# Inductive Programming
# Lecture 3
# One-shot induction and Bias reformulation

Stephen Muggleton
Department of Computing
Imperial College, London and
University of Nanjing

16th October, 2023

## Papers for this lecture

**Paper3.1:** S.H. Muggleton, W-Z. Dai, C. Sammut, A. Tamaddoni-Nezhad, J. Wen, and Z-H. Zhou. Meta-interpretive learning from noisy images. Machine Learning, 107:1097-1118, 2018.

**Paper3.2:** D. Lin, E. Dechter, K. Ellis, J.B. Tenenbaum, and S.H. Muggleton. Bias reformulation for one-shot function induction. In Proceedings of the 23rd European Conference on Artificial Intelligence (ECAI 2014), pages 525-530.

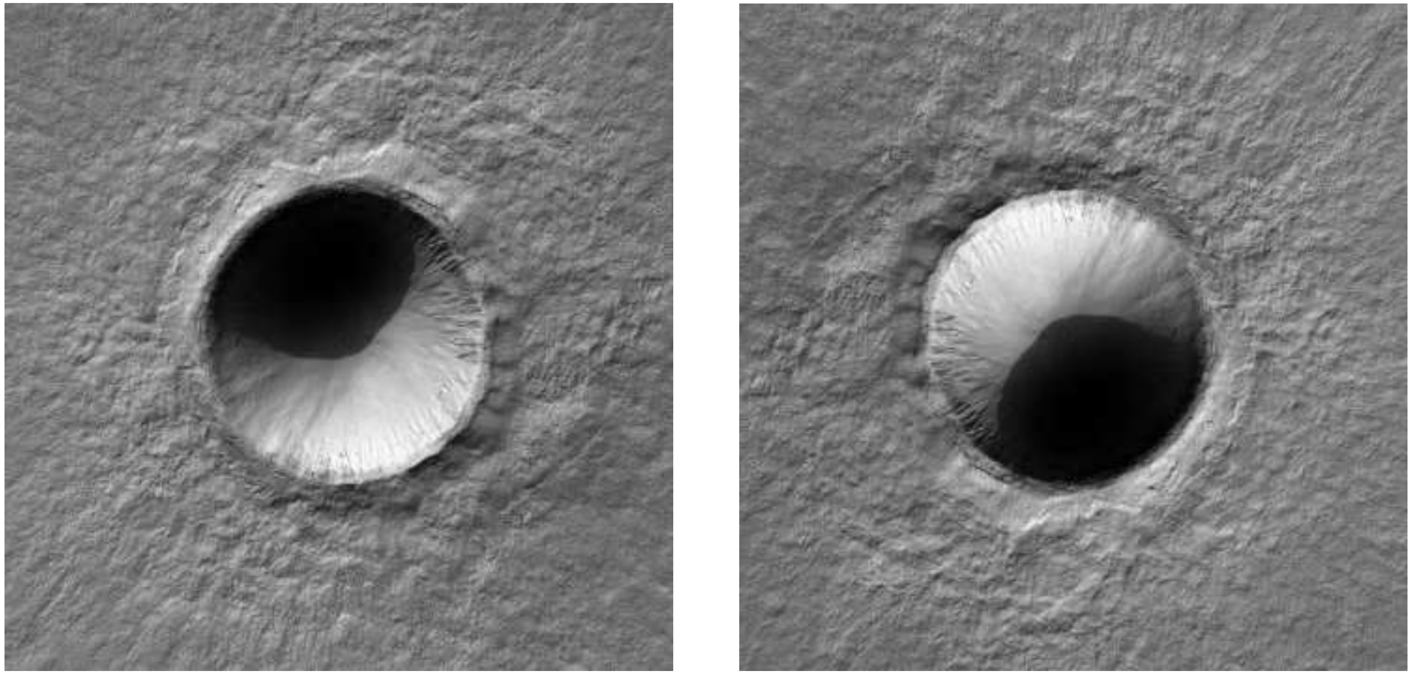
## Motivation

- 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

