An Opening Exercise

Consider the following sentence:

> "Last month, the stock market was like a roller coaster."

Suppose that you have a program that knows about roller coasters.

After reading and understanding the above sentence, what should it know about the stock market?

How will it figure that out?

On Learning

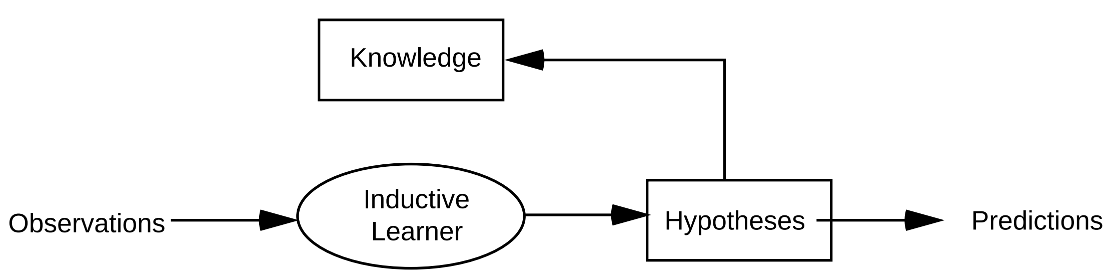
The forms of learning that we have studied so far, induction and reinforcement learning, both operate on the principle of finding meaningful patterns in a set of data.

Today, let's consider some of the shortcomings of such learning, with induction as our running example.

The basic form of inductive learning is:

$$\underbrace{\text{Hypothesis}}_{\textit{goal}} \cap \underbrace{\text{Descriptions} \Rightarrow \text{Classifications}}_{\textit{inputs to the learner}}$$

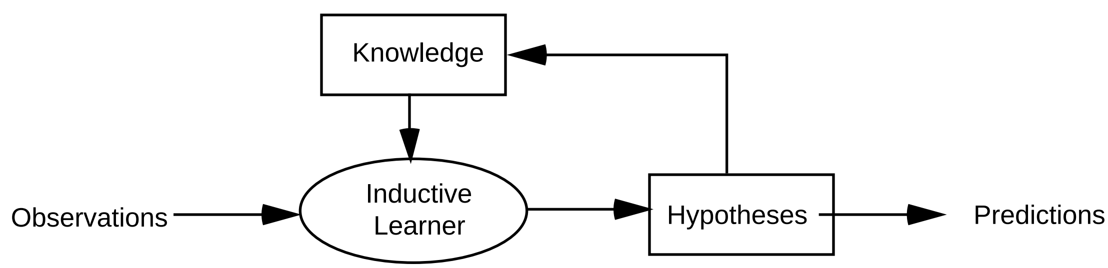
... which looks like this:



On Learning

Why shouldn't the learner use what it learns on previous trials as a basis for future learning?

In practice, an agent is more likely to do induction as:



The use of background knowledge allows much faster — and more *consistent* — learning than one can expect from a batch induction program.

If the knowledge is "good", then it can drastically cut the search space for future learning...

What if the learned knowledge is "bad"?

Another Exercise

Identify as many ways as possible that we could use knowledge to improve our induction algorithm.

For each suggestion you make, provide:

1. the step in the algorithm being improved

2. the kind of knowledge being used

3. a description of how using this knowledge will make the algorithm learn better

Some possibilities:

- compute better default values
- handle cases of incomplete data
- handle conflicting cases
- redefine "best discriminates"

Toward Knowledge-Based Learning...

Leaping to a general conclusion based on one observation

Gary Larson once drew a cartoon in which a geek caveman named Zog is roasting his lizard on the end of a pointed stick. Watching him in amazement is a crowd of his less intellectual peers, who have been using their bare hands to hold their food over the fire. This enlightening experience convinces the onlookers of the general principle of painless cooking.

Explanation-Based Learning

The cavemen are able to learn the principle of painless cooking because they can **explain to themselves** *why* what they see works.

