Informatics 1 CG – Tutorial 6

Chris Lucas

Week 7

1. Category-based induction

In the lectures we covered category-based induction, or generalising on the basis of category-level information. For instance, we can use our beliefs about different kinds of animals to make judgments about whether a newly discovered feature will be present in a new animal or category of animals. In this tutorial, we will go through several examples of category-based induction problems in the style of Osherson et al. (1990), to (1) develop our intuitions for some of the underlying phenomena; (2) understand what we might want from a good theory of category-based induction; and (3) connect category-based induction to ideas about uncertainty and probability.

Single-premise inferences

Consider the following premise/conclusion pairs, in the syntax used by Osherson et al. and others.

```
Owls have protein P        Owls have protein P
--------------------       ----------------------
Hawks have protein P       Rabbits have protein P


Owls have protein P        Birds have protein P
--------------------       --------------------
Birds have protein P       Owls have protein P


Owls have protein P        Owls have protein P
----------------------     --------------------------
Animals have protein P     Sea sponges have protein P
```

```
Owls have protein P
-----------------------------------
Mathilda the pygmy owl has protein P
```

Exercises

1. For each of the premise/conclusion pairs above, how strongly does the premise support the conclusion? Rank each premise/conclusion pair according to your own intuitions.
    a. Do your rankings align with the phenomena summarised in Osherson et al. (1990)? Were there any surprising or interesting points of disagreement, or aspects Osherson et al. didn't address?
    b. How did you interpret the premises and the conclusions? For example, do you take "Owls have protein P" to mean that every owl must have protein P? If so, how and when are these premise/conclusion pairs useful – are they relevant to our everyday intuitions about animals and their characteristics? If not, explain how you interpret a conclusion like "Owls have protein P."
2. How might you use a similarity-based model to make judgments about a conclusion given its premises? There are several possible answers to this; sketch out an approach that makes sense to you.

Knowledge effects

Consider the premise/conclusion pairs below, filling in X as appropriate (see the exercises for details).

```
Sea eagles X      Rabbits X
--------------    -----------
Owls X            Owls X


Ostriches X       Kittens X
--------------    -----------
Owls X            Owls X
```

Exercises

1. For each of the premise/conclusion pairs above, how strongly does the premise support the conclusion? Rank them, filling in X in the following ways:
    a. What if X is "carry the parasite Q"?
    b. What if X is "produce the hormone H"?
    c. What if X is "consume 10-20 units of vitamin B12 on an average day"?
2. Did your ranks differ for different X? Why or why not?

2. Theory of mind

"Theory of mind" is an overarching term for beliefs or knowledge about the mental states of agents, and how those mental states relate to an agent's behaviour. Mental state reasoning is a classic example of an inductive problem; it involves using overt behavior to make inferences about variables that are hidden from direct observation, and using those inferences in turn to predict future behaviour. These inferences are often uncertain and subject to disagreement. For example, people might disagree about whether a politician's statements reflect his or her actual beliefs.

Exercises

1. Give an example of a scenario where two different people might disagree about what a particular behaviour reveals about a third person's mental states, where the behaviour is:
    a. A choice among different options, e.g., on a restaurant menu, among different means of transportation, or when purchasing a retail good.
    b. A statement in natural language.
    c. A physical action, e.g., walking to a location, looking at something, or a gesture.
2. For each of the above examples, suggest a piece of evidence that would reduce the ambiguity or potential for different interpretations.