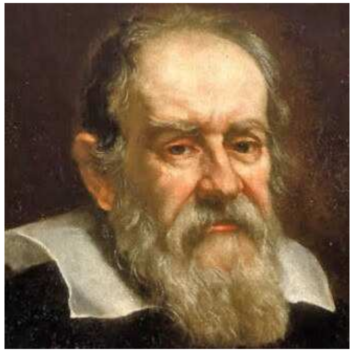
| **Characteristic** | **Human** | **Statistical** |
|---|---|---|
| Examples per concept | Few ($\approx 1$) [Tenenbaum, 2011] | Many ($\geq 10K$) |
| Concepts to learn | Many ($\geq 10K$) [Brown et al, 2008] | Few ($\approx 1$) |
| Background knowledge | Large [Brown, 2000] | Small |
| Structure |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

## Human Visual Bias - Mars Crater/Mountain



(Credit: NASA/JPL/University of Arizona)

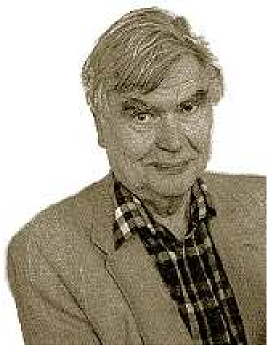
**Logic, Learning and Vision**



Galileo(1610)

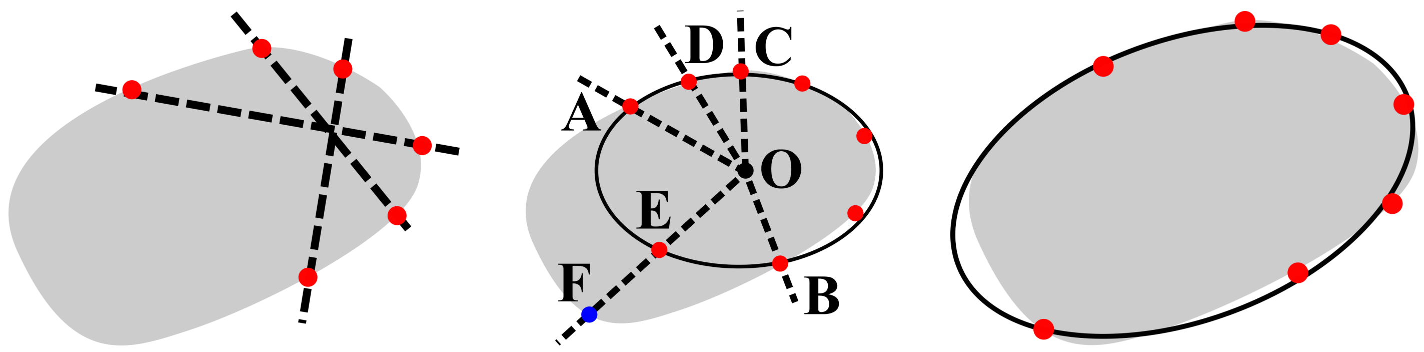
Helmholtz(1867)

Turing(1950)

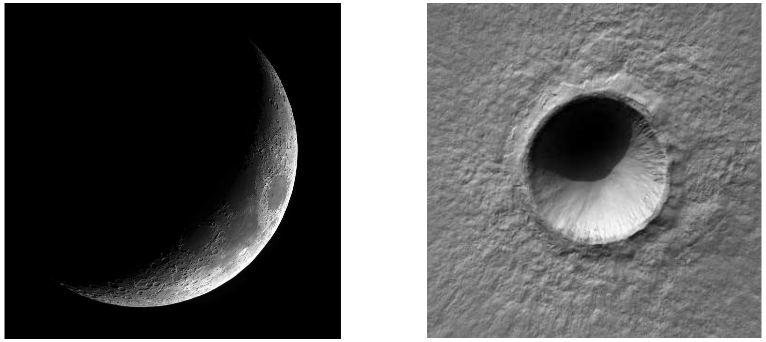


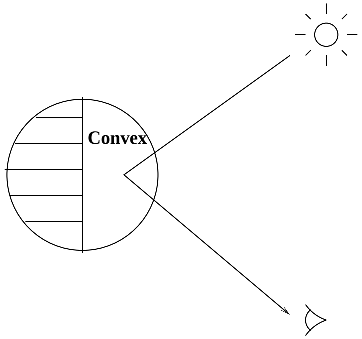
Gregory(1966)

