# AI and Machine Learning
# 2 hour primer

Igor Krawczuk

**BayesMill**

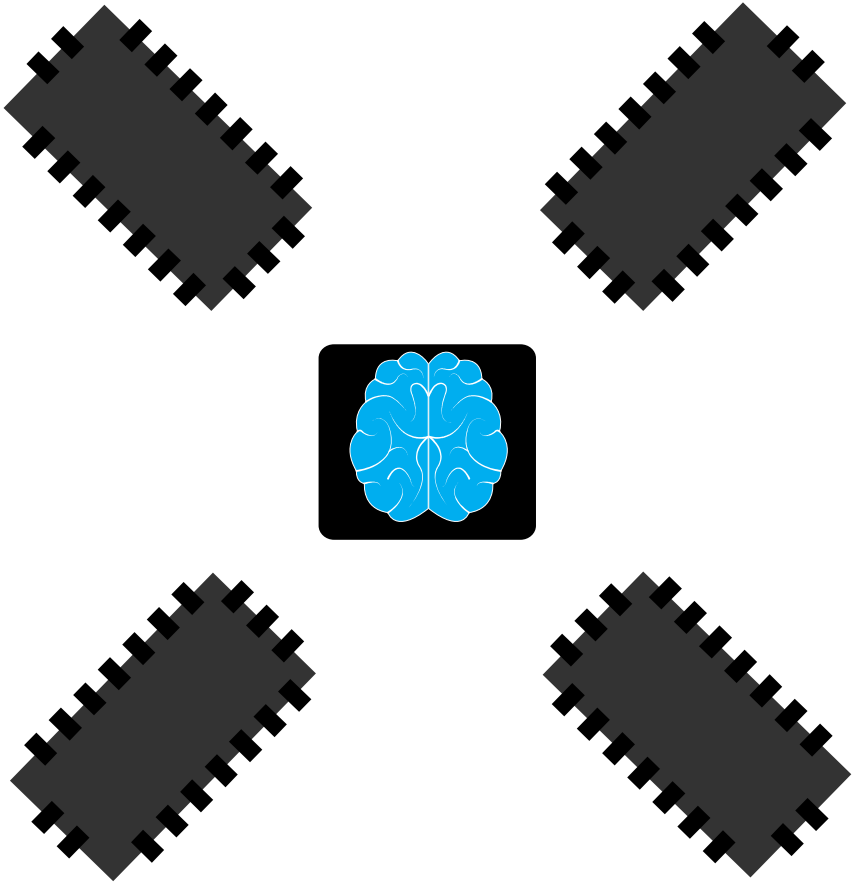
Consulting for:

Data Science

Analytics

Artificial Intelligence

inquiries@bayesmill.com



# Who am I?

- Igor Krawczuk
- MSc. Electrical and Computer Engineering
- Specialization: AI, Machine Learning and neuromorphic Hardware for Machine Learning
- Consulting:
  - Data Science (Modeling and Pipelines)
  - Machine Learning Prototyping
  - AI and Automation advisory
  - AI,Machine Learning and Coding Training

# What this workshop...

**doesn‘t offer:**

- In depth examination
- Academically rigorous definitions
- In depth insight into State of the art(SOA)
- Deep,Hidden knowledge – you can google most of this stuff

**offers:**

- Quick overview
- Intuitive explanations
- Foundations and small tastes of SOA
- A chance to ask somebody who already did the googling

# Toys!

- Python & JavaScript examples to follow along the workshop under:
- https://github.com/igor-krawczuk/mm-aiprimer
- Developed mostly by people smarter than me :-)
- Feel free to explore during the talky bits

  - Install guide in same repository
  - Notebooks today also accessible at: http://163.172.188.213:28801
    - Username+Password: manage+more
    - => don‘t kill my server please, prefer local/JavaScript

# Plan

1)Quick history and definitions

2)Presentation non-machine learning AI

3)Machine Learning

4)Discussion (iff time)

- Applications
- Business considerations
- AI risk
- Technical questions

# Rules

1) Ask questions!

2) Feel free to play around with the examples while I‘m talking/discussing !

3) Interrupt me! Raise your hands and ask questions!

# What is AI?

- „The automation of machines that perform functions that require intelligence when performed by people“(Bellman, 1978)
- „The study of computations that make it possible to perceive, reason and act“(Winston,1992)
- „The art of creating machines that perform function that require intelligence when performed by people.“(Kurzweil,1990)
- „Computational Intelligence is the study of the design of intelligent agents.“(Poole,et.al.,1998)

# What is machine learning?

- The sub-field of AI concerned with learning systems

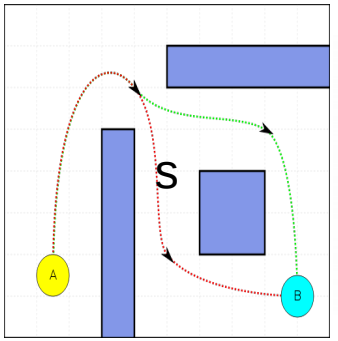
- Machine learning is the sub-field of computer science that "gives computers the ability to learn without being explicitly programmed" (Arthur Samuel, 1959)

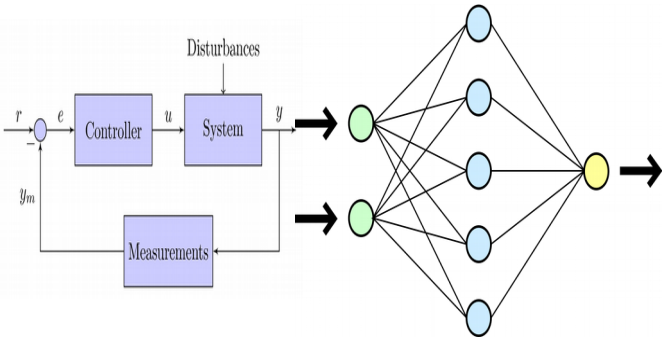
# AI Categories

| Reasoning | Knowledge | Perception | Planning | Control | Learning |
|---|---|---|---|---|---|
| • Logic<br>• symbolic Inference | • Storing<br>• Querying<br>• Sharing | • Vision<br>• Hearing<br>• Reading | • Plan actions<br>• Find routes<br>• Scheduling | • Reactions<br>• Stabilization<br>• Noise canceling | • Information extraction<br>• Estimation<br>• Prediction |

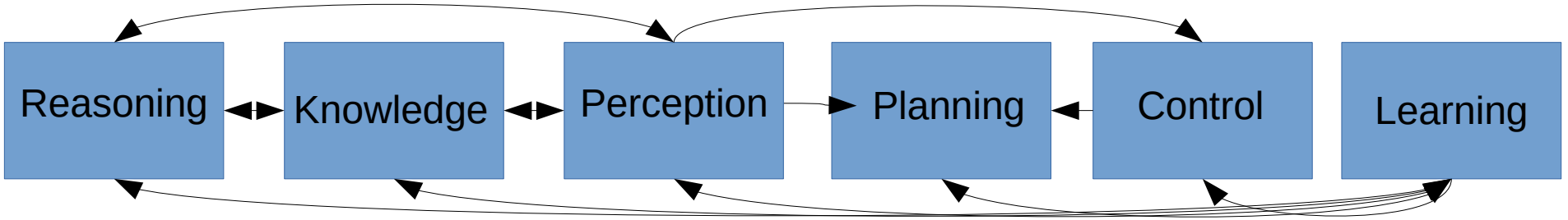
$$\neg(p \wedge \neg p)$$

SWI Prolog

# AI Hierarchy




- Logic
- symbolic Inference

- Storing
- Querying
- Sharing

- Vision
- Hearing
- Reading

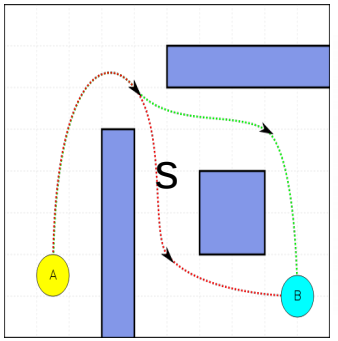
- Plan actions
- Find routes
- Scheduling

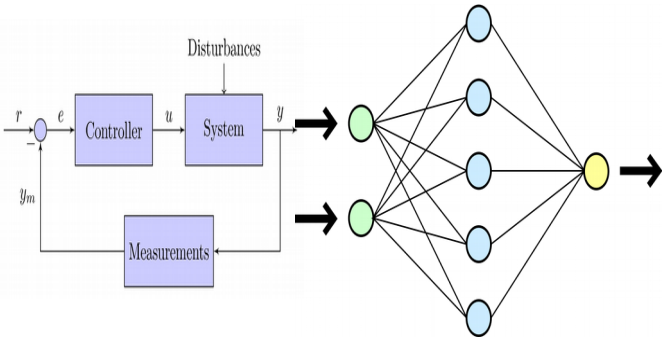
- Reactions
- Stabilization
- Noise canceling

