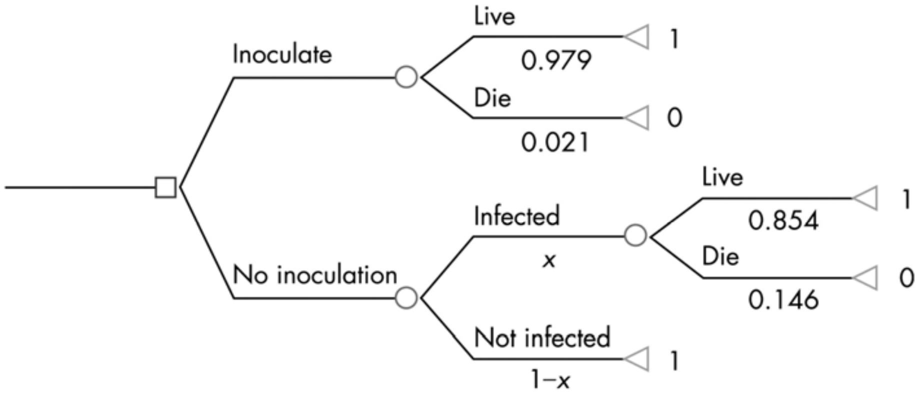
Decision Tree (this is known since Hippocrates!)



Helps to make rational decisions (risks vs. success)



Example: Clinical Decision Tree (CDT)



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.*

Remember: 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

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

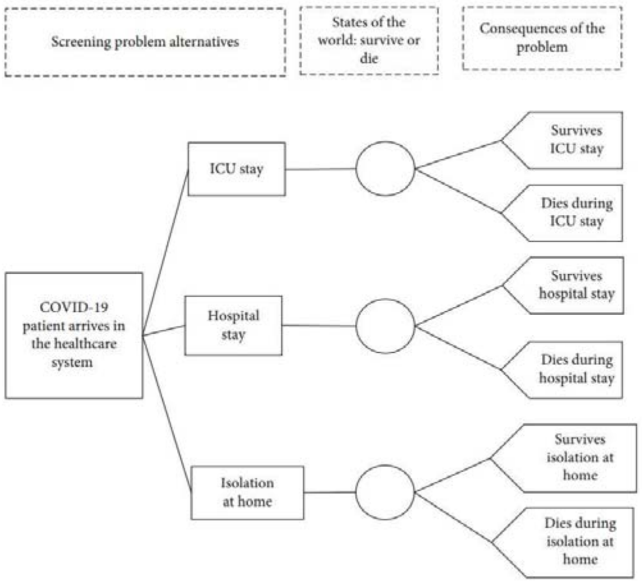
An optimal single decision is the decision $D\ =\ d_{max}$ whose expected utility is maximal:

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

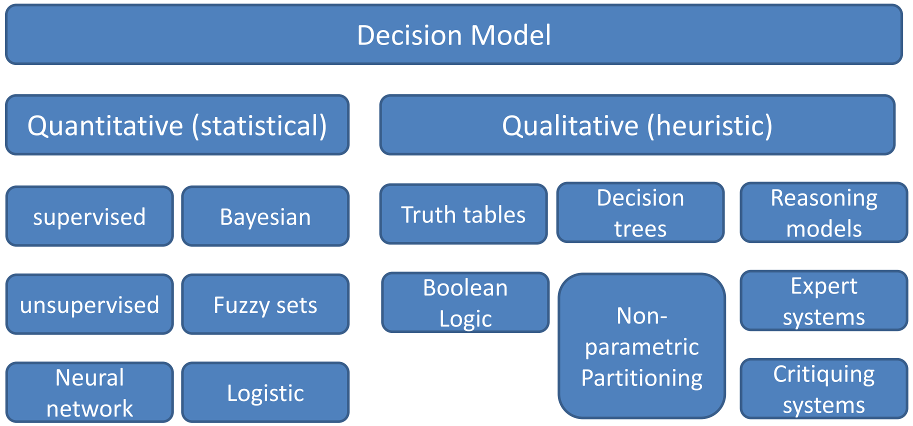
John Von Neumann & Oskar Morgenstern (1944). *Theory of games and economic behavior,* Princeton, Princeton University Press.

Expectation Maximization Algorithm

Lucia Reis Peixoto Roselli, Eduarda Asfora Frej, Rodrigo José Pires Ferreira, Alexandre Ramalho Alberti & Adiel Teixeira De Almeida (2020). Utility-Based Multicriteria Model for Screening Patients under the COVID-19 Pandemic. Computational and Mathematical Methods in Medicine, 2020, 9391251, doi:10.1155/2020/9391251.



What are the “classic” Decision Support Models ?



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

03 Human Information Processing

- 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,
- 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.

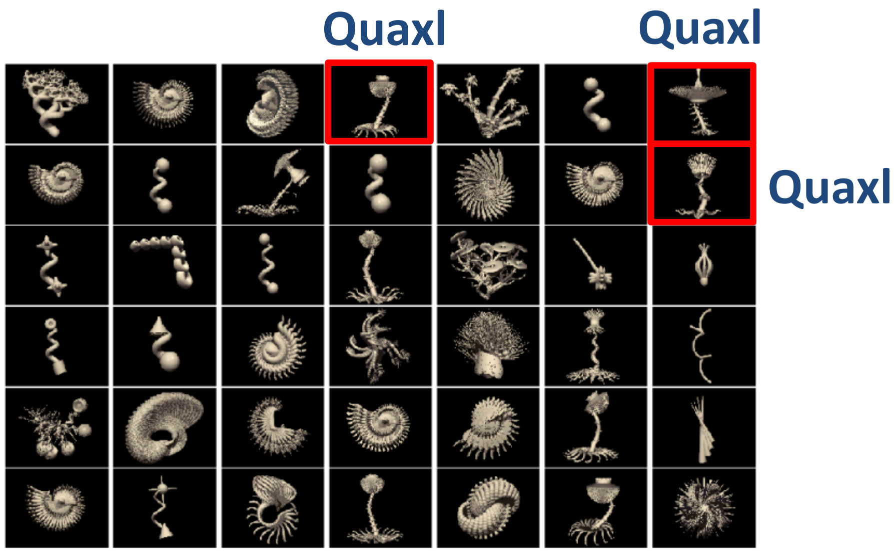
(Sometimes – **not** always!) humans are able ...

