Andreas Holzinger

185.A83 Machine Learning for Health Informatics

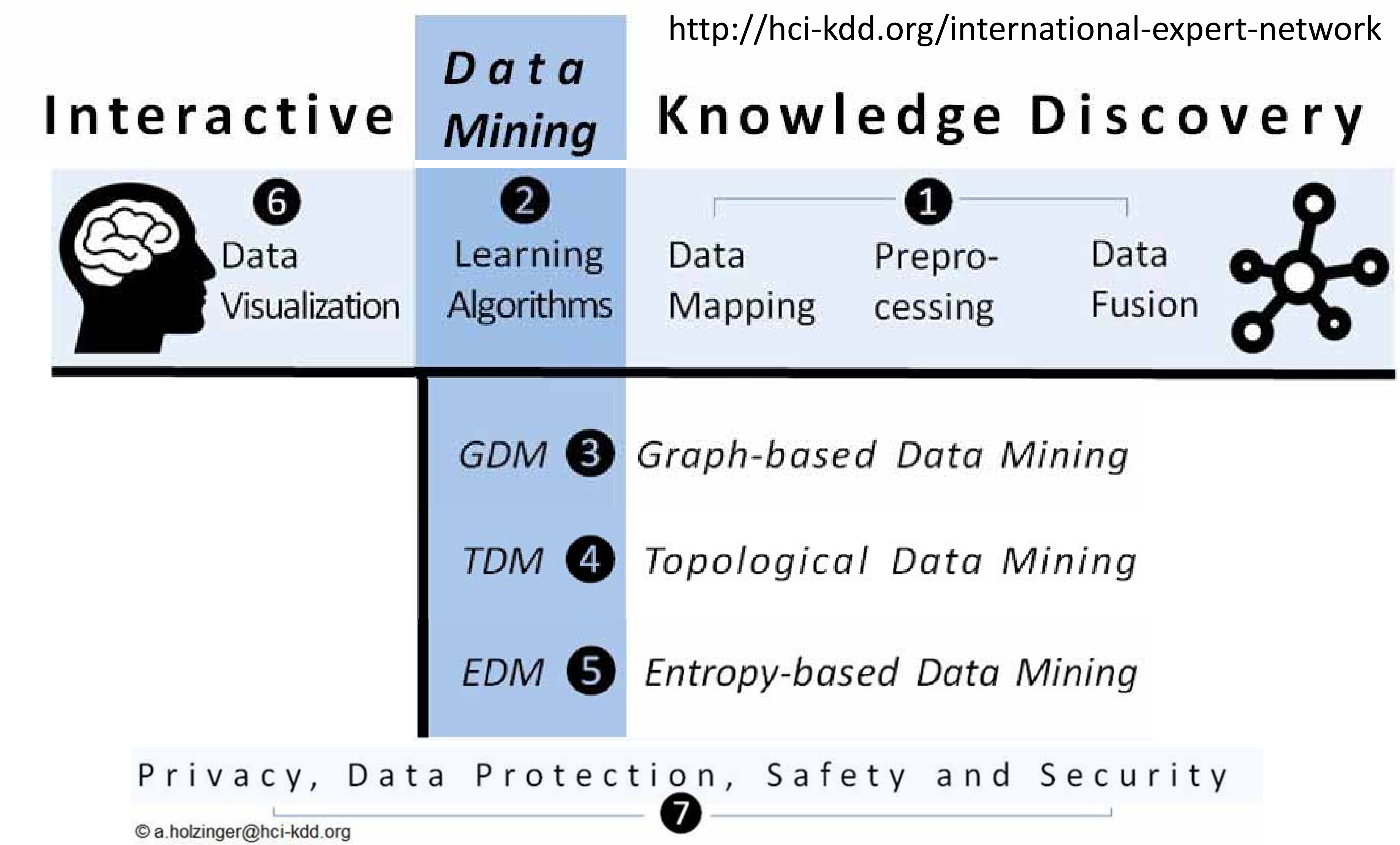
2016S, VU, 2.0 h, 3.0 ECTS

Week 22 - 01.06.2016 17:00-20:00

Evolutionary Computing for solving Health informatics problems - Part 1

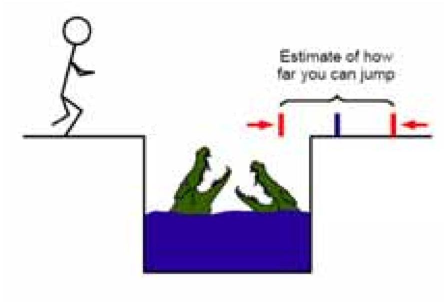
a.holzinger@hci-kdd.org

http://hci-kdd.org/machine-learning-for-health-informatics-course

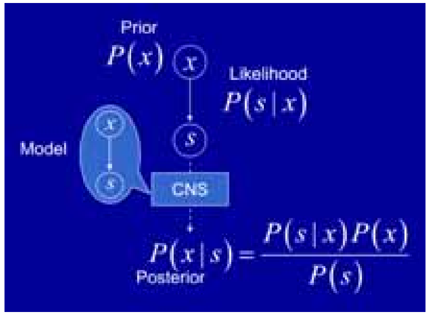


Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: **Cognitive Science meets Machine Learning.** IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

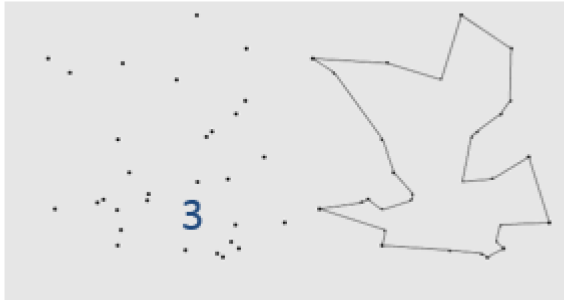
Let us start with a warm-up Quiz (solutions -> last page)

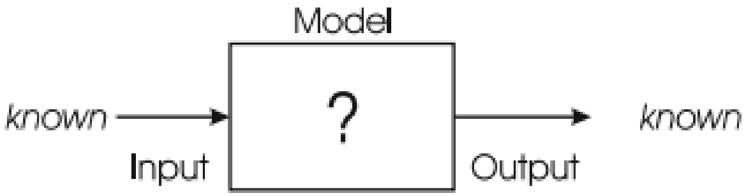


1



2

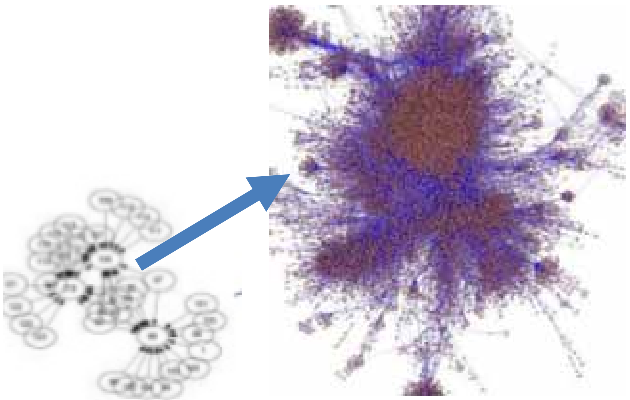




4

5

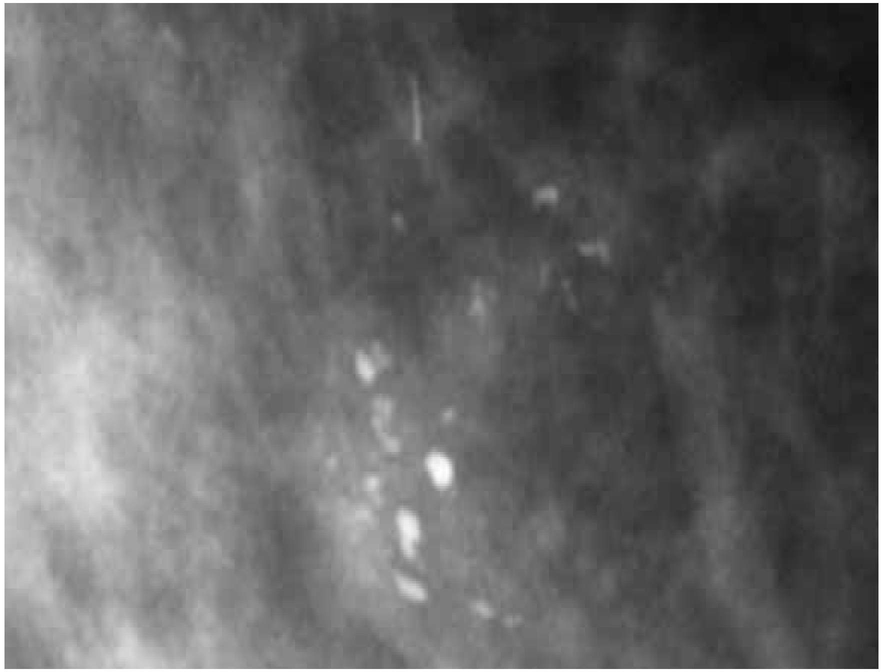
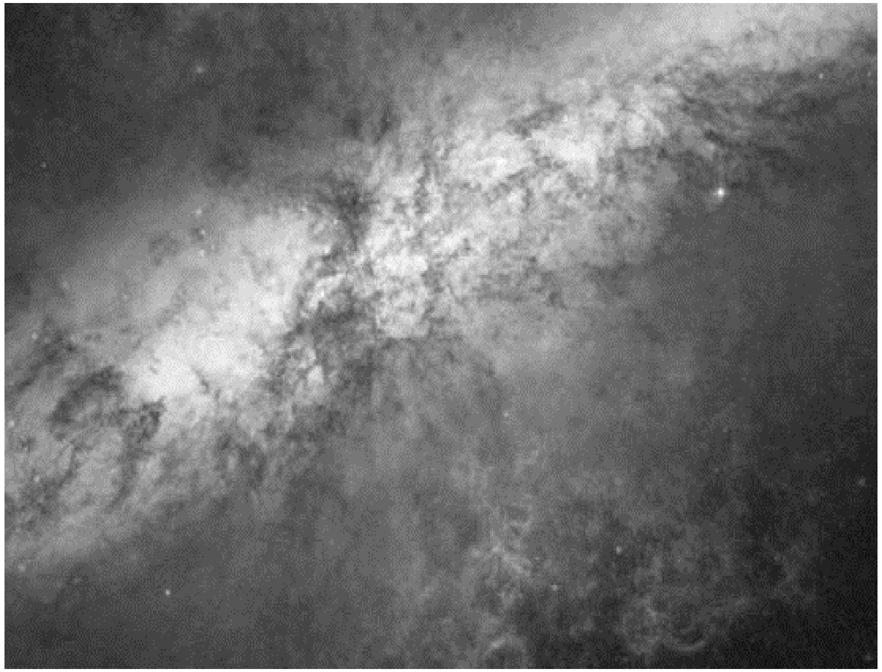
6