- Information extraction
- Estimation
- Prediction

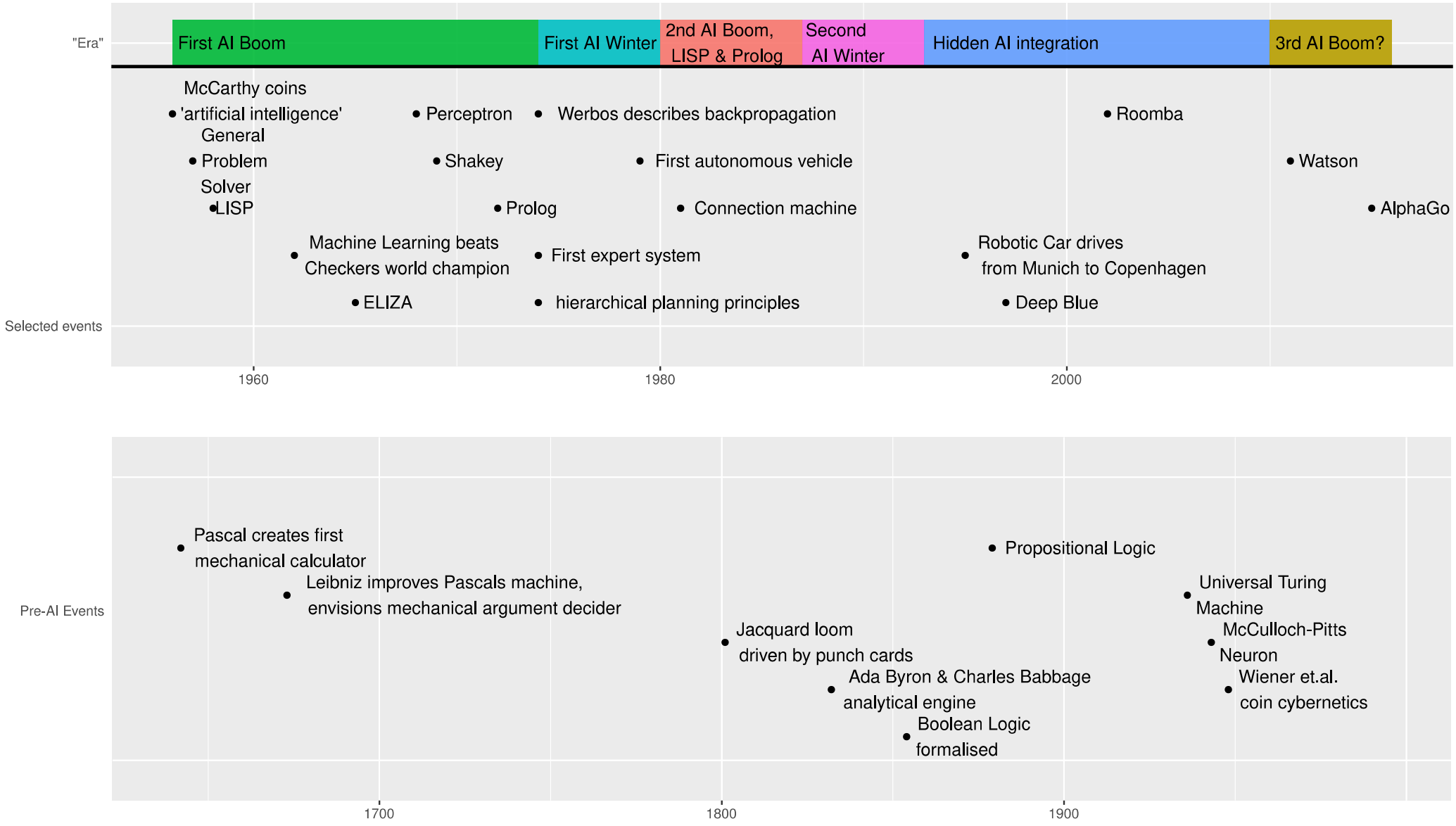
$$\neg(p \wedge \neg p)$$

SWI Prolog

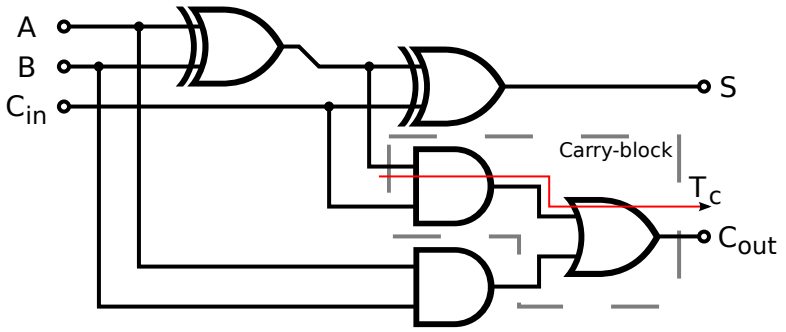
# History of AI



# Reasoning:Boolean Logic

- Basic Logic in Computers
  - We know 1,0 and the operations AND, OR, NOT
  - Can (probably) express *every* possible computation through this(Church-Turing Thesis)
- Example:
  - Addition: XOR and Carry

| A | B | OR | AND |
|---|---|---|---|
| 1 | 0 | **1** | 0 |
| 1 | 1 | **1** | **1** |
| 0 | 0 | 0 | 0 |

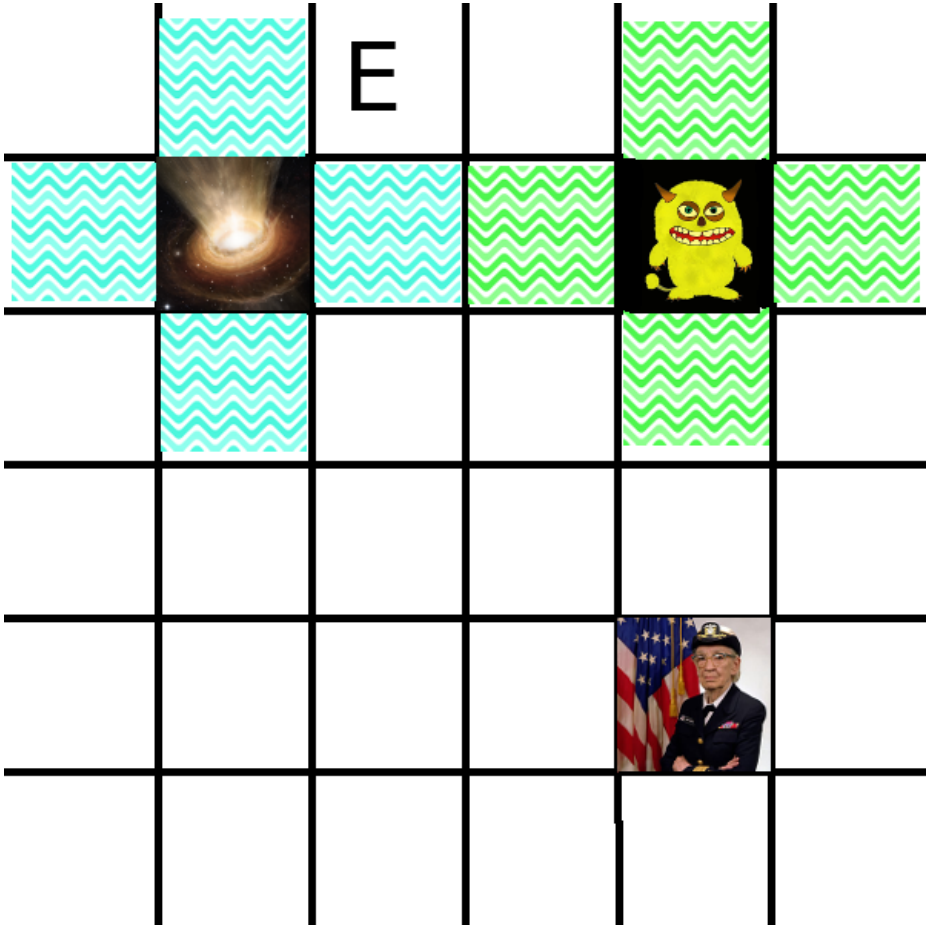


# Reasoning: Symbolic Logic

- Propositional Logic: And, OR,XOR,NOT,iff
  - Formal system, defined by set of axioms
    - P => ¬ Q
    - P
    - => ¬Q
    - What cannot be determined => False(Closed World assumption)
- First Order Logic: add quantifiers
  - ∀Q∀E:¬ Q ∧ E
  - Either is tested (false if Q ∧ E found) or restricts choices for other solutions

# Reasoning Example: Wumpus Hunt

- Wumpus
  - Eats player if in same chamber
  - Stinks=> can be smelled from adjacent rooms
- Pits
  - Kill player if he enters room
  - Drafty => can be felt from adjacent room
- Goal: Escape without dying



# Reasoning: Limits to Symbolic Knowledge

- Combinatorial explosion
- Real world not formally specifiable
  - Fuzzy states
  - Unknown variables (open world)
  - Gödels incompleteness theorem

# Knowledge Representation(KR)

- Prolog encodes knowledge as propositions(facts) and inferences(rules)
  - friends(a,b).
  - friends(b,c).
  - friendOfFriend(X,Y):-friends(X,Z) ,friends(Z,Y).
  - =>friendOfFriend(a,X)=> X=c
- Problems:
  - Ambiguity and Redundancy: friends(X,Y) != friends(Y,X)
  - Definitions and hierarchies: friends <=> acquaintance <=> relationship?
  - Scopes: Pizza is food or fast-food?

- Main Problem: no consistent representation

# Knowledge Representation(KR)

- Solution: Ontologies
  - Collections of knowledge in standardized formats
  - Formats: RDF (Resource Description Framework), OWL (Web Ontology Language)
  - Example: **Cyc**
- Problems:
  - Ambiguity, Scope, Definitions and hierarchy still there...but at least only once
  - LOTS of manpower necessary to create and maintain
  - Challenge: Ontology unification (compatibility)
- =>Humans don‘t need this (as much). We‘ll talk about ontology **learning** later

# KR: Cyc Project

- Ontology of „everyday common sense knowledge“
- 500e3 concepts,

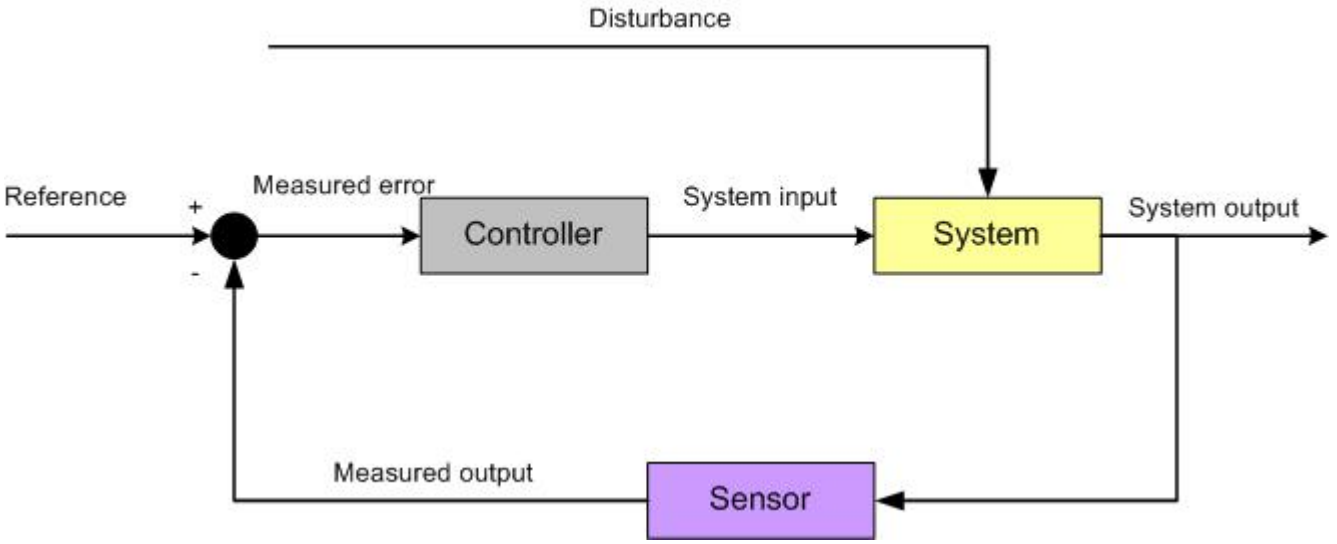
  17e3 types of relations,

  7e6 relations
- Further reading:

