Imperial College London

Developing robust Synthetic Biology designs using automatically chosen microfluidic experiments

Stephen Muggleton
Department of Computing
Imperial College, London
April, 2009

Overview

Inductive Logic Programming (ILP)

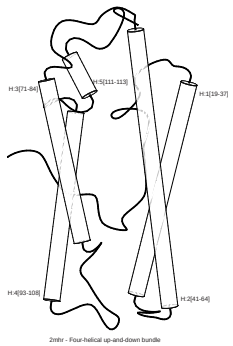
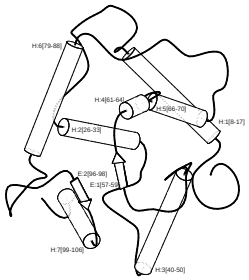
Probabilistic ILP (PILP)

The Robot Scientist project

Synthetic Biology

Microfluidic Robot Scientist

Inducing a logical model for “4-helical up-and-down bundle”

Positive(12)	Negative(12)
	

The protein P has fold class “Four-helical up-and-down bundle” if it contains a long helix H1 at a secondary structure position between 1 and 3, and H1 is followed by a second helix H2.

[JMB, 2001; MLJ, 2001]

Inductive Logic Programming

Background knowledge. Protein sequence, partial grammar, incomplete biological network.

Examples. Molecules, annotated sentences, temporal traces of up/down regulation.

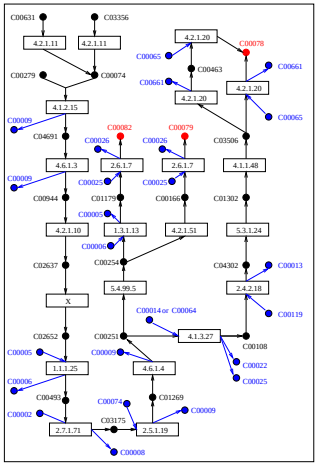
Hypothesis. Explanation of molecular 3-D shape, new clauses in a grammar, extra network annotation.

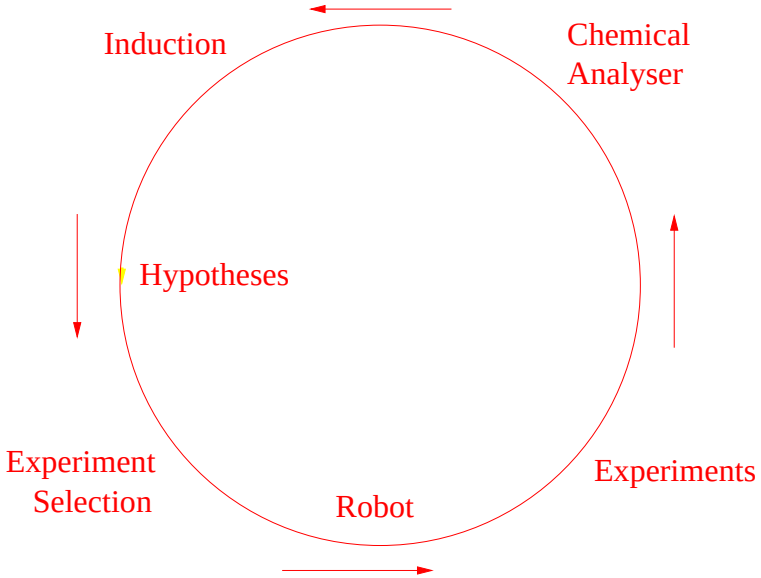
Probabilistic ILP

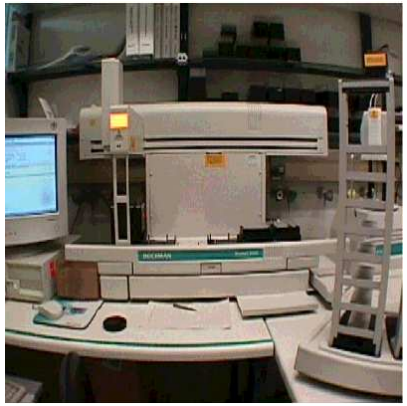
- Active area of research [Getoor and Taskar, 2007; DeRaedt et al 2008]
- SRL/PILP 10 year research challenges [Dietterich et al 2008]
- Experiment design and active learning [King et al, 2004]
- Chemical Turing Machines [Muggleton, 2006]
- Abduction maximising p(H)p(E|H) [Arvanitis+Muggleton, 2006]
- Use of probabilistic examples [Chen, 2008]
- Learnability models for PILP [Watanabe, 2008]