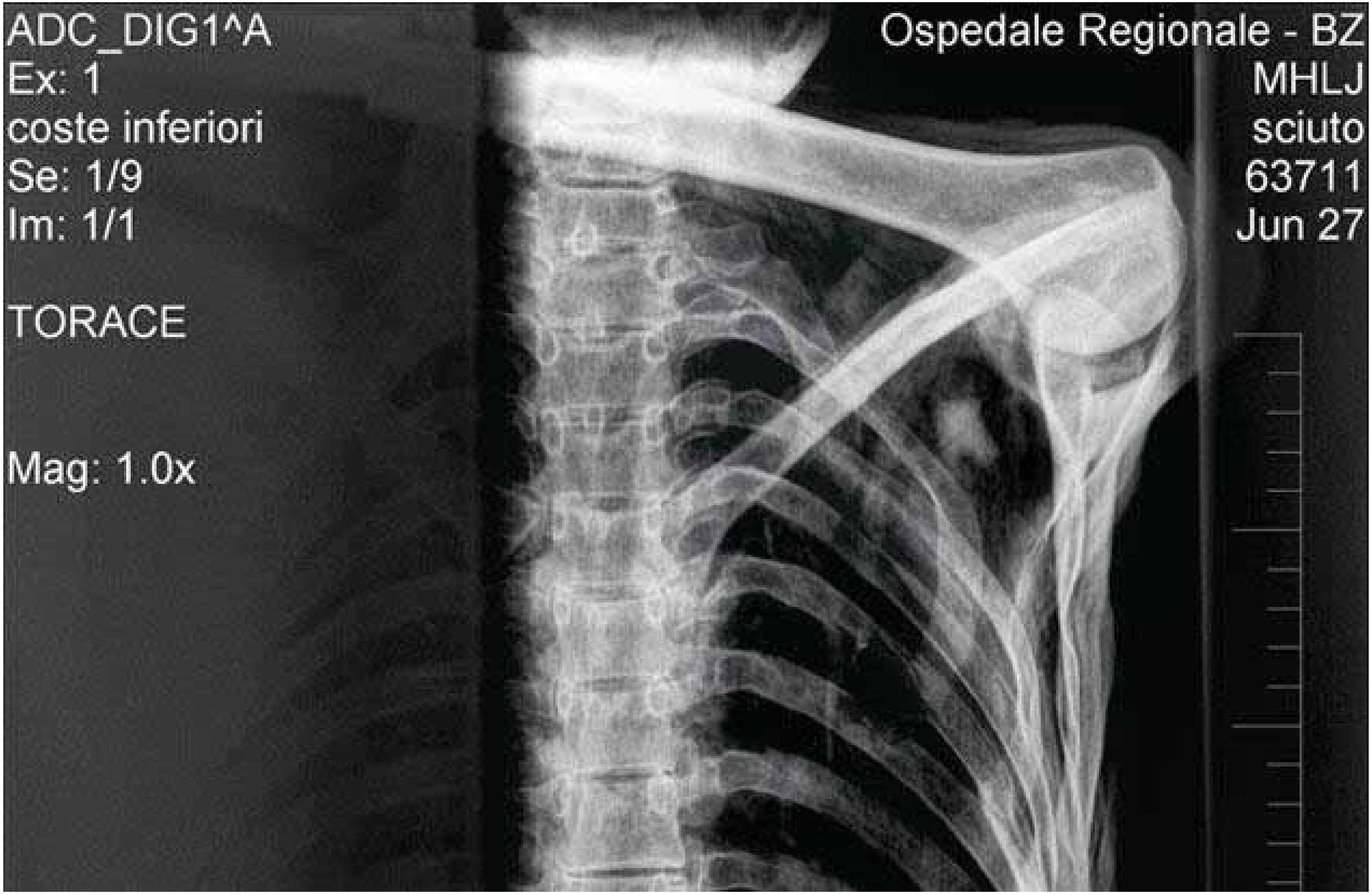
στόχος

7

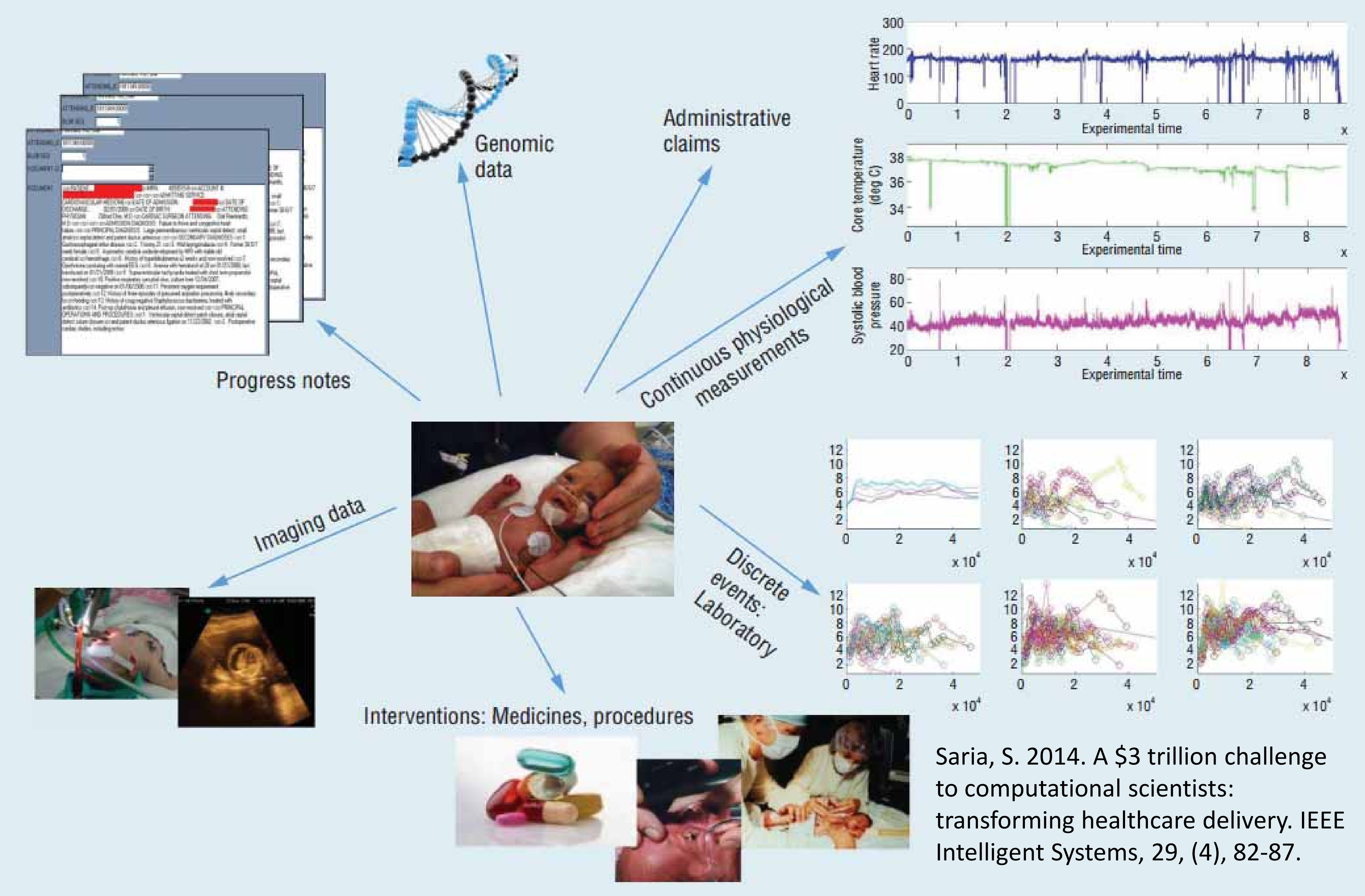
Quiz: Which is from medicine, which from astronomy?

Quiz: Is there an anomaly in this image?



Quiz: What are the key challenges in such a scenario:



Saria, S. 2014. A $3 trillion challenge to computational scientists: transforming healthcare delivery. IEEE Intelligent Systems, 29, (4), 82-87.

- 1) Medical Decision Making as Search Problem
- 2) Evolutionary Principles and Applications
- 3) Evolutionary Computing
- 4) Special Case: Genetic Algorithms

- I) Machine Learning: Evolutionary computation is a key concept in ML [1]
- II) Health Informatics: Evolutionary computation is widely applied in medical problem solving [2]
- Whenever a **decision** is required, it is possible to apply evolutionary techniques, e.g.
  - 1) Learning, Knowledge Discovery, Mining, … applied to both diagnosis and prognosis (=prediction)
  - 2) Medical imaging, signal processing, … and
  - 3) Planning and scheduling

[1] Zhang, J., Zhan, Z.-H., Lin, Y., Chen, N., Gong, Y.-J., Zhong, J.-H., Chung, H. S., Li, Y. & Shi, Y.-H. 2011. Evolutionary computation meets machine learning: A survey. Computational Intelligence Magazine, IEEE, 6, (4), 68-75
[2] Pena-Reyes, C. A. & Sipper, M. 2000. Evolutionary computation in medicine: an overview. Artificial Intelligence in Medicine, 19, (1), 1-23, doi:10.1016/S0933-3657(99)00047-0.

- Study of the design of **intelligent agents**
- Set of *nature-inspired* methodologies to solve complex real-world problems, when traditional methods might be useless, because:
- 1) the processes might be too **complex** for mathematical reasoning within the given time,
- 2) the problem contains a lot of **uncertainties**
- 3) the problem/process is **stochastic** in nature

Kruse, R., Borgelt, C., Klawonn, F., Moewes, C., Steinbrecher, M. & Held, P. 2013. Computational Intelligence: A Methodological Introduction, Heidelberg, New York, Springer.
Online in both German and English: http://www.computational-intelligence.eu/

IFIP WG 12.9 http://www.ifip.org/bulletin/bulltcs/memtc12.htm

Define (2/3): Evolutionary Computing (EC) :=

- Subfield of CI which studies evolutionary algorithms [1] based on **evolutionary principles** (e.g. Darwin, Baldwin, Lamarck, Mendel [2]),
- **Trial-and-error problem solvers**, considered as
- **Global optimization** methods with metaheuristic or stochastic optimization character – mostly applied for black-box problems (with exception of interactive machine learning approaches, where the black box is opened to a glass box [3])

[1] Eiben, A. E. & Smith, J. E. 2015. *Introduction to evolutionary computing. Second Edition,* Berlin, Springer. Online: http://www.evolutionarycomputation.org/

[2] Holzinger, K., Palade, V., Rabadan, R. & Holzinger, A. 2014. Darwin or Lamarck? Future Challenges in Evolutionary Algorithms for Knowledge Discovery and Data Mining. In: Lecture Notes in Computer Science LNCS 8401. Berlin: Springer, pp. 35-56, doi:10.1007/978-3-662-43968-5_3.

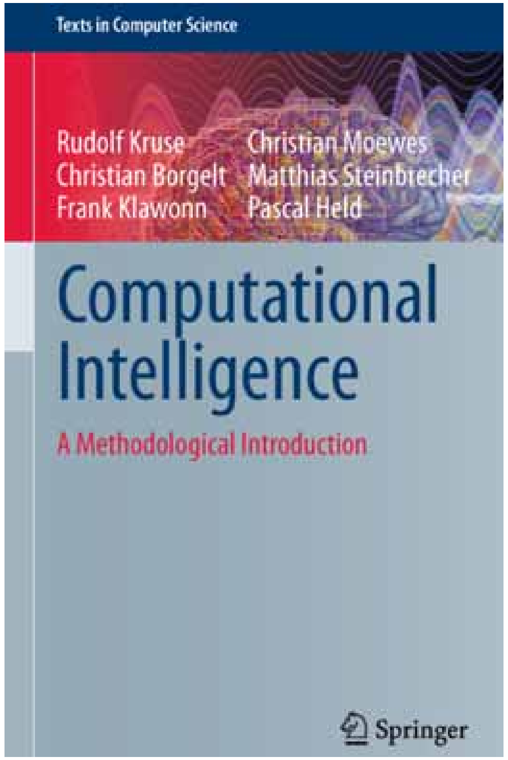
[3] Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? Brain Informatics, 3, (2), 119-131, doi:10.1007/s40708-016-0042-6.

Define (3/3): Genetic Algorithms (GA) :=

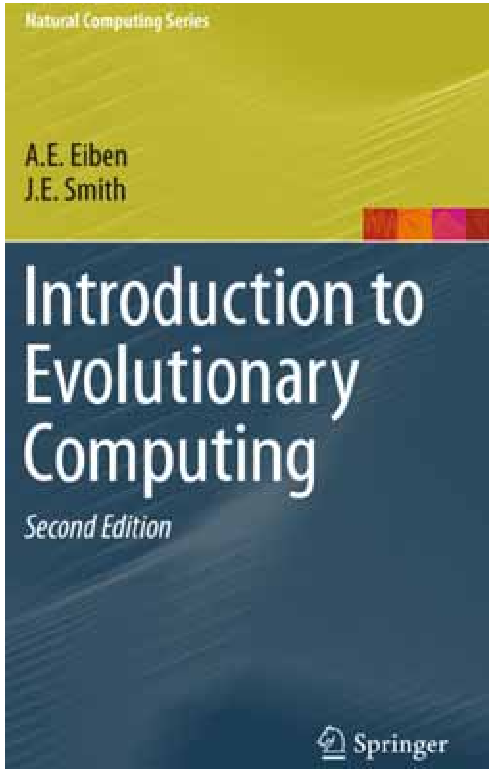
- search heuristic mimicking the process of **natural selection** used to generate useful solutions to optimization and search problems [1];
- particularly making use of techniques inspired by natural evolution (competing for resources), such as inheritance, reproduction, recombination, mutation, selection, inversion and crossover, according to a
- **fitness function (evaluation function)** [2].

[1] Mitchell, Melanie (1996). An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press
[2] Kallel, L., Naudts, B. & Reeves, C. 2001. Properties of fitness functions and search landscapes. In: Kallel, L. (ed.) Theoretical Aspects of Evolutionary Computing. Heidelberg: Springer, pp. 175-206.

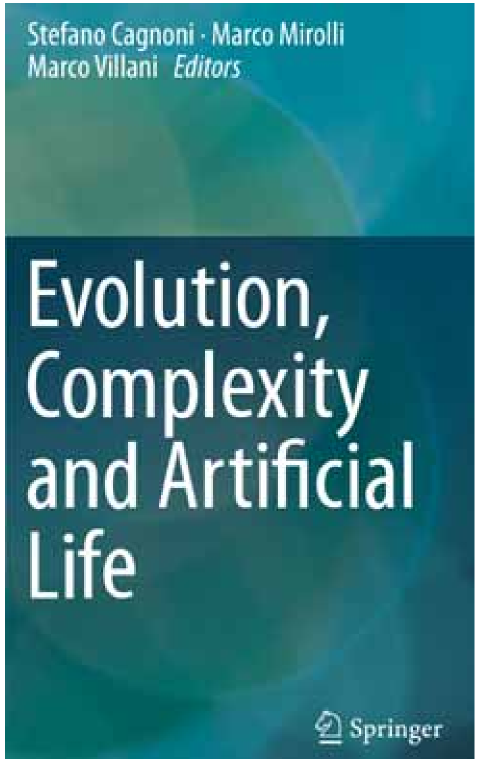
Recommended Books



Kruse, R., Borgelt, C., Klawonn, F., Moewes, C., Steinbrecher, M. & Held, P. 2013. Computational Intelligence: A methodological Introduction, Heidelberg, New York, Springer.



Eiben, A. E. & Smith, J. E. 2010. Introduction to evolutionary computing, Springer Berlin.

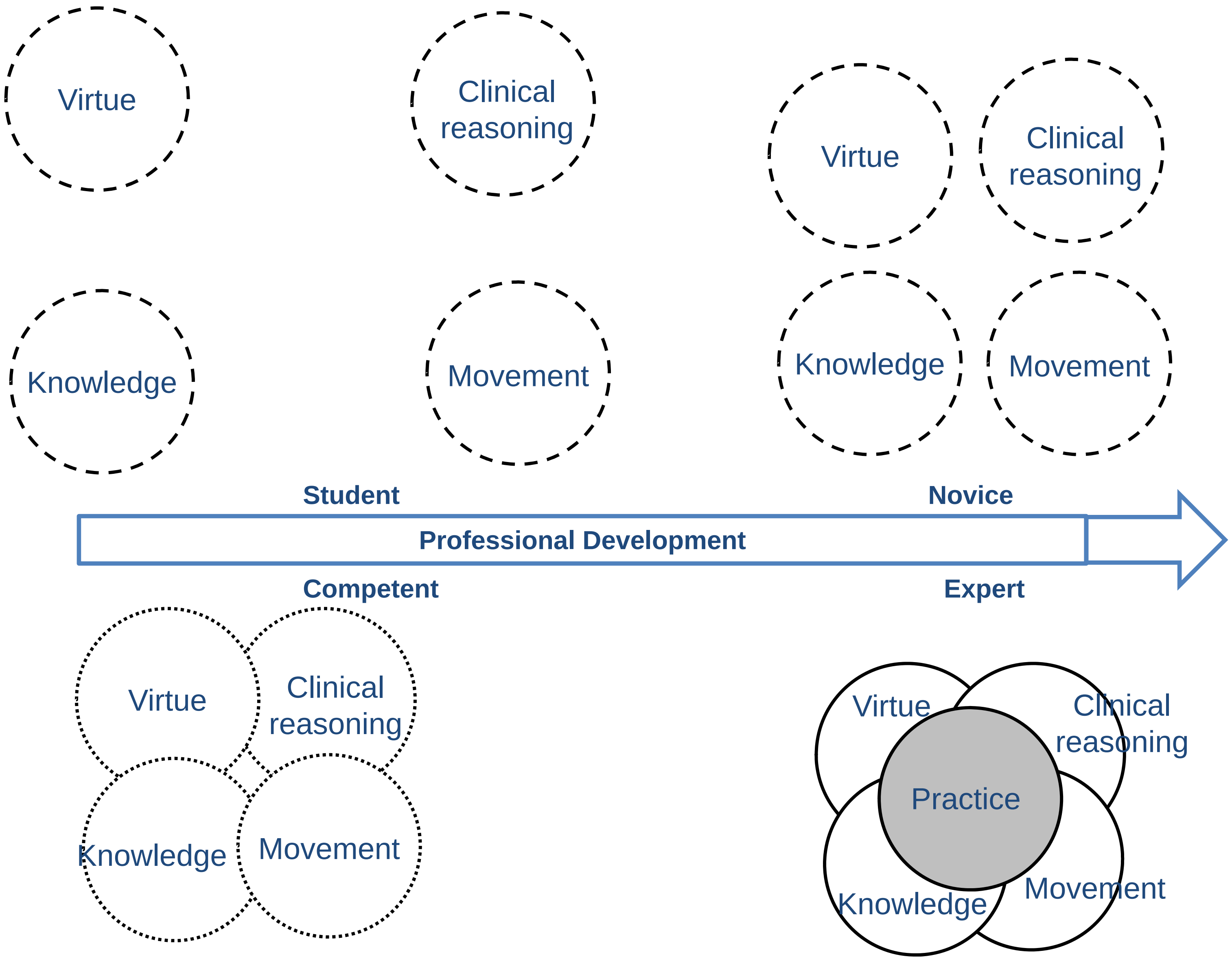


Cagnoni, S., Mirolli, M. & Villani, M. 2014. *Evolution, Complexity and Artificial Life*, Springer.