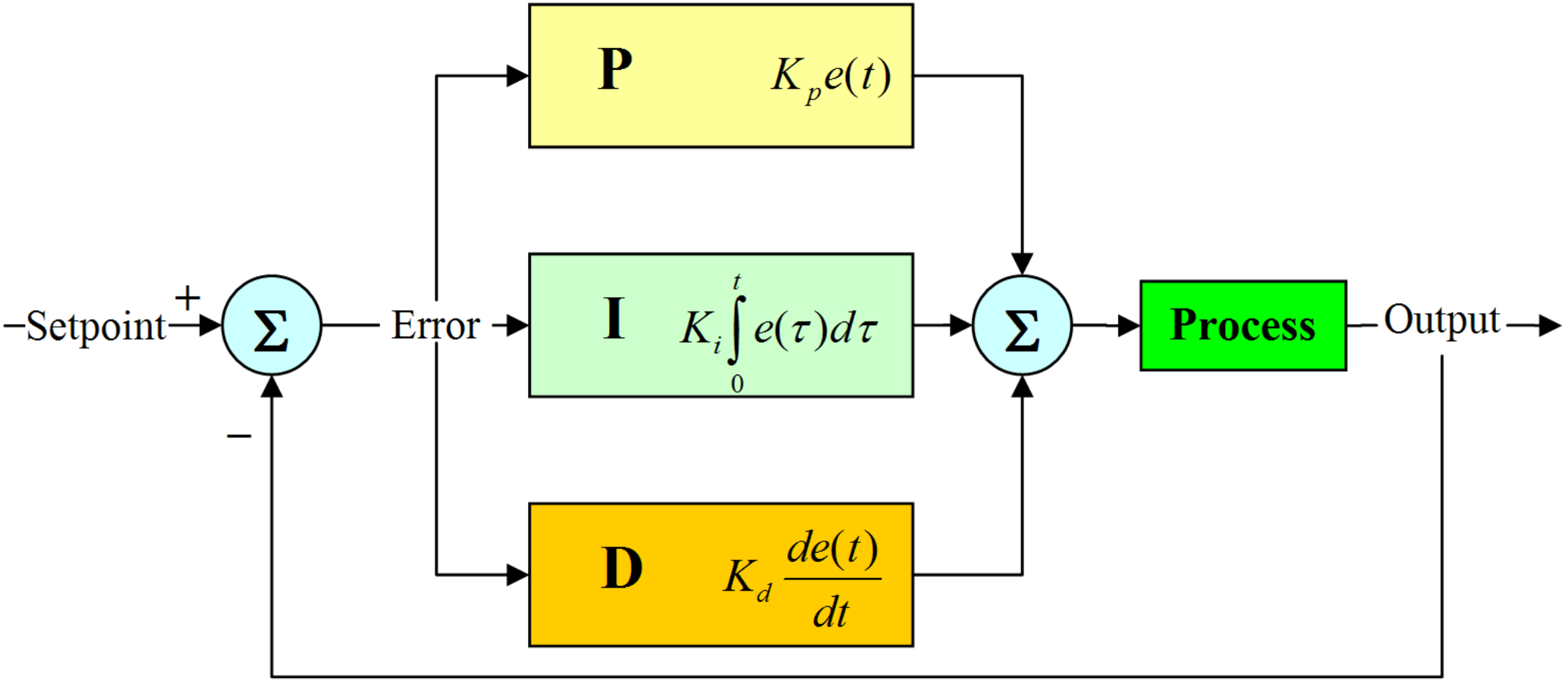
  Survey of Ontologies

# Control

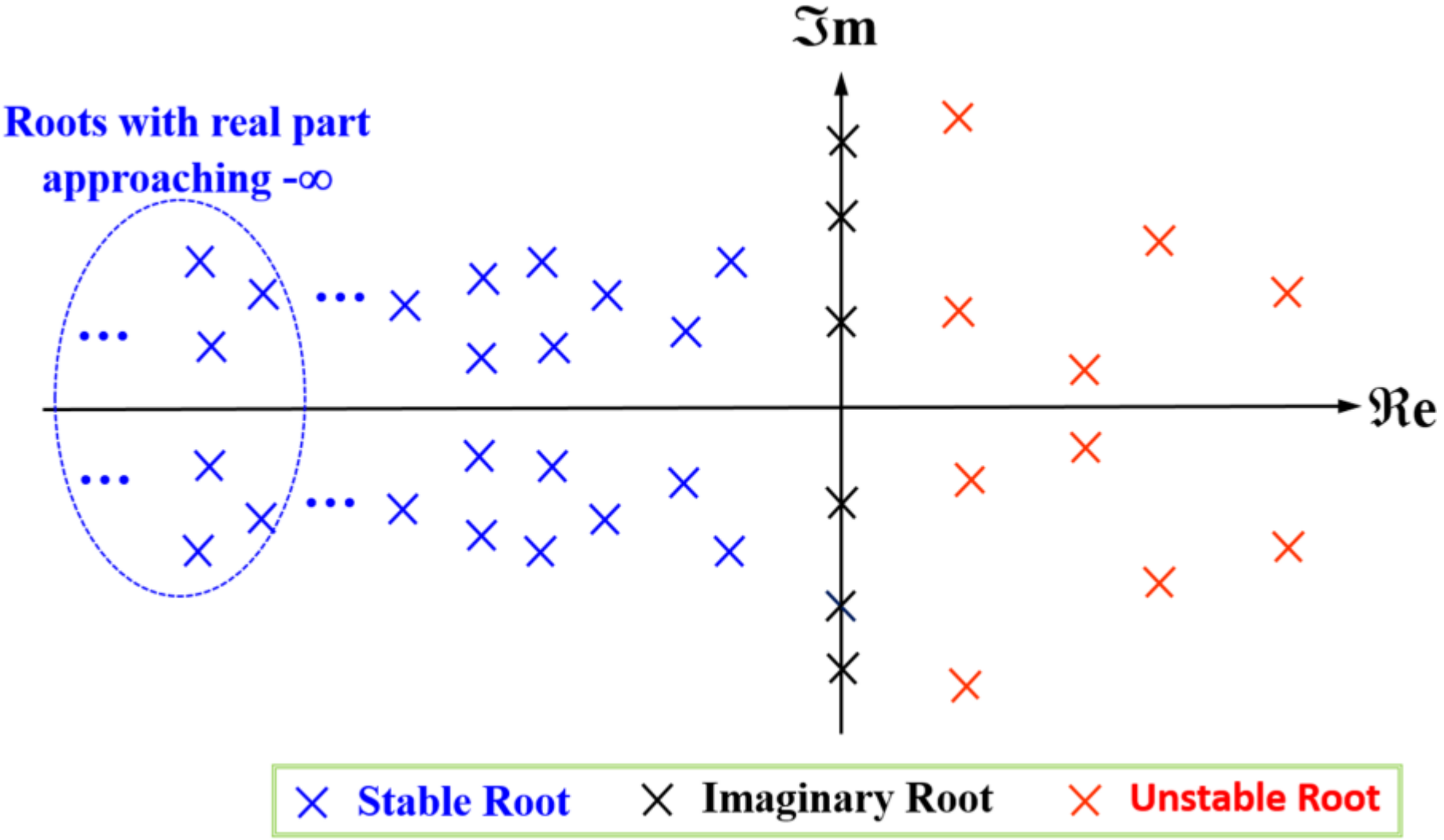
- Real time decision making based on imperfect data to achieve a goal



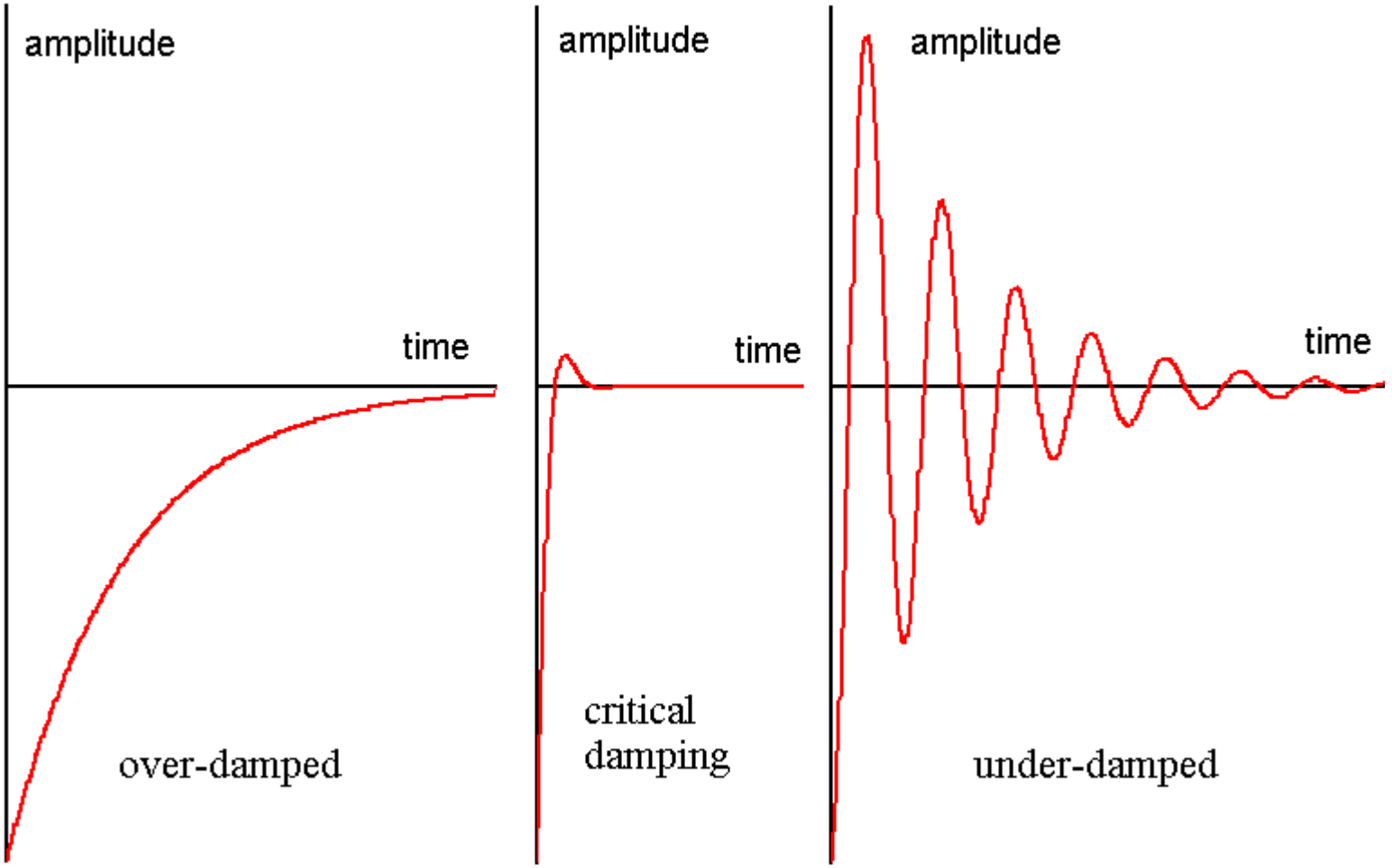
# Control Theory 101 in two Slides: 1. PID



