Learning Classifier Systems: Recent Trends

Albert Orriols Puig
aorriols@salle.url.edu

Grup de Recerca en Sistemes Intel·ligents
Enginyeria i Arquitectura La Salle
Universitat Ramon Llull

Outline

1. What was an LCS meant to be?

1.1. Michigan envision of LCSs

1.2. Parallel developments: Pitts-style LCS

2. What are LCSs nowadays

2.1. General schema

2.2. What kind of Machine Learning problems can we solve?

2.3. What we are good at and bad at

2.4. Current applications

3. Little discussion about LCSs future

3.1. Are we happy with our LCSs?

3.2. Future applications

What was an LCS meant to be?

- **Holland's envision: Cognitive Systems**
  - *Create true artificial intelligence itself*
  - *True intelligence requires adaptive behavior in the face of changing circumstances (Holland & Reitman, 1978)*
  - *Hollands vision going back to late 50s and early 60s of roving bands of computer programs.*

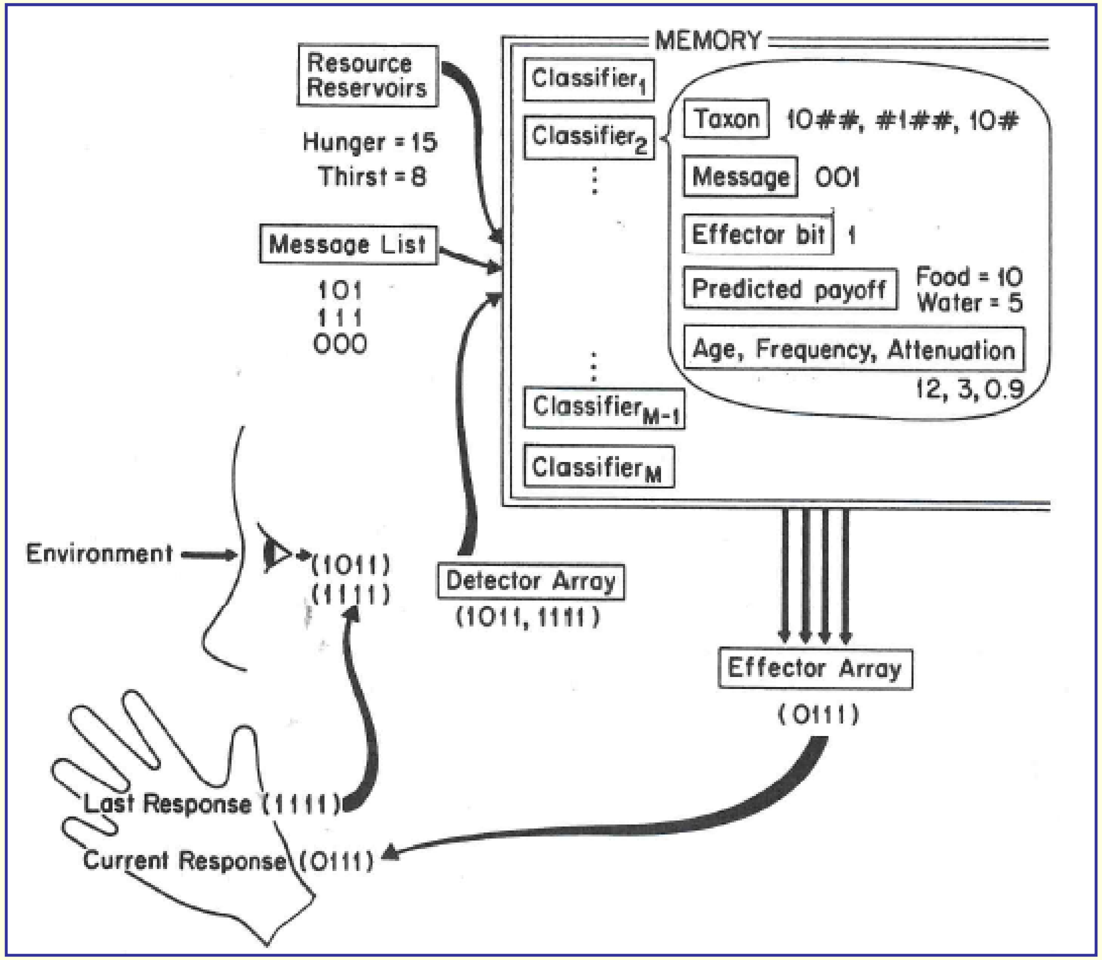
> Holland's notion of genetic search as program searching (1962)
>
> The free generation procedure. . . Requires the generators (and combinations of generators) to "shift" and "connect" at random in the computer…two or more generators occupying adjacent modules ("in contact") may become connected. Such connected sets of generators are to shift as a unit.