- to understand the context
- to make inferences from little, noisy, incomplete data sets
- to learn relevant representations
- to find shared underlying explanatory factors,
- in particular between $P(x)$ and $P(Y|X)$, with a causal link between $Y \rightarrow X$

Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths & Noah D. Goodman 2011. How to grow a mind: Statistics, structure, and abstraction. *Science,* 331, (6022), 1279-1285, doi:10.1126/science.1192788.

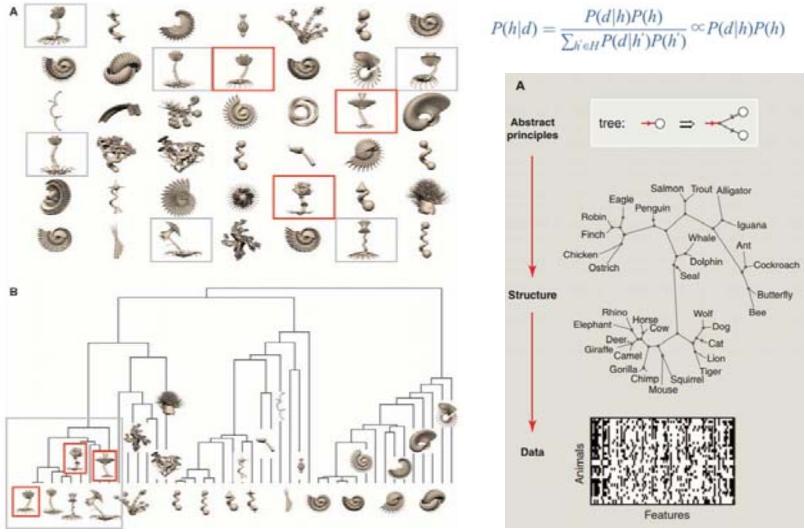
Salakhutdinov, R., Tenenbaum, J. & Torralba, A. 2012. One-shot learning with a hierarchical nonparametric Bayesian model. Journal of Machine Learning Research, 27, 195-207.

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 ?



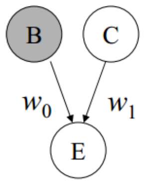
Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285.

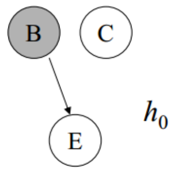
One of the unsolved problems in human concept learning

- 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

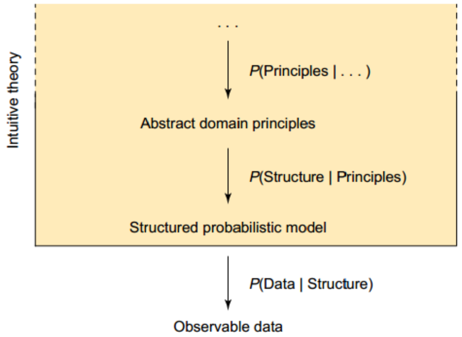




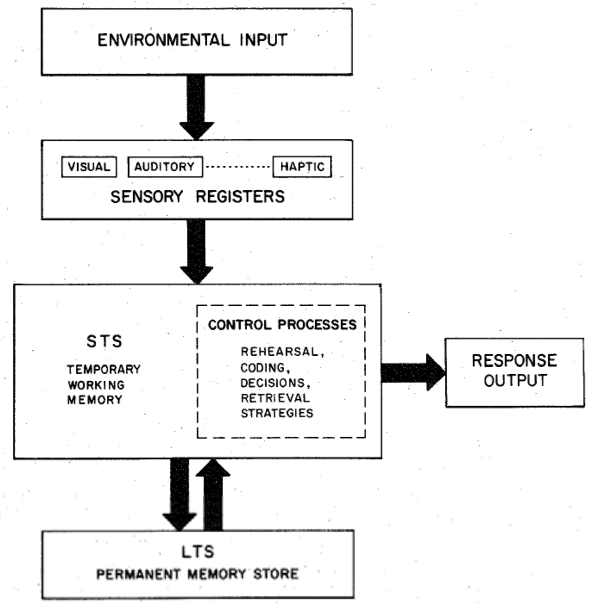
- Cognition as probabilistic inference
  - Visual perception, language acquisition, motor learning, associative learning, memory, attention, categorization, reasoning, causal inference, decision making, theory of mind
- Learning concepts from examples
- Learning and applying intuitive theories (balancing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

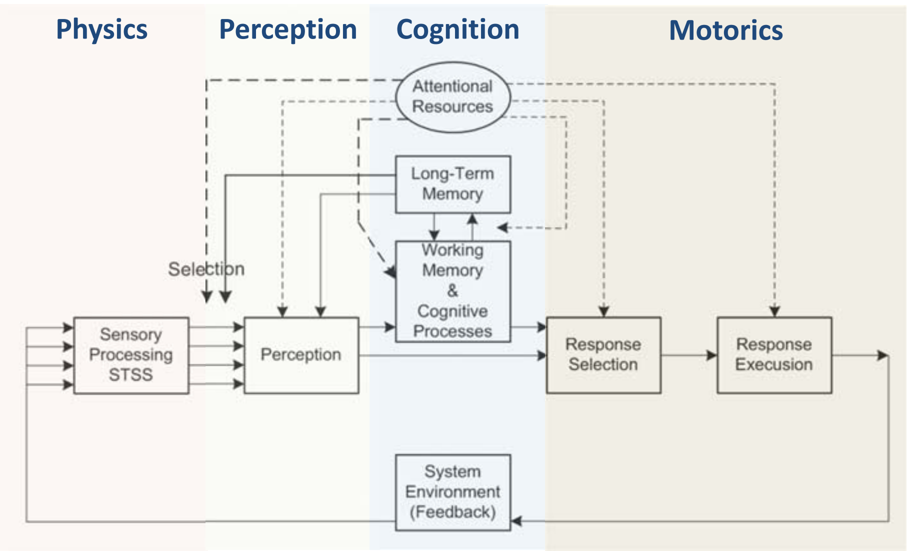
- Similarity
- Representativeness and evidential support
- Causal judgement
- Coincidences and causal discovery
- Diagnostic inference
- Predicting the future

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.