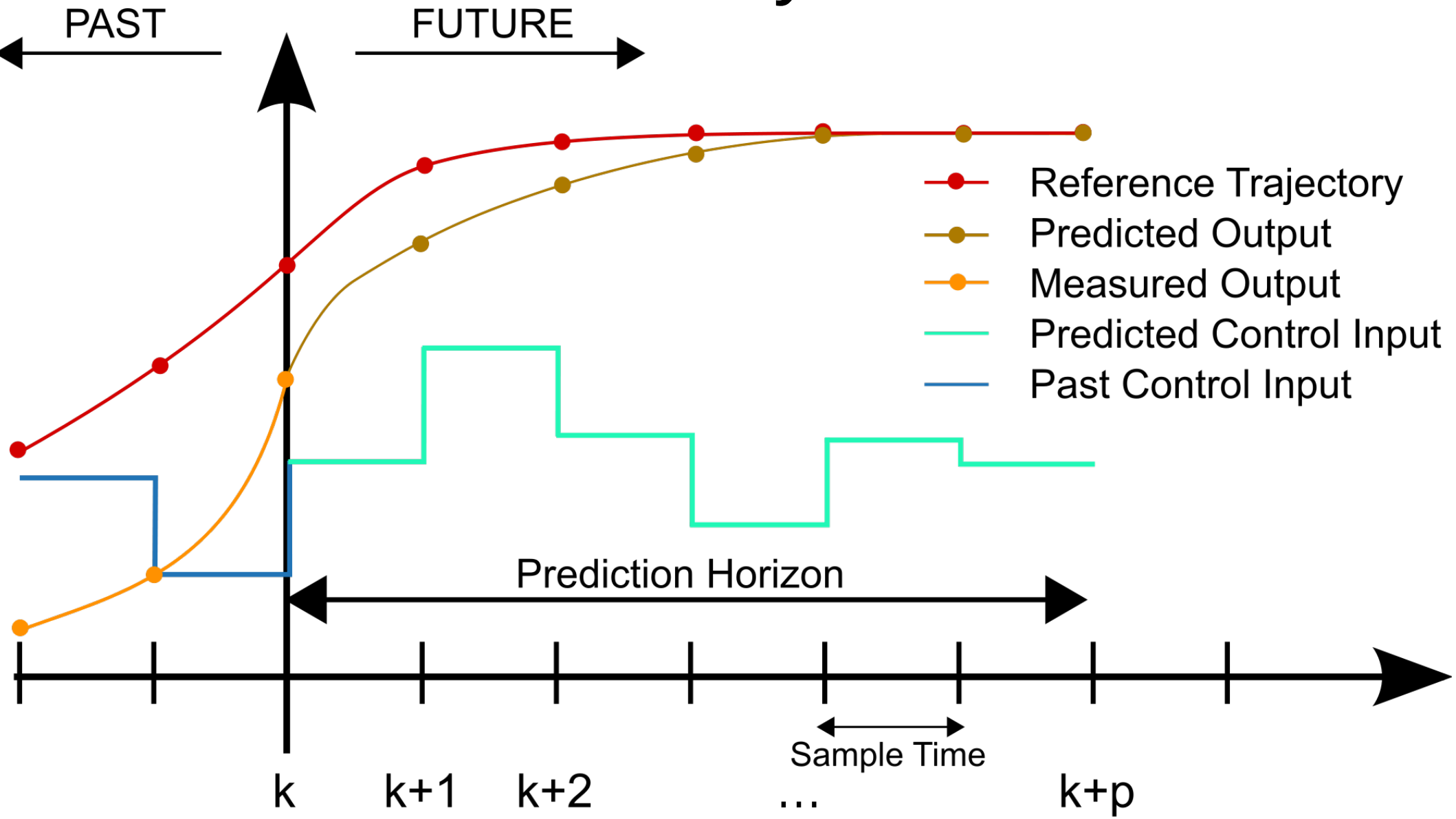
# Control Theory 101 in three Slides: 2. Stability



# Control Theory 101 in three Slides: 3. Dampening

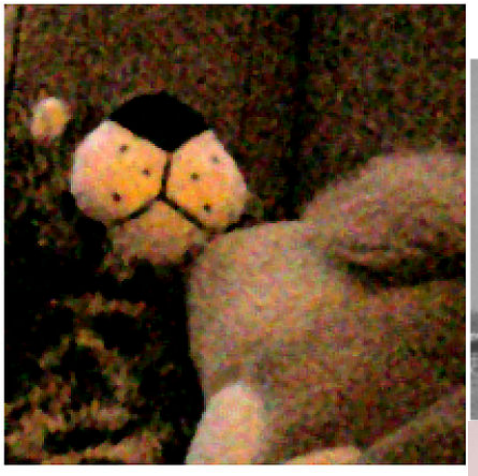


# Control Theory 201:MPC

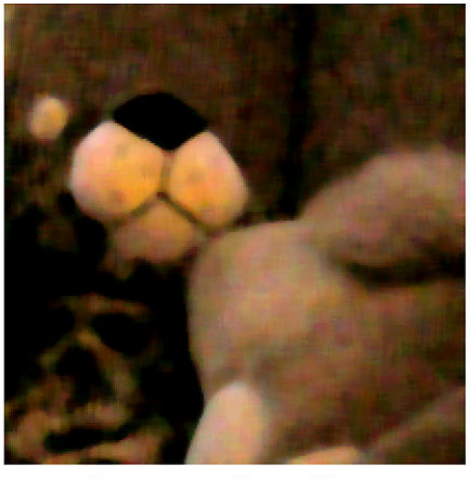


# Computer Vision:Processing

original image



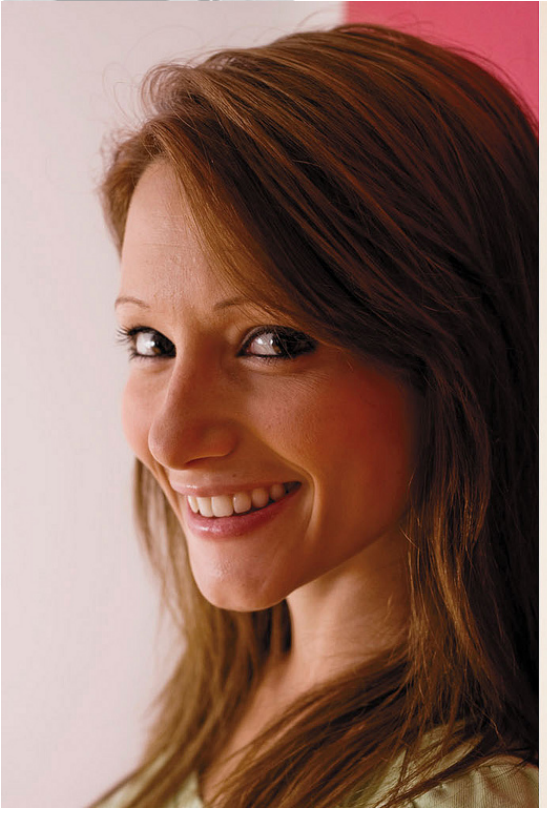
1px median filter



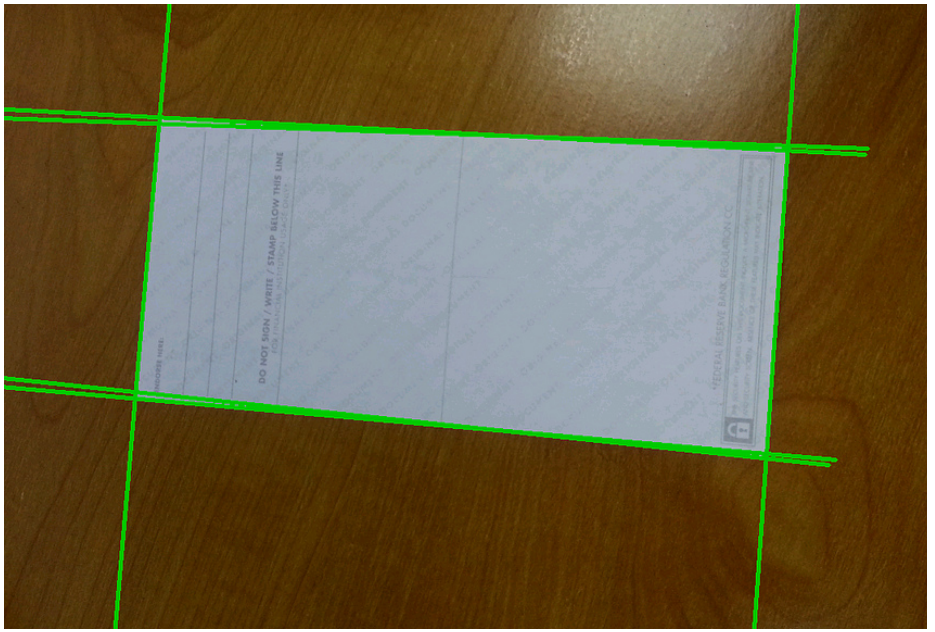
3px median filter

10px median filter



# Computer Vision: Recognition



# Computer Vision:Motion

https://inspirit.github.io/jsfeat/sample_oflow_lk.html
=> Optic flow

https://inspirit.github.io/jsfeat/sample_haar_face.html

=> Face Tracking

# Perception: Computer Audition

No time for details sorry :(

- Speech recognition
  - Identify phonemes
  - Machine transcription
- Source separation
  - Identify different speakers, instruments, noises...
- Musical analysis
  - Detect beat
  - Detect voices, chords..