  - *From stimulus-response to internal states and modifiable detectors and effectors*

What was an LCS meant to be?

- CS-1 (Holland & Reitman, 1978)

- **Post-production system**
- **General memory containing classifiers**
- **Process:**
  - Code the situation and find in memory the actions that are appropriate to both CS-1 goal and situation
  - Store in memory the consequences of these actions (learning)
  - Generate new good productions (classifiers) to endure.



Population of classifiers → Current system knowledge

Performance component → short term behavior of the system

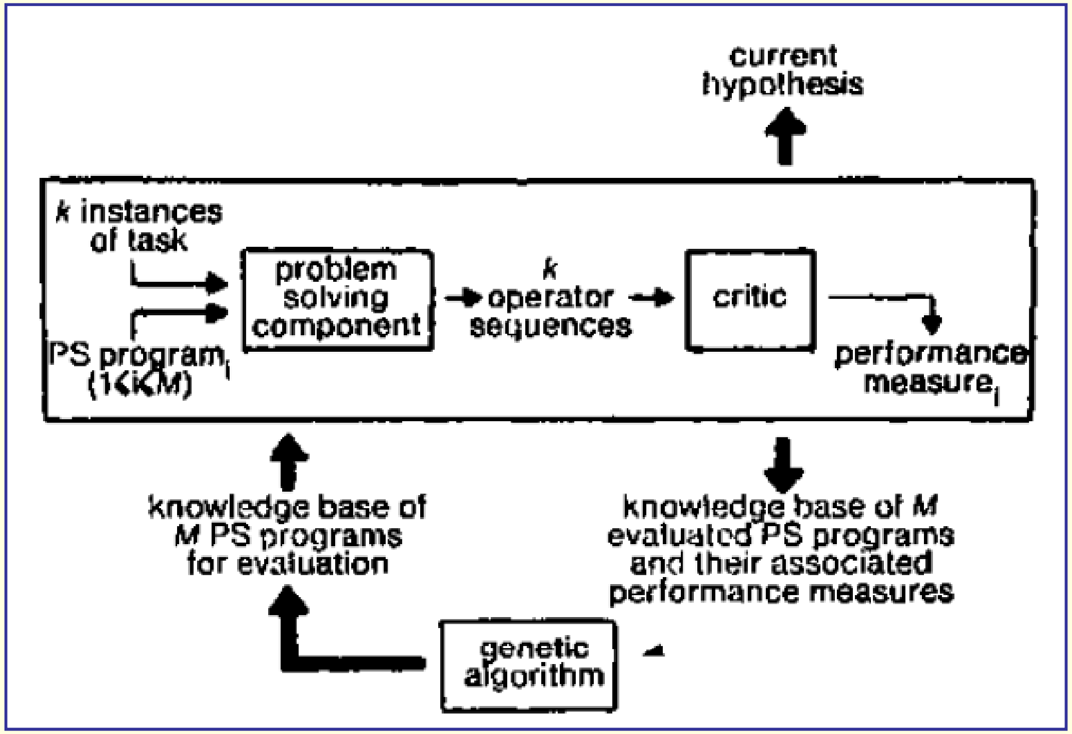
Rule discovery component → get new promising rules

