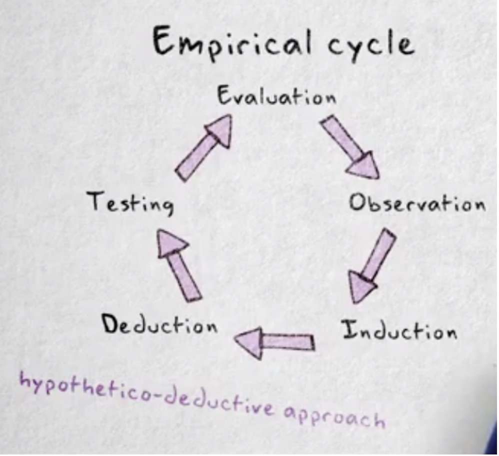
Research in Computer Science

Javier Larrosa

UPC Barcelona Tech

.

Hypothetic-Deductive approach



Pure vs Applied Science

Pure Science

When the goal is to **increase** the knowledge. Scientific contributions are a new **fact**, a new **law**

Good Research: It has to be significant and potentially useful

However, see "The usefulness of useless knowledge" (Maxwell and Hertz work on magnetism made Marconi's development of the radio inevitable)

Pure vs Applied Science

Applied Science

When the goal is to **apply** existing knowledge to interfere in the world. Scientific contributions are technological **inventions**, new **methodologies**, ... (e.g. a new drug, a new artifact, a new programming language,...)

Research in Computer Science

Research in Computer Science

It has ingredients from Engineering, Mathematics, Psychology,... which makes its research methodology very diverse

Types of Research

Rough classification (many research works combine many of them):

- **Formal:** Proof facts
- **Build:** Building an artifact
- **Process:** Understand the process followed by users to accomplish tasks in CS
- **Model:** Defining abstract models for real systems

In all of them there may be an **experimental** part.

Formal (similar to Mathematics research)

Proof properties about paradigms, algorithms, systems or architectures.

Concerned with:

- **capturing the essence** of the problem under study
- Identify fundamental **possibilities** and **limitations**
- Requires **formal description** (formal languages, logic)

Examples:

- **Formal specifications** of software components (for automatic verification)
- analyze time and space **complexities**
- **correctness**, quality of solution of algorithms
- **unification** between disparate areas
- **Understanding quantum computing**

Build (similar to Engineering)

Build an artifact (algorithm, software, architecture)

Concerned with:

- **Proof of concept** (demonstrate that it is possible)
- **Novelty** (the artifact must be new or it must include new features not demonstrated before).
- **Good design** (modular approach, re-use components), documentation
- **Good Evaluation** (standard benchmarks, report architecture-independent measures)

Example:

- Artificial *Go* player
- **Killer application for quantum computing**

Process (similar to Psicology)

Understand the **processes** used to accomplish **tasks** in Computer Science

Concerned with:

- Human (irrational and unpredictable) behavior
- Needs to deal with unprecise concepts
- Maximum scientific rigor
  - **Research Questions** (empiricably testable, falsifiable)
  - **Experiment Design** (replicable, survey design)
  - **Results Interpretation** (logically consistent)
- Good (i.e, precise) **research setup**:
  - define concepts, measurement instruments, procedures, sample,...

Process: Example

- **Software Engineering:** Design and construction of software systems by team workers
  - improve communication, put more discipline, raise quality
  - e.g. SCRUM
- **Man-Machine Interaction:** Design more friendly interfaces
  - improve communication, comfort,...
  - e.g. Support for Color-blind users
- **Tools for quantum programming**

Model

Defining an **abstract model** for a real system

Concerned with:

- less complex than the real system, while preserving fundamental aspects
- Simulation: experiments based on a model
- Model Checking: verify properties on the model

Example:

- Balancing Bike Sharing Systems

Example 1 (boolean SATisfiability)

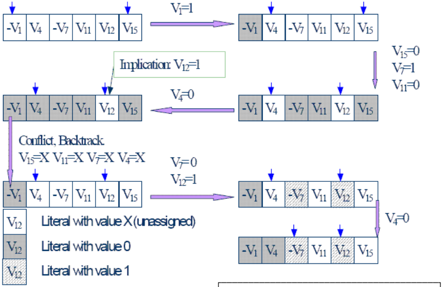
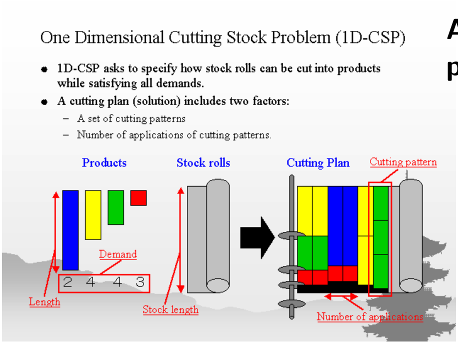

