

Meta-Interpretive Learning

Lecture 8.2

One-shot induction and Bias reformulation

Stephen Muggleton
University of Nanjing
and Imperial College, London

12th April, 2022

Papers for this lecture

Paper8.2.1: Meta-Interpretive Learning from noisy images, MLJ
(2018)

Paper8.2.2: Bias reformulation for one-shot function induction,
ECAI2014

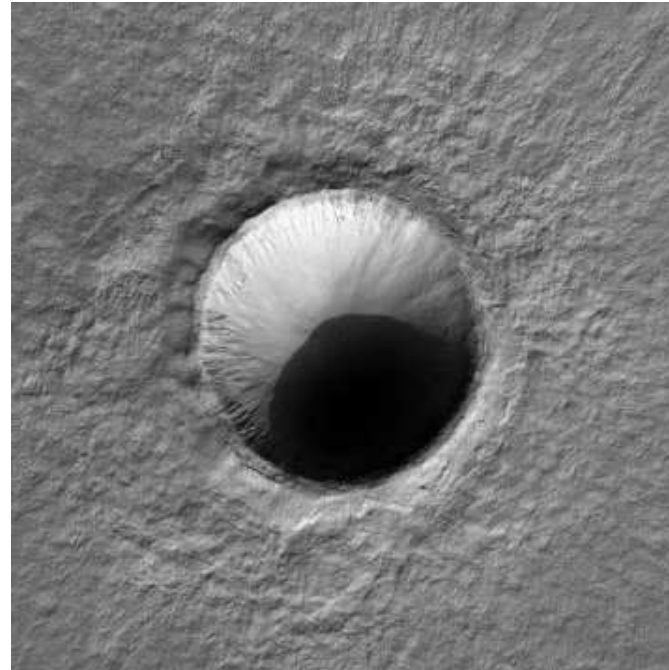
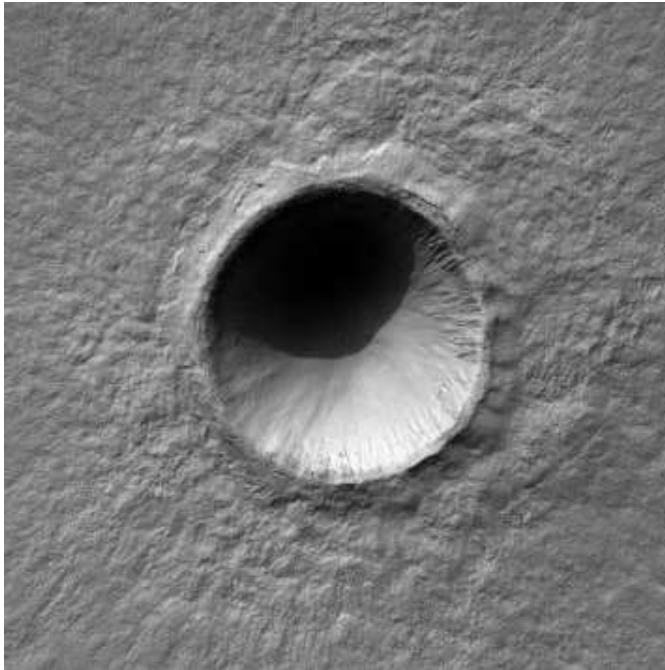
Motivation

- Inductive Programming
- Simple repetitive programs
- Human provides examples and checks results
- Requires data efficient inference
- PAC, Blumer bound, Strong Learning Bias
- Compatible with Human Bias? - Images and Text

Human versus Machine Learning

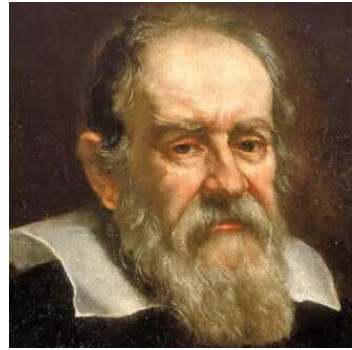
Characteristic	Human	Statistical
Examples per concept	Few (≈ 1) [Tenenbaum, 2011]	Many ($\geq 10K$)
Concepts to learn	Many ($\geq 10K$) [Brown et al, 2008]	Few (≈ 1)
Background knowledge	Large [Brown, 2000]	Small
Structure	Modular, re-useable [Omrod et al, 2004]	Monolithic

