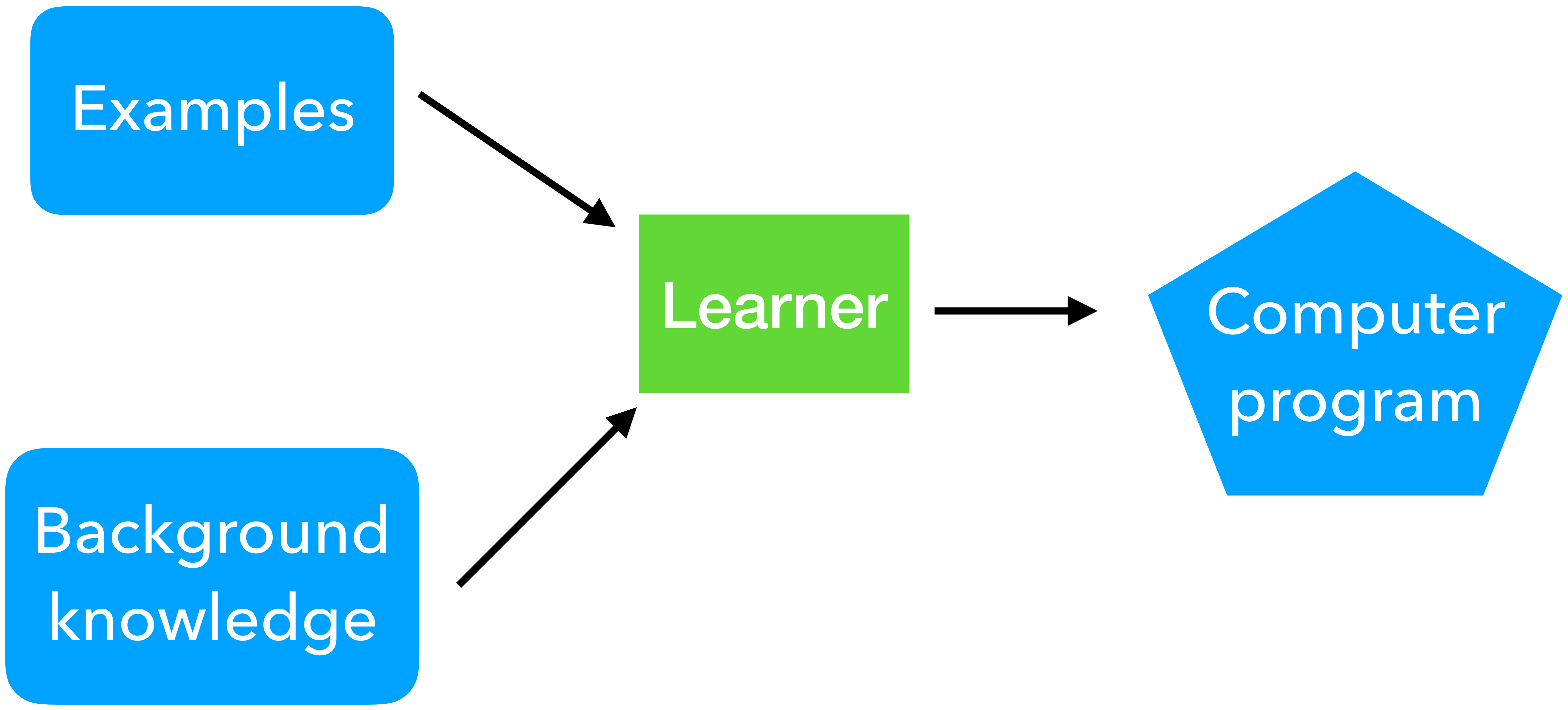
Logical reduction of metarules

Andrew Cropper & Sophie Tourret

ILP



Mode declarations

Progol
ILASP
Aleph
XHAIL
TILDE
...

Metarules

Metagol

MIL-Hex

∂ILP

ProPPR

Clint

MOBAL

...

(almost all neural-ILP approaches)

Metarules

$$\mathbf{P}(A,B) \leftarrow \mathbf{Q}(A,C),\mathbf{R}(C,B)$$

Metarules

P(A,B) ← **Q**(A),**R**(A,B)

Metarules

P(A,B) ← **Q**(A),**R**(A,B)

Input

```
% background
parent(alice,bob).
parent(bob,charlie).

% example
grandparent(alice,charlie).

% metarule
```

$\mathbf{P}(A,B) \leftarrow \mathbf{Q}(A,C),\mathbf{R}(C,B)$

Output

% subs

Subs = {**P**\grandparent,**Q**\parent,**R**\parent}

% program

grandparent(A,B) ← parent(A,C), parent(C,B)

Where do we get metarules from?