1) Medical Decision Making as a Search Problem

Search in an arbitrarily high-dimensional space < 5 min. !

Medical Action is constantly reasoning/decision making



Resnik, L. & Jensen, G. M. 2003. Using clinical outcomes
to explore the theory of expert practice in physical therapy. Physical Therapy, 83, (12), 1090-1106.

Remember: two types of Decision Making

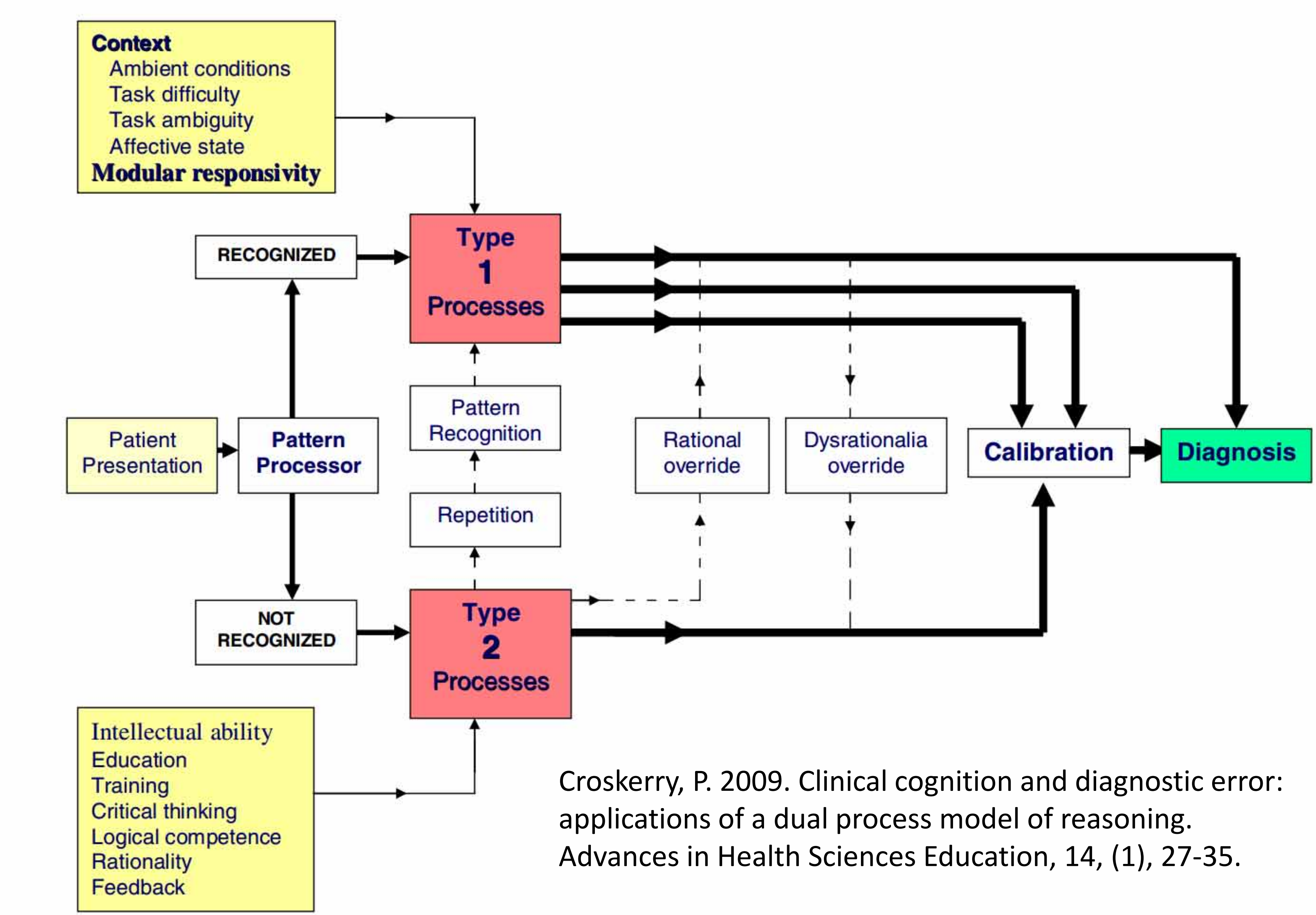
Characteristic	Type 1 Heuristic Intuitive	Type 2 Systematic Analytic
Cognitive Style	Heuristic associative (experience-based) Inductive reasoning	Bounded rationality (Hypothetico-deductive) Normative reasoning
Cost (high/low)	Low	
Automaticity(high/low)		Low
Rate (fast/slow)		Low
Reliability (high/low)	Low	
Errors (high/low)		Low
Effort (high/low)	Low	
Predictive Power (high/low)	Low	
Emotional Component		Low
Scientific Rigor (high/low)	Low	
Context (high/low)		Low
Cognitive Awareness	Low	

Remember: 2 types of Decision Making Croskerry 2009

Characteristic	Type 1 Heuristic Intuitive	Type 2 Systematic Analytic
Cognitive Style	Heuristic associative (experience-based) Inductive reasoning	Bounded rationality (Hypothetico-deductive) Normative reasoning
Cost	Low	High
Automaticity	High	Low
Rate	Fast	Slow
Reliability	Low	High
Errors	High	Low
Effort	Low	High
Predictive Power	Low	High
Emotional Component	High	Low
Scientific Rigor	Low	High
Context	High	Low
Cognitive Awareness	Low	High

Croskerry, P. 2009. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in Health Sciences Education, 14, (1), 27-35.

Model for diagnostic reasoning



Croskerry, P. 2009. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in Health Sciences Education, 14, (1), 27-35.

Most (if not all) medical decisions can be formulated as a search in a huge search space [1]

Medical Decision Making is searching for an optimal (“good”*) solution within a search space

***) Attention in clinical practice:**
“Good intentions are the opposite of good”
in German: “Gut gemeint ist das Gegenteil von gut”

[1] Pena-Reyes, C. A. & Sipper, M. 2000. Evolutionary computation in medicine: an overview. Artificial Intelligence in Medicine, 19, (1), 1-23.

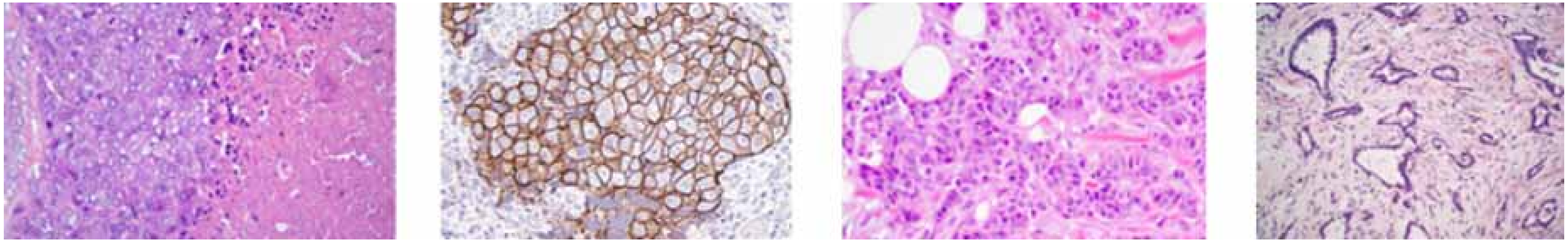
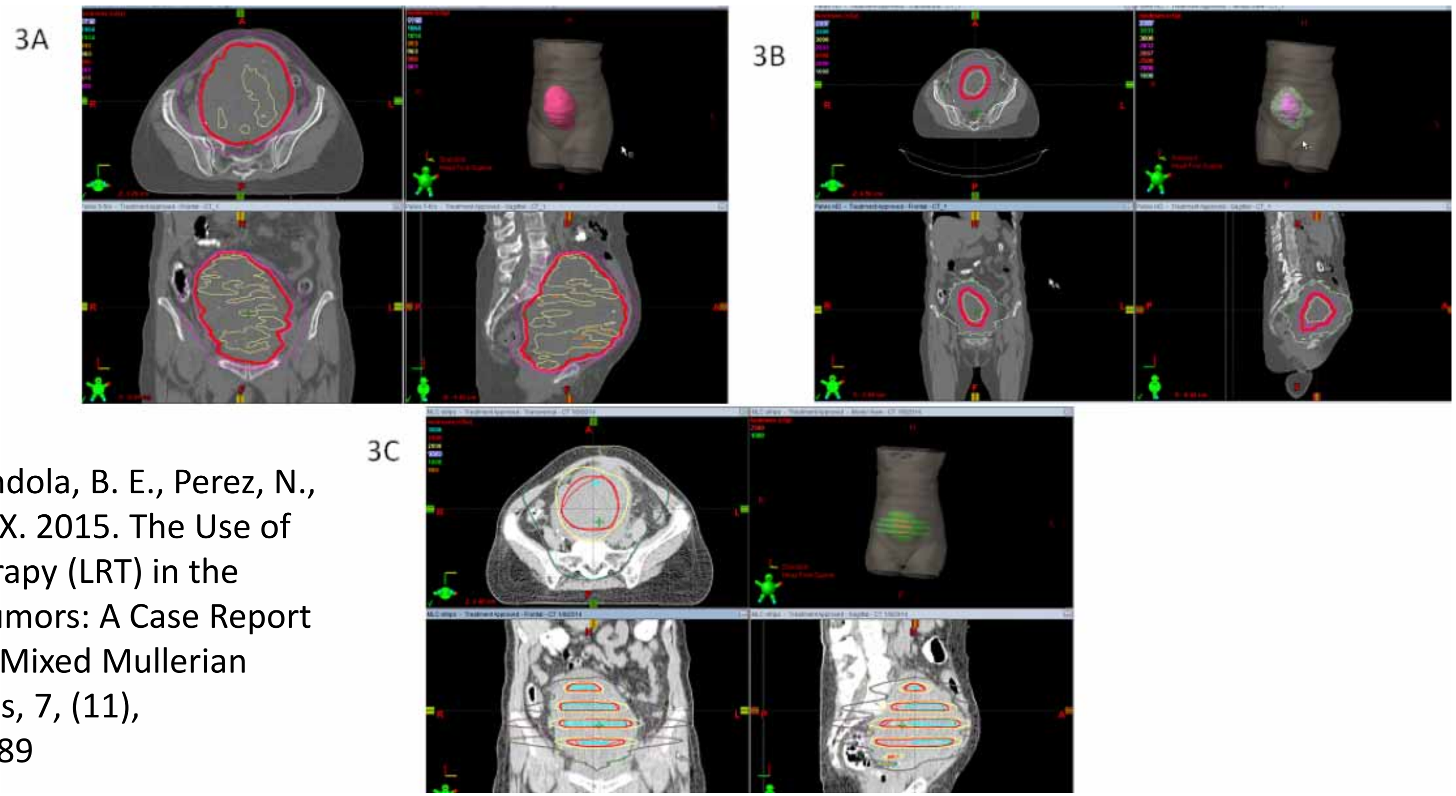


Image Source: https://blogforbreastcancer.wordpress.com/2015/06/30/biopsy-basics-prediction-prognistics-pathology/

- Example 1: a pathologist analyzing biopsies to decide whether they are malignant or not.
- The pathologist is searching in the space of all possible cell features for a set of features permitting to provide a clear diagnosis

Pena-Reyes, C. A. & Sipper, M. 1999. A fuzzy-genetic approach to breast cancer diagnosis. *Artificial intelligence in medicine,* 17, (2), 131-155.

Why EC for health applications? Example 2



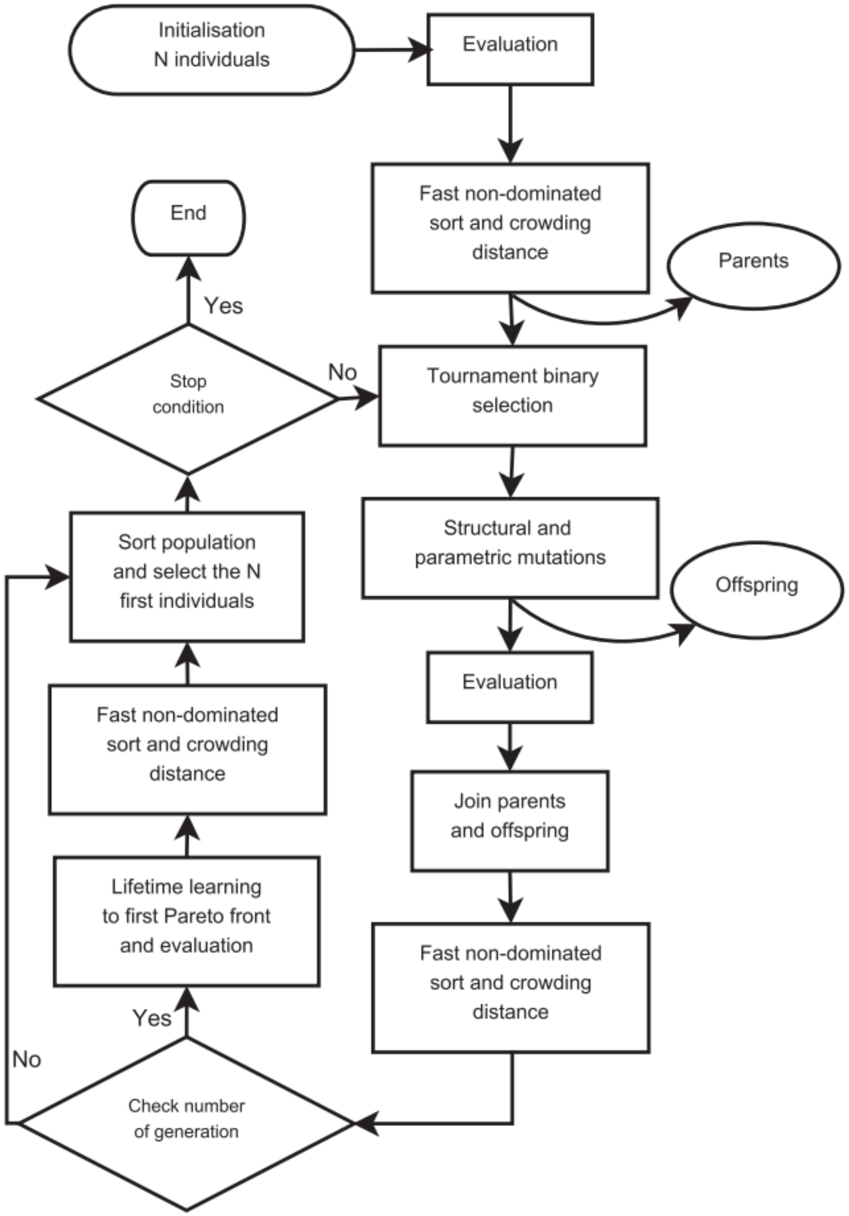
Suarez, J. M. B., Amendola, B. E., Perez, N., Amendola, M. & Wu, X. 2015. The Use of Lattice Radiation Therapy (LRT) in the Treatment of Bulky Tumors: A Case Report of a Large Metastatic Mixed Mullerian Ovarian Tumor. Cureus, 7, (11), doi:10.7759/cureus.389

- Example 2: A radiologist planning a sequence of radiation doses is searching for the best treatment in the space of all possible treatments

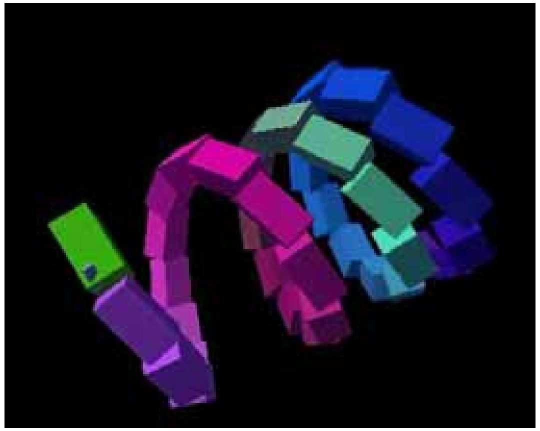
Why EC for health applications? Example 3

The optimal allocation of organs in liver transplantation is a problem that can be resolved using machine-learning techniques. Classical methods of allocation included the assignment of an organ to the first patient on the waiting list without taking into account the characteristics of the donor and/or recipient.

