Logic-based and Probabilistic Symbolic Learning

Lecture 1: Human vs Statistical Learning

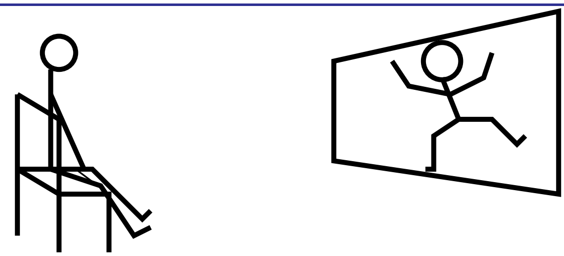
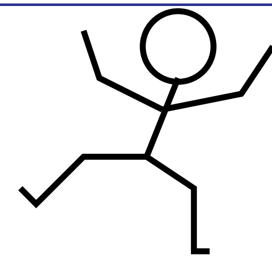
Stephen Muggleton
Department of Computing
Imperial College, London

Human vs Statistical Learning

UK EPSRC Priority 2016-2021 - Human-like Computing

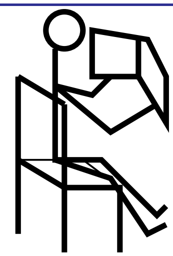
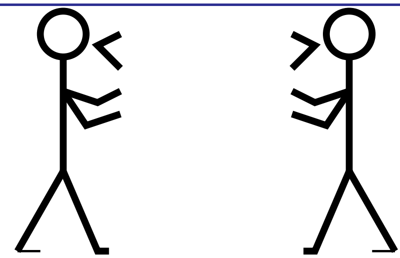
Characteristic	Human	Statistical
Examples per concept	Few (≈ 1) [Tenenbaum, 2011]	Many ($\geq 10K$)
Concepts	Many ($\geq 10K$) [Brown et al, 2008]	Few (≈ 1)
Background knowledge	Large [Brown, 2000]	Small
Structure	Modular, re-useable [Omrod et al, 2004]	Monolithic

Example 1: Dance Routine

	
Observe Visual perception	Perform Motor program

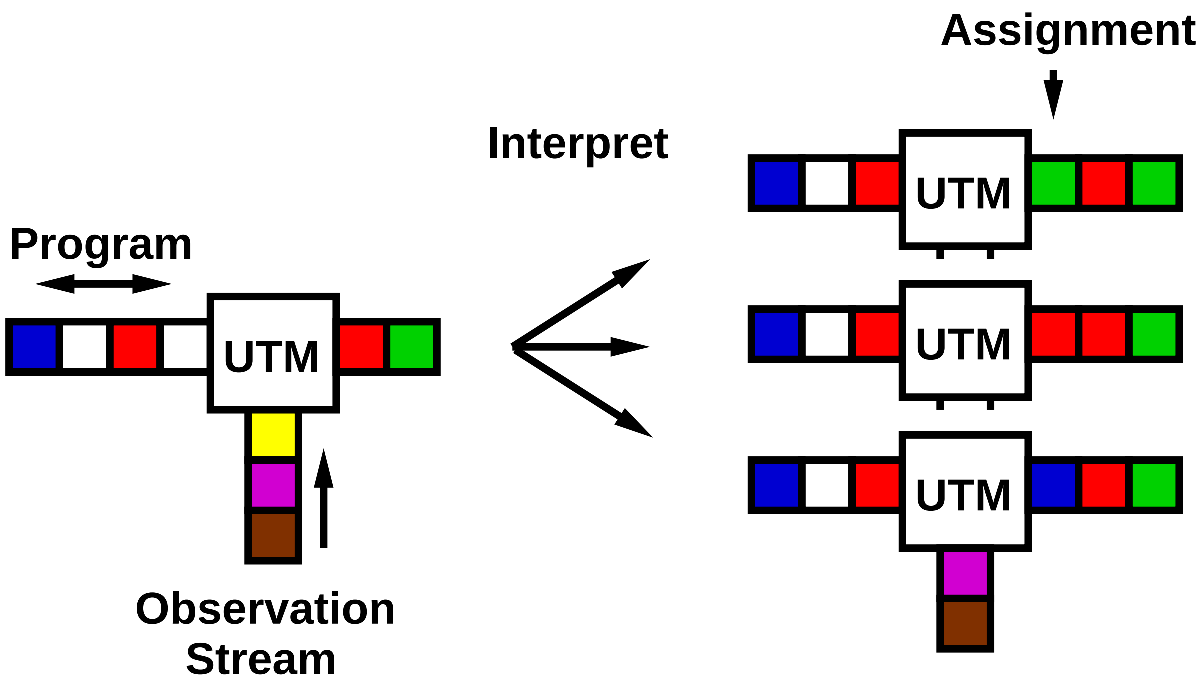
- A girl watches a dance routine on television.
- Afterwards she reproduces the routine.
- The new dance moves are incorporated into her repertoire.
- Subsequent improvisation allows re-use of parts of routines.