Robot Scientist (Nature, 2004)

Aberystwyth, Manchester, York, Imperial







Active Learning

Optimal Experimentation (Fedorov, 1972)

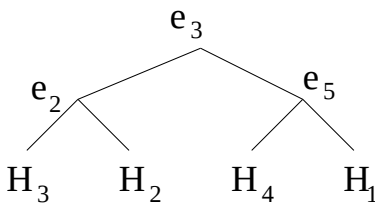
ASE-Progol

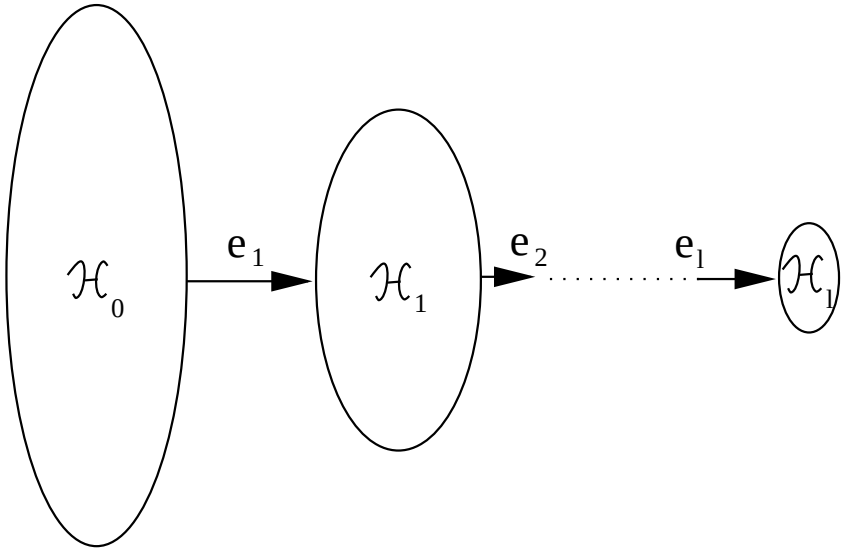
$$X = \{x_1, .., x_m\}, \mathcal{H}_0 = \{H_1, .., H_n\}$$

a)

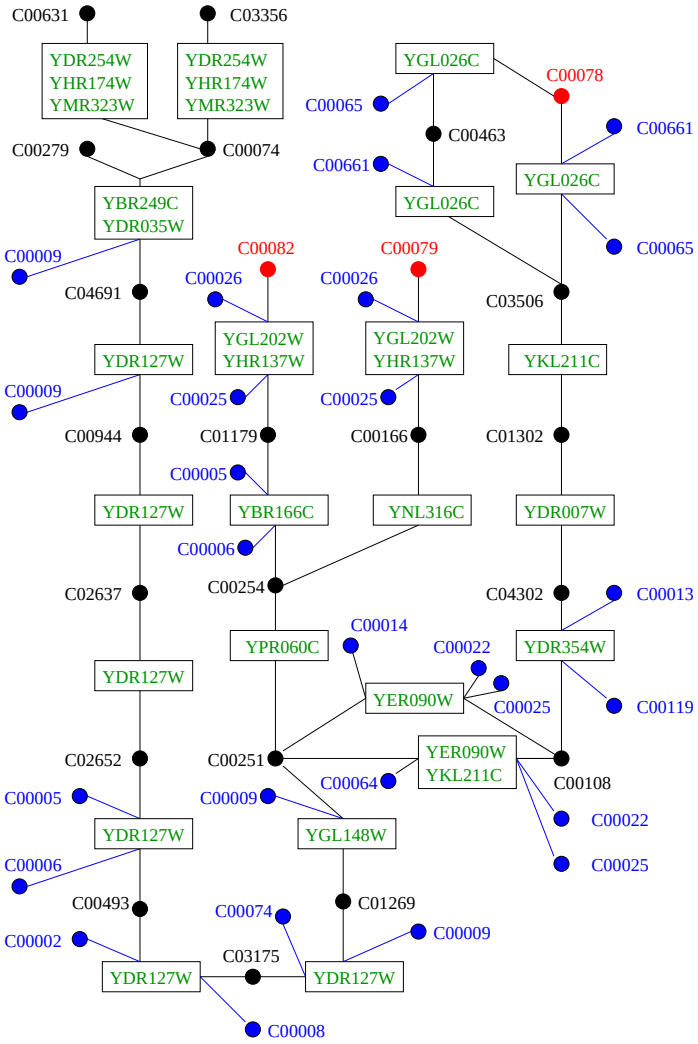
	e_1	...	e_m
H_1			
...			
H_n			

b)