Cruz-Ramírez, M., Hervás-Martínez, C., Fernandez, J. C., Briceno, J. & De La Mata, M. 2013. Predicting patient survival after liver transplantation using evolutionary multi-objective artificial neural networks. Artificial intelligence in medicine, 58, (1), 37-49, doi:doi:10.1016/j.artmed.2013.02.004.

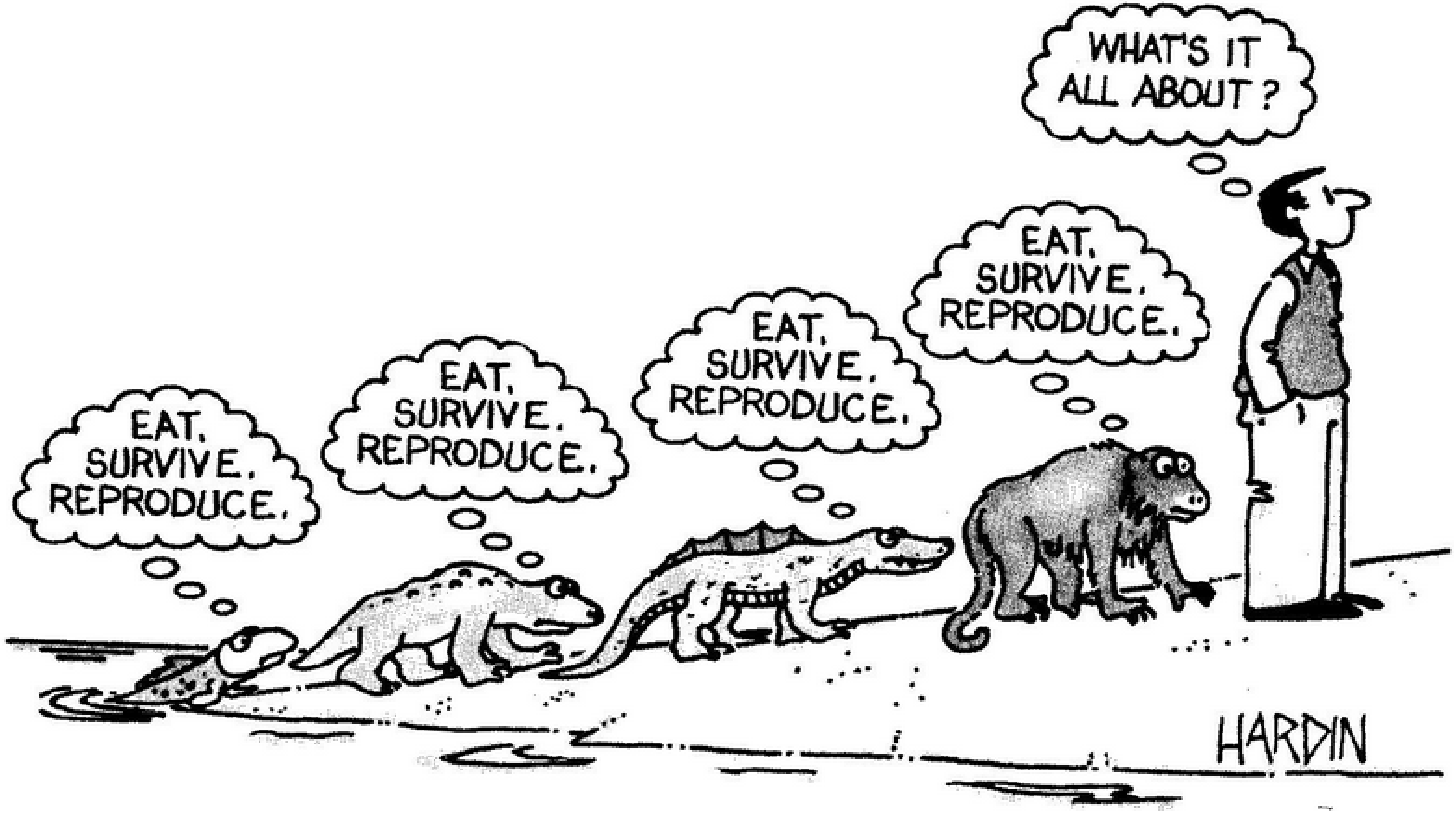


2) Evolutionary Principles



http://www.interaliamag.org/audiovisual/thomas-ray-aesthetically-evolved-virtual-pets/

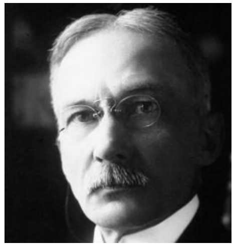
"Evolution is the natural way to program"
Thomas S. Ray, University of Oklahoma,
http://life.ou.edu/



Knoll, A. H. & Bambach, R. K. 2000. Directionality in the history of life: diffusion from the left wall or repeated scaling of the right? Paleobiology, 26, 1-14.

- **Jean Baptiste de Lamarck,** 1801. Theory of Inheritance of Acquired Characteristics, Paris

- **Charles Darwin,** 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, London, John Murray.



- **James M. Baldwin,** 1896. A New Factor in Evolution. The American Naturalist, 30, (354), 441-451, doi:10.2307/2453130.

- **Gregor Mendel,** 1866. Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brunn 4: 3, 44.

- The goal of aML is to build systems that learn and make decisions *without* the human.
- Early aML efforts, e.g. the perceptron [1], had been truly inspired by human intelligence.
- Today, probabilistic modelling has become the cornerstone of aML [2], with applications in neural processing [3] and human learning [4].

[1] McCulloch, W. S. & Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biology, 5, (4), 115-133, doi:10.1007/BF02459570.
[2] Doya, K., Ishii, S., Pouget, A. & Rao, R. 2007. Bayesian brain: Probabilistic approaches to neural coding, Boston (MA), MIT press.
[3] Deneve, S. 2008. Bayesian spiking neurons I: inference. Neural computation, 20, (1), 91-117.
[4] 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.

- Based on the evolutionary theories of **Darwin**, **Lamarck, Baldwin, Mendel.**
- Since the 1980s, EAs have been used for **optimization** problems
- Exploring the possibility of **optimizing** machine learning algorithms rather recently [1]

[1] Z. Zhang, G. Gao, J. Yue, Y. Duan, and Y. Shi, “Multi-criteria optimization classifier using fuzzification, kernel and penalty factors for predicting protein interaction hot spots,” Applied Soft Computing, vol. 18, no. 0, pp. 115–125, 2014.

Terminology of EC [1]

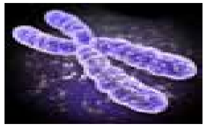
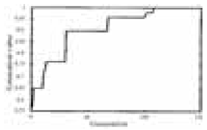
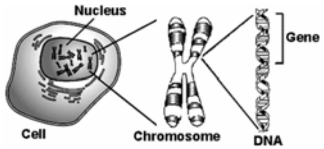
Evolutionary algorithms

- **Genetic algorithms**
- **Evolutionary programming**
- **Genetic programming, Neuroevolution, …**

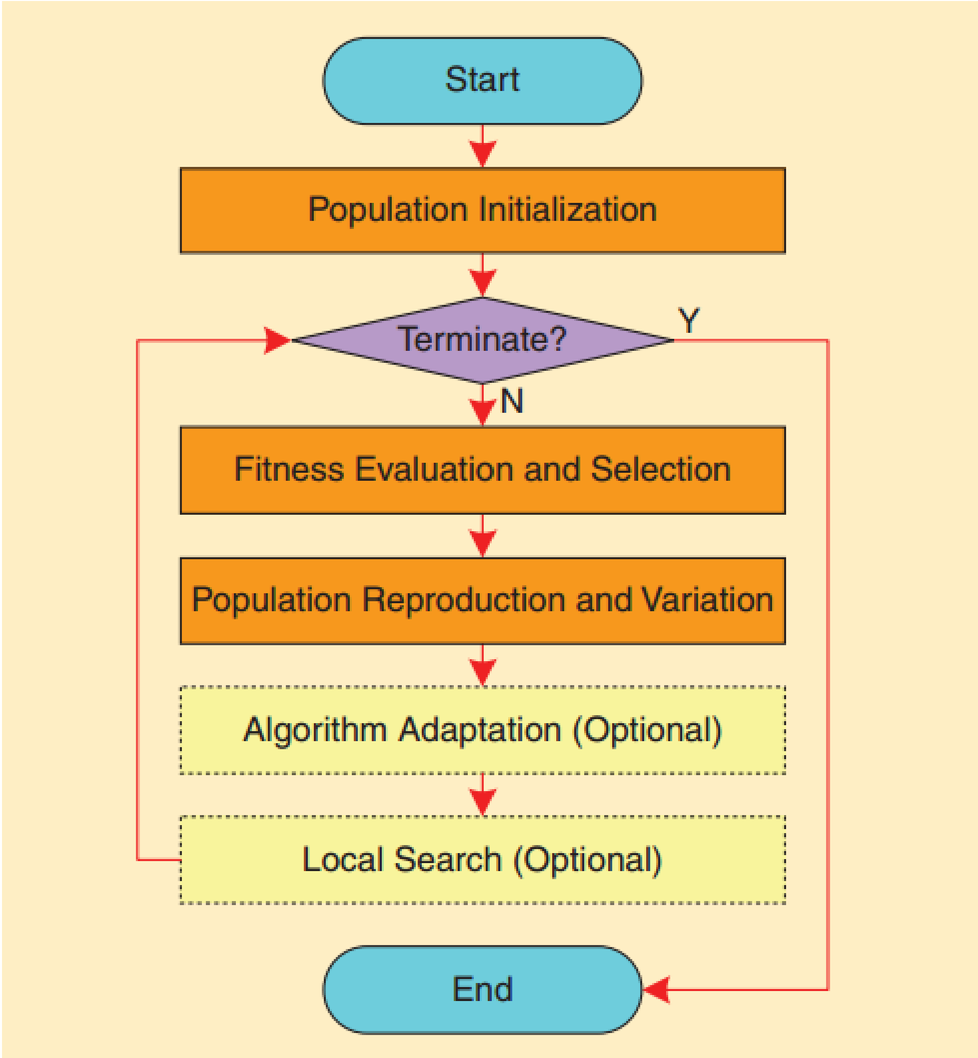
Particle swarm optimization,
Artificial Bee algorithm,
Invasive Weed Optimization
Intelligent Water Drops
Ant colony optimization,
…

[1] Michalewicz, Z. 1996. Genetic algorithms + data structures = evolution programs, New York, Springer.

Biological Universe vs. Computational Universe

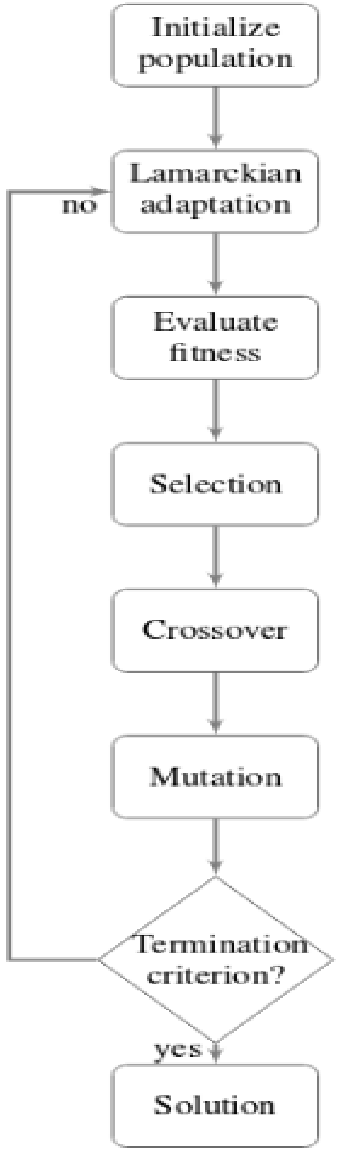
NOTION	*BIOLOGICAL UNIVERSE*	*COMPUTATIONAL UNIVERSE*
Chromosome 	DNA, protein, and RNA sequence in cells	Sequence of information objects
Fitness 	Determines chances of survival and reproduction	Determines chances of survival and reproduction
Gene  	Part of a Chromosome, determines a (partial) characteristic of an individual	Information object, e.g. a bit, a character, number etc.
Generation	Population at a point in time	Population at a point in time
Individual	Living organism	Solution candidate
Population	Set of living organisms	Bag or multi-set of Chromosomes

Holzinger, K., Palade, V., Rabadan, R. & Holzinger, A. 2014. Darwin or Lamarck? Future Challenges in Evolutionary Algorithms for Knowledge Discovery and Data Mining. *In: LNCS 8401.* Heidelberg, Berlin: Springer, pp. 35-56.



[1] Zhang, J., Zhan, Z.-H., Lin, Y., Chen, N., Gong, Y.-J., Zhong, J.-H., Chung, H. S., Li, Y. & Shi, Y.-H. 2011. Evolutionary computation meets machine learning: A survey. *Computational Intelligence Magazine, IEEE, 6, (4), 68-75.*

- Modify chromosomes to adapt to the environment
  - can be used additionally or instead of mutation process
- A local search optimization is applied (e.g. Hill Climbing)
- **Baldwin** uses only pseudo adaptation



[1] B. J. Ross, "A lamarckian evolution strategy for genetic algorithms," Practical handbook of genetic algorithms: complex coding systems, vol. 3, pp. 1–16, 1999.

- **Naive Bayes** is a very effective classifier
- EAs need parameters that can be modified
- A **Weighted Naive Bayesian** (wnb) [1] classifier offers the possibility of easy optimization:

$$p(a_1, a_2, \cdots, a_n|c) = \prod_{i=1}^{n} p(a_i|c). \qquad V_{nb}(E) = \arg\max_{c} p(c) \prod_{i=1}^{n} p(a_i|c)$$

$$V_{wnb}(E) = \arg\max_{c} p(c) \prod_{i=1}^{n} p(a_i|c)^{w_i}$$

[1] Zhang, H. & Sheng, S. Learning weighted naive Bayes with accurate ranking. Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on, 2004. IEEE, 567-570.