Human Visual Bias - Mars Crater/Mountain



(Credit: NASA/JPL/University of Arizona)

Logic, Learning and Vision



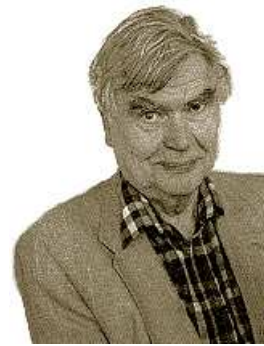
Galileo(1610)



Helmholtz(1867)

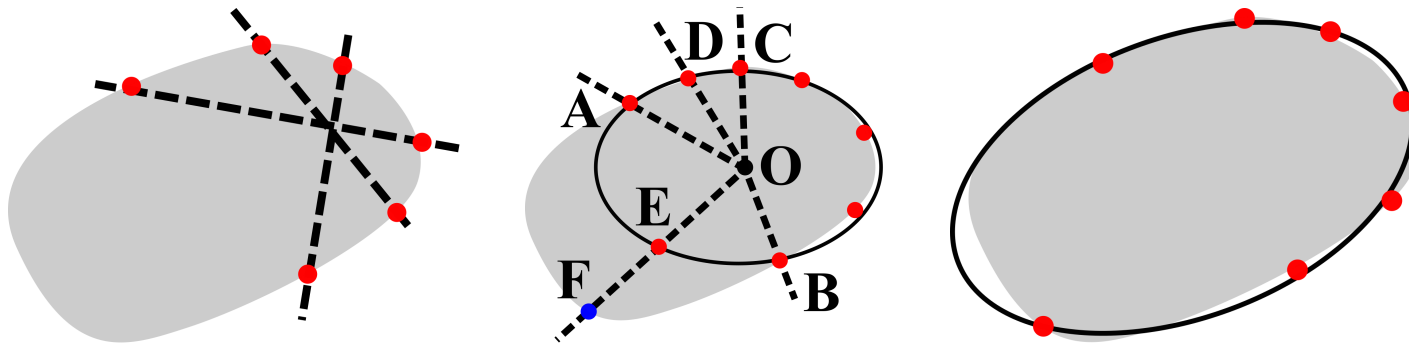


Turing(1950)

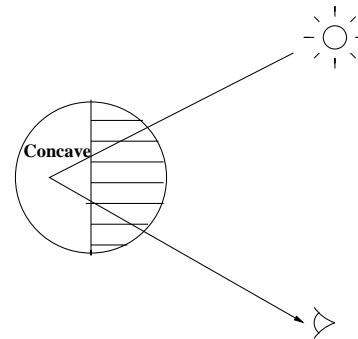
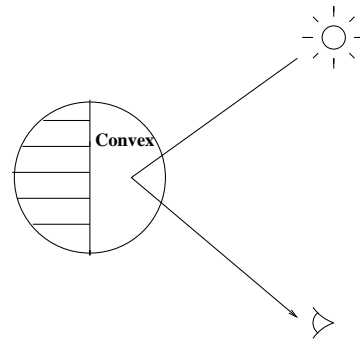
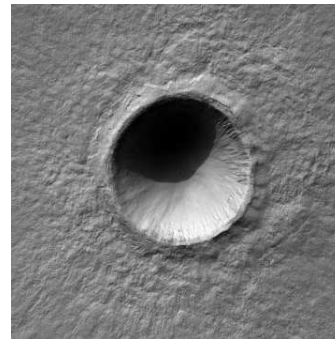


Gregory(1966)