Meanwhile, in Pittsburgh University…

- **Smith's interpretation of Holland's GA envision**

Smith's notion of learning as adaptive search (1980, 1983)

LS-1: "*Learns a set of heuristics, represented as production system programs, to govern the application of a set of operators in performing a particular task*"



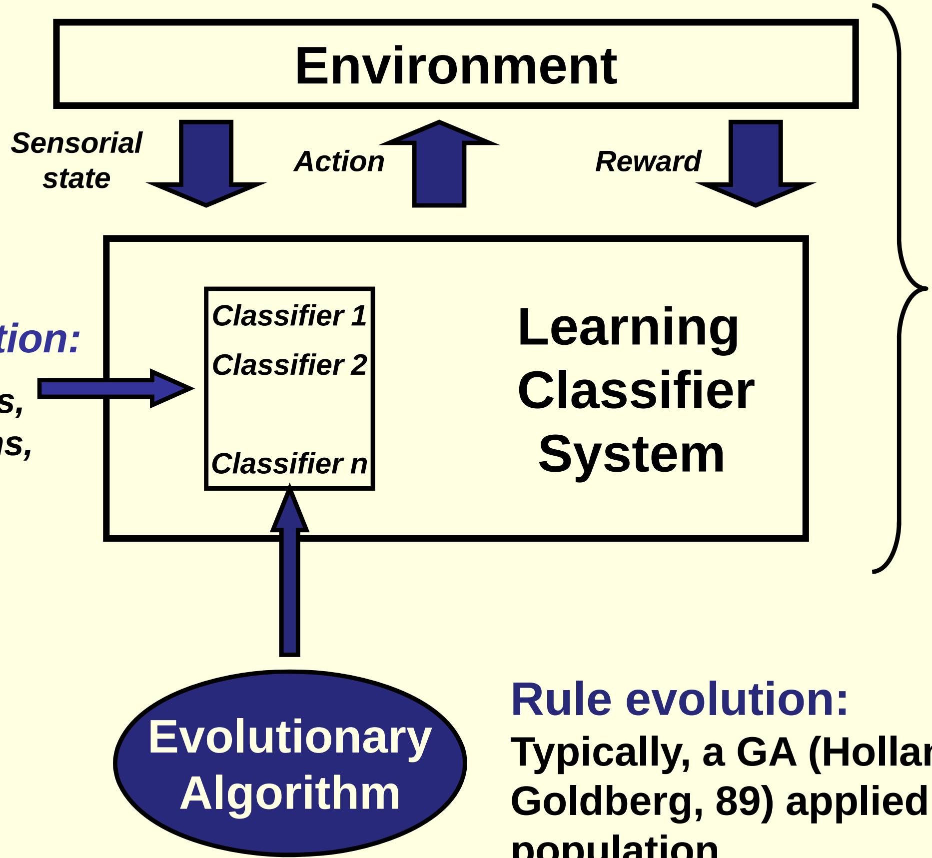
Great success! LS-1 took Waterman's poker player to the cleaners (not bluffing)

Outline

1. What was an LCS meant to be?
   1.1. Michigan envision of LCSs
   1.2. Parallel developments: Pitts-style LCS
2. What are LCSs nowadays
   2.1. General schema
   2.2. What kind of Machine Learning problems can we solve?
   2.3. What we are good at and bad at
   2.4. Current applications
3. Little discussion about LCSs future
   3.1. Are we happy with our LCSs?
   3.2. Future applications

General schema of LCSs

- Focused on Michigan-style LCSs: *Incremental Learning*



Any Representation:

production rules, genetic programs, perceptrons, SVMs

Online rule evaluator:

- **XCS: Q-Learning** (Sutton & Barto, 1998)
  Uses *Widrow-Hoff delta rule*

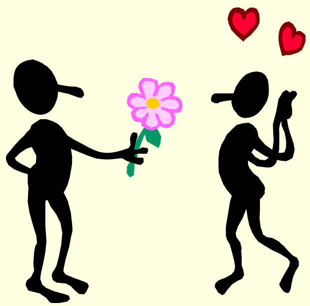
Rule evolution:

Typically, a GA (Holland, 75; Goldberg, 89) applied on the population.