## Low-level object detection - elipse model



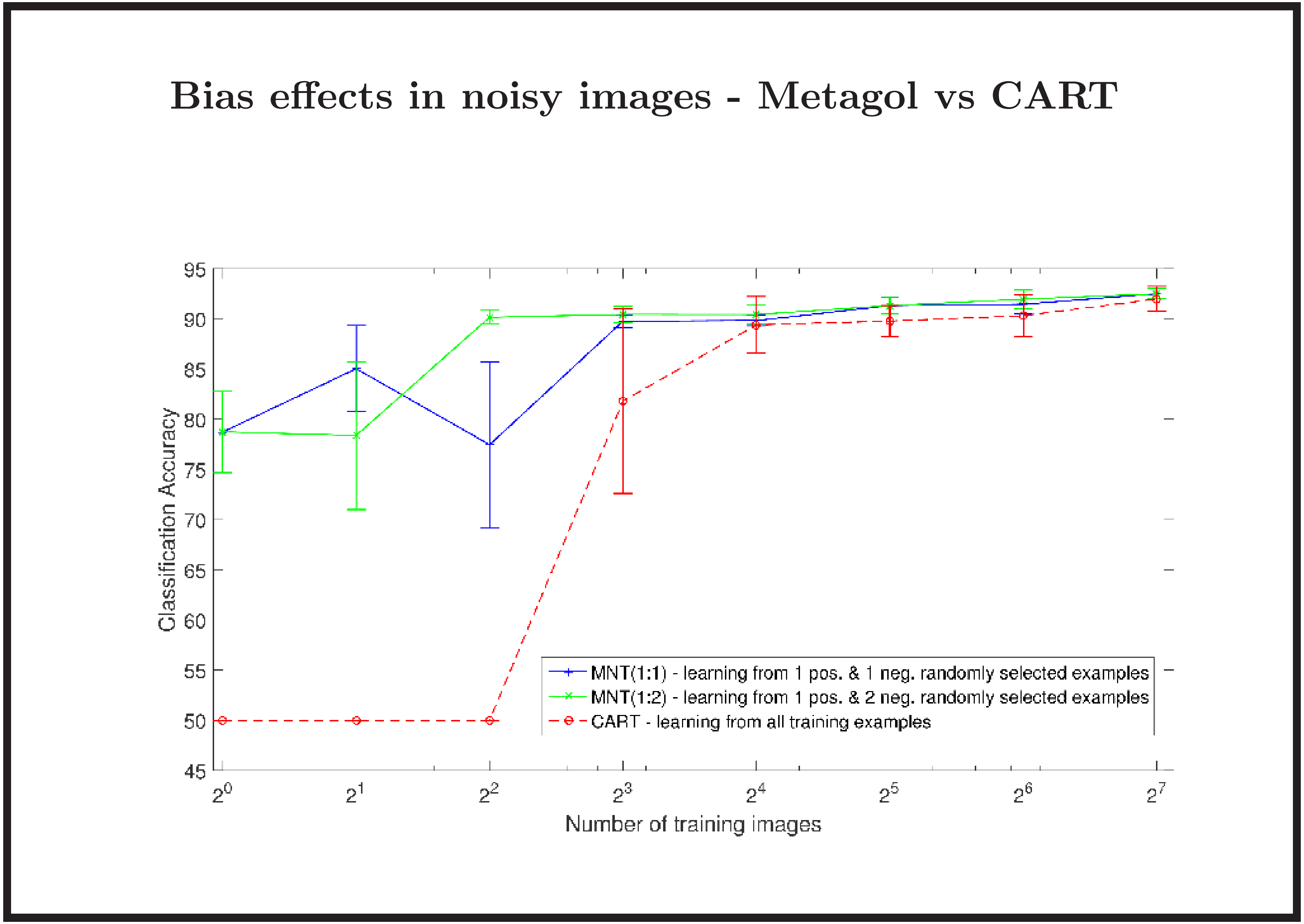
# High-level Visual Bias and Background Knowledge






$$light(X, X).$$

$$light(X, Y) :- reflect(X, Z), light(Z, Y).$$


Bias effects in noisy images - Metagol vs CART
Classification Accuracy
95
90
85
80
75
70
65
60
55
50
45
$2^0$
$2^1$
$2^2$
$2^3$
$2^4$
$2^5$
$2^6$
$2^7$
Number of training images
MNT(1:1) - learning from 1 pos. & 1 neg. randomly selected examples
MNT(1:2) - learning from 1 pos. & 2 neg. randomly selected examples
CART - learning from all training examples