Low-level object detection - ellipse model



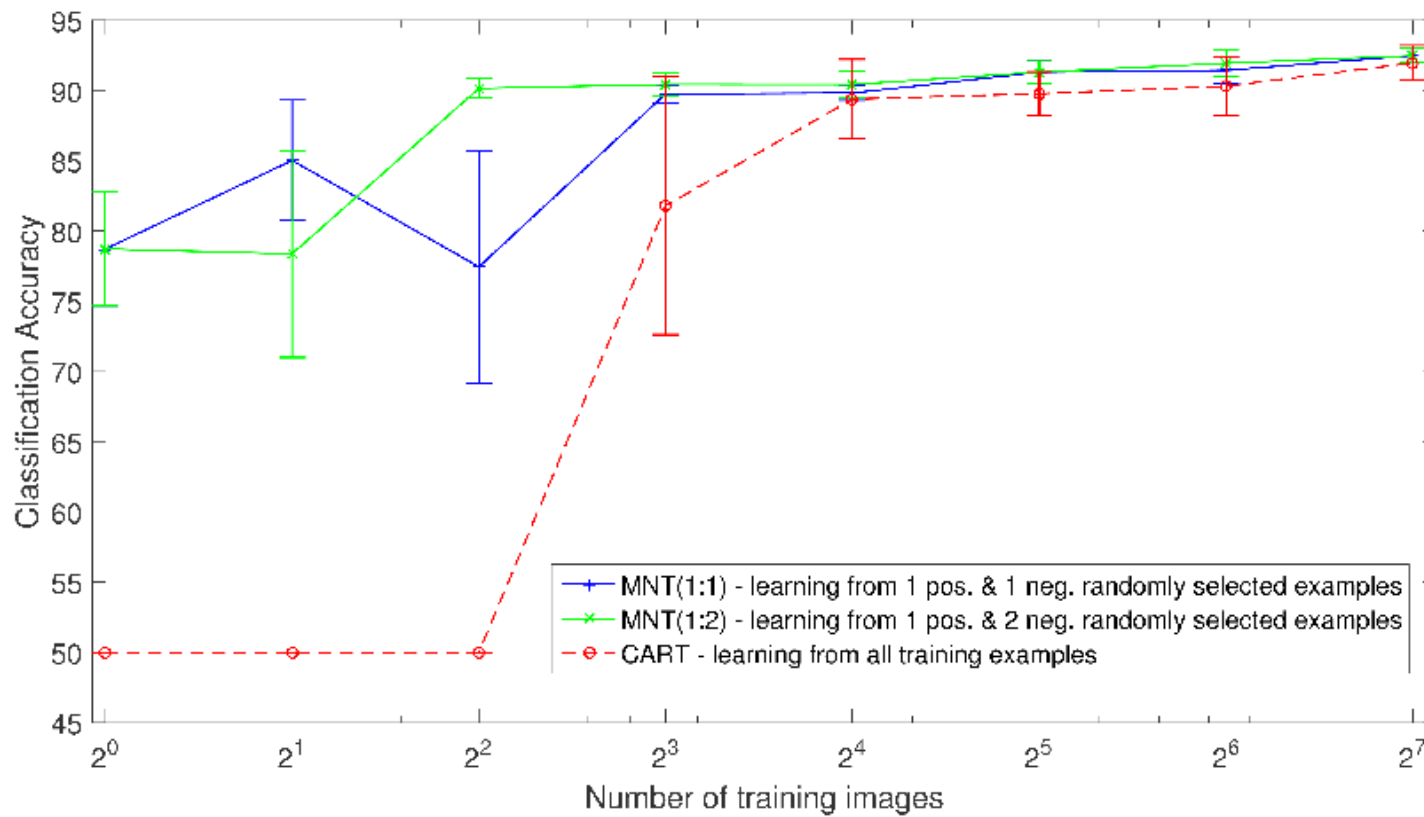
High-level Visual Bias and Background Knowledge



$light(X, X).$

$light(X, Y) : \neg reflect(X, Z), light(Z, Y).$

Bias effects in noisy images - Metagol vs CART



Textual Bias
Analogy problem

bob	BOB
alice	?

Alternative 1

bob	BOB
alice	BOICE

Alternative 2

bob	BOB
alice	ALICE

Harder tasks

Task1	miKe dwIGHT	Mike Dwight
Task2	European Conference on Artificial Intelligence	ECAI
Task3	My name is John.	John

Question

How can this human text transformation bias be learned by a computer?