What kind of ML problems can we solve?

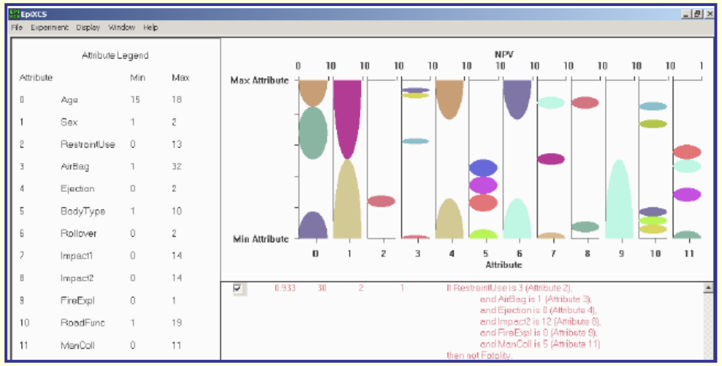
- **Supervised learning**
  - Data Classification **(Bernadó & Garrell, 2003; Barcadit & Butz, 2004; Orriols & Bernadó., 2008)**
  - Function Approximation **(Wilson, 2002; Butz et al. 2007)**
- **Reinforcement learning**
  - Simple MDP problems **(Lanzi et al. 2006, Lanzi et. al 2007, Butz et al. 2005)**
- **Unsupervised learning**
  - Clustering algorithms **(Tamee et al. 2006, Tamee et al. 2007)**
- **Data stream processing**
  - Adapting to concept drift **(Abbass et al. 2004)**
- **Anticipatory LCS (ACS)**
  - Model the effect of our actions **(Butz, 2000)**

Therefore, LCSs represent a really broad architecture!

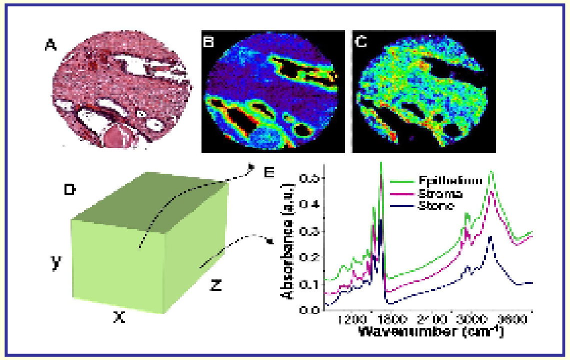


Current real-life applications

- **Data mining**
  - Most important application domain of LCSs
  - John H. Holmes
    - Epidemiologic study by means of LCSs
    - Hidden relationships among variables discovered by LCSs