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.*





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

04

d … *data*

h … *hypotheses*

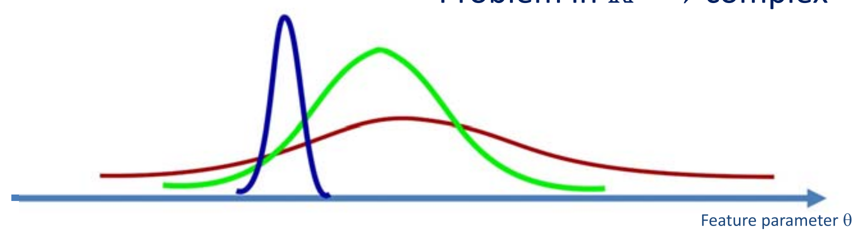
$\mathcal{H}$ … $\{H_1, H_2, \ldots, H_n\}$ $\forall\, h, d$ …

Likelihood Prior Probability

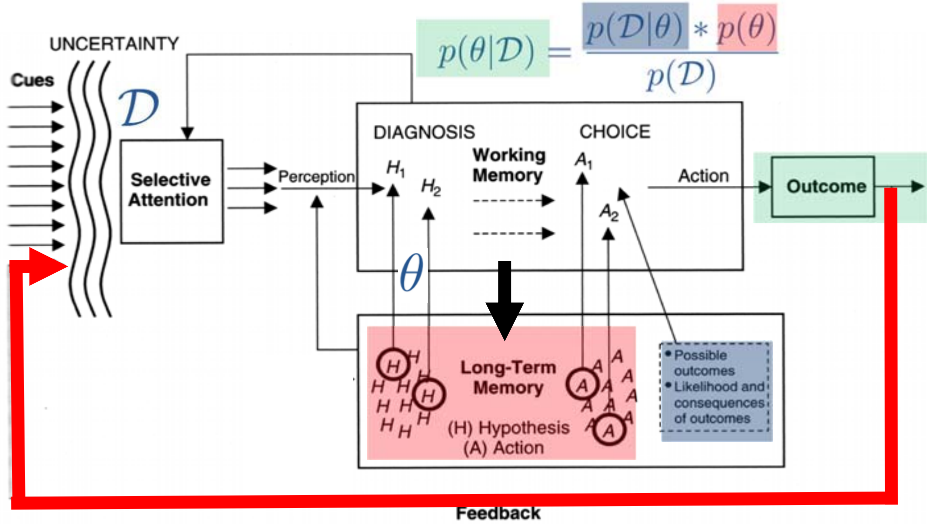
$$p(h|d) = \frac{p(d|h) * p(h)}{\sum_{h \in H} p(d|h')\, p(h')}$$

Posterior Probability

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



Connection to Cognitive Science: Decision Making



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

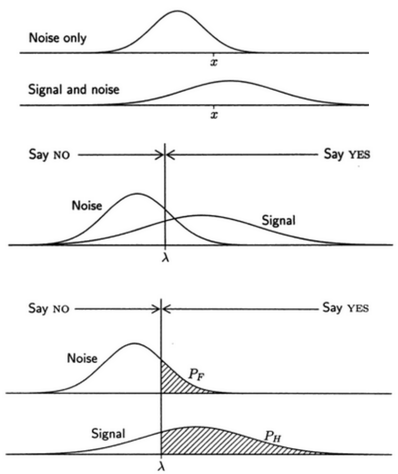
06 How to make decisions in an domain of uncertainty

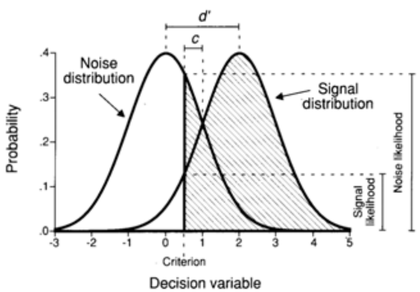
04 Probabilistic Decision Making

"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

What is Signal Detection Theory (SDT) doing ?

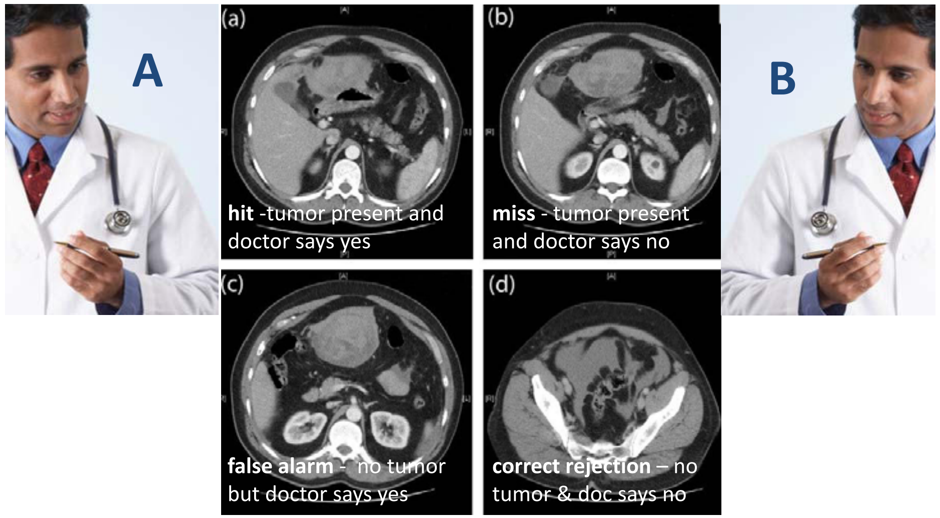
Image source: Staffordshire University Computing Futures Museum http://www.fcet.staffs.ac.uk/jdw1/sucfm/malvern.htm





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

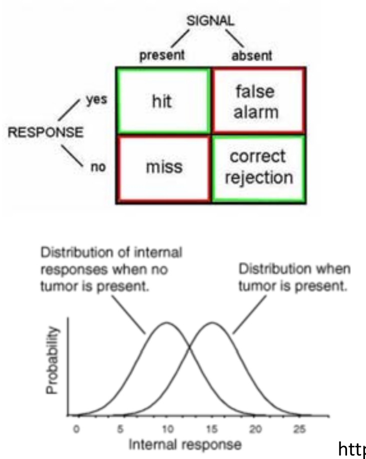
How is Signal Detection Theory applied in medicine ?