Example 2: Learning words in a language

	
Observe	Perform
Reading	Talking

- Average undergraduate knows 20K words.
- Learning rate = $\frac{20000}{20 \times 365} = 2.7$ new words per day since birth.
- Presentations new word before assimilation ≈ 1 [Zipf's Law].
- Word assimilation involves visual, auditory, sense and context recognition of associated concept.

Learning as Interpretation



Write-once, Non-deterministic **Universal Turing Machine**

Computation = Learning = Interpretation = Perception

Meta-Interpretive Learning [IJCAI 2013]

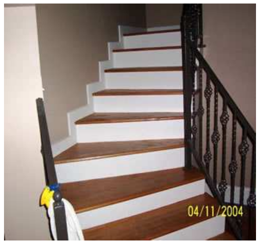
Prolog Meta-Interpreter implements Learning as Interpretation.

Input to Meta-Interpreter: 1) Observations, 2) Meta-Rules, 3) Background Knowledge assignments (substitutions).

Output from Meta-Interpreter: Hypothesised assignments.

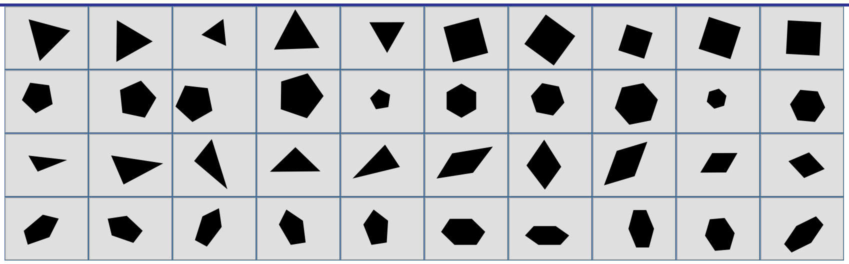
Metagol supports Problem decomposition by Predicate Invention and Learning recursion [MLJ 2015], Single example multi-task learning [ECAI 2014], Program Induction with resource and time-complexity optimisation [IJCAI 2015].

Vision applications



Staircase

ILP 2013



Regular Geometric

ILP 2015

stair(X,Y) :- a(X,Y).

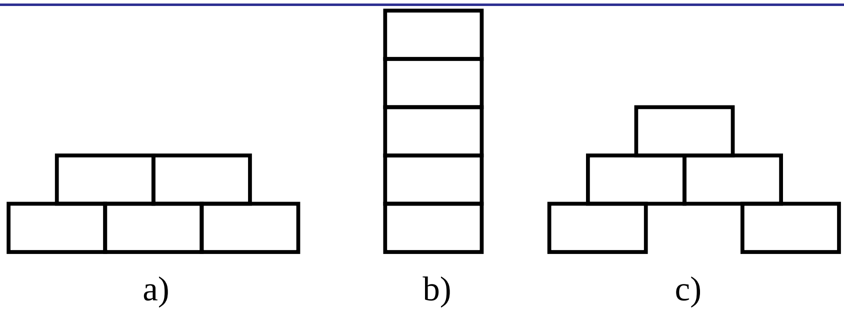
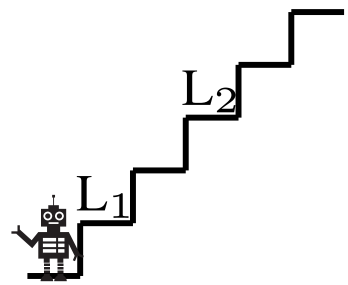
stair(X,Y) :- a(X,Z), **stair(Z,Y)**.

a(X,Y) :- vertical(X,Z), horizontal(Z,Y).

Learned in 0.08s on laptop from single image.

Note Predicate invention and **recursion**.