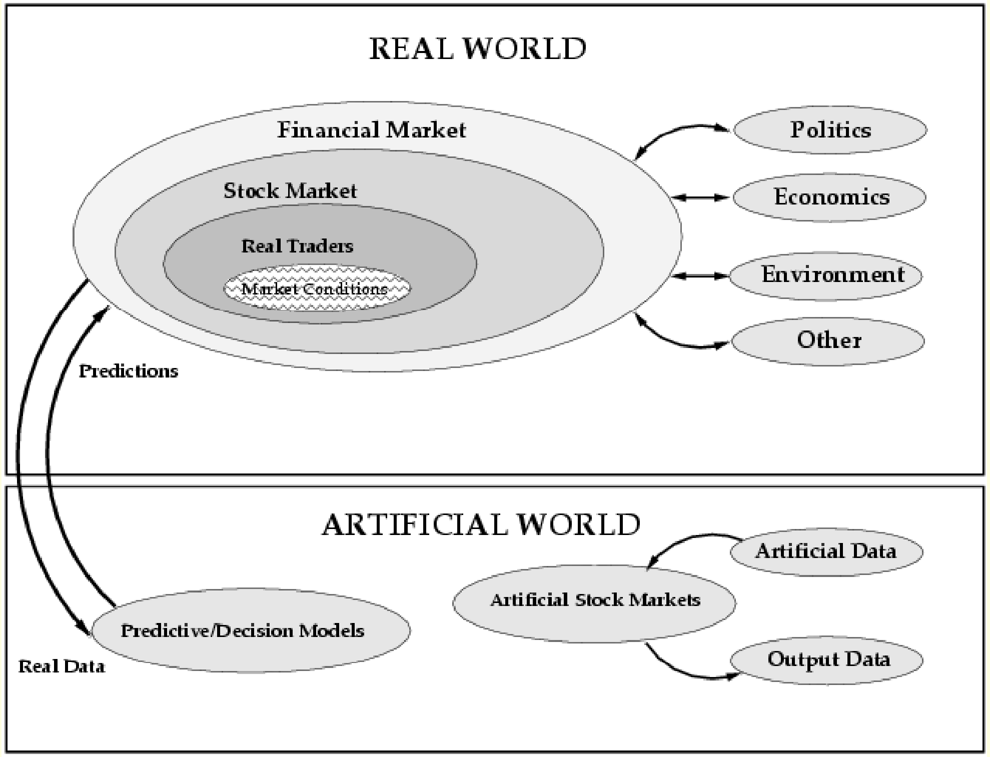
  - Xavier Llorà et al.
    - *Better than Human Capability in Diagnosing Prostate Cancer Using Infrared Spectroscopic imaging*



  - Many other applications and miners:
    - GALE, GAX, GAssist, UCS, MIPS …
  - See: *Bull, Bernadó-Mansilla & Holmes (eds) Learning Classifier Systems in Data mining. Springer (2008)*

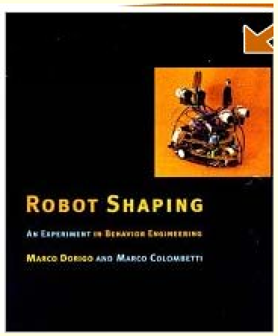
Current real-life applications

- **Modeling market traders**
  - LETS project: Evolving artificial traders for successful market trading ***(Sonia Schulenburg et al, 2007)***
  - Evolutionary economics:
    - **Create trend followers and value investors**
    - **Let them interact**
    - **Evolve a population of strategies**



Current real-life applications

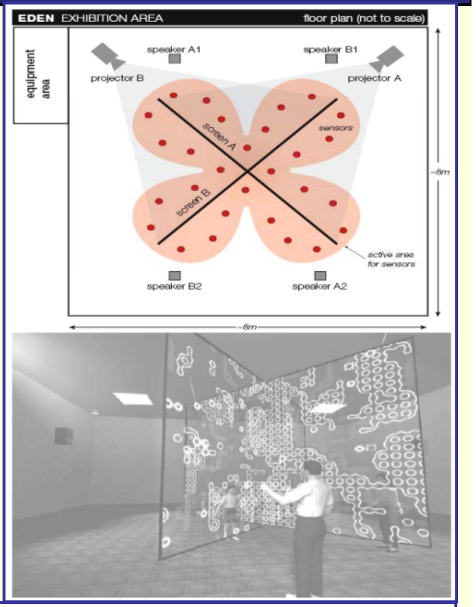
- **Autonomous Robotics**
  - Robot shaping: Early efforts of Marco Dorigo and Marco Colombetti (1997)
    - Small mobile robots equipped with sensors and motors
    - Robots connected in real time by various sorts of modem cable
    - Robots controlled by LCS, ICS, running on desktop computers
    - Constant stream of positive/negative rewards (bucket brigade)
  - Tasks solved:
    - Following lights
    - Gather food and run home
    - Hunt around for a light hidden behind and obstacle
  - Impressive results, high performance
  - Recent applications to model robotic problems performed in the University of West England