- **Dataset**: Pima Indians Diabetes dataset [8]
  - 768 instances (patients)
  - 8 attributes
  - 2 classes
- **Fitness** of an chromosome determined by: number of correctly classified instances in training set
- **Performance** was compared to algorithms in Weka

[8] K. Bache and M. Lichman, "UCI machine learning repository," 2013.[Online]. Available: http://archive.ics.uci.edu/m

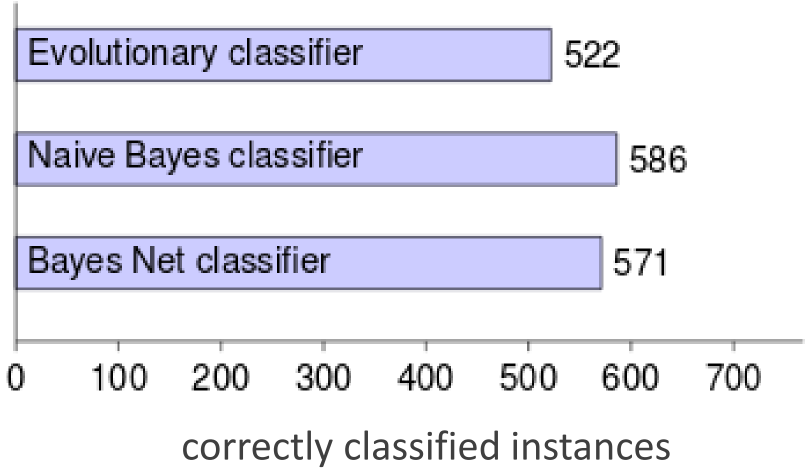
Fitness function

Algorithm 1 Fitness function

```
procedure FITNESS FUNCTION(weightings[ ], List trainingSet)
    for all instances of trainingset do
        for i = 1 to NumberOfClasses do
            for all attribute to MaxNumberAttributes do
                probability[i] *= NORMDISTRIBUTION(attribute + weightings[attribute])

        index ←INDEX OF MAX(probability[])

        if index == CLASS OF(instance) then
            INCREMENT(fitness)
        else
            DECREMENT(fitness)

    RETURN fitness
```

Results and Encountered Problems

Holzinger, A., Blanchard, D., Bloice, M., Holzinger, K., Palade, V. & Rabadan, R. Darwin, Lamarck, or Baldwin: Applying Evolutionary Algorithms to Machine Learning Techniques. In: Ślęzak, D., Dunin-Kęplicz, B., Lewis, M. & Terano, T., eds. IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 Warsaw, Poland. IEEE, 449-453, doi:10.1109/WI-IAT.2014.132.



- Advantages:
  - Fast to train and fast to classify
  - Not sensitive to irrelevant features
  - Handles real and discrete data
- Disadvantages:
  - Assumes independence of features

Conclusion

- Offers many possibilities to improve machine learning algorithms, but finding the right parameters is a difficult task
- Not many machine learning algorithms are suitable for **direct function optimization**
- Implementation of EA:
  - straightforward
  - simple
- EAs are suitable for many tasks in health informatics beyond function optimization

Holzinger, A., Blanchard, D., Bloice, M., Holzinger, K., Palade, V. & Rabadan, R. Darwin, Lamarck, or Baldwin: Applying Evolutionary Algorithms to Machine Learning Techniques. In: Ślęzak, D., Dunin-Kęplicz, B., Lewis, M. & Terano, T., eds. IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 Warsaw, Poland. IEEE, 449-453, doi:10.1109/WI-IAT.2014.132.

- Improvement of function optimization strategy
- Use EAs in different **fields**
  - Graph Optimization
  - Text Mining [1]
  - Feature selection
- Usage of novel **evolutionary strategies**
  - Intelligent Water Drops
  - Invasive Weed
  - Ant Colony with humans-in-the-loop (Super-Ants)

[1] Mukherjee, Indrajit, et al. Content analysis based on text mining using genetic algorithm. In: Computer Technology and Development (ICCTD), 2010 2nd International Conference on. IEEE, 2010. S. 432-436.2

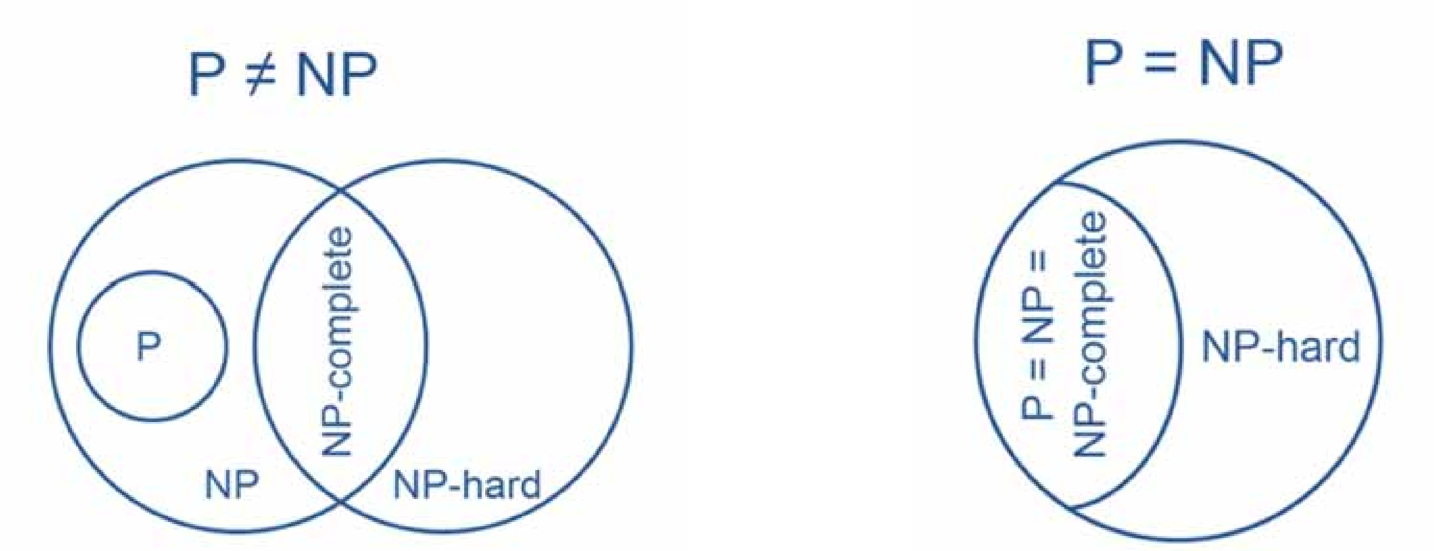
- Text mining with EAs on unstructured information:
  - Doctors/Nurse reports
  - Different Medical Records
  - …
- Sample applications:
  - Categorizing Texts into subject groups [1]
  - Mining “interesting” details [2] like:
    - Gender
    - Age
    - Addresses
    - Occupation

[1] Mukherjee, Indrajit, et al. Content analysis based on text mining using genetic algorithm. In: Computer Technology and Development (ICCTD), 2010 2nd International Conference on. IEEE, 2010. S. 432-436.2

[2] Deepankar B. and Suneet S. Text Mining Technique using Genetic Algorithm. *IJCA Proceedings on International Conference on Advances in Computer Application 2013* ICACA 2013: 7-10, Feb. 2013. Pub.: Foundation of Computer Science, N.Y., USA.

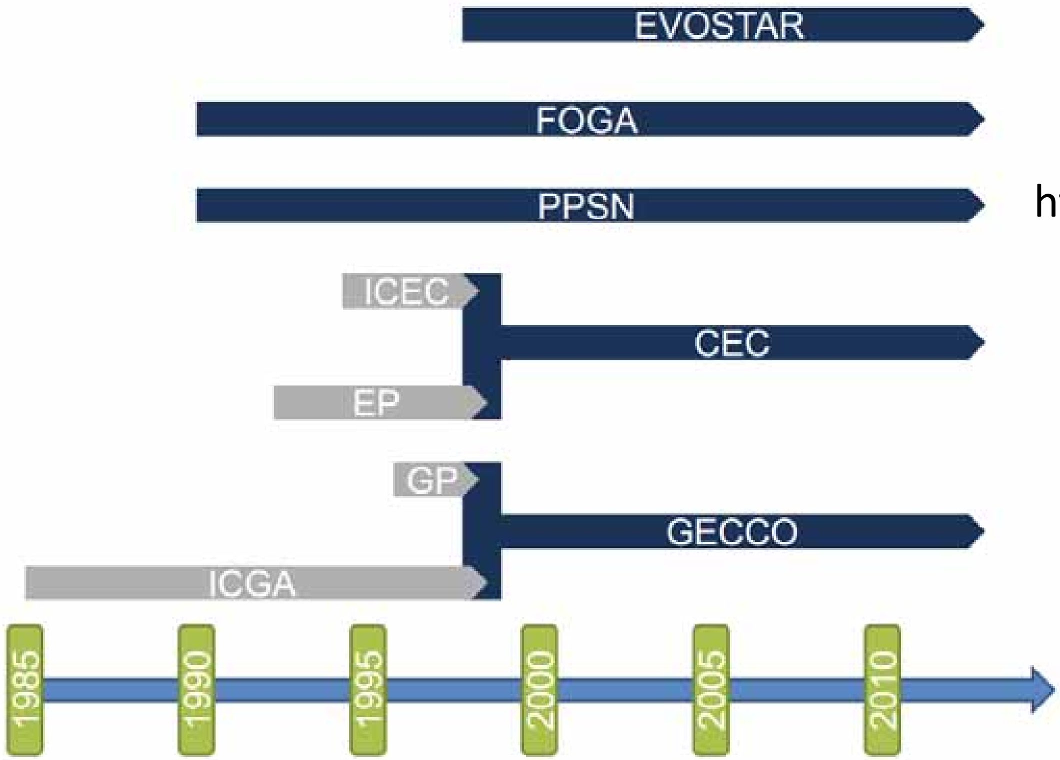
3) Evolutionary Computing

- **P:** algorithm can solve the problem in polynomial time (worst-case running-time for problem size n is less than F(n))
- **NP:** problem can be solved and any solution can be verified within polynomial time ($P \subseteq NP$)
- **NP-complete:** problem belongs to class NP and any other problem in NP can be reduced to this problem
- **NP-hard:** problem is at least as hard as any other problem in NP-complete but solution cannot necessarily be verified within polynomial time



- 1948 Alan Turing:
  “genetical or evolutionary search”
- 1962 Hans-Joachim Bremermann:
- optimization through evolution and recombination
- 1964 Ingo Rechenberg:
- introduces evolution strategies
- 1965 Lawrence J. Fogel, Owens and Walsh:
- introduce evolutionary programming
- 1975 John Holland:
- introduces genetic algorithms
- 1992 John Koza:
- introduces genetic programming

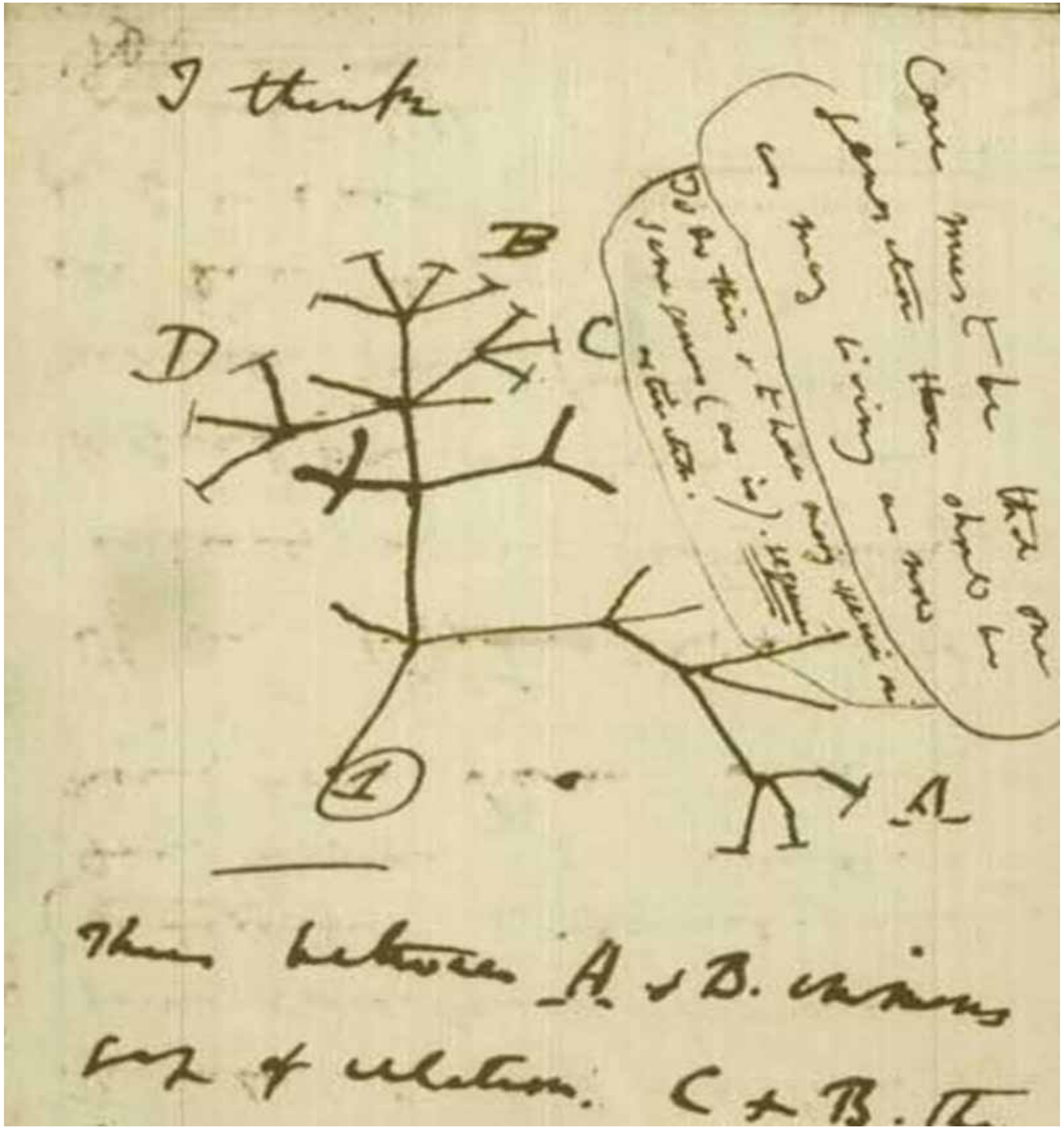
Evolution of evolutionary computing conferences