# Perception: Natural Language Processing (NLP)

- Understanding of unstructured text
- Unstructured vs structured (table vs description of data)
- Sample Tasks:
  - Summarization
  - Machine translation
  - Machine reading (extraction of facts form text: concepts, named entities, events, relationships…)
  - Natural language generation
- Approaches:
  - Grammars and parsing
  - Statistical approaches

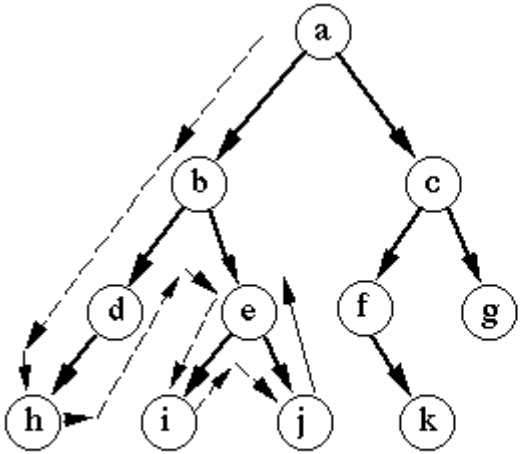
https://demos.explosion.ai/displacy/

http://nlp-compromise.github.io/website/

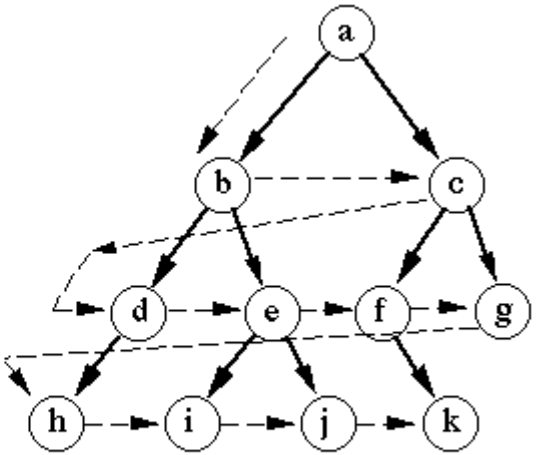
# Planning

- Want to find „good“ solution quickly
- Can‘t use inference because of
  - Combinatorial explosion
  - Fuzzy goals
  - Unknown states (Open world)
- 2 „Dominant“ models: Search & Optimization
  - Can (almost always) be transformed into each other
- Search: given a graph structure, find a path, or a certain node
- Optimization: given a number of hard(must have) and soft(nice to have) objectives, increase a target score