**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 Rodny |
| 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) ← ep04_1(A,C), ep04_2(C,B).

ep04_1(A,B) ← ep04_3(A,C), ep04_4(C,B).

ep04_2(A,B) ← ep04_3(A,C), skiprest(C,B).

ep04_3(A,B) ← make_uppercase(A,C), copyword(C,B).

ep04_4(A,B) ← 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) ← ep02_1(A,C), write_dot(C,B).

ep02_1(A,B) ← make_upper(A,C), copyword(C,B).

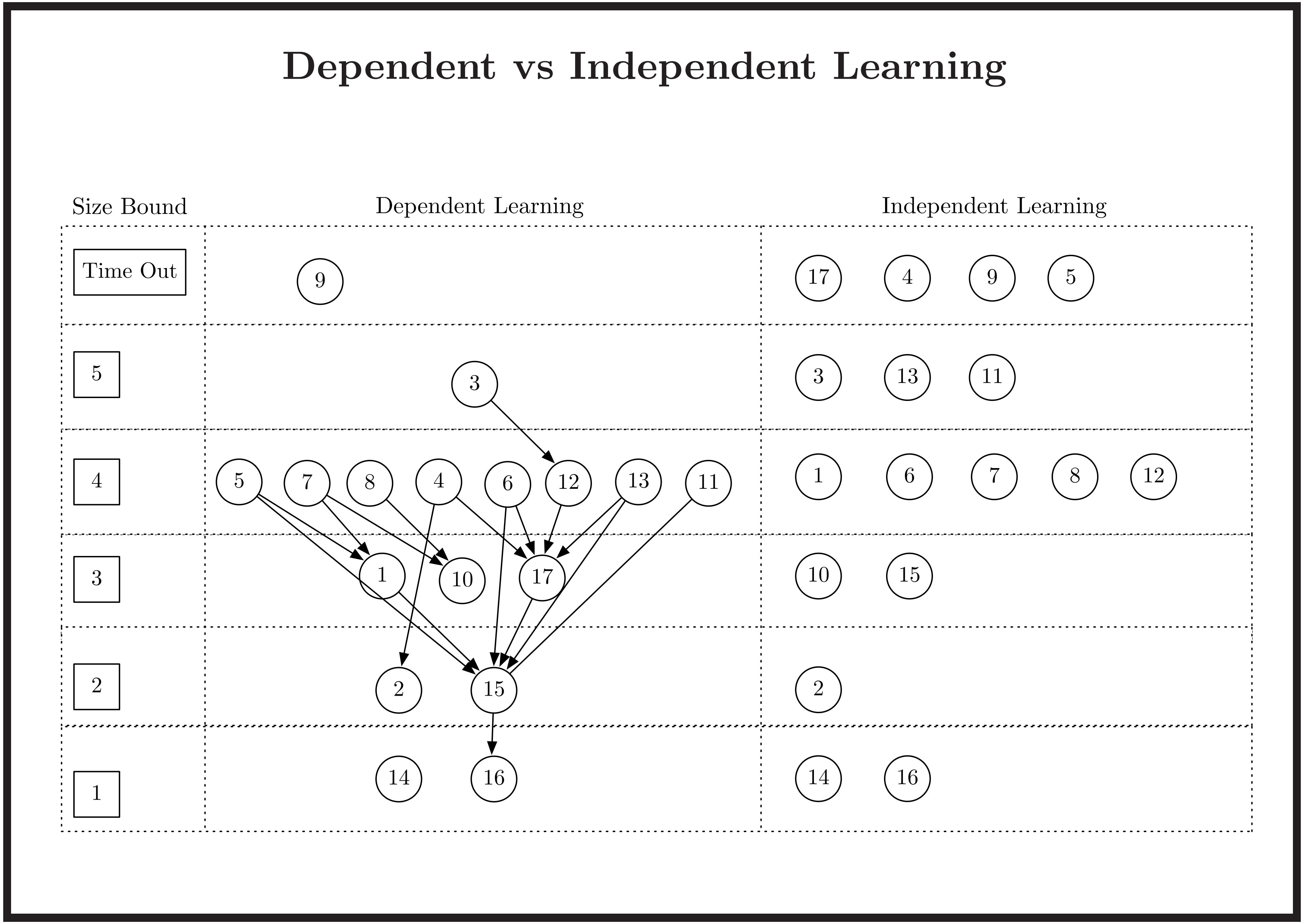
ep04(A,B) :- ep04_2(A,C), ep04_3(C,B).