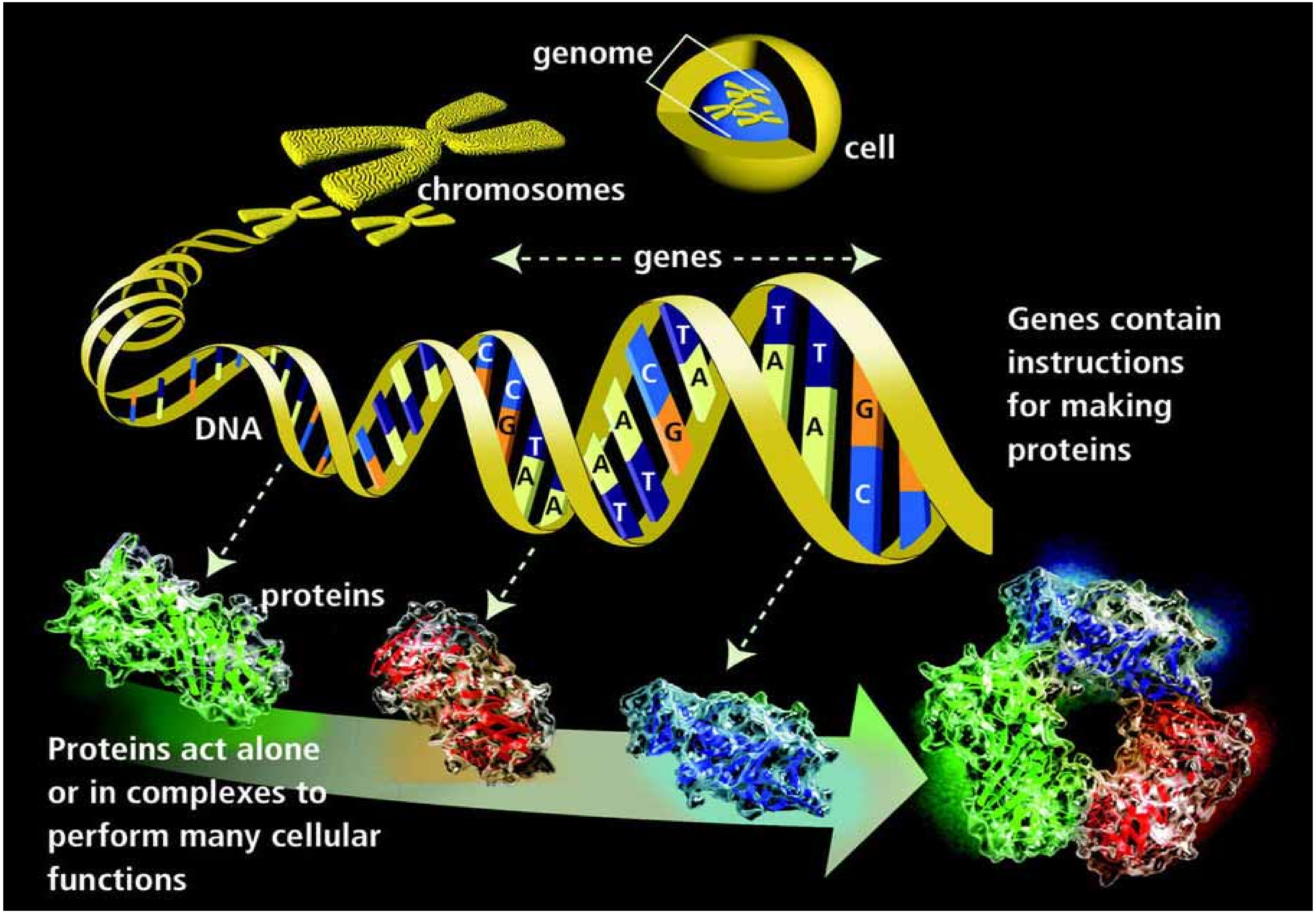
http://dblp2.uni-trier.de/db/conf/evoW/

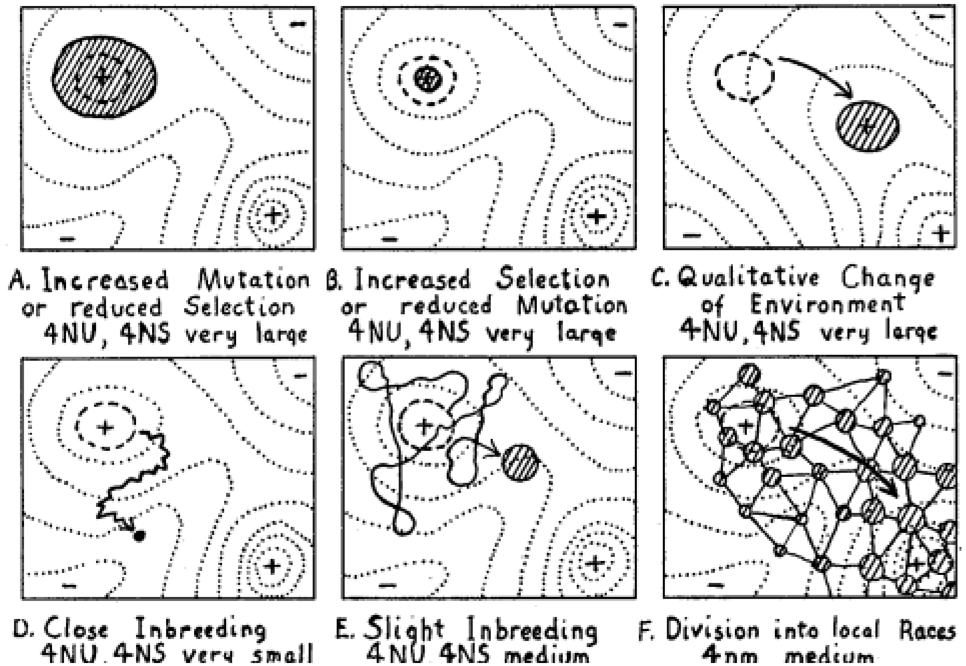
http://dblp.uni-trier.de/db/conf/cec/

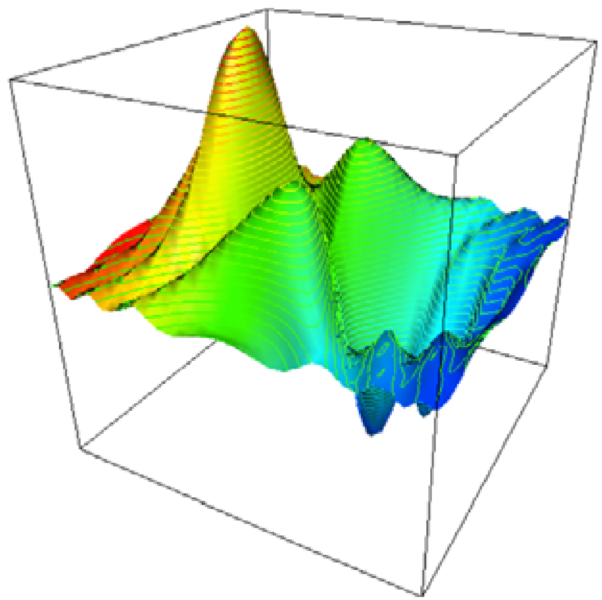
http://dblp.uni-trier.de/db/conf/gecco/


I think
B
C
D
1
A

Microscopic View on Natural Evolution





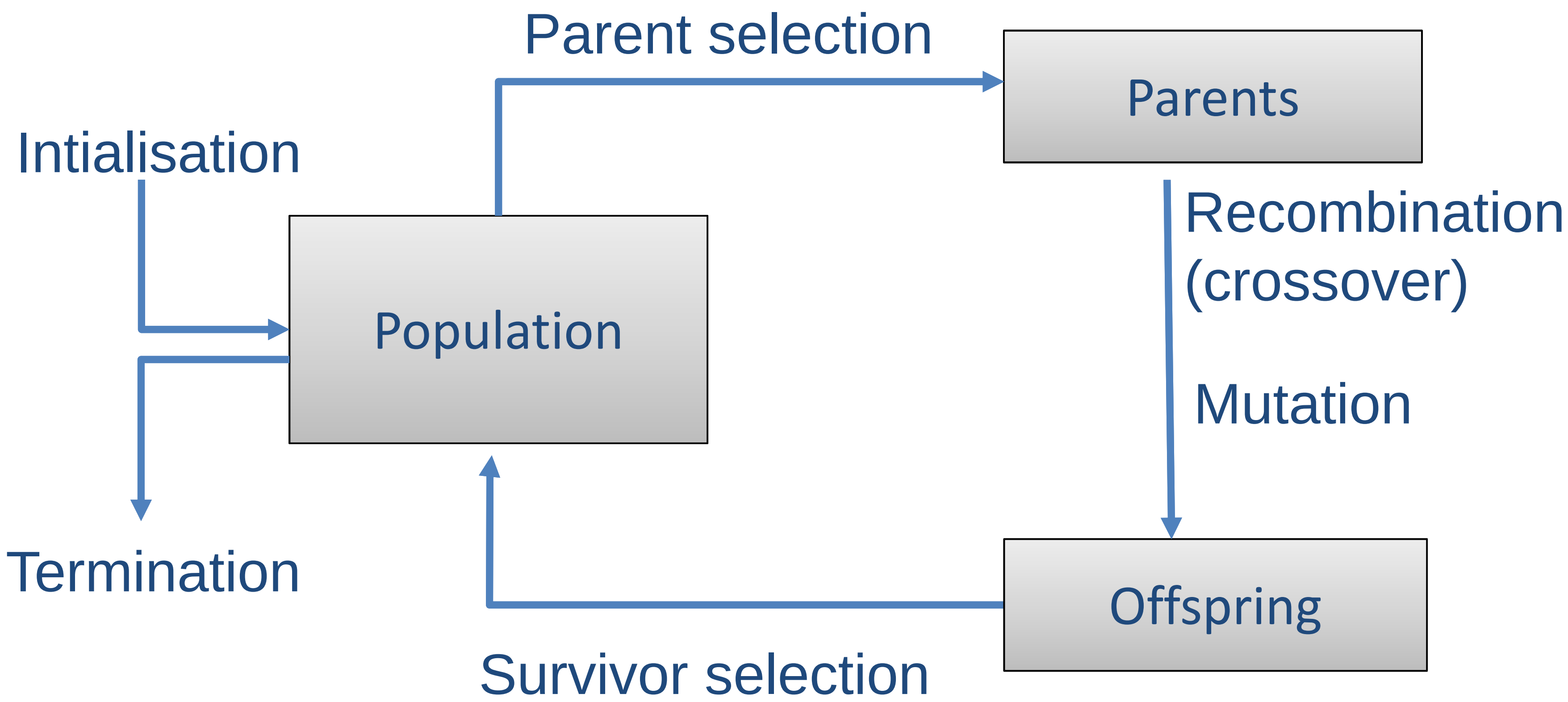


An evolving population is conceptualized as moving on a surface whose points represent the set of possible solutions = search space

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. 6th International Congress on Genetics. Ithaca (NY). 356-366.

General Scheme of an Evolutionary Algorithm


```
BEGIN
  INITIALISE population with random candidate solutions;
  EVALUATE each candidate;
  REPEAT UNTIL ( TERMINATION CONDITION is satisfied ) DO
    1 SELECT parents;
    2 RECOMBINE pairs of parents;
    3 MUTATE the resulting offspring;
    4 EVALUATE new candidates;
    5 SELECT individuals for the next generation;
  OD
END
```



- Population of individuals
- Each individual has a fitness function
- Variation operators: crossover, mutation, …
- Selection towards higher fitness by "survival of the fittest" and "mating of the fittest"

Neo Darwinism:

Evolutionary progress towards higher life forms

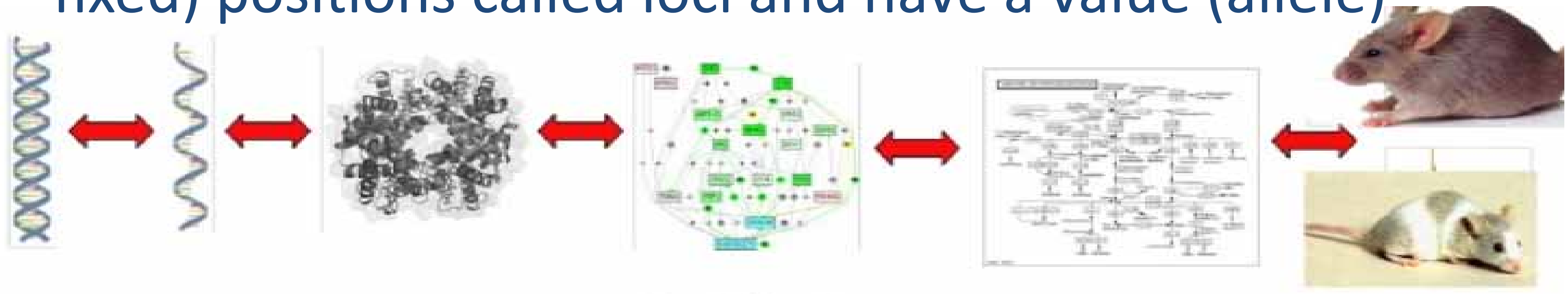
=

Optimization according to some fitness-criterion
(optimization on a fitness landscape)

- 1) **Increasing population diversity** by genetic operators (e.g. mutation, recombination, …)
  Push towards creating **novelty**
- 2) **Decreasing population diversity** by selection of parents and survivors
  Push towards **quality**

- Role: provides code for candidate solutions that can be manipulated by variation operators, and leads to two levels of existence:
  - **phenotype**: object in original problem context (outside)
  - **genotype:** code to denote that object, the inside (chromosome, "digital DNA")
- Implies two mappings:
- Encoding: phenotype → genotype (not necess. 1:1)
- Decoding: genotype → phenotype (must be 1:1)

Chromosomes contain genes, which are in (usually fixed) positions called loci and have a value (allele)



Phenotype -> Genotype (integers -> binary code)

- In order to find the global optimum, every feasible solution must be represented in the genotype space

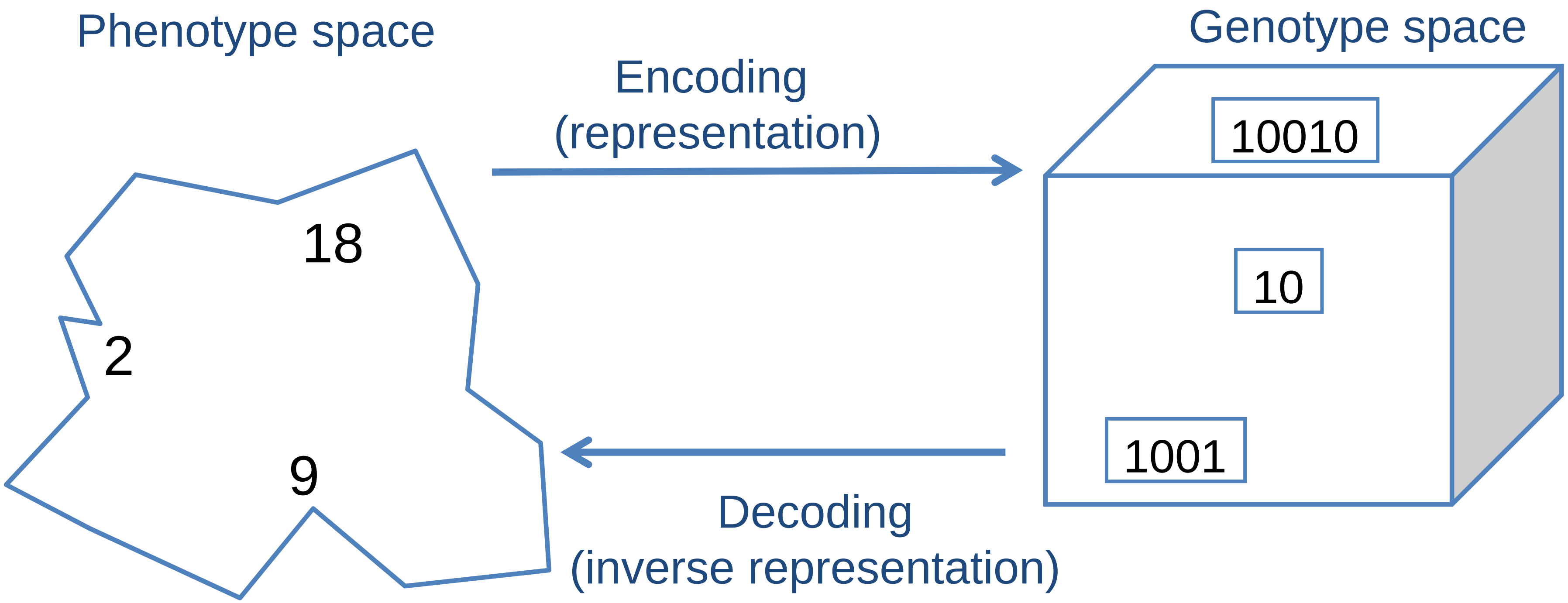


Image credit: Eiben, A. E. & Smith, J. E. 2015. Introduction to evolutionary computing. Second Edition, Berlin, Springer.