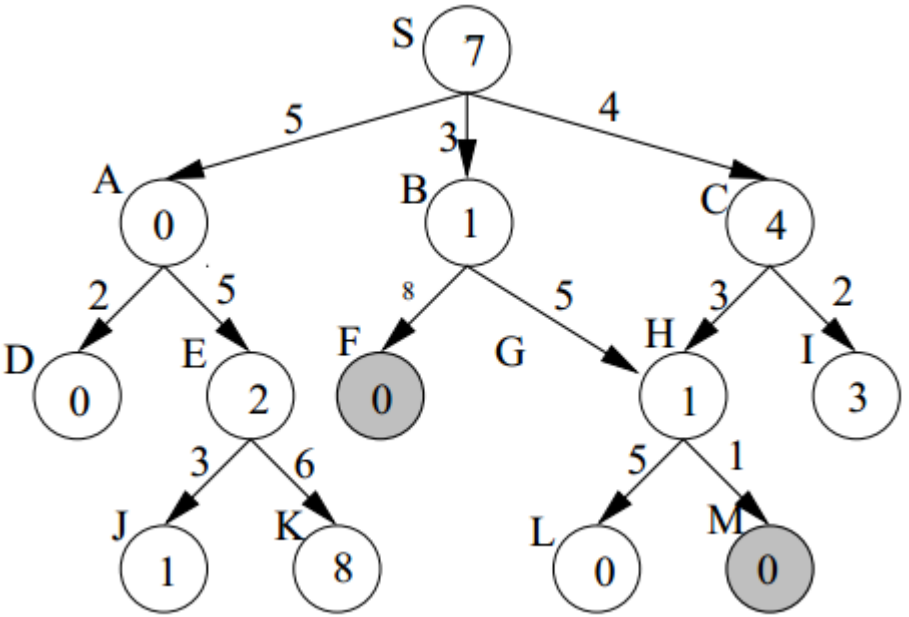
# Planning:Search



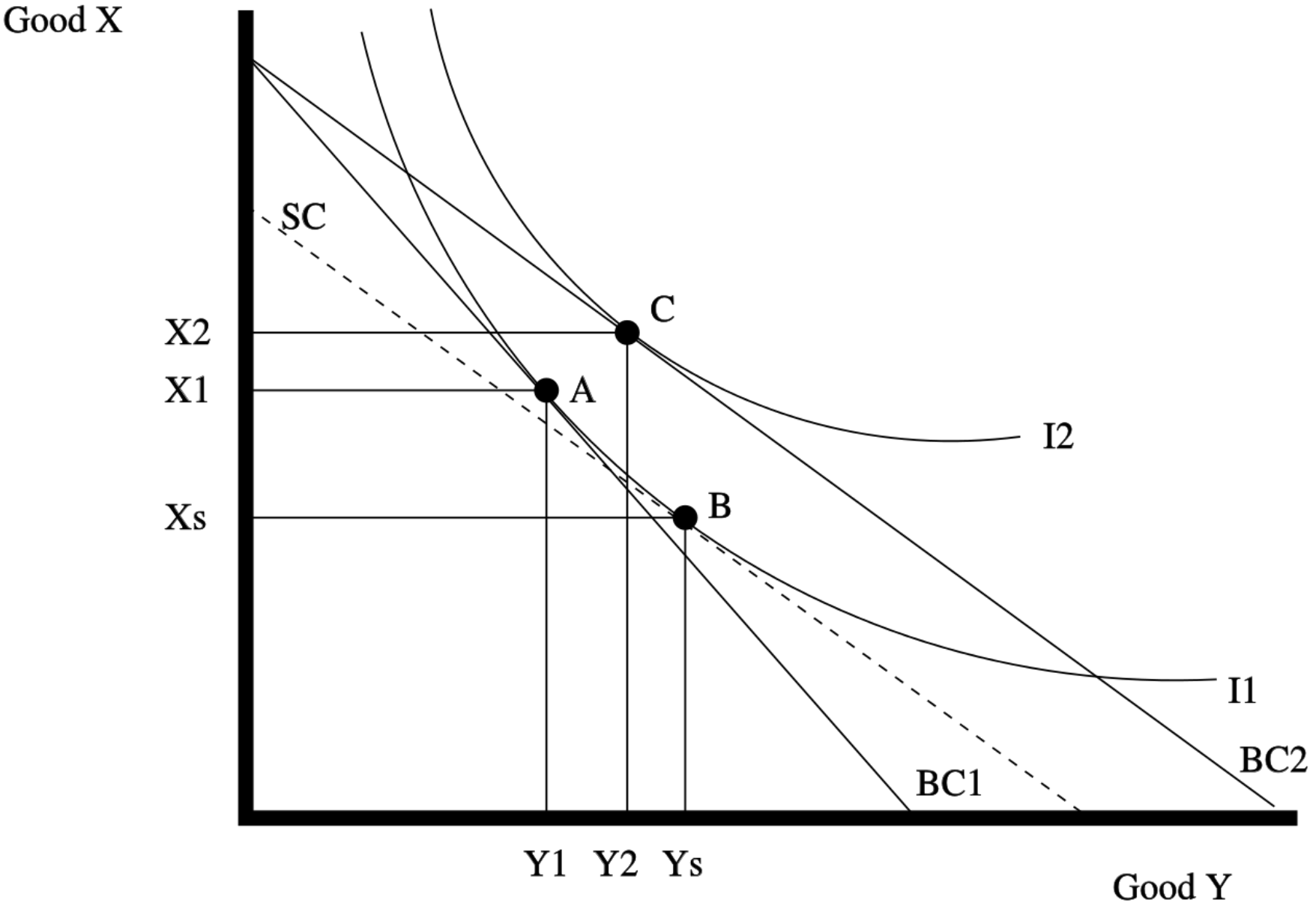

Depth-first search



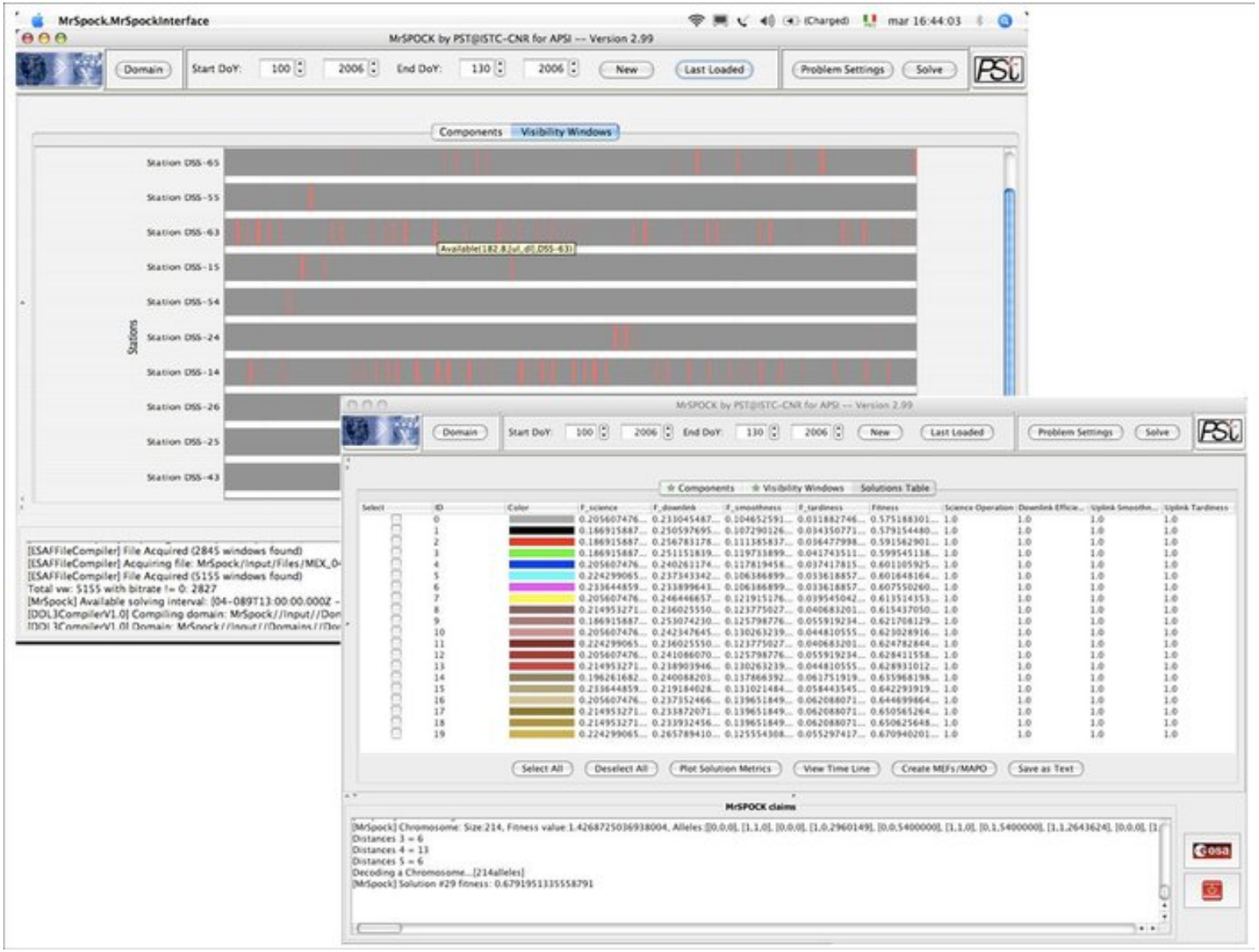

Breadth-first search



# Planning:Optimization



# Planning: Optimization at ESA



# Learning

- DADADA-DAM!

# Learning

- All other fields have „expert knowledge“ encoded
- This is how systems can get that knowledge with explicitly encoding
- Personal opinion: we work the same way
- Spectrum: Supervised <=> Unsupervised

# Learning: „Classical“ Machine Learning

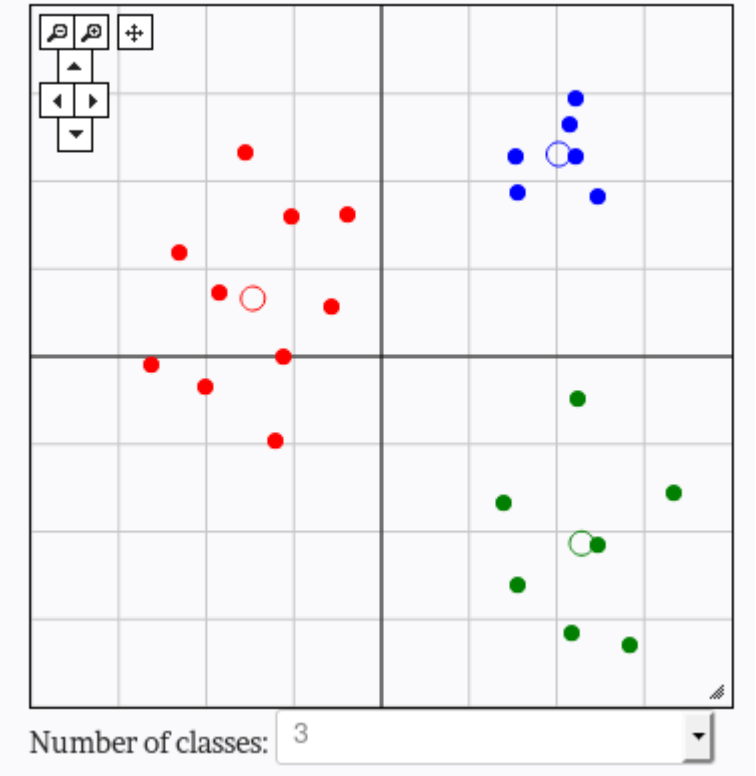
- Clustering:
  - K means
  - Self organizing/Kohonen maps
  - Flame…
- Classification:
  - SVM
  - Random Forests

- Parameter Learning:
  - Bayesian Network
  - Conditional Random Fields
- Dimensionality Reduction
  - **Technically** not learning per se but uses
  - LDA
  - SVD

Expert knowledge:

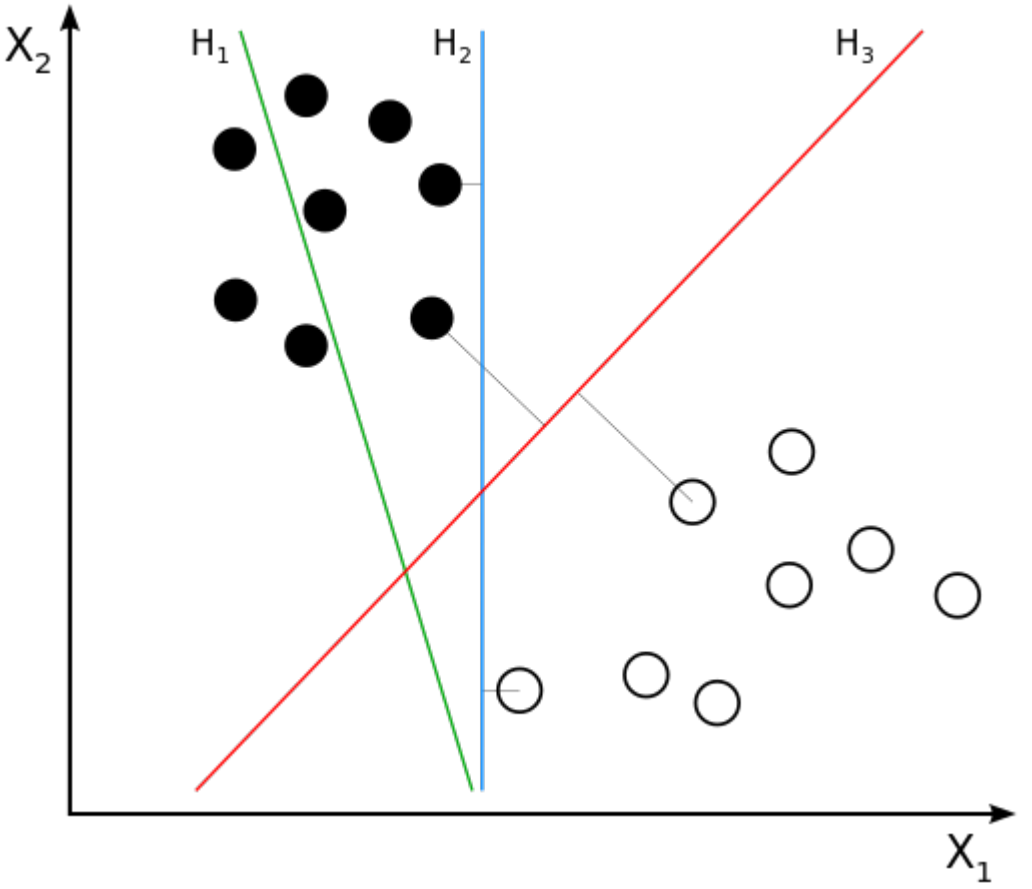
- Preprocessing/labeling data
- Model/method selection very important