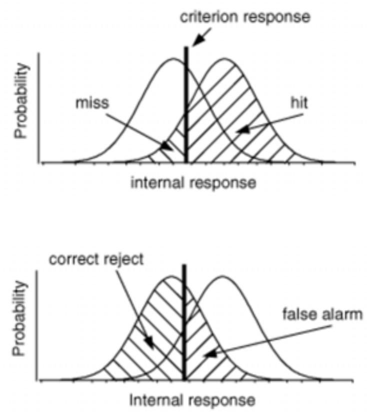


Two doctors, with equally good training, looking at the same CT scan, will have the same information … but they may have a **different bias/criteria!**

How do you measure the quality in SDT ?

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*.

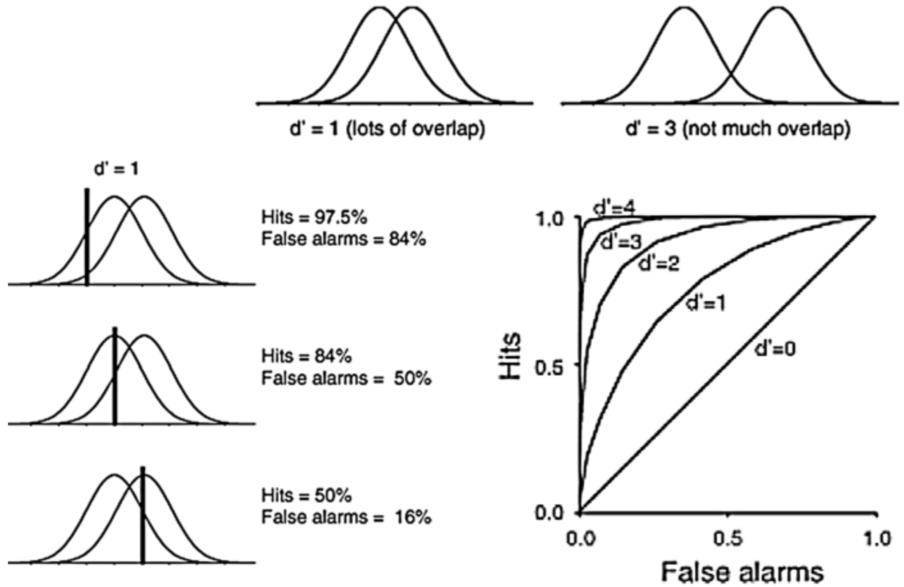




http://www-psych.stanford.edu/~lera/psych115s/notes/signal

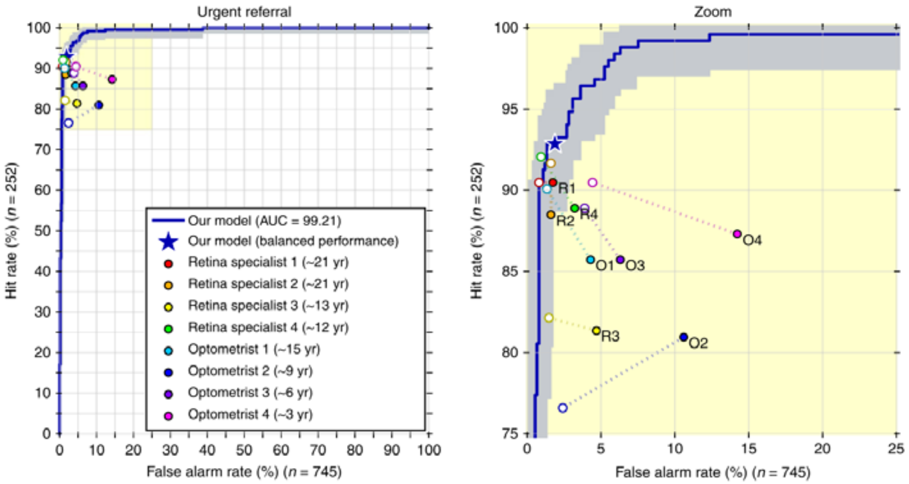
For an example see: Braga & Oliveira (2003) Diagnostic analysis based on ROC curves: theory and applications in medicine. *Int. Journal of Health Care Quality Assurance, 16, 4, 191-198.*

What does the Receiver Operating Characteristics tell ?



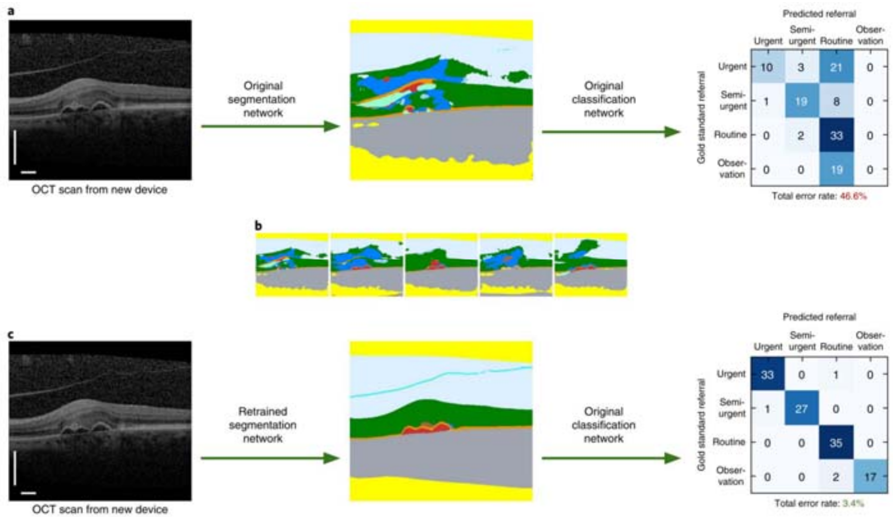
http://gru.stanford.edu/doku.php/tutorials/sdt

How do you measure AI performance ?