Evaluation function = Fitness function

- Role:
  - Represents the task to solve, the requirements to adapt to (can be seen as "the environment")
  - Enables selection (provides basis for comparison)
  - e.g., some phenotypic traits are advantageous, desirable, e.g. big ears cool better, these traits are rewarded by more offspring that will expectedly carry the same trait
- A.k.a. *quality* function or *objective* function
- Assigns a single real-valued fitness to each phenotype which forms the basis for selection
  - So the more discrimination (different values) the better
- Typically we talk about fitness being maximised
  - Some problems may be best posed as minimisation problems, but conversion is trivial

Population

- Role: holds the candidate solutions of the problem as individuals (genotypes)
- Formally, a population is a multiset of individuals, i.e. repetitions are possible
- Population is the basic unit of evolution, i.e., the population is evolving, not the individuals
- Selection operators act on population level
- Variation operators act on individual level
- Some sophisticated EAs also assert a spatial structure on the population e.g., a grid
- Selection operators usually take whole population into account i.e., reproductive probabilities are relative to current generation
- **Diversity** of a population refers to the number of different fitness / phenotypes / genotypes present (note: not the same thing)

Role:

- Identifies individuals
  - to become parents
  - to survive
- Pushes population towards higher fitness
- Usually probabilistic
  - high quality solutions more likely to be selected than low quality
  - but not guaranteed
  - even worst in current population usually has non-zero probability of being selected
- This *stochastic* nature can aid escape from local optima

- Most EAs use fixed population size so need a way of going from (parents + offspring) to next generation
- Often deterministic (while parent selection is usually stochastic)
- Fitness based : e.g., rank parents + offspring and take best
- Age based: make as many offspring as parents and delete all parents
- Sometimes a combination of stochastic and deterministic (elitism)

- Role: to generate new candidate solutions
- Usually divided into two types according to their arity (number of inputs):
- Arity 1 : mutation operators
- Arity >1 : recombination operators
- Arity = 2 typically called crossover
- Arity > 2 is formally possible, seldom used in EC
- There has been much debate about relative importance of recombination and mutation
- Nowadays most EAs use both
- Variation operators must match the given representation

- Role: causes small, random variance
- Acts on one genotype and delivers another
- Element of randomness is essential and differentiates it from other unary heuristic operators
- Importance ascribed depends on representation and historical dialect:
- Binary GAs – background operator responsible for preserving and introducing diversity
- EP for FSM’s / continuous variables – only search operator
- GP – hardly used
- May guarantee connectedness of search space and hence convergence proofs

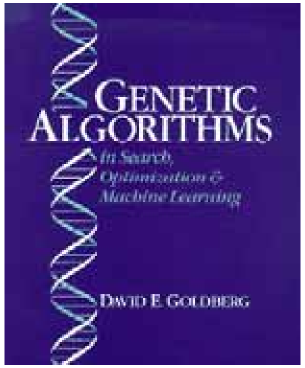
- Role: merges information from parents into offspring
- Choice of what information to merge is stochastic
- Most offspring may be worse, or the same as the parents
- Hope is that some are better by combining elements of genotypes that lead to good traits
- Principle has been used for millennia by breeders of plants and livestock

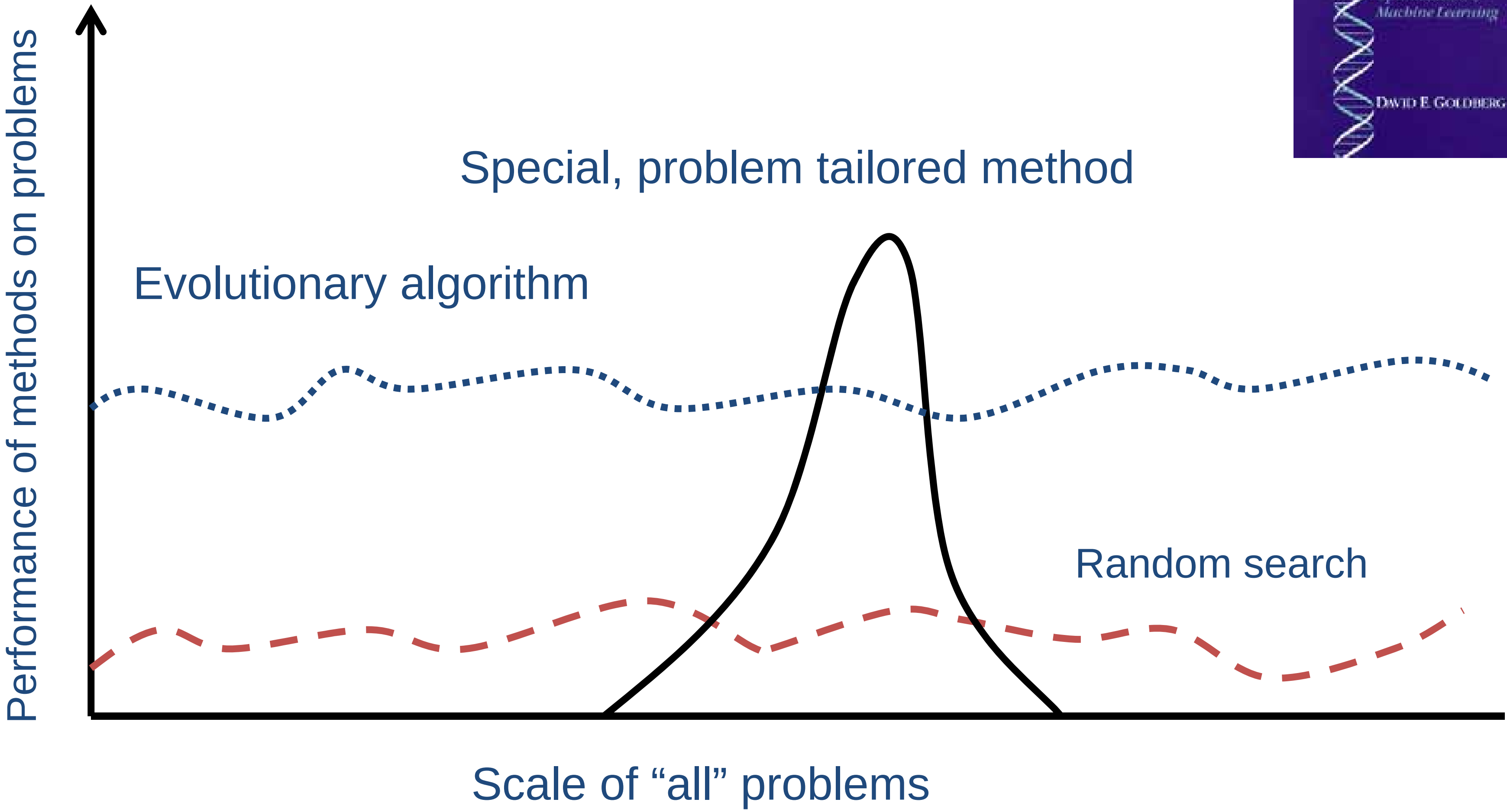
- Initialisation usually done at random,
- Need to ensure even spread and mixture of possible allele values
- Can include existing solutions, or use problem-specific heuristics, to "seed" the population

- Termination condition checked every generation
- Reaching some (known/hoped for) fitness
- Reaching some maximum allowed number of generations
- Reaching some minimum level of diversity
- Reaching some specified number of generations without fitness improvement

- Historically different EAs have been associated with different data types to represent solutions
- Binary strings : Genetic Algorithms
- Real-valued vectors : Evolution Strategies
- Finite state Machines: Evolutionary Programming
- LISP trees: Genetic Programming
- These differences are largely irrelevant, best strategy
- choose representation to suit problem
- choose variation operators to suit representation
- Selection operators only use fitness and so are independent of representation

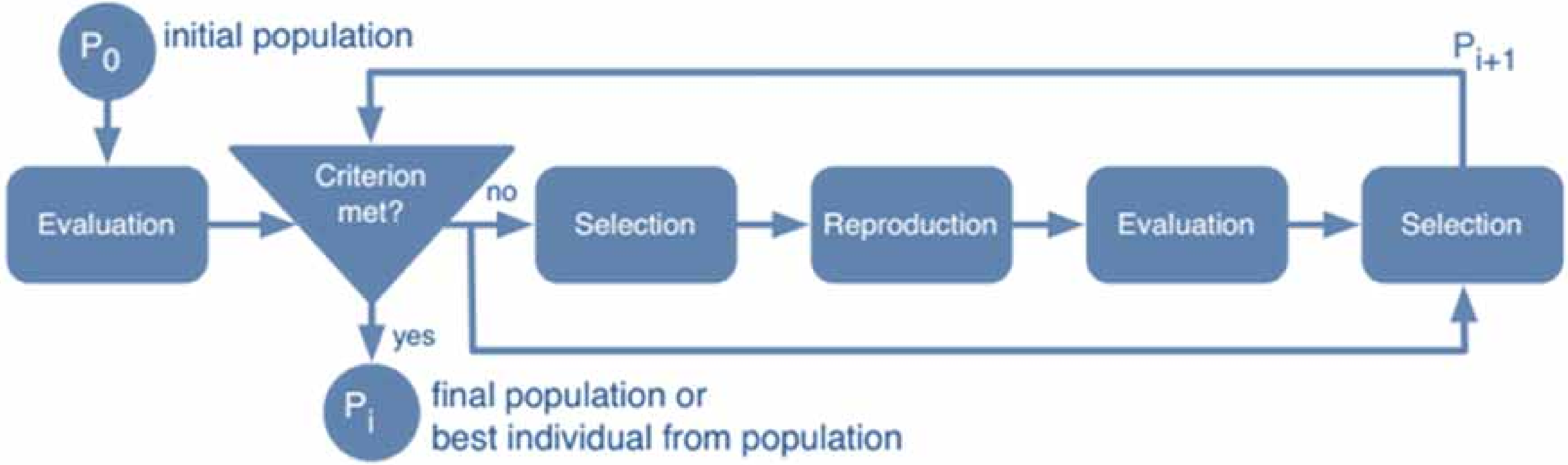
Typical EA behaviour





Goldberg, D. E. 1989. Genetic algorithms in search, optimization, and machine learning, Reading (MA), Addison-Wesley

Image credit to Sascha Lange, University of Freiburg



- **Individuals:** hypothesis x from a hypothesis space X
- **Population:** collection P of μ hypotheses $P = \{x_i \mid i = 1, \ldots, \mu\}$
- **Evaluation:** $f : X \rightarrow R$ (fitness function) to all individuals
- **Selection mechanism:** selects individuals $x \in P_i$ for reproduction (mating); selects individuals from off-springs and P_i to form the new population $P_i + 1$
- **Reproduction:** combination of two or more individuals (Crossover) and random alteration (Mutation).

Fitness function

Algorithm 1 Fitness function

```
procedure FITNESS FUNCTION(weightings[ ], List trainingSet)
    for all instances of trainingset do
        for i = 1 to NumberOfClasses do
            for all attribute to MaxNumberAttributes do
                probability[i] *= NORMDISTRIBUTION(attribute + weightings[attribute])

        index ←INDEX OF MAX(probability[])

        if index == CLASS OF(instance) then
            INCREMENT(fitness)
        else
            DECREMENT(fitness)

    RETURN fitness
```

4) Genetic Algorithms

- Similar to stochastic optimization
- Iteratively trying to improve a possibly large set of candidate solutions
- Few or no assumptions about the problem (need to know what is a good solution)
- Usually finds good rather than optimal solutions
- Adaptable by a number of adjustable parameters