# Learning: Clustering



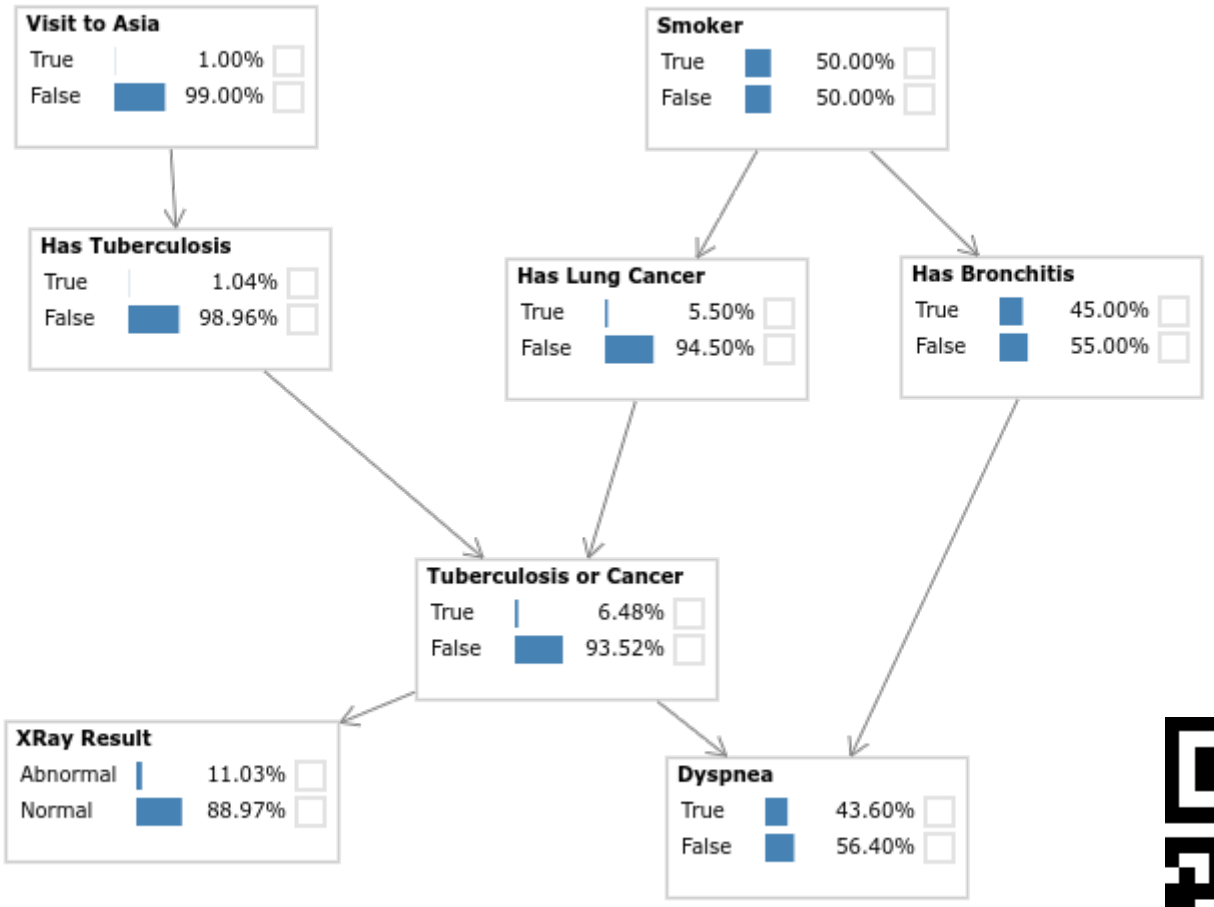

http://util.io/k-means

# Learning: Classification




http://vision.stanford.edu/teaching/cs231n/linear-classify-demo/

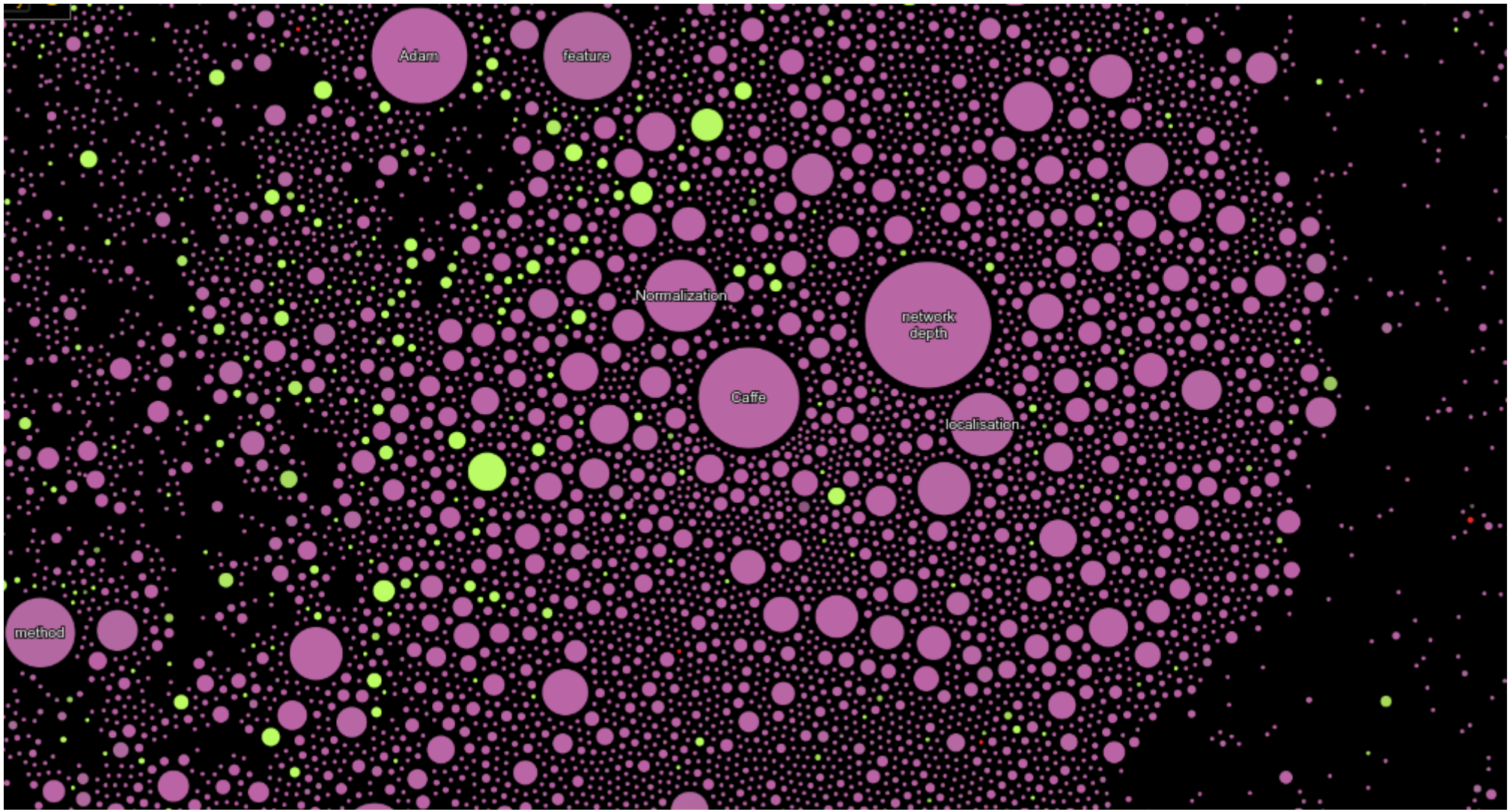
# Learning: Model Learning



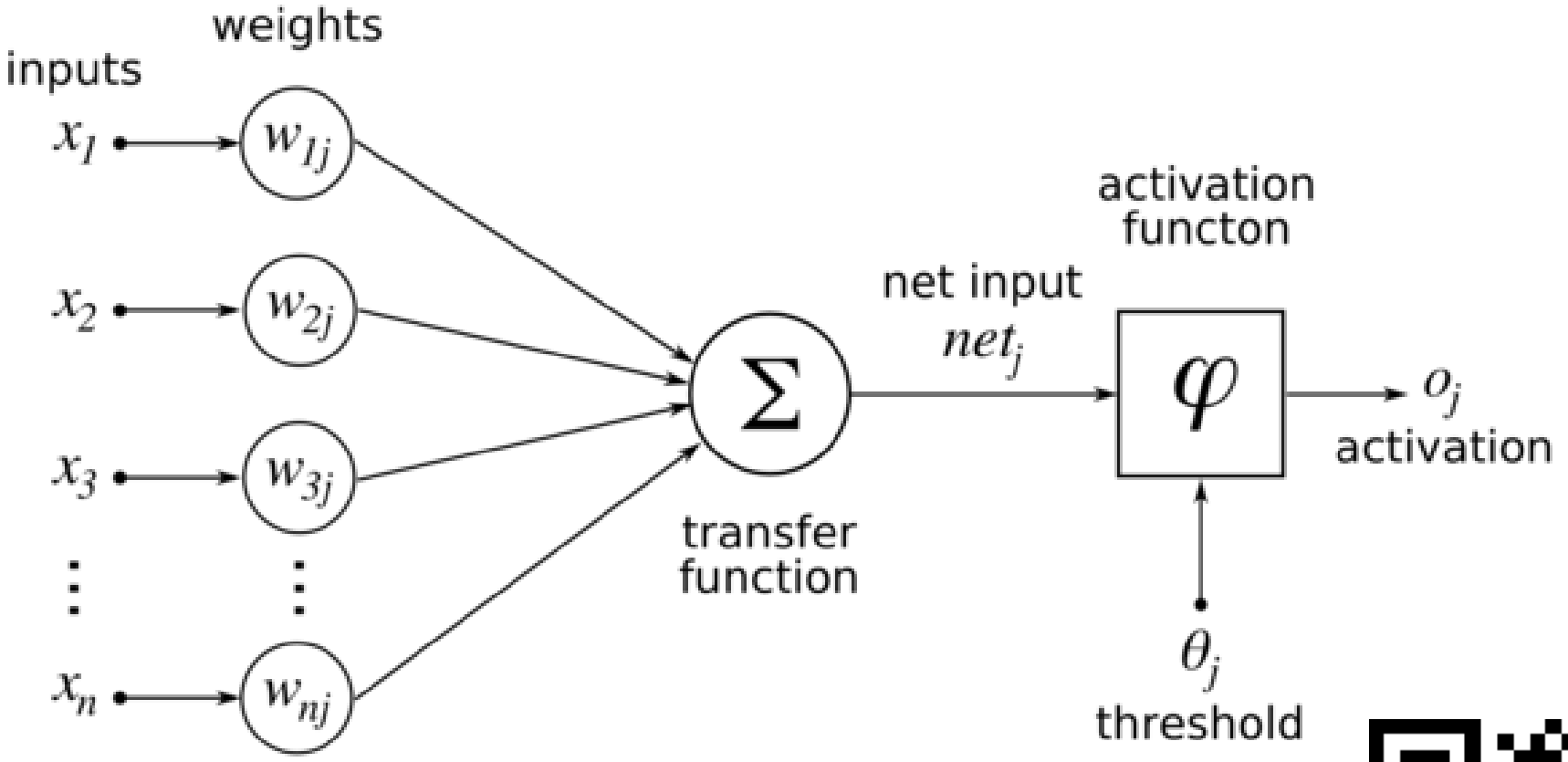

http://www.bayesserver.com/Live.aspx

# Learning: Dimensionality Reduction

View of „Computer Science“ on http://paperscape.org/



# Learning:Neural Networks




http://playground.tensorflow.org

# Learning:Neural Networks

Quick intro to backpropagation

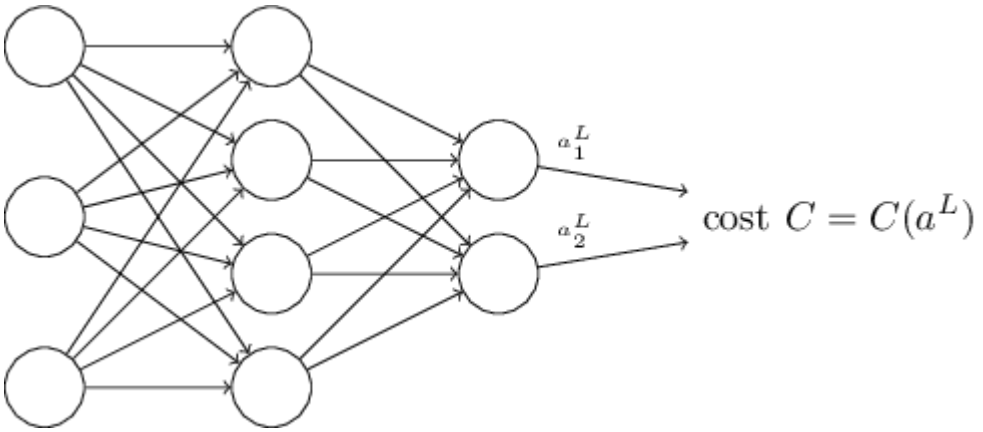
$$\delta^L = \nabla_a C \odot \sigma'(z^L) \tag{BP1}$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \tag{BP2}$$