Jeffrey De Fauw et al. 2018. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine,* 24, (9), 1342-1350

What is the confusion matrix showing you ?



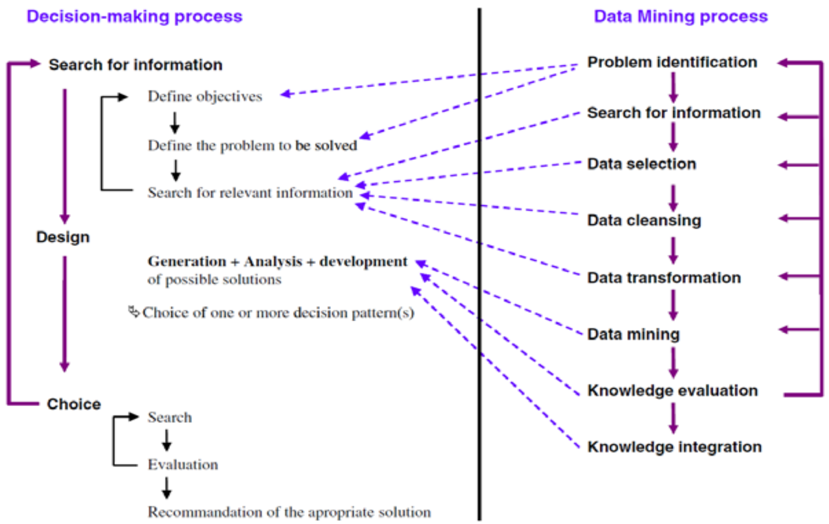
Jeffrey De Fauw et al. 2018. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine,* 24, (9), 1342-1350

What are 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

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)

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

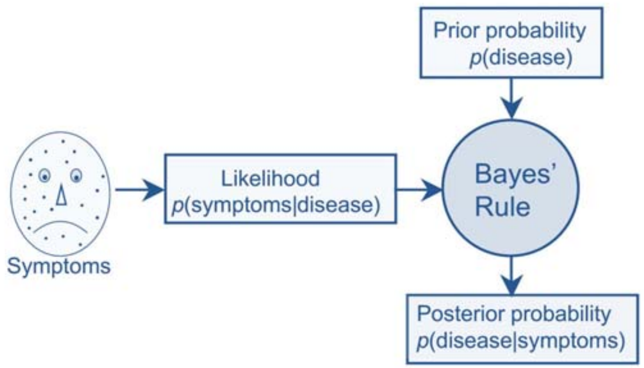
$P(d)=\sum_h P(d\mid h)P(h)$: data evidence (marginal probability of the data)

$P(h\mid d)$: 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!**

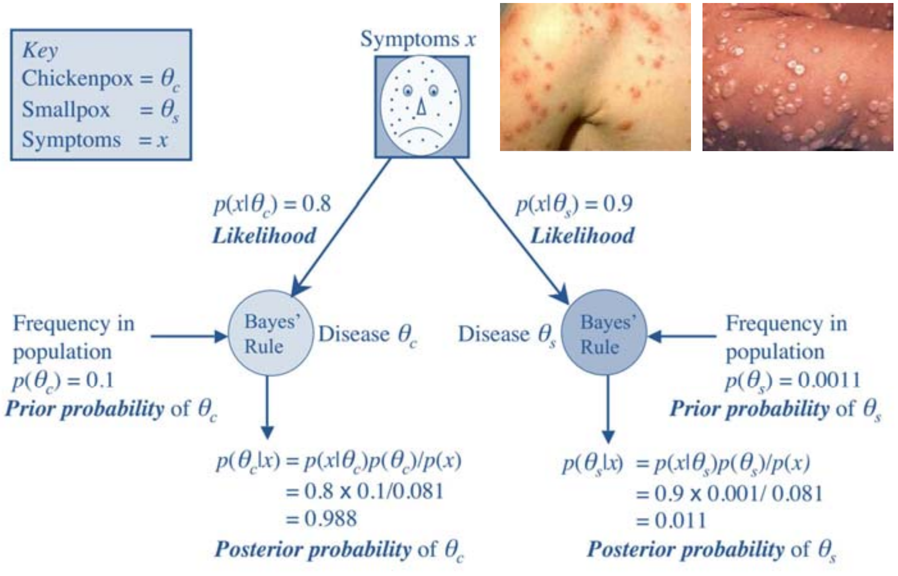
Bayes Rule for Medical Diagnosis



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

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 if you are infected (T)
- Good news: It is a rare disease, striking 1 in 10,000 (D)
- **How worried would you now be?**

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

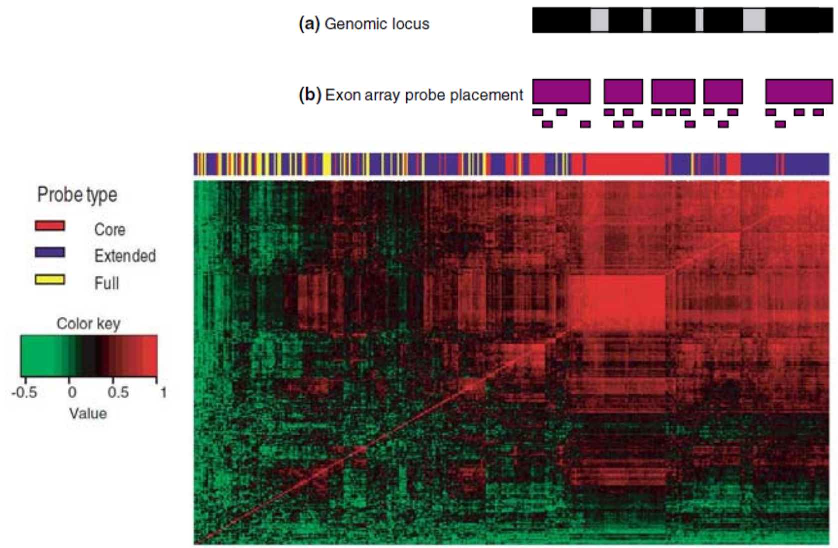
$$p(T = 1|D = 1) = p(d|h) = 0{,}99\ \ and$$
$$p(D = 1) = p(h) = 0{,}0001$$