**Option1: Hardcode bias as DSL [Gulwani 2011,2012]
FlashFill, Excel 2013**

brent.harold@hotmail.com	Brent Harold
matthew.rosman@yahoo.com	
jim.james@fas.harvard.edu	
ruby.clinton@mit.edu	
josh.smith@gmail.com	

**Option1: Hardcode the bias [Gulwani 2011,2012]
FlashFill, Excel 2013**

brent.harold@hotmail.com	Brent Harold
matthew.rosman@yahoo.com	Matthew Rosman
jim.james@fas.harvard.edu	Jim James
ruby.clinton@mit.edu	Ruby Clinton
josh.smith@gmail.com	Josh Smith

Option1: Non-intuitive FlashFill error

IaN RoDny	Ian Rodney
StaNleY TRAVis	Itanley Rley travis

Option2: Learn bias using variant of Meta-Interpretive Learning (Paper2.1)

brent.harold@hotmail.com

Brent Harold

ep04(A,B) \leftarrow ep04_1(A,C), ep04_2(C,B).

ep04_1(A,B) \leftarrow ep04_3(A,C), ep04_4(C,B).

ep04_2(A,B) \leftarrow ep04_3(A,C), skiprest(C,B).

ep04_3(A,B) \leftarrow make_uppercase(A,C), copyword(C,B).

ep04_4(A,B) \leftarrow skip1(A,C), write_space(C,B).

Learning time = **9.3 seconds**

Sequential episodes

ep02	james	James.
------	-------	--------

ep04	brent.harold@hotmail.com	Brent Harold
------	--------------------------	--------------

ep02(A,B) \leftarrow ep02_1(A,C), write_dot(C,B).

ep02_1(A,B) \leftarrow make_upper(A,C), copyword(C,B).

ep04(A,B) :- ep04_2(A,C), ep04_3(C,B).

ep04_2(A,B) :- ep04_4(A,C), skip1(C,B).

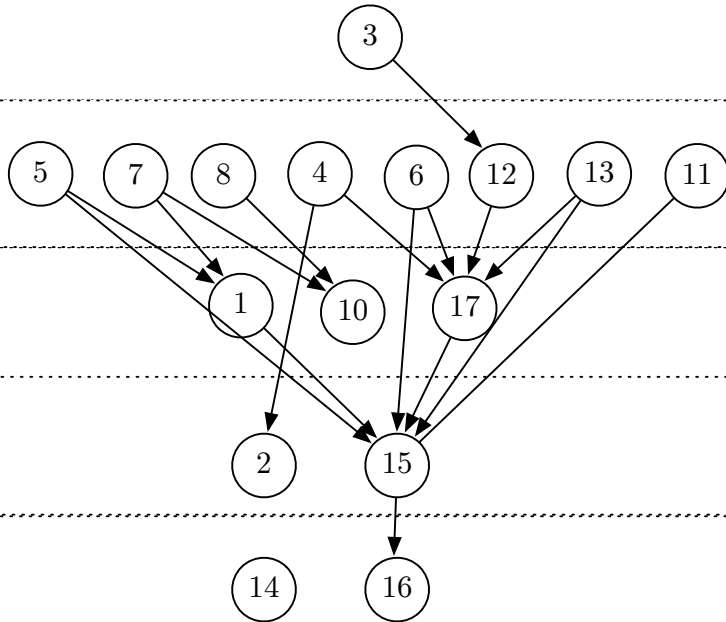
ep04_3(A,B) :- ep02_1(A,C), skiprest(C,B).

ep04_4(A,B) :- ep02_1(A,C), write_space(C,B).