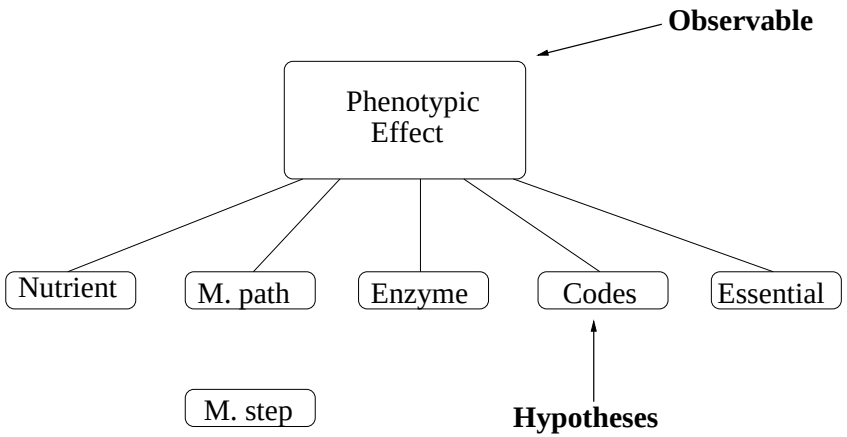




Aromatic amino acid pathways for yeast



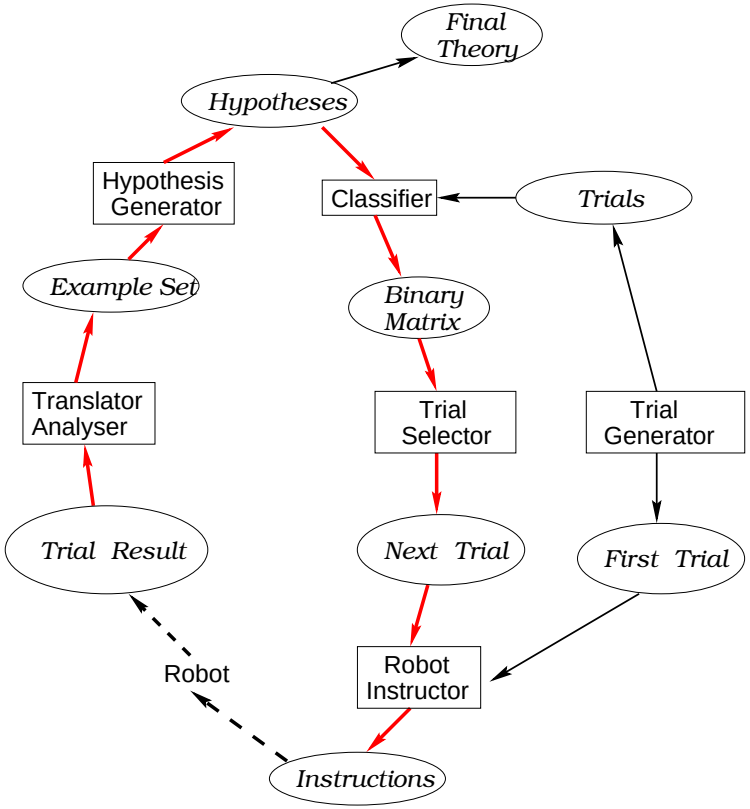
Calling diagram



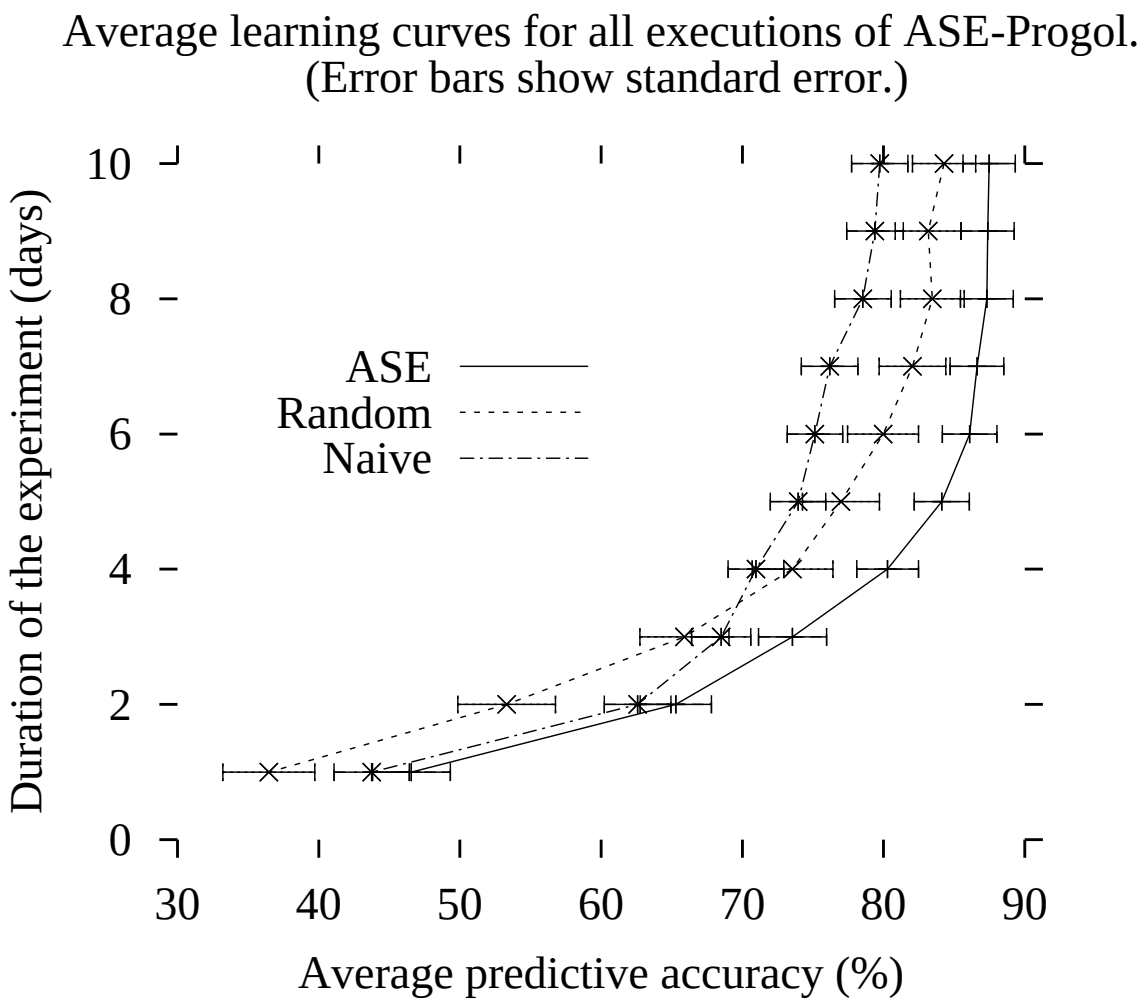
Prolog representation

```
phenotypic_effect(G, Medium):-
          nutrient_in(Nutrient, Medium),
          metabolic_path(Nutrient, Mi),
          enzyme(E, Mi, Mj),
          codes(G,E),
          metabolic_path(Mj, Mn),
          essential_molecule(Mn).
```

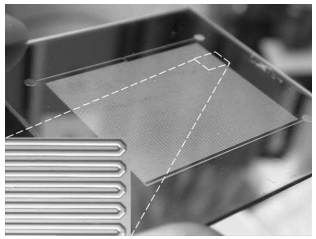
ASE-Progol

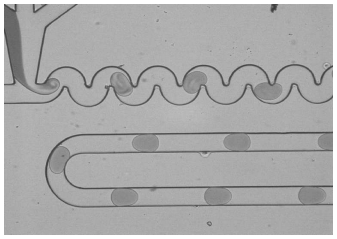

Final Theory
Hypotheses
Hypothesis Generator
Classifier
Trials
Example Set
Binary Matrix
Translator Analyser
Trial Selector
Trial Generator
Trial Result
Next Trial
First Trial
Robot
Robot Instructor
Instructions

Learning curves - time



Microfluidic Chemical Turing Machine [Nature, 2006]