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
3
13
11
4
5
7
8
4
6
12
13
11
1
6
7
8
12
3
1
10
17
10
15
2
2
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), *skip*1(C,B).

$f_{12\_2}$(A,B) :- $f_{12\_3}$(A,C), *write*1(C,B,'.').

$f_{12\_3}$(A,B) :- *copy*1(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).

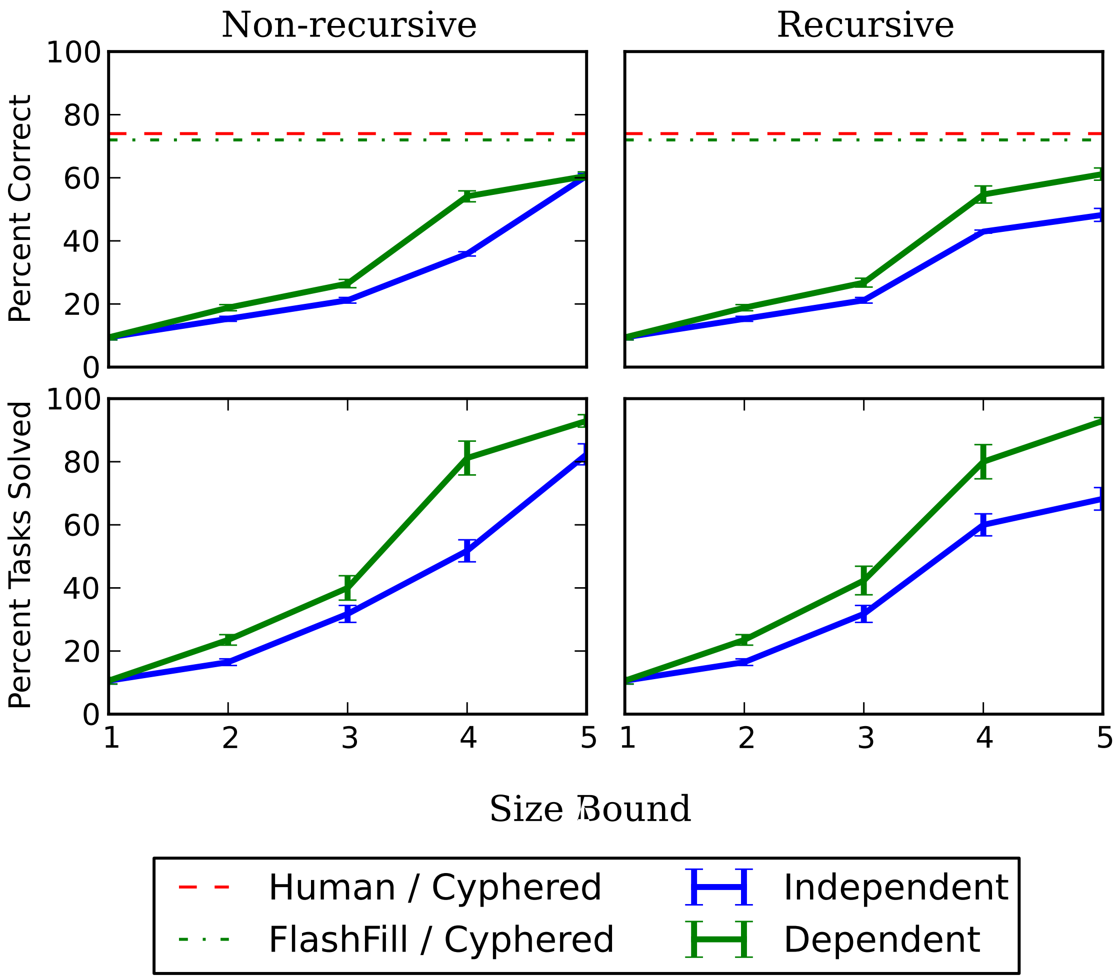
$f_{17\_1}$(A,B) :- *skipalphanum*(A,B).