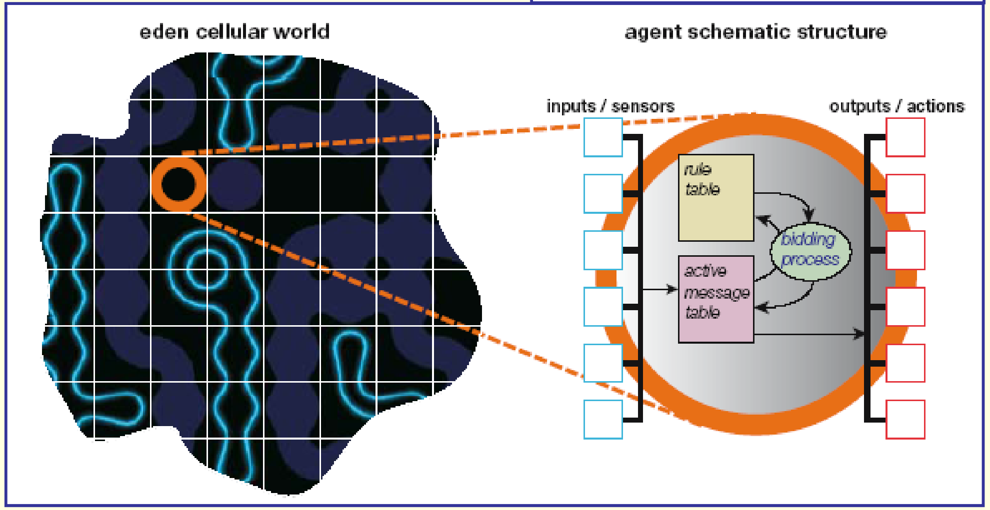


Current real-life applications

Modeling Artificial Ecosystems

- Eden: Artificial Life environment ***(Jon McCromak, 2004)***
- Model of an environment where evolvable rule-based classifier systems drive agent behavior.
- Autonomous LCSs or agents compete for limited resources.
- Agents can:
  - Make and listen to sounds
  - Forage for food
  - Encounter predators
  - Mate with each other
- Goal: Maintain the audience in tension without fitness needing the audience explicitly perform fitness selection



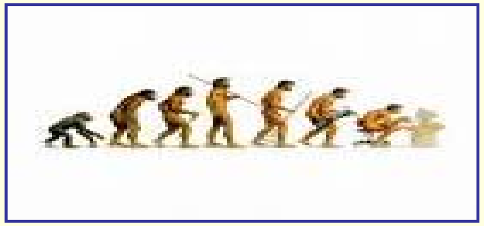


What are we good at?

- General evaluation/reinforcement architecture. Suitable for supervised/non-supervised/semi-supervised/reinforcement learning
- Representation-blind
- ***On-line learning***:
  - Knowledge evolution through data streams processing
  - Really important for current industrial processes
- Population-based search: parallel search toward the objective
- Able to solve complex problems
- Able to optimize several objectives (e.g., Pitt-style LCSs)

What are we bad at?