Completeness

Cannot learn grandparent/2 with only `P(X)←Q(X)`

Efficiency

More metarules = larger hypothesis space

Idea: find logically minimal sets

Entailment redundant metarules [Cropper and Muggleton, ILP14]

The clause C is **entailment redundant** in the clausal theory $T \cup \{C\}$ when $T \models C$

Entailment redundancy

$C1 = p(A,B) \leftarrow q(A,B)$

$C2 = p(A,B) \leftarrow q(A,B), r(A)$

$C3 = p(A,B) \leftarrow q(A,B), r(A), s(B,C)$

Entailment redundancy

$C1 = p(A,B) \leftarrow q(A,B)$

$\cancel{C2 = p(A,B) \leftarrow q(A,B), r(A)}$

$\cancel{C3 = p(A,B) \leftarrow q(A,B), r(A), s(B,C)}$

$$\{C1\} \models \{C2, C3\}$$

Entailment reduction

$P(A,B) \leftarrow Q(A,B)$

$P(A,B) \leftarrow Q(B,A)$

$P(A,B) \leftarrow Q(A,C), R(B,C)$

$P(A,B) \leftarrow Q(A,C), R(C,B)$

$P(A,B) \leftarrow Q(B,A), R(A,B)$

$P(A,B) \leftarrow Q(B,A), R(B,A) \quad \longrightarrow \quad ?$

$P(A,B) \leftarrow Q(B,C), R(A,C)$

$P(A,B) \leftarrow Q(B,C), R(C,A)$

$P(A,B) \leftarrow Q(C,A), R(B,C)$

$P(A,B) \leftarrow Q(C,A), R(C,B)$

$P(A,B) \leftarrow Q(C,B), R(A,C)$

$P(A,B) \leftarrow Q(C,B), R(C,A)$

Entailment reduction

$P(A,B) \leftarrow Q(A,B)$

$P(A,B) \leftarrow Q(B,A)$

$P(A,B) \leftarrow Q(A,C),R(B,C)$

$P(A,B) \leftarrow Q(A,C),R(C,B)$

$P(A,B) \leftarrow Q(B,A),R(A,B)$

$P(A,B) \leftarrow Q(B,A),R(B,A)$

$P(A,B) \leftarrow Q(B,C),R(A,C)$

$P(A,B) \leftarrow Q(B,C),R(C,A)$

$P(A,B) \leftarrow Q(C,A),R(B,C)$

$P(A,B) \leftarrow Q(C,A),R(C,B)$

$P(A,B) \leftarrow Q(C,B),R(A,C)$

$P(A,B) \leftarrow Q(C,B),R(C,A)$

$\longrightarrow$

$P(A,B) \leftarrow Q(B,A)$

$P(A,B) \leftarrow Q(A,C),R(C,B)$

Entailment reduction problem

$$C1 = P(A,B) \leftarrow Q(A,B)$$

$$C2 = P(A,B) \leftarrow Q(A,B), R(A)$$

$$C3 = P(A,B) \leftarrow Q(A,B), R(A,B)$$

$$C4 = P(A,B) \leftarrow Q(A,B), R(A,B), S(A,B)$$

Entailment reduction problem

$C1 = P(A,B) \leftarrow Q(A,B)$

$\cancel{C2 = P(A,B) \leftarrow Q(A,B),R(A)}$

$\cancel{C3 = P(A,B) \leftarrow Q(A,B),R(A,B)}$

$\cancel{C4 = P(A,B) \leftarrow Q(A,B),R(A,B),S(A,B)}$

$$\{C1\} \models \{C2,C3,C4\}$$

Entailment reduction problem

$C1 = P(A,B) \leftarrow Q(A,B)$

$\cancel{C2 = P(A,B) \leftarrow Q(A,B),R(A)}$

$\cancel{C3 = P(A,B) \leftarrow Q(A,B),R(A,B)}$

$\cancel{C4 = P(A,B) \leftarrow Q(A,B),R(A,B),S(A,B)}$

$$\{C1\} \models \{C2,C3,C4\}$$

father(A,B) ← parent(A,B), male(A) ✖

Derivation redundancy
[Cropper and Tourret, ILP18, JELIA19]

The clause C is **derivationally redundant** in the theory $T \cup \{C\}$ when $T \vdash C$

Derivation redundancy

The clause C is **derivationally redundant** in the theory $T \cup \{C\}$ when $T \vdash C$