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

= 0,0098

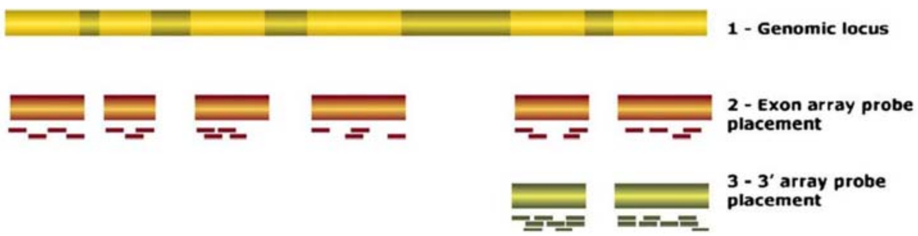
05 Example: P4-Medicine

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.*

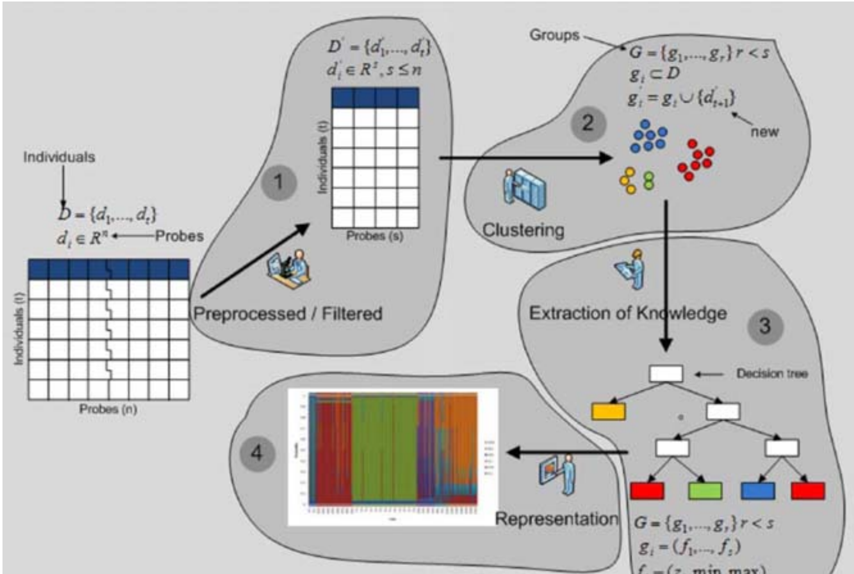
Computational leukemia cancer detection 1/6



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.*

Computational leukemia cancer detection 2/6

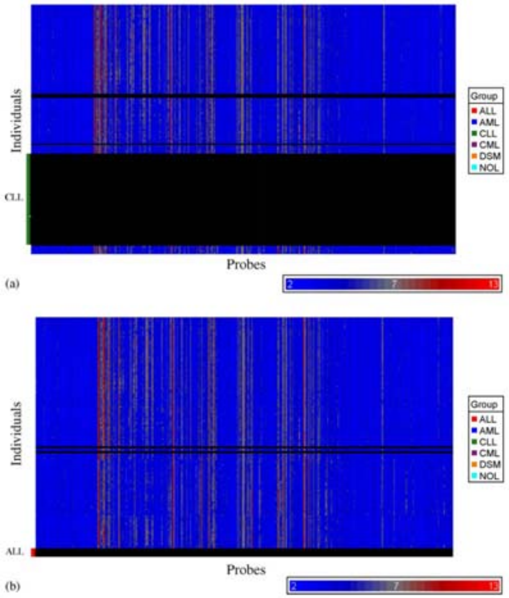


Corchado et al. (2009)

Computational leukemia cancer detection 3/6

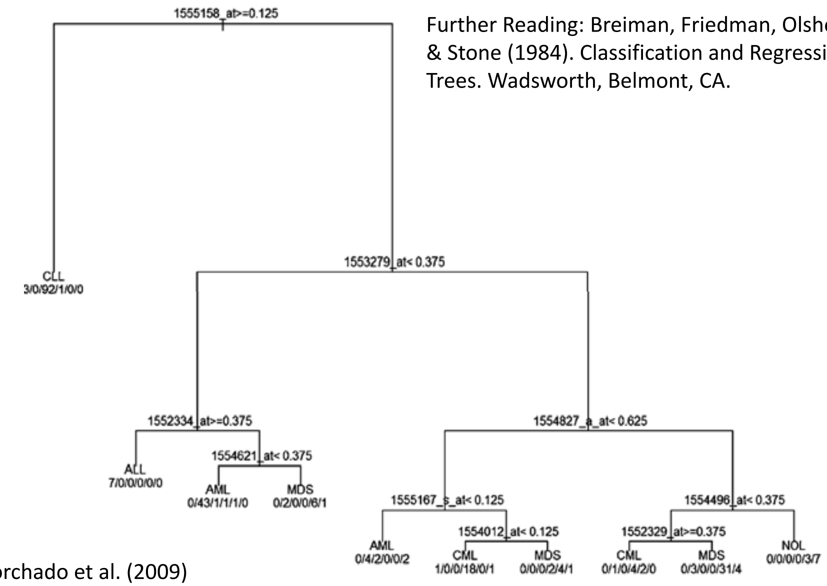
A = acute, C = chronic,
L = lymphocytic, M = myeloid

- **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)