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

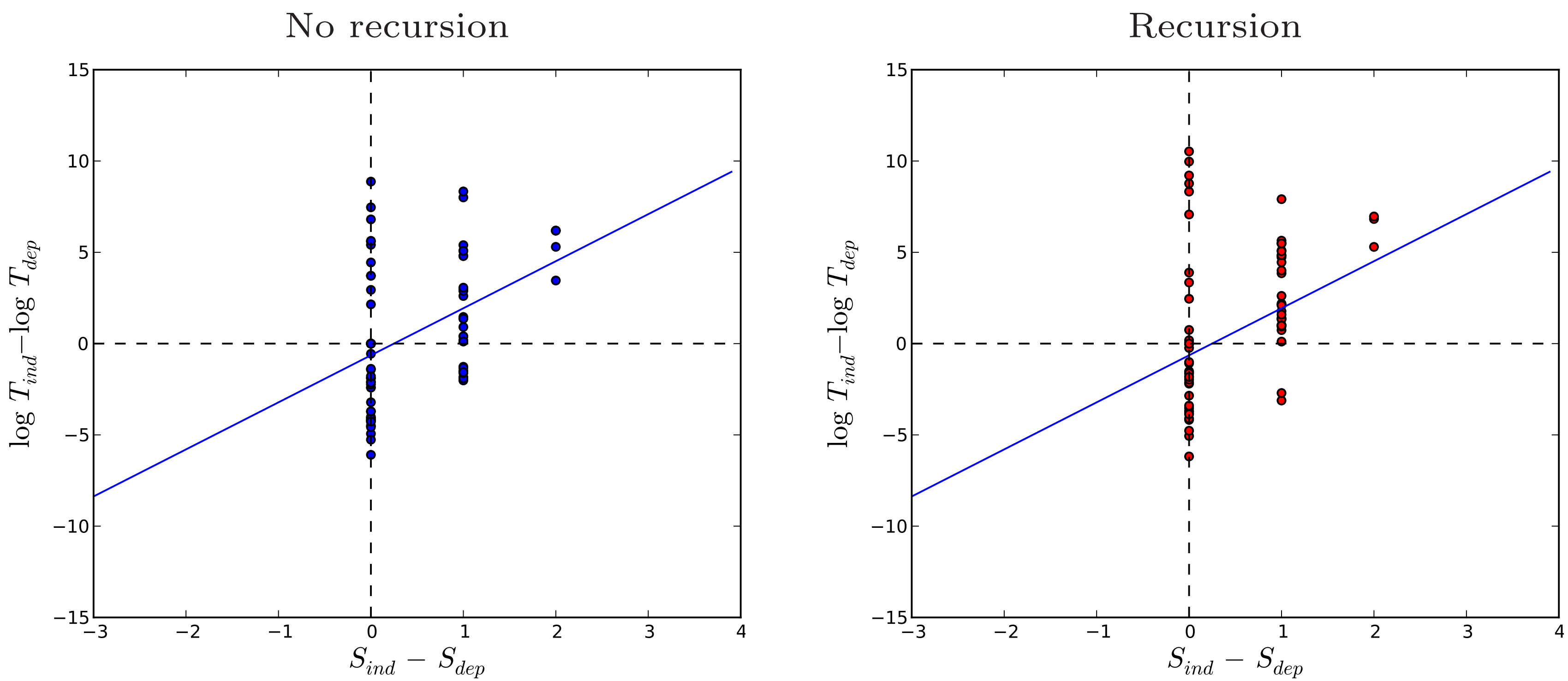
$f_{15\_1}$(A,B) :- *skipalphanum*(A,C), *skip*1(C,B).

$f_{16}$(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