SLD-resolution in this work

Derivation redundancy

$C1 = P(A,B) \leftarrow Q(A,B)$

$C2 = P(A,B) \leftarrow Q(A,B), R(A)$

$C3 = P(A,B) \leftarrow Q(A,B), R(A,B)$

$C4 = P(A,B) \leftarrow Q(A,B), R(A,B), S(A,B)$

Derivation redundancy

$C1 = P(A,B) \leftarrow Q(A,B)$

$C2 = P(A,B) \leftarrow Q(A,B), R(A)$

$C3 = P(A,B) \leftarrow Q(A,B), R(A,B)$

$\cancel{C4 = P(A,B) \leftarrow Q(A,B), R(A,B), S(A,B)}$

$$\{C1, C2, C3\} \vdash \{C4\}$$

father(A,B) ← parent(A,B), male(A) ✔

MLJ Paper

We compare subsumption, entailment, and derivation reduction

We theoretically show whether infinite fragments of metarules can be logically reduced to finite sets

We run the reduction algorithms on finite sets of metarules to identify minimal sets

We experimentally compare the learning performance of Metagol when supplied with reduced sets of metarule

Theoretical questions

Q. Can we reduce M to a fragment with only two body literals?

Q. Can we reduce M to a fragment with finitely many body literals?

Q. If M has a finite reduction, what is that fragment?

Idea

1. Generate big sets of metarules
2. Run the reduction algorithms on the sets
3. Study the results.

$\mathbf{H}^{\mathbf{a}}_{\mathbf{m}}$

maximum arity **a**
maximum body literals **m**

Can the connected fragment C^2_∞ be reduced to two body literals?

Can the connected fragment C^2_∞ be reduced to two body literals?

Arity	S	E	D
1	✓	✓	✓
2	✓	✓	×
>2	✓	✓	×

Can the connected fragment C^2_∞ be reduced to two body literals?

Arity	S	E	D
1	✓	✓	✓
2	✓	✓	×
>2	✓	✓	×

C^2_∞ cannot be derivationally reduced to C^2_2

Can the connected fragment C^2_∞ be reduced to two body literals?