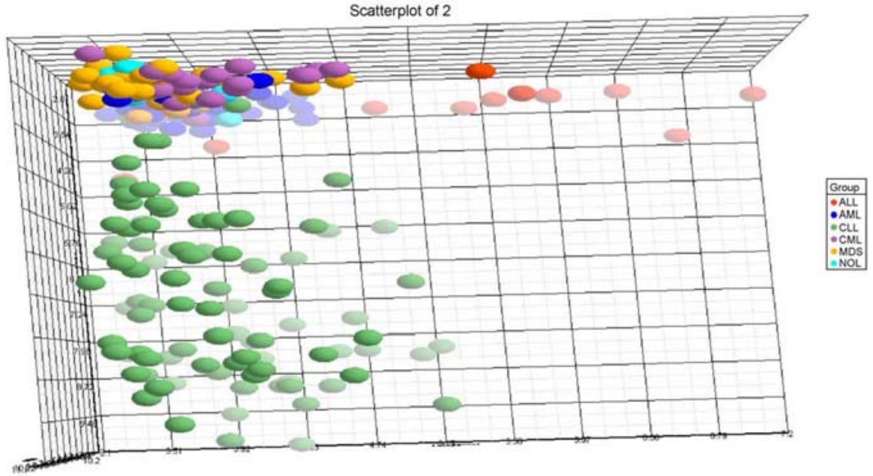
Computational leukemia cancer detection 4/6



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

Corchado et al. (2009)

Classification CLL—ALL. Representation of the probes of the decision tree which classify the CLL and ALL to 1555158_at, 1553279_at and 1552334_at



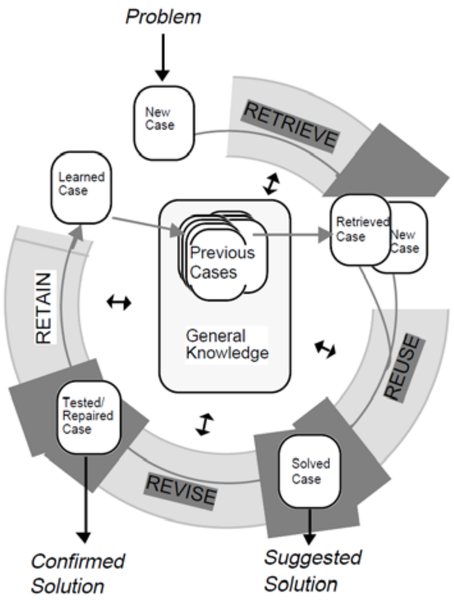
Corchado et al. (2009)

- 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

06 Example: Case Based Reasoning (CBR)

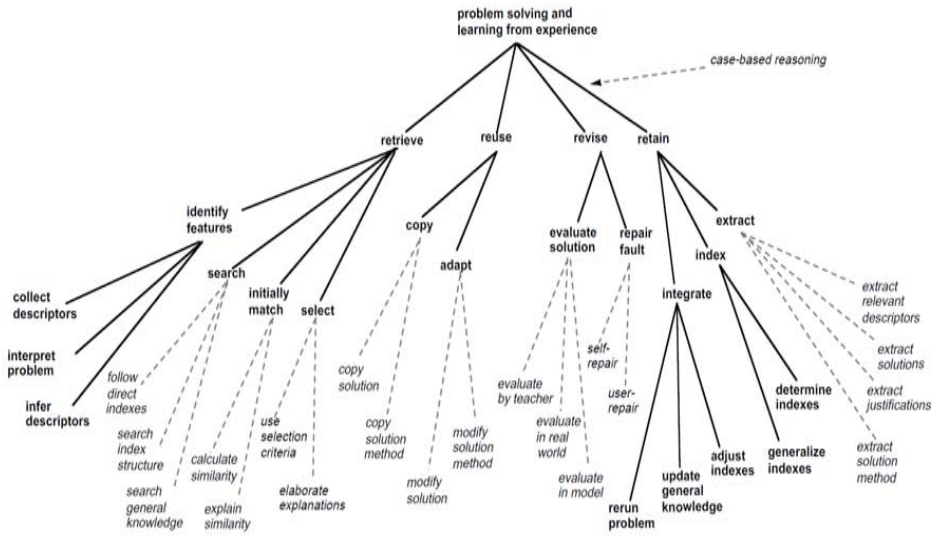
Thinking – Reasoning – Deciding – Acting

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.*

The task-method decomposition of CBR

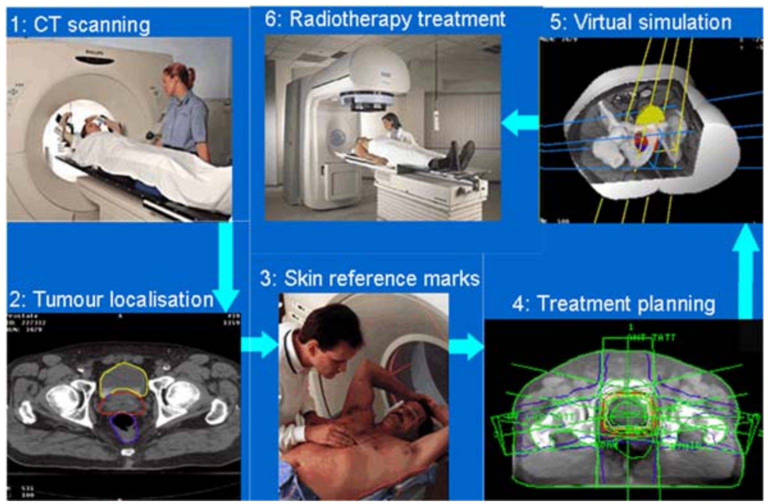


Aamodt & Plaza (1994)

CBR Example: Radiotherapy Planning 1/6

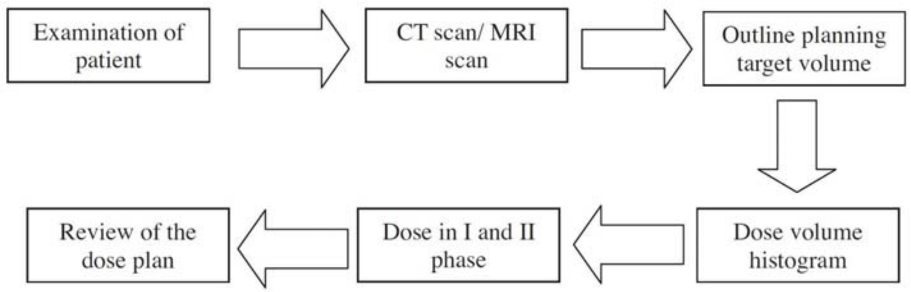
Source: http://www.teachingmedicalphysics.org.uk