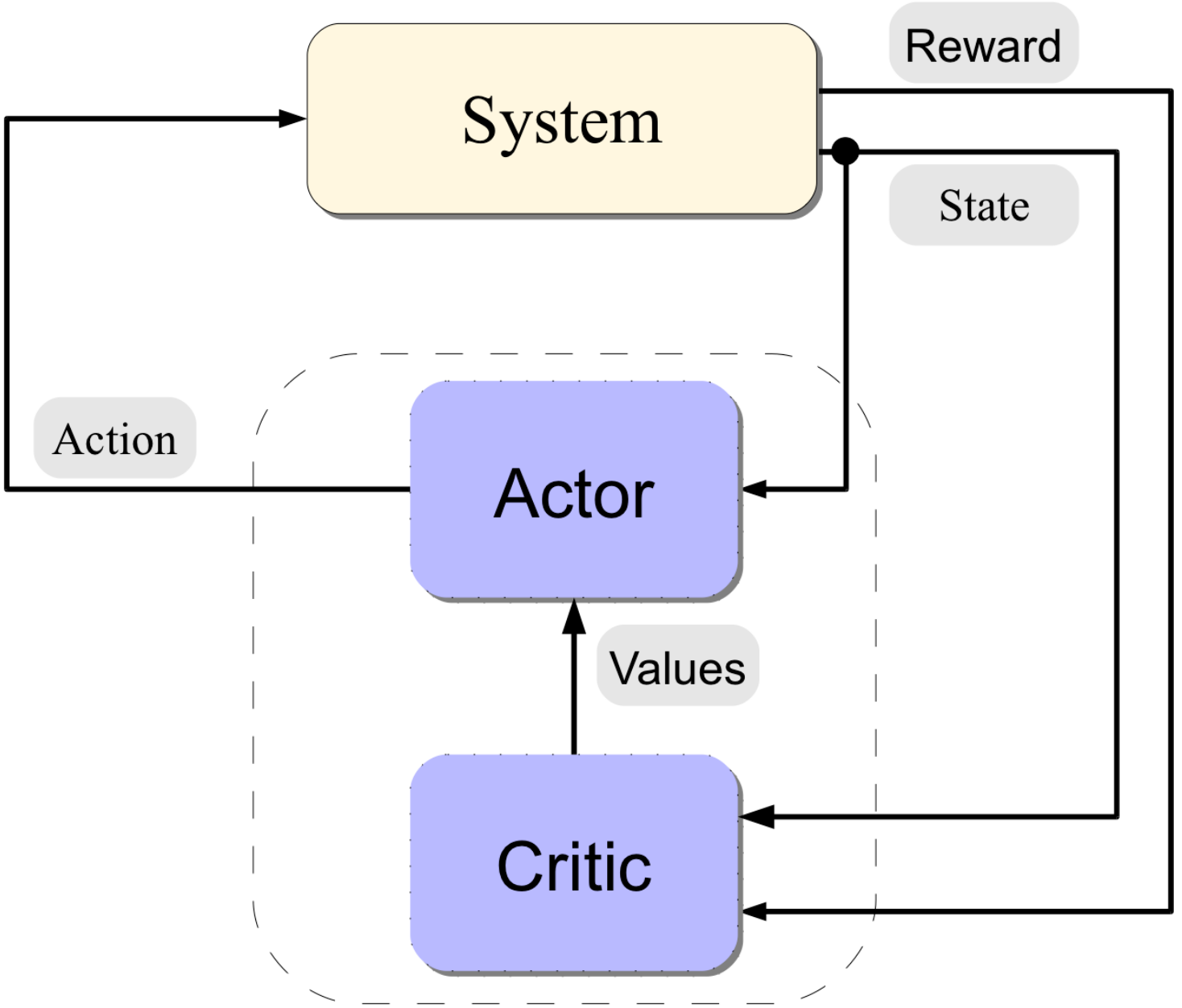
$$\frac{\partial C}{\partial b^l_j} = \delta^l_j \tag{BP3}$$

$$\frac{\partial C}{\partial w^l_{j\,k}} = a^{l-1}_k \delta^l_j \tag{BP4}$$

Easy right? ;-)



# Learning: Reinforcement Learning



# Learning: Reinforcement Learning

- Iterative Reward Function update
- Almost **always** approximated somehow

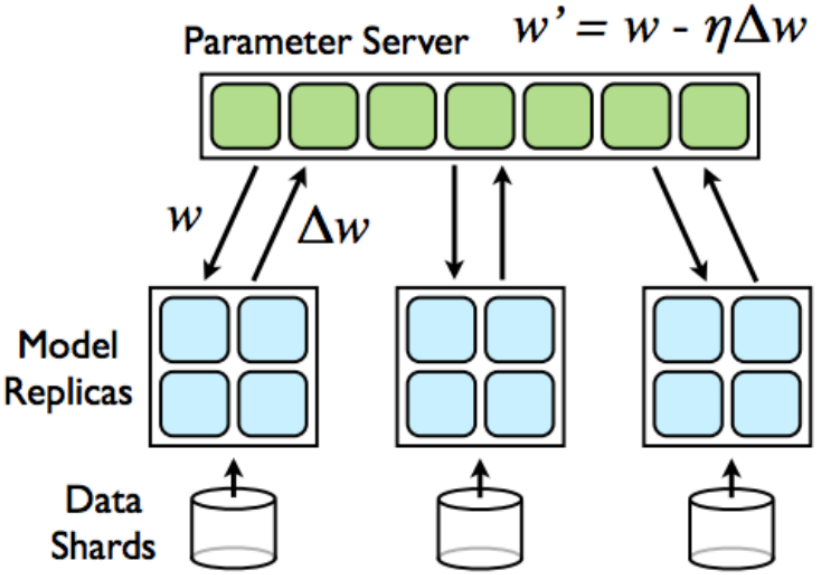
$$Q(s_t, a_t) \leftarrow \underbrace{Q(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha}_{\text{learning rate}} \cdot \left( \overbrace{\underbrace{r_{t+1}}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1}, a)}_{\text{estimate of optimal future value}}}^{\text{learned value}} - \underbrace{Q(s_t, a_t)}_{\text{old value}} \right)$$

- Keyword: Deep Q Learning

http://cs.stanford.edu/people/karpathy/reinforcejs/index.html

# Learning: Where is the hard part?

- Basic deal always:
  - Error minimization through backpropagation
  - Need to:
    - Avoid over-fitting
    - Build surrounding infrastructure
    - Model problems in very indirect ways
- Need LOADS of data (or trials in RL)
- Need processing power (and/or time
- Model might be large (RAM!)
- Probably needs to be distributed
  - Distributed gradient
  - Data and code parallelism



# Hungry Machine learning

- Learning is eating NLP, Knowledge, Computer Vision,Planning...with conditions
  - Needs loads of processing power and good data
  - The more complex the learning task, the better the data needs to be
  - Domain knowledge required for best models
- If any condition poses restrictions, „classic“ AI wins
  - PID Controllers => manual tuning, but established and optimized
  - Template NLP => faster and better than training...for now

# Thanks for attending!

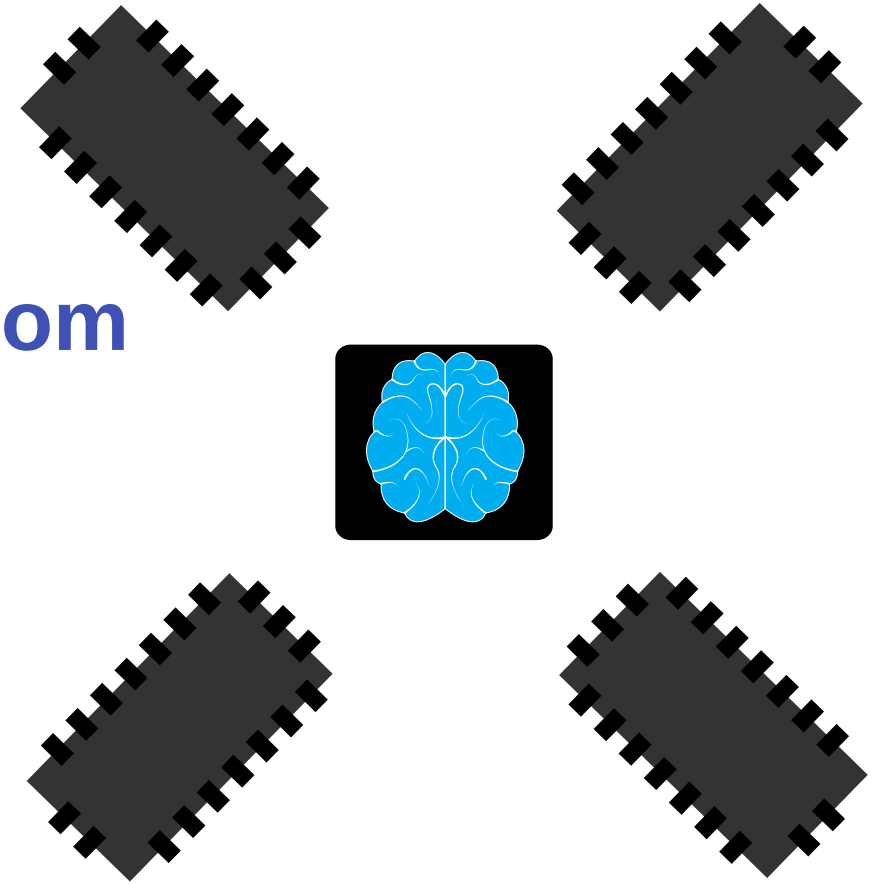
Feel free to contact

me at

**inquiries@bayesmill.com**

with any questions or

if you need consulting



# Sources

- „Artificial Intelligence: A modern Approach“(Pears,3rd Edition,Russel,Norvig et.al)
- „An Introduction to Machine Learning“(Cambridge University Press,Smola and Vishwanathan)
- Computer Vision:Algorithms and Applications (Richard Szeliski, http://szeliski.org/Book)