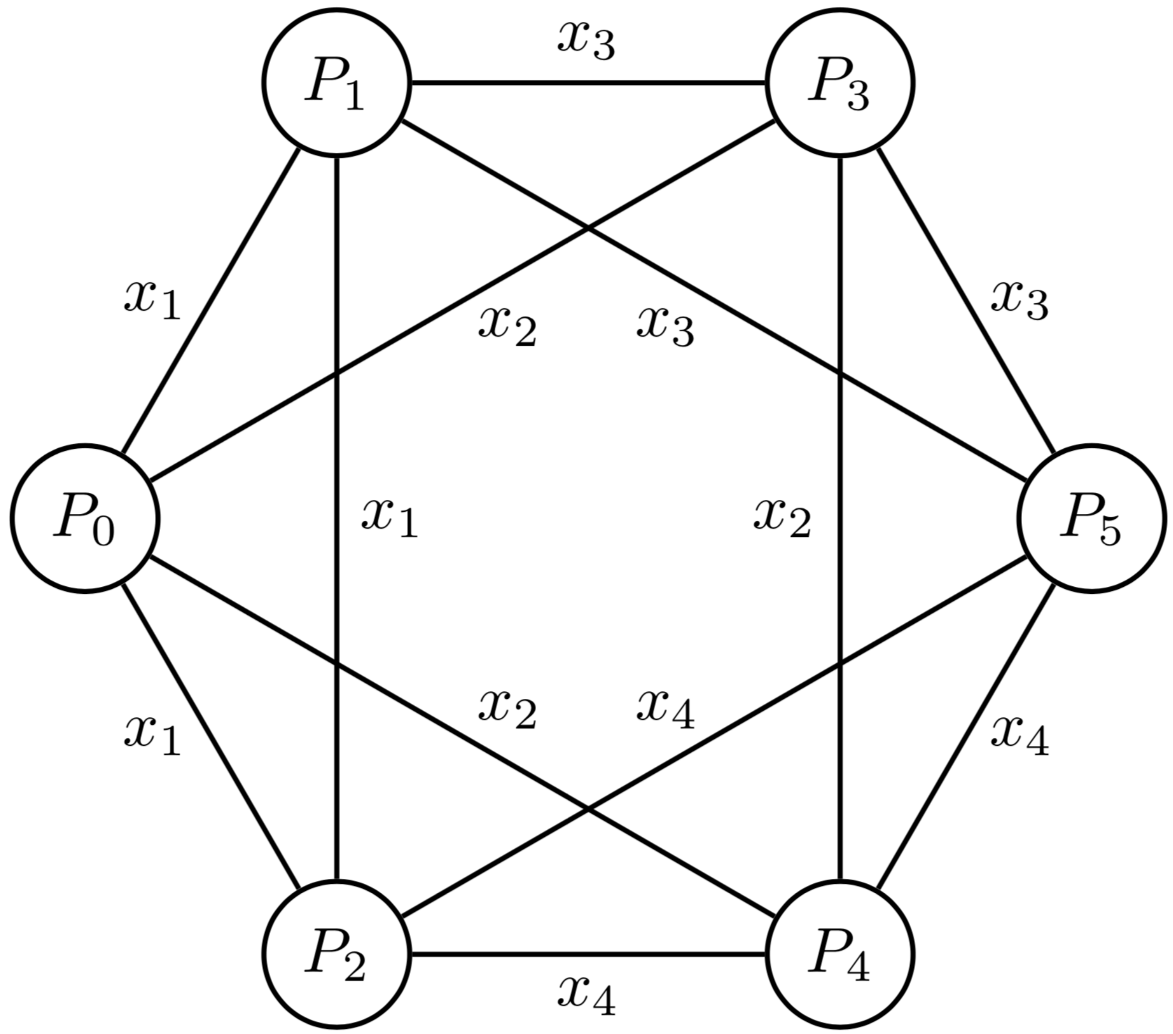
Arity	S	E	D
1	✓	✓	✓
2	✓	✓	×
>2	✓	✓	×

C^2_∞ cannot be derivationally reduced to a finite fragment!

Derivation reduction of connected fragment

$$P(A) \leftarrow Q(B,A)$$
$$P(A,A) \leftarrow Q(B,A)$$
$$P(A,B) \leftarrow Q(B)$$
$$P(A,B) \leftarrow Q(B,A)$$
$$P(A,B) \leftarrow Q(B,B)$$
$$P(A,B) \leftarrow Q(A,B), R(A,B)$$
$$P(A,B) \leftarrow Q(A,C), R(B,C)$$
$$P(A,B) \leftarrow Q(A,C), R(A,D), S(B,C), T(B,D), U(C,D)$$

Why not?



$$P_0(x_1, x_2) \leftarrow P_1(x_1, x_3), P_2(x_1, x_4), P_3(x_2, x_3), P_4(x_2, x_4), P_5(x_3, x_4)$$

Can the Datalog D^2_∞ fragment be reduced to two body literals?

Can the Datalog D^2_∞ fragment be reduced to two body literals?