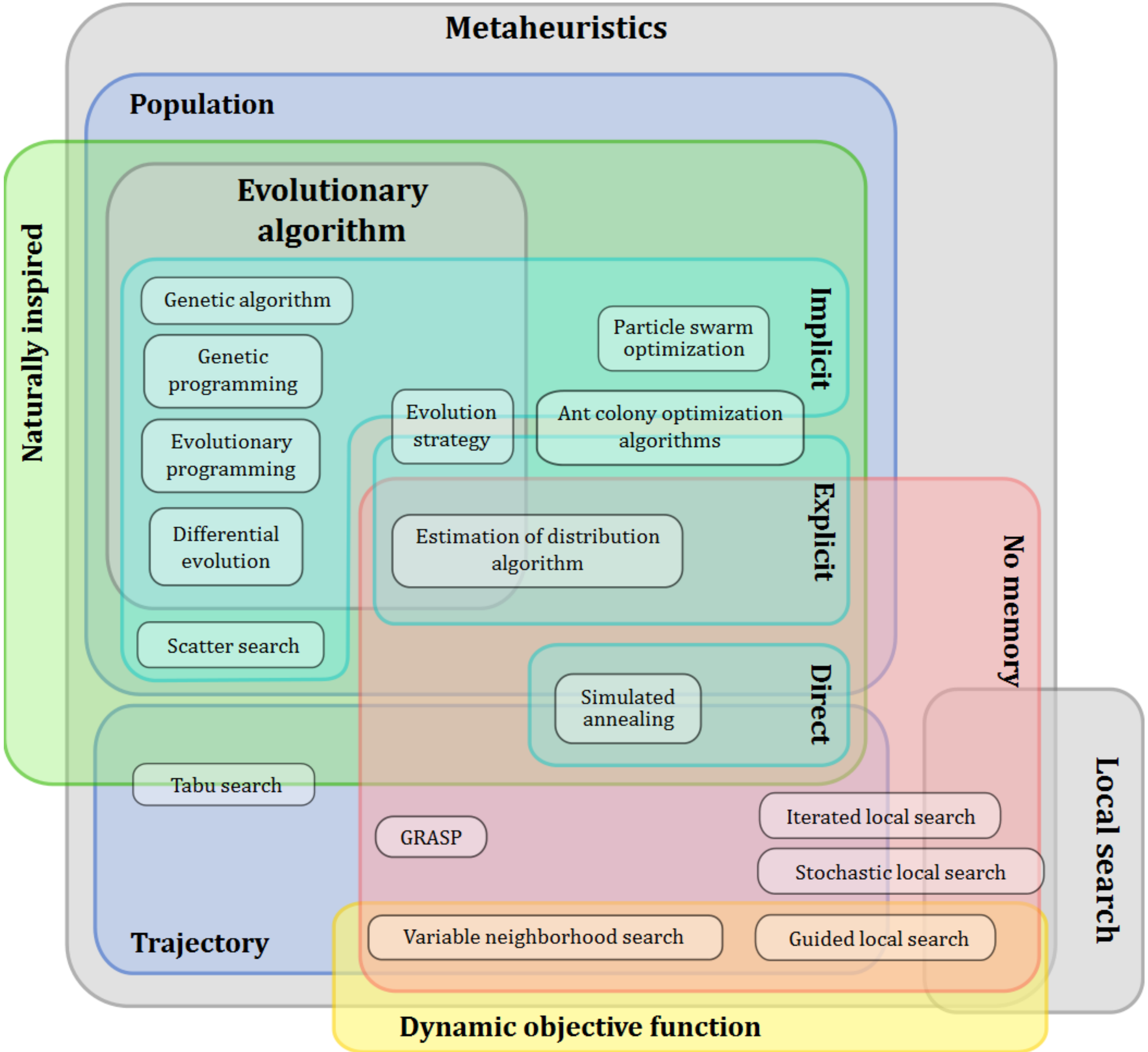
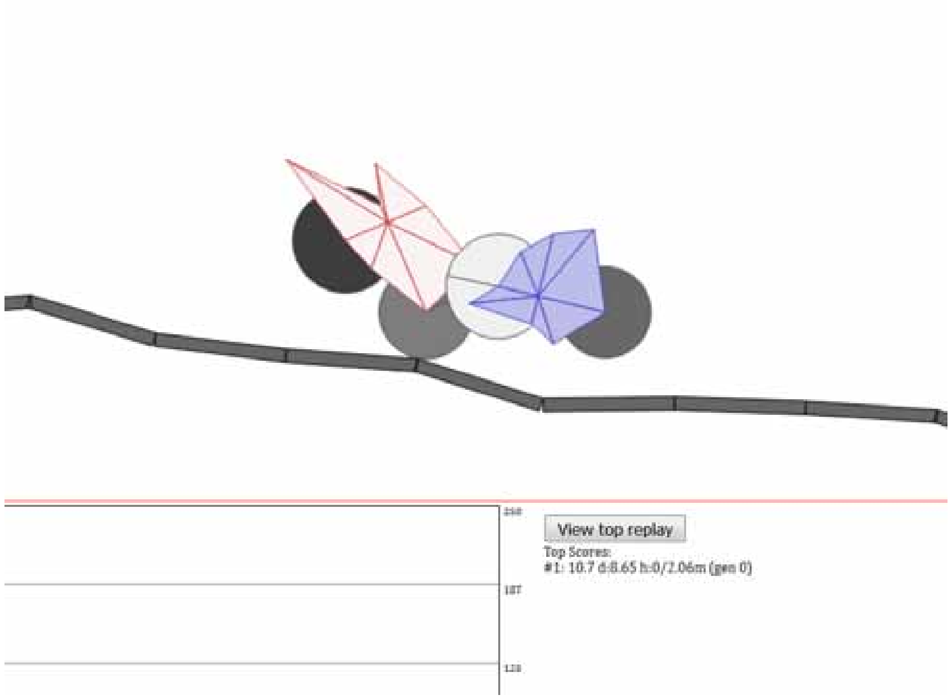
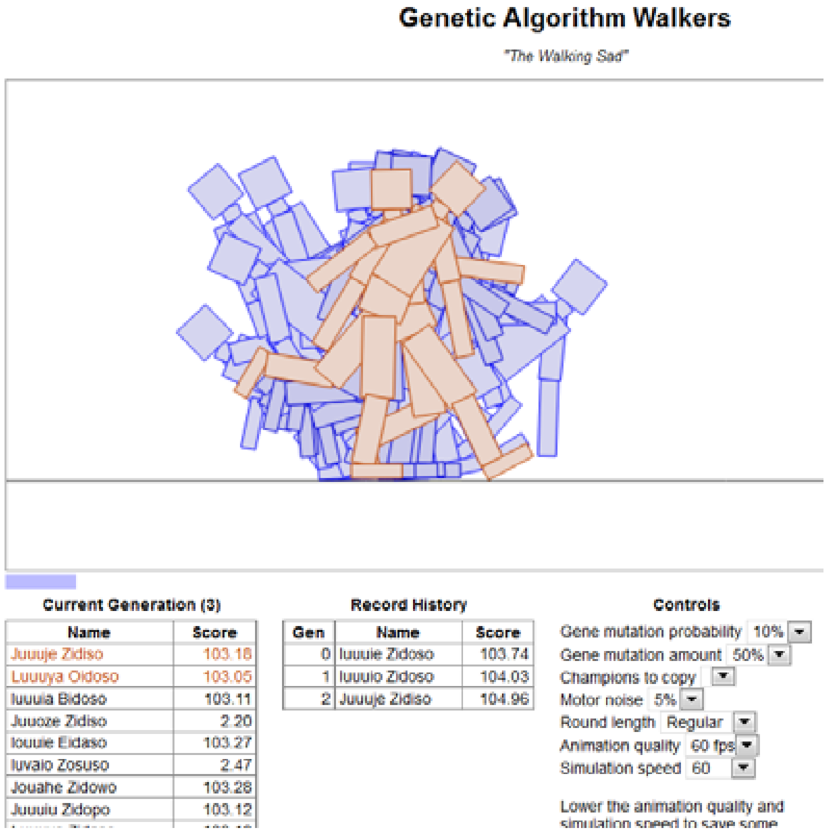


Image Credit to Johann Dréo, Caner Candan - Metaheuristics classification CC BY-SA 3.0 https://commons.wikimedia.org/w/index.php?curid=16252087

Sample genetic algorithm

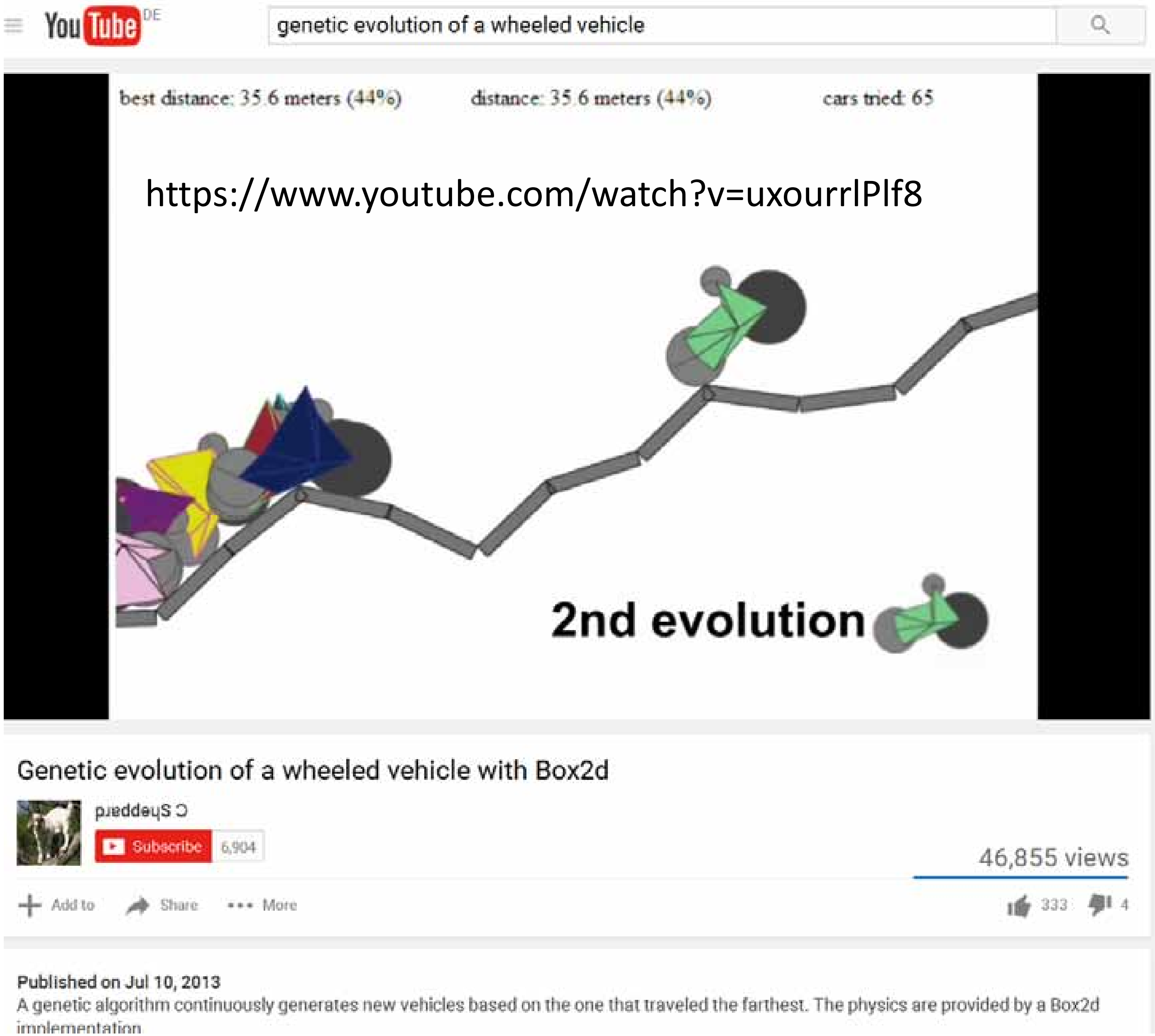




http://rednuht.org/genetic_cars_2/

http://rednuht.org/genetic_walkers/

Video Sample of Genetic Algorithms



K= 2 ⟶ **Two-armed bandit problem**:

Arm 1: award μ_1 with variance σ_1^2

Arm2: award μ_2 with variance σ_2^2

$\mu_1 > \mu_2$

Question: Which arm (left/right) is which index 1, 2?

Can be used for motivation of the Schema Theorem by John Holland (1975): is widely taken to be the foundation for explanations of the power of genetic algorithms: low-order schemata with above-average fitness increase exponentially in successive generations.

Holland, J. H. 1975. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, U Michigan Press (as of 01.06.2016 49,320 citations !)

- N = total number of trials
- $b = \dfrac{\sigma_1}{\mu_1 - \mu_2}$
- Conclusion:
  Expected loss is minimal if approximately:

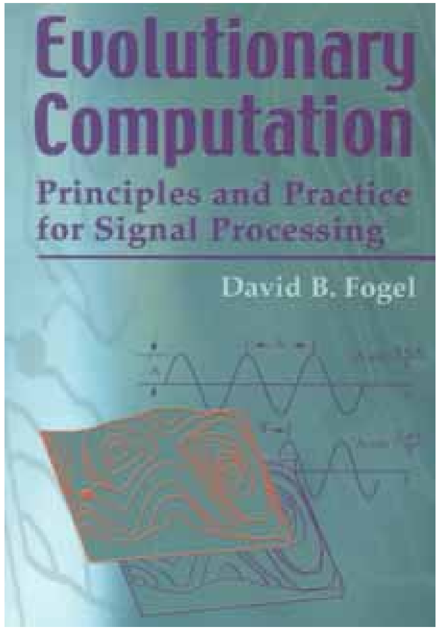
$$n^* \approx b^2 \cdot \ln\left(\frac{N^2}{8\pi \cdot b^4 \cdot \ln(N^2)}\right)$$

- Consequently, trials are allocated to the observed worst arm

$$N - n^* \approx \sqrt{8\pi \cdot b^4 \cdot \ln(N^2)} \cdot \exp\left(\frac{n^*}{2b^2}\right)$$

- The trials are allocated to the **observed** best arm
- This 2-arm bandit can be generalized to a k-armed bandit, resulting in:
- A) Generalized corollary: The optimal strategy is to allocate an exponentially increasing n of trials to the **observed** best arm
- B) This links-up to Genetic Algorithms because: Minimizing expected losses from k-armed bandits ≈ Minimizing expected losses while sampling from order $\log_2(k)$ schemata (=GA’s allocate trials opt.)

- Why would this be optimal for global optimization?
- Minimizing expected losses does not always correspond to maximizing potential gains.



Science **is to test crazy ideas –**

Engineering is put these ideas into Business

Lucky Students ☺

Thank you!

Solutions to the Quiz

- 1=our daily life is decision making! The metaphor "estimate how far you can jump" – shall demonstrate that uncertainty matters – particular in clinical medical decisions!
- 2= The Bayesian brain – our brain as Bayesian statistical inference machine: i.e. when we perceive our physical world, make a decision, and take an action: we are always uncertainties – Bayesian networks help to understand how our brain works;
- 3= Travelling salesman problem – NP-hard – here the human-in-the-loop can help as we will see in the next lecture
- 4= Modeling or system identification problems – typical in machine learning – problem in aML is that all these are black-box approaches and iML fosters a glass-box approach for direct interaction with the algorithm itself
- 5=shows again the complexity of natural-language and the context-dependency!
- 6=In graph theory, an isomorphism of graphs G and H is a bijection between the vertex sets of G and H Find the matches - > graph matching -> very important in proteins -> subgraph isomorphism -> NP hard
- 7=grch. Stokhos ("aim") -> stochastic – in medicine we are constantly confronted with random variables over time. It is the counterpart to deterministic processes;
- 8= Image right: Starburst galaxy, Messier 82 (M82) in the center of milky way (with Hubble telescope); Left: Cluster of benign microcalcifications
- 9= The famous "Ötzi" – the radiologists needed 10 years to discover the arrow in the chest of the prehistoric man. Example for decision making
- 10= The grand challenge is in data integration, to fuse the heterogeneous data sets, sampled from very diverse sources and time-dependend data collected over time; this also needs temporal models; 3 Billion USD per year are spend alone in the US for health (320 Mill Inhabitants);

Sample Questions (1)

- What is the general idea of evolutionary algorithms?
- What is the difference between CI, EC, and GA?
- Why are EC relevant for health informatics?
- What are the main differences in the ideas of Lamarck, Darwin, Baldwin, and Mendel?
- Please explain the general scheme of an evolutionary algorithm and explain the components!
- Sketch the pseudocode of a fitness function!

Appendix: Who is it …