$-V_1$ V_4 $-V_7$ V_{11} V_{12} V_{15}
$V_1=1$
Implication: $V_{12}=1$
$V_{15}=0$
$V_7=1$
$V_{11}=0$
$V_4=0$
Conflict, Backtrack.
$V_{15}=X$ $V_{11}=X$ $V_7=X$ $V_4=X$
$V_7=0$
$V_{12}=1$
$V_4=0$
V_{12} Literal with value X (unassigned)
V_{12} Literal with value 0
V_{12} Literal with value 1

Figure 1: BCP using two watched literals

Invent a new algorithm or data structure that is better than previous alternatives

- **Better:** always or for a particular problem/domain (in that case show relevance of the problem)
- **Better:** theoretically, empirically (in that case well designed experiments), simpler

Example 2 (cutting stock)



One Dimensional Cutting Stock Problem (1D-CSP)

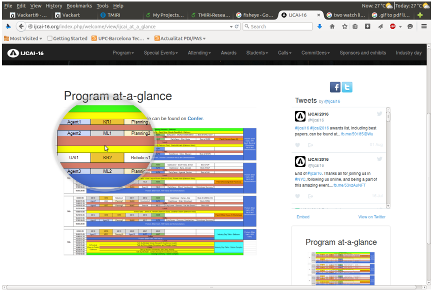
- 1D-CSP asks to specify how stock rolls can be cut into products while satisfying all demands.
- A cutting plan (solution) includes two factors:
  - A set of cutting patterns
  - Number of applications of cutting patterns.

Products
Stock rolls
Cutting Plan
Cutting pattern
Demand
2 4 4 3
Length
Stock length
Number of applications

Apply known technology to a problem for the first time

- show relevance of the problem
- the application must be non-trivial
- the application must yield better results (theoretically, empirically, simpler)

Example 3 (file navigation in software development)


File Edit View History Bookmarks Tools Help
Now: 27 °C Today: 27 °C
Vackart® · … Vackart TMIRI My Projects… TMIRI-Resea… fisheye - Go… IJCAI-16 two watch li… .gif to pdf li…
ijcai-16.org/index.php/welcome/view/ijcai_at_a_glance
Search
Most Visited Getting Started UPC-Barcelona Tec… Actualitat PDI/PAS
IJCAI-16
Program Special Events Attending Awards Students Calls Committees Sponsors and exhibits Industry day
Agent1 KR1 Planning
Agent2 ML1 Planning2
UAI1 KR2 Robotics1
Agent3 ML2 Planni

Program at-a-glance

…e can be found on Confer.

Tweets by @ijcai16

IJCAI 2016 @ijcai16
#ijcai16 #ijcai2016 awards list, including best papers, can be found at... fb.me/59185IBWu
01 Aug

IJCAI 2016 @ijcai16
End of #ijcai16. Thanks all for joining us in #NYC, following us online, and being a part of this amazing event.... fb.me/53vzAuNFT
16 Jul

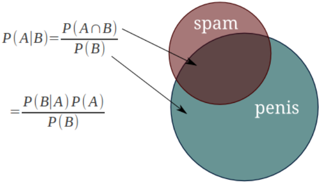
Embed View on Twitter

Program at-a-glance

Propose a new interface for a user's task (e.g. fisheye)

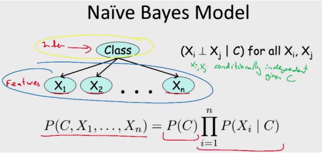
- show relevance of the task
- show novelty
- show advantages (friendlyness, quality, efficiency)
- proof-of-concept, statistics (based on surveys,...)

Example 4 (spam filters)



$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

$$=\frac{P(B|A)P(A)}{P(B)}$$



Naïve Bayes Model

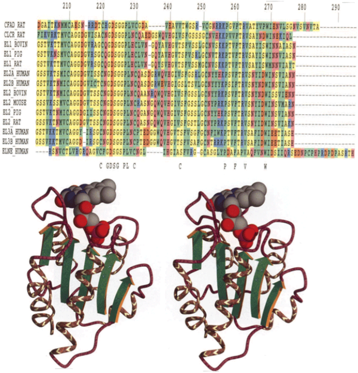
$(X_i \perp X_j \mid C)$ for all X_i, X_j

$$\underline{P(C,X_1,\ldots,X_n)} = \underbrace{P(C)}\underbrace{\prod_{i=1}^{n} P(X_i \mid C)}$$

Propose a new model for an interesting problem

- show relevance of the task
- show novelty
- show advantages (computationally simpler, more accurate,...)

Example 5 (protein alignment)



Propose a new model for an interesting problem

- show relevance of the task
- show novelty
- show advantages (computationally simpler, more accurate,...)