Learning time = **3.1 seconds**

Dependent vs Independent Learning

Size Bound	Dependent Learning	Independent Learning
Time Out	(9)	(17) (4) (9) (5)
5	(3) → (12)	(3) (13) (11)
4	(5) (7) (8) (4) (6) (12) (13) (11) → (1) (10) (17)	(1) (6) (7) (8) (12)
3	(1) (10) (17) → (2) (15)	(10) (15)
2	(2) (15) → (15)	(2)
1	(14) (16)	(14) (16)



Chain of programs from dependent learning

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

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

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

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

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

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

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

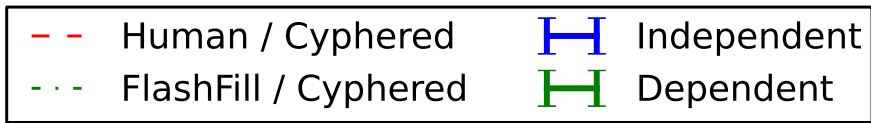
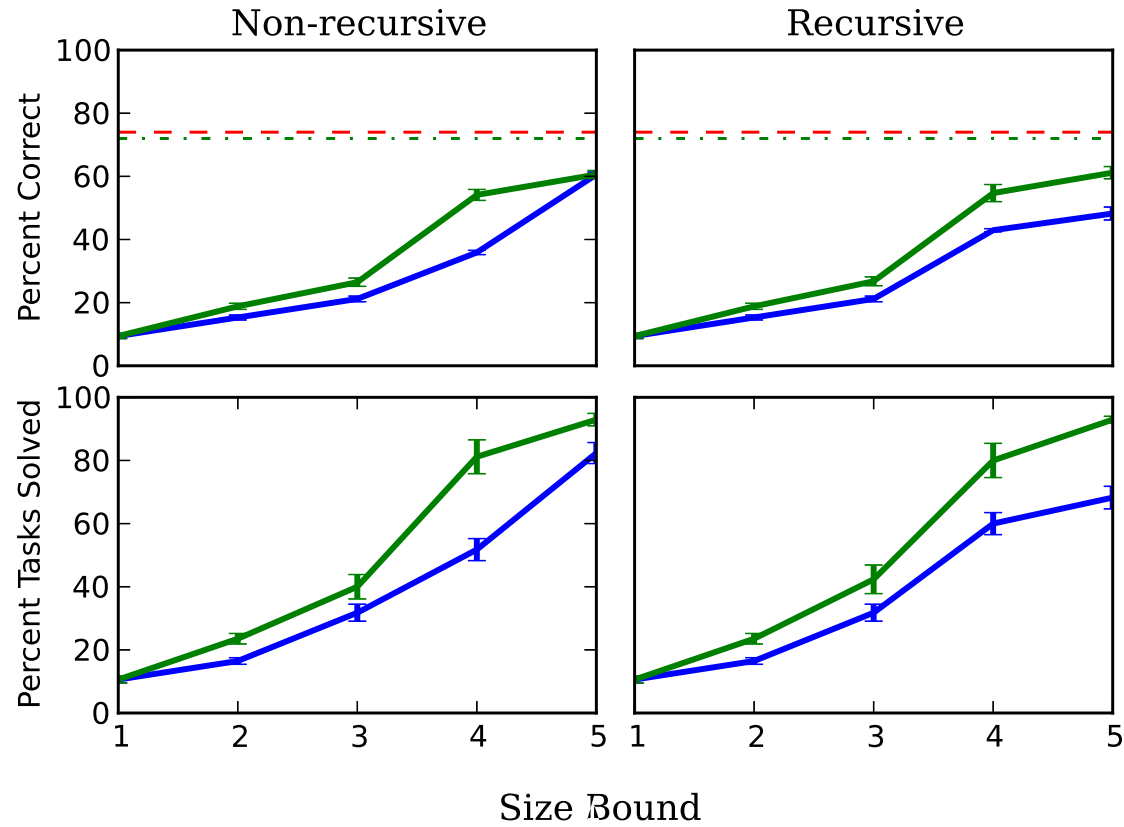
$f_{17_1}(A,B) :- skipalphanum(A,B).$