- Several configuration parameters to tune
  - Tuning is problem-specific
  - Difficult for non-expert users. We want a plug and play LCS
- Still not full understanding of the processes that guide LCS search
  - Facet-wise analysis has provided good insights
  - LCSs community towards principle design
- Computational cost
- Many knowledge gathered during evolution is overlooked
  - Processing generation by generation
  - Evolution could explain us a lot



Outline

1. What was an LCS meant to be?

1.1. Michigan envision of LCSs

1.2. Parallel developments: Pitts-style LCS

2. What are LCSs nowadays

2.1. General schema

2.2. What kind of Machine Learning problems can we solve?

2.3. What we are good at and bad at

2.4. Current applications

3. Little discussion about LCSs future

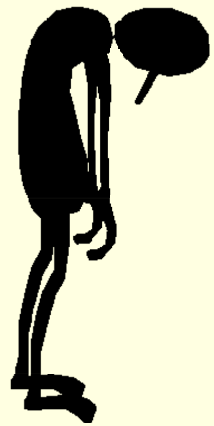
3.1. Are we happy with our LCSs?

3.2. Future applications

Little discussion about LCSs future

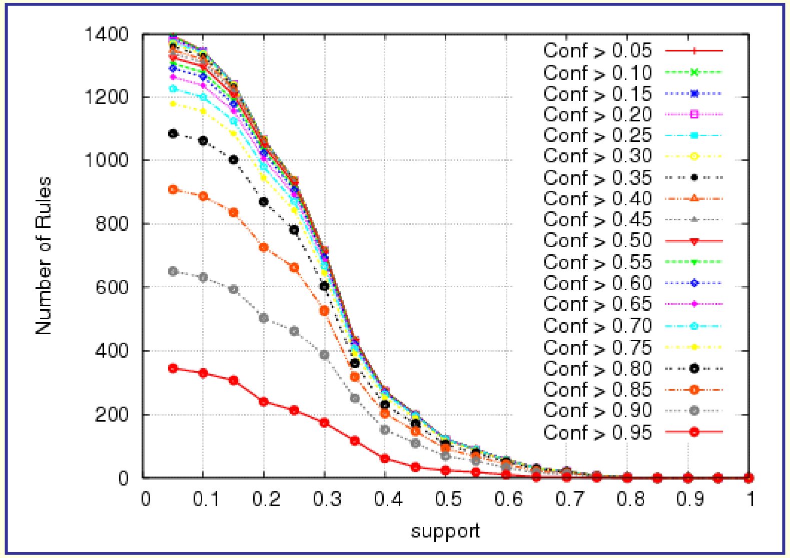
- **Are we happy with our LCSs?**
  - They can solve increasingly complex, relevant real-life problems
  - We have a better understanding of their underlying processes

- **However…**
  - In general, solving problems which were already solved
  - Not accepted in ML community. *Not theoreticians, but innovators*
  - Have we met Holland's original idea?
    - **Some criticisms**
      - LCSs do not seem very smart
      - Not autonomous in a serious sense
      - We are discussing technical issues, not fundamental ones



Future Applications towards meeting ML community interests

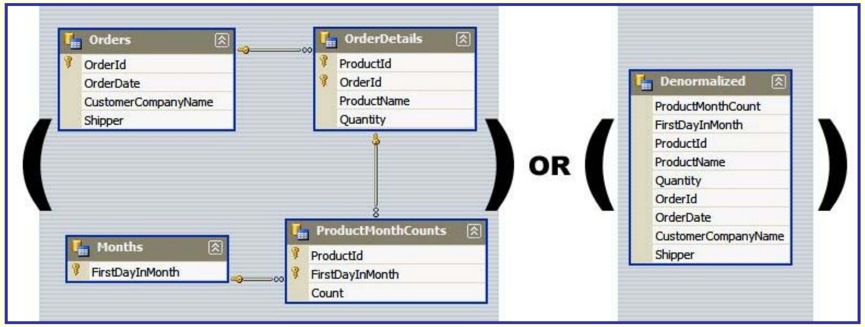
- **Exploiting on-line learning architecture**
  - On-line mining of fuzzy-rules
  - Adapt the rule-based to *association drift*
  - Many industrial apps. where data is continuously generated
  - First early efforts in (Orriols, Casillas & Bernadó, 2008)