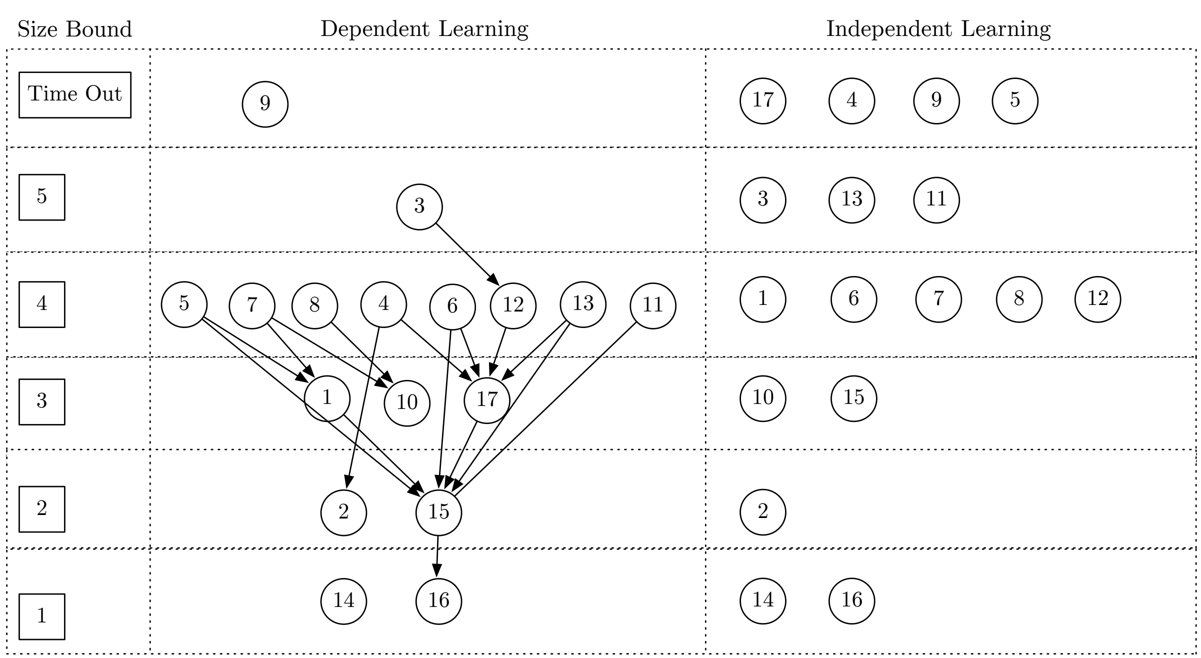
Robotic applications

	
Building a Stable Wall IJCAI 2013	Learning Efficient Strategies IJCAI 2015

Language applications

Formal grammars [MLJ 2014]

Dependent string transformations [ECAI 2014]



Chain of programs from dependent learning

f_{03}(A,B) :- f_{12_1}(A,C), f_{12}(C,B).

f_{12}(A,B) :- f_{12_1}(A,C), f_{12_2}(C,B).

f_{12_1}(A,B) :- f_{12_2}(A,C), $skip1$(C,B).

f_{12_2}(A,B) :- f_{12_3}(A,C), $write1$(C,B,'.').

f_{12_3}(A,B) :- $copy1$(A,C), f_{17_1}(C,B).

f_{17}(A,B) :- f_{17_1}(A,C), f_{15}(C,B).

f_{17_1}(A,B) :- f_{15_1}(A,C), f_{17_1}(C,B).

f_{17_1}(A,B) :- $skipalphanum$(A,B).

f_{15}(A,B) :- f_{15_1}(A,C), f_{16}(C,B).

f_{15_1}(A,B) :- $skipalphanum$(A,C), $skip1$(C,B).

f_{16}(A,B) :- $copyalphanum$(A,C), $skiprest$(C,B).

Other applications

Learning proof tactics [ILP 2015]

Learning data transformations [ILP 2015]

Conclusions and Challenges

- New form of Declarative Machine Learning [De Raedt, 2012]
- H^2_2 is tractable and Turing-complete fragment of High-order Logic
- Knuth-Bendix style ordering guarantees termination of queries
- Beyond classification learning - strategy learning

Challenges

- Generalise beyond Dyadic logic
- Deal with classification noise
- Active learning
- Efficient problem decomposition
- Meaningful invented names and types

Bibliography

- A. Cropper, S.H. Muggleton. Learning efficient logical robot strategies involving composable objects. IJCAI 2015.
- W-Z Dai, S.H. Muggleton, Z-H Zhou. Logical vision: Meta-interpretive learning for simple geometrical concepts. ILP 2015.
- S.H. Muggleton, D. Lin, A. Tamaddoni-Nezhad. Meta-interpretive learning of higher-order dyadic datalog: Predicate invention revisited. Machine Learning, 2015.
- D. Lin, E. Dechter, K. Ellis, J.B. Tenenbaum, S.H. Muggleton. Bias reformulation for one-shot function induction. ECAI 2014.