CBR Example: Radiotherapy Planning 2/6



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

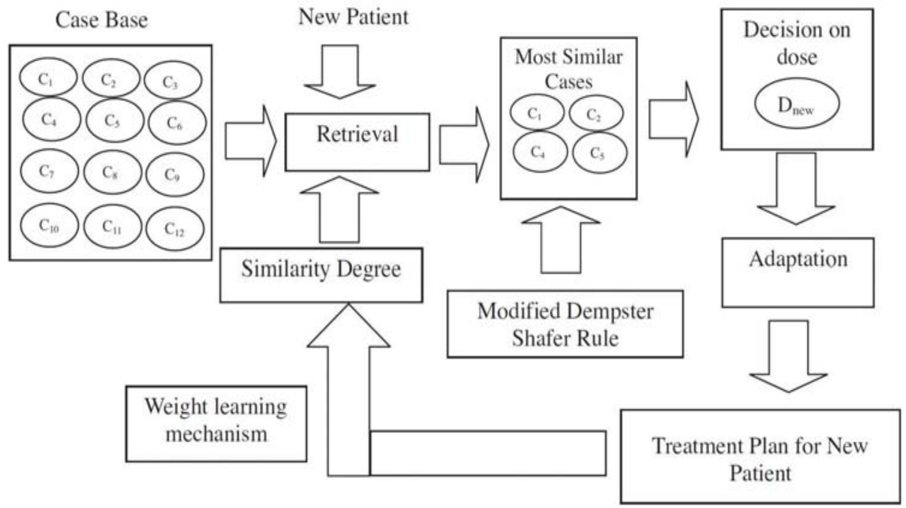
CBR Example: Radiotherapy Planning 3/6



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 %)

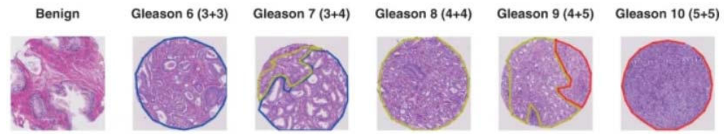
Petrovic, S., Mishra, N. & Sundar, S. (2011) A novel case based reasoning approach to radiotherapy planning. *Expert Systems With Applications, 38, 9, 10759-10769.*

CBR System Architecture 4/6



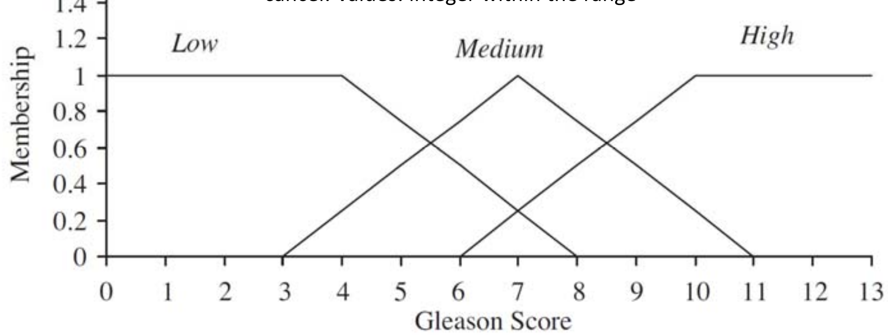
Petrovic, S., Mishra, N. & Sundar, S. (2011) A novel case based reasoning approach to radiotherapy planning. *Expert Systems With Applications, 38, 9, 10759-10769.*

Membership funct. of fuzzy sets Gleason score 5/6



Eirini Arvaniti et al (2018). Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Scientific Reports, 8, (1), doi:10.1038/s41598-018-30535-1.

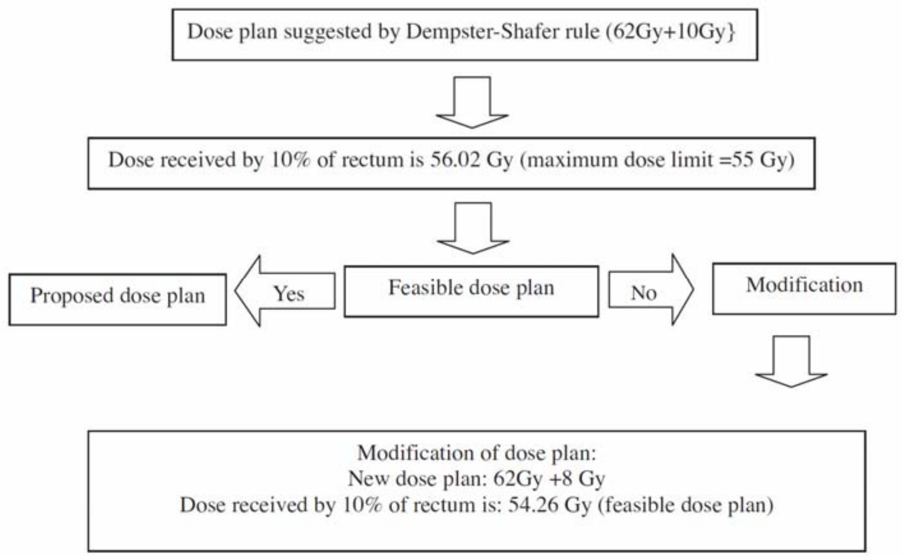
Gleason score evaluates the grade of prostate cancer. Values: integer within the range



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

Example: Case Based Reasoning 6/6

Petrovic et al. (2011)



Conclusion

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?

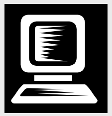
**) Computational intelligence, Artificial Intelligence/soft computing/ML

- **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



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.

Conclusion

- Computers are incredibly fast, accurate and stupid,
- humans are incredibly slow, inaccurate and brilliant,
- together they are powerful beyond imagination

(Einstein never said that)

https://www.benshoemate.com/2008/11/30/einstein-never-said-that

Thank you!