

Which background knowledge is relevant?

Stassa Patsantzis, Stephen Muggleton

Imperial College London

ep2216@ic.ac.uk, s.Muggleton@imperial.ac.uk

Which background knowledge is relevant?

- Background knowledge in ILP is typically selected by human experts
- How to select appropriate background knowledge automatically?

Which background knowledge is relevant?

- What is "relevant" background knowledge?
- Intuitively, a background predicate is irrelevant to a target concept if a hypothesis about the concept can be learned without that predicate.
 - More formally:
$$\text{irrelevant}(p) : H \cup B \setminus \{p\} \models E^+ \wedge H \cup B \setminus \{p\} \not\models E^-$$
- A relevant predicate is one that is not irrelevant!

Which background knowledge is relevant?

- Naive algorithm:
 - Start with some set of background knowledge
 - Remove a predicate and try to learn a hypothesis
 - If a hypothesis can be learned, continue with another predicate
 - Else put the predicate back and repeat
- Very inefficient!

Which background knowledge is relevant?

- Is there another way to figure out the relevance of a predicate?

- Predicate generality:

$$g(Q/n) = \Pr(Q(x_1, \dots, x_n) \mid \text{random} \{x_1, \dots, x_n\})$$

- Informally: the probability that a random atom from the Herbrand base of Q/n is true.

Which background knowledge is relevant?

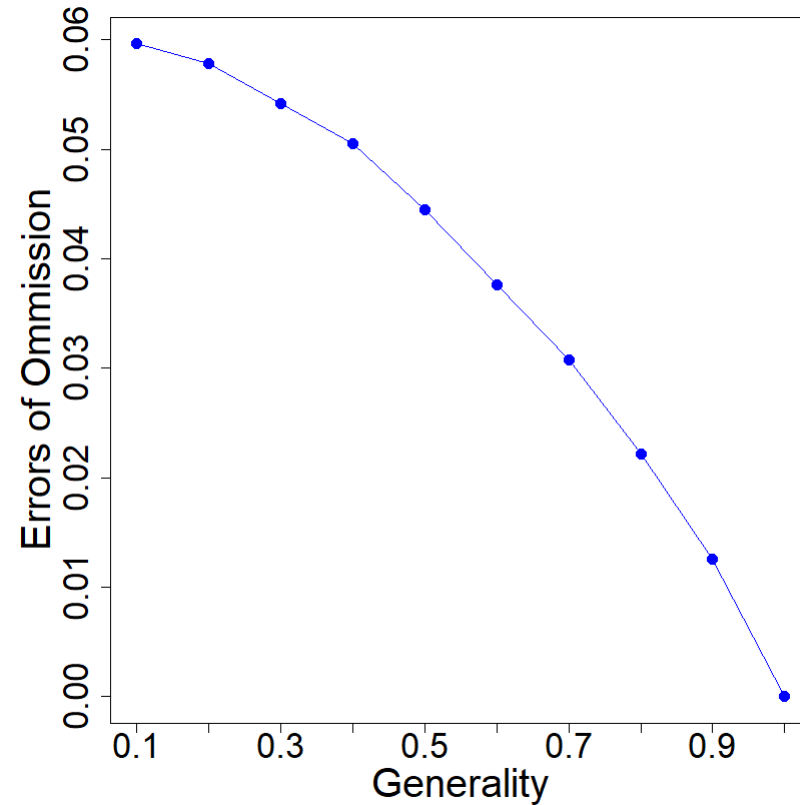
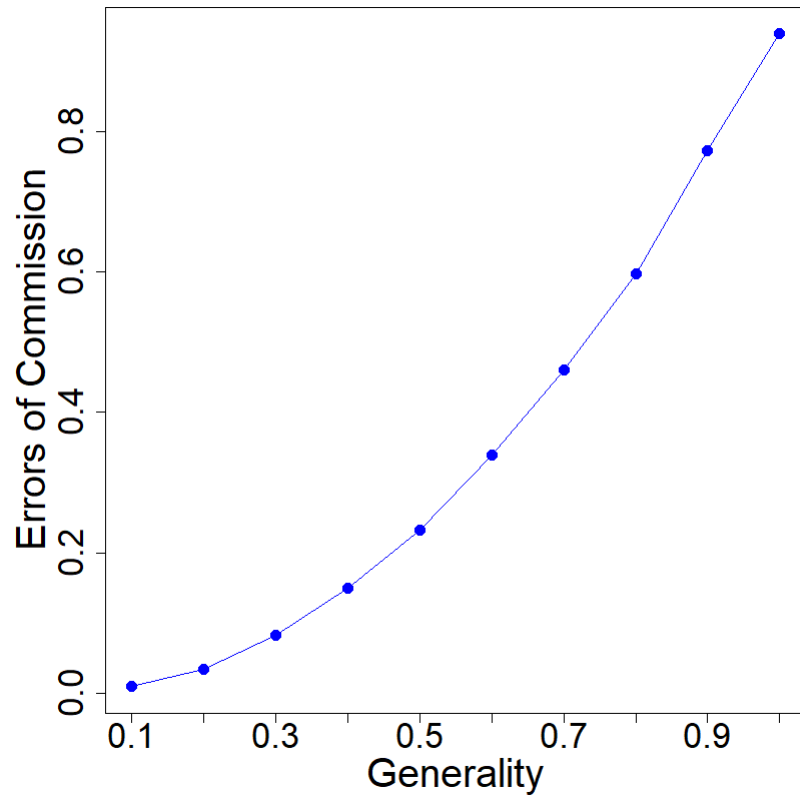
- Why is generality informative of relevance?
 - A predicate of generality 0 is a contradiction- it can't be used to learn a valid hypothesis (with Metagol).
 - A predicate of generality 1 is a tautology- it is not needed to learn a valid hypothesis.
- The generality of the most relevant predicates should be somewhere in the middle.

Which background knowledge is relevant?

- How does background generality affect hypothesis performance?
- Experiments:
 - We generated random atoms with a fixed generality (by sampling from their Herbrand base).
 - Sampled atoms from a background predicate with probability equal to the desired generality class:
 $\langle 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 \rangle$
 - Background predicate with a randomly generated symbol.
 - Atoms sampled with probability equal to desired generality.
 - Background knowledge: [father/2, mother/2, <random predicate>]
 - Learn grandfather/2 from 0.2 sample of positive examples
 - Evaluate hypothesis on remaining positive and all negative examples.

Which background knowledge is relevant?

- As generality increases, Errors of Commission (rate of false positives) increases.
- At the same time, Errors of Omission (rate of false negatives) decreases.



Which background knowledge is relevant?

- Experiment results suggest a search procedure:
 - Order the hypothesis space by generality class of background predicates
 - Try the hypotheses with least generality first
- Coupled with iterative deepening over hypothesis size this should find the shortest, least general hypotheses first.
- Current direction of our research: generality ordered search in Metagol.
 - Work in progress!!

Which background knowledge is relevant?

- Thank you!
- Please feel free to ask questions!