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

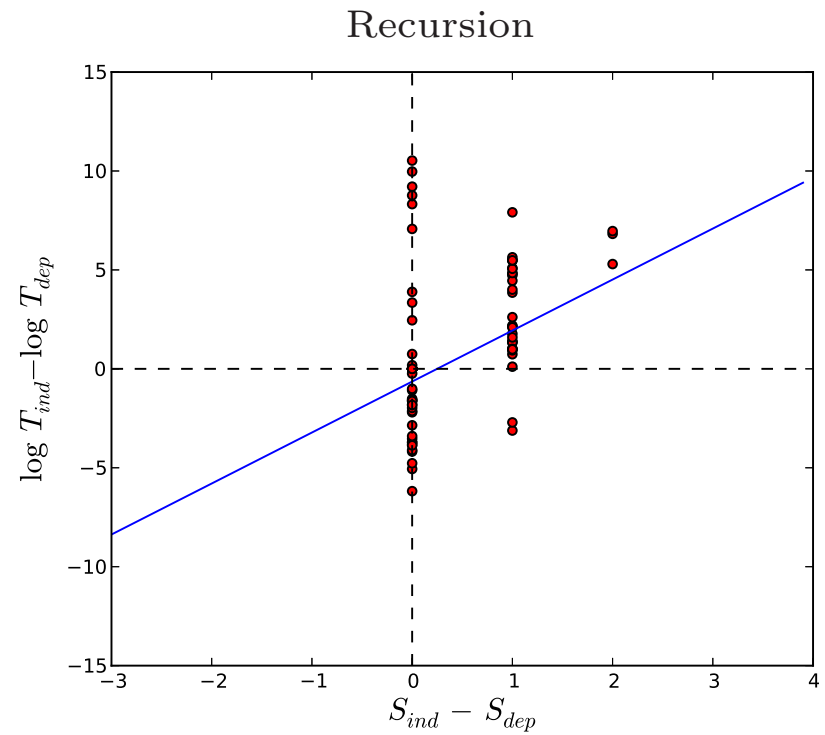
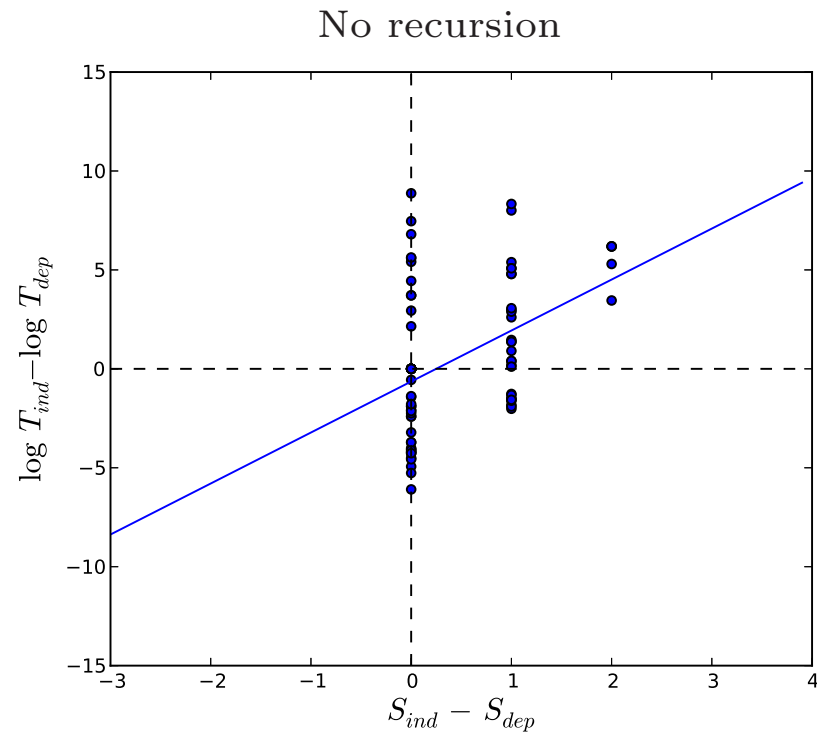
$f_{15_1}(A,B) :- skipalphanum(A,C), skip1(C,B).$

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

Performance graphs



Running times Dependent vs Independent



Summary

- Human perceptual bias in Images and Text
- Bias can be learned by transfer in multi-task learning
- Effect is consistent increase in data and running efficiency
- Complex programs can be built by sharing invented predicates