Arity	S	E	D
1	✓	✓	✓
2	✓	✓	×
>2	×	✓	×

D^2_∞ cannot be derivationally reduced to a finite fragment

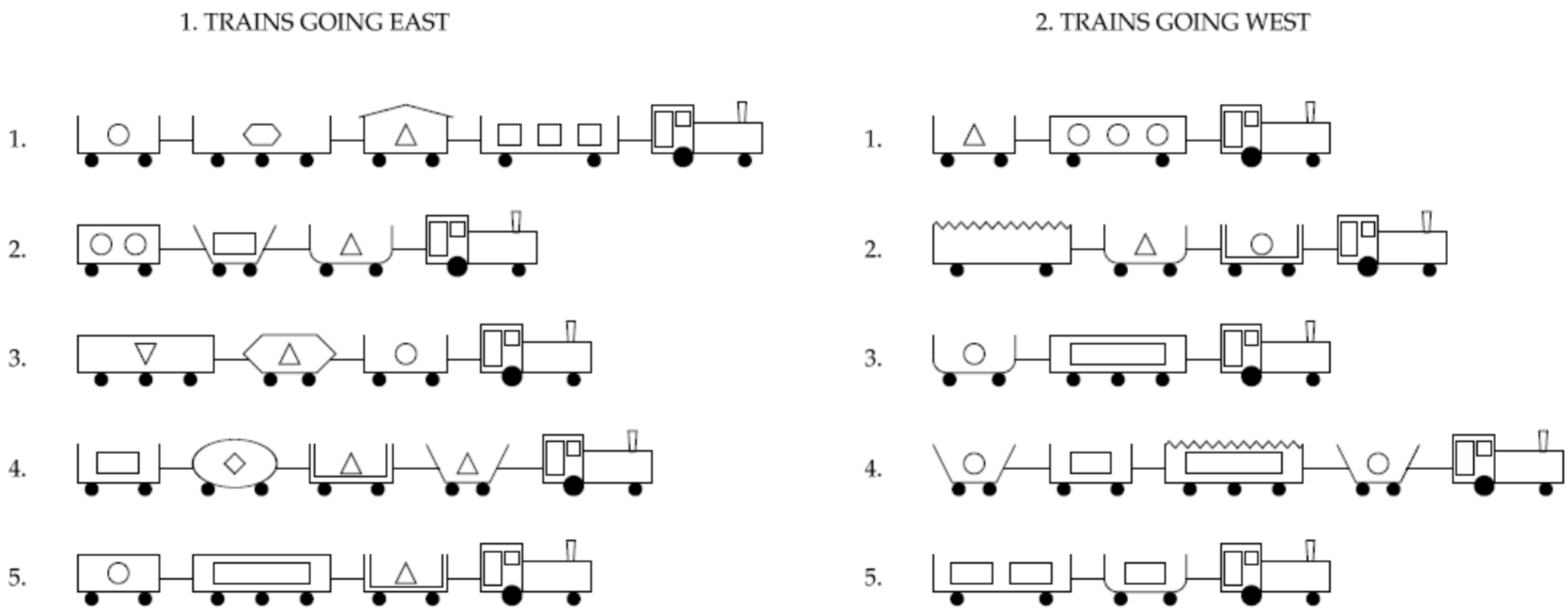
Reduction summary

Arities	$\mathscr{C}^a_\infty$			$\mathscr{D}^a_\infty$			$\mathscr{K}^a_\infty$			$\mathscr{U}^a_\infty$		
a	S	E	D	S	E	D	S	E	D	S	E	D
1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	✓	✓	×	✓	✓	×	×	✓	×	×	✓	×
>2	✓	✓	×	×	✓	×	×	✓	×	×	✓	×

Does it matter?

Is there any difference in learning performance when using different reduced sets of metarules?

Trains



Accuracies

Task	**S**	**E**	**D**	**D***
T_1	**100** ± 0	**100** ± 0	**100** ± 0	**100** ± 0
T_2	**100** ± 0	**100** ± 0	**100** ± 0	**100** ± 0
T_3	68 ± 5	62 ± 5	**100** ± 0	**100** ± 0
T_4	75 ± 6	75 ± 6	**100** ± 0	**100** ± 0
T_5	92 ± 4	78 ± 6	78 ± 6	**100** ± 0
T_6	52 ± 2	50 ± 0	70 ± 6	**100** ± 0
T_7	95 ± 3	65 ± 5	82 ± 5	**100** ± 0
T_8	55 ± 3	52 ± 2	72 ± 6	**98** ± 2
mean	80 ± 1	73 ± 2	88 ± 2	**100** ± 0

Learning times

Task	**S**	**E**	**D**	**D***
T1	**0** ± 0	**0** ± 0	**0** ± 0	**0** ± 0
T2	**0** ± 0	**0** ± 0	**0** ± 0	**0** ± 0
T3	424 ± 59	461 ± 56	**0** ± 0	**0** ± 0
T4	322 ± 64	340 ± 61	**0** ± 0	**0** ± 0
T5	226 ± 48	320 ± 59	361 ± 59	**5** ± 2
T6	583 ± 17	600 ± 0	429 ± 51	**7** ± 2
T7	226 ± 44	446 ± 55	243 ± 61	**6** ± 1
T8	550 ± 35	570 ± 30	361 ± 64	**183** ± 40
mean	292 ± 16	342 ± 17	174 ± 16	**25** ± 5

String transformations

Input	Output
Arthur Joe Juan	AJJ
Jose Larry Scott	JLS
Kevin Jason Matthew	KJM
Donald Steven George	DSG
Raymond Frank Timothy	RFT

String transformations

	S	**E**	**D**	**D***
Mean predictive accuracy (%)	22 ± 0	22 ± 0	32 ± 0	**56 ± 1**
Mean learning time (seconds)	467 ± 1	467 ± 1	407 ± 3	**270 ± 3**

Inducing game rules

GT attrition	GT chicken
GT prisoner	Minimal decay
Minimal even	Multiple buttons and lights
Scissors paper stone	Untwisty corridor

Inducing game rules

	S	**E**	**D**	**D***
Balanced accuracy (%)	66	66	72	**73**
Learning time (seconds)	316	316	327	**296**

Conclusions

New form of logical reduction

Negative theoretical results (especially for MIL)

Little impact on practical performance

Todo

Overcome negative theoretical result

Expand results to higher-arities

Identify domain-specific sets of metarules

Identify optimal sets of metarule

Logically reduce BK