Future Applications towards meeting ML community interests

cost	quality	avabity	quantity	respect	prestige	experie	popula
1	3	4	6	7	2	4	5
2	3	4	3	4	6	7	6
4	5	6	7	7	2	3	4
3	4	5	6	7	3	5	4
2	5	5	5	6	2	4	5
3	4	6	7	7	4	3	5
2	3	6	4	5	4	4	4
1	3	4	5	6	3	3	4
3	3	5	6	6	4	4	3
4	4	5	6	7	4	3	4
2	3	6	7	5	4	4	4
2	3	5	7	6	3	3	3

- **Processing new types of data to face new types of problems**
  - Up to now, flat data sets
  - New challenging problems may have more complex data structures:
    - Hierarchical data
    - Multi-instance learning
    - Imprecise data



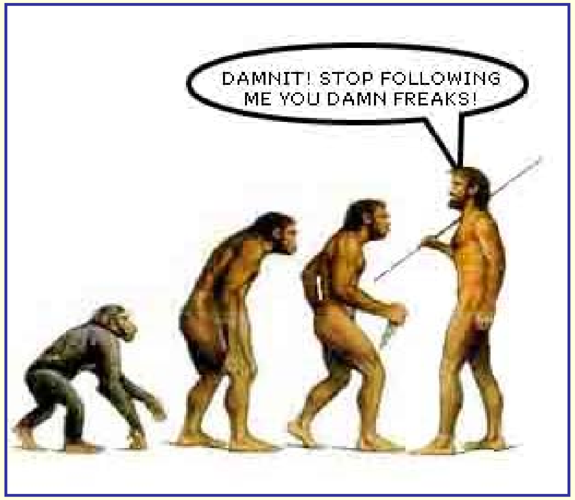
  - Example: Consumer behavior modeling with GFS (Casillas & Martínez, 2007) and Fuzzy LCSs (Orriols, Casillas & Martínez, 2008)

Future Applications towards meeting ML community interests

- **Scalability to deal with very large data sets**
  - Incremental learning seems suitable for extremely large data sets
  - Llorà met the challenge: Extremely high-dimensional data processed by means of Incremental Rule Learning plus efficiency enhancement techniques (Llorà et al., 2007)
  - LCSs evolve distributed sub-solutions: prepared for parallelization (Bull et al., 2007)
  - Many other contributions to bioinformatics (Bacardit et. al, 2007)

Future Applications towards meeting ML community interests

- **Domain of Competence of LCSs**
  - Dots partially connected: Found relationships between LCS's performance and measurements of geometrical complexity in classification problems (Bernadó & Ho, 2005)
  - Learning classifier systems are ensembles that evolve distributed solutions, so:
    - Can we co-evolve different representations?
    - Can introduce our knowledge about *geometrical complexity* into evolution to improve learning or inference?



Future Applications towards meeting ML community interests

- **Improve usability**
  - Wouldn't it be nice to have a self-configuring, self-adaptive machine?
    - Heuristics for automatic configuration of the systems
    - In data classification: use complexity metrics
    - In more complex environments: auto-adapt your configuration if drifting

Future Applications towards meeting Holland's original envision

- **Do you thing that LCSs are very clever?**
  - What would I want from an intelligent machine?
    - Consciousness, Causality and Intentionality
  - Goldberg's thoughts: *"Shooting for consciousness is not completely crazy"*
  - What's missing in LS-1 schema?
    - *Are LCSs intentional? Third, hunger cause desire, intention.*
    - *Collective intentionality*
    - **Consciousness?**
      - I'm conscious because my consciousness exists
      - Is it a multiple system sending things to each other?
      - Anyway, it comes unified
      - It connects us to world (tie to intentionality)
      - Decompose the problem and later put the pieces together

Learning Classifier Systems: Recent Trends

Albert Orriols Puig
aorriols@salle.url.edu

Grup de Recerca en Sistemes Intel·ligents
Enginyeria i Arquitectura La Salle
Universitat Ramon Llull