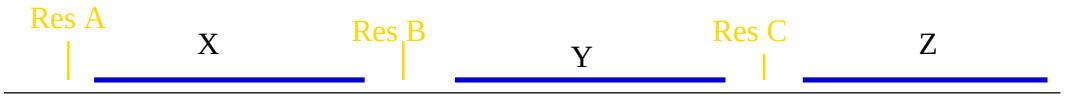


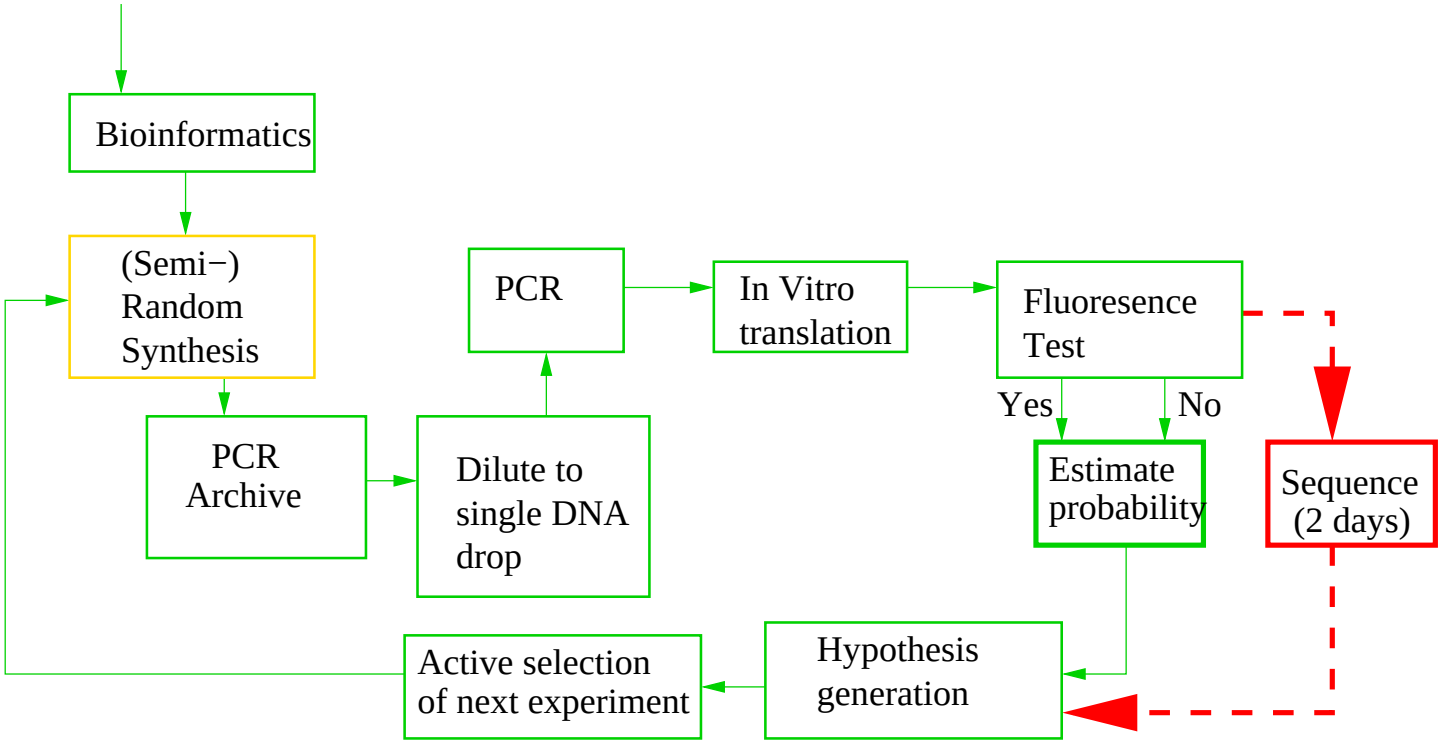
- Existing - custom designed special-purpose machines.
- Etched on silicon and poly-dimethylsiloxane (PDMS).
- Ducts, valves, ionic pumps and fluorescence tests.
- CUTM - addressable storage as micro-droplets($\approx 5\mu m$).
- Computer interface - “robot” replacement.
- Collaboration with Profs de Mello (Chemical Nanosciences), Freemont (Structural Biology), Sternberg (Structural Bioinformatics).

Synthetic Biology application

- MIT BioBrick competition IGem 2007
- Imperial College entry - Kitney+Freemont
- Cell-free biofilm infection detector
- Aim: improve robustness of design
- Microfluidic experiments and reconfiguration

Closed Loop Microfluidic Experiments





Summary and conclusion

- Priolog development
- Microfluidic machine
- Synthetic biology application
- PILP reduces experimental bottlenck
- High-speed cycle
- Comprehensible theories