The general rule follows from background knowledge:

$$\text{Hypothesis} \cap \mathit{Descriptions} \Rightarrow \mathit{Classifications}$$
$$\boldsymbol{\mathit{Background}} \Rightarrow \textbf{Hypothesis}$$

By modus ponens, we see that:

$$\boldsymbol{\mathit{Background}} \cap \mathit{Descriptions} \Rightarrow \mathit{Classifications}$$

The agent learns nothing factually new! The agent could have derived the rule from what it already knew. However, to do so might have required either an inordinate amount of computation or dumb luck.

Toward Knowledge-Based Learning...

Drawing one general conclusion but not drawing another

Lifelong Iowa resident Ida Mae travels to Brazil, where she meets her first Brazilian.

When she hears him speak Portuguese, she concludes immediately that Brazilians speak Portugese.

When she discovers that his name is Fernando, she does not conclude that all Brazilians are named Fernando.

Relevance-Based Learning

Our traveller is able to learn the common language but not the common name because she knows that some features are common to groups, and some are not.

Sometimes, a general rule follows from background knowledge plus the classification itself:

$$\text{Hypothesis} \cap \mathit{Descriptions} \Rightarrow \mathit{Classifications}$$

$$\boldsymbol{Background} \cap \boldsymbol{Descriptions} \cap \boldsymbol{Classifications} \Rightarrow \textbf{Hypothesis}$$

Again, the agent learns nothing factually new! The agent could have derived the rule from what it already knew. The process here might be more complex, but the reasoning would be the same.

Relevance-Based Learning in Science

RBL plays an important role in any scientific endeavor.

Consider a physics student doing a lab experiment.

- She measures the density and conductance of a sample of copper at a particular temperature and confidently generalizes those values to all pieces of copper.

- She measures its mass and doesn't even consider the hypothesis that all pieces of copper have the same mass.

  On the other hand, it would be quite reasonable to make such a generalization over the set of all pennies!

How can science be objective when it benefits from the scientist using her knowledge in the process of building theories? (Or, worse, *requires* her to use her knowledge!)

Learning "Nothing New"

In both explanation-based learning and relevance-based learning, the agent learns nothing "new". She deduces a new rule from what she already knows.

However, the deduction is **guided by observations**, making it:

- computationally feasible where it otherwise might not have been
- worth deducing and storing, because the agent can use the knowledge in practice

These days, EBL and RBL are viewed as techniques for converting knowledge of "first principles" into knowledge that is **useful** for a set of tasks.

This process of converting basic knowledge into applied knowledge is sometimes called ***compilation***. EBL and RBL are practical examples of how an agent can learn something by reasoning deductively.

Toward Knowledge-Based Learning...

Filling in a gap in a reasoning process

Jane, a medical student, has mastered her courses on how to do diagnosis but has dropped the ball on learning her pharmacology. She observes a session between a patient and an expert internist. After a series of questions and answers, the expert tells the patient to take a particular antibiotic.

The medical student infers the general rule that this particular antibiotic is effective for a particular type of infection.

Knowledge-Based Inductive Learning

This student is able to learn a new rule out of the "noise" of observations because she has knowledge of the process that is being observed and because she is aware of a gap in her knowledge.

The general rule must be inductively learned, but the learner uses background knowledge in the process:

$$\textbf{\textit{Background}} \cap \textbf{Hypothesis} \cap \textbf{\textit{Descriptions}} \Rightarrow \textbf{\textit{Classifications}}$$

KBIL's use of background knowledge constrains the hypothesis search space to those hypotheses that are consistent with what is already known.

KBIL's use of background knowledge means that the hypotheses need not explain the whole observation — and thus can be correspondingly smaller.

The Objectivity of Learning

Even within a single theory, any number of hypotheses may explain a sequence of data.

So: Every learning agent needs a <u>bias</u>.

How objective is learning?

How objective is science?

Toward Knowledge-Based Learning...

Reasoning from examples rather than rules

Last month, the stock market was like a roller coaster.

An agent that reasons by analogy to other situations can learn in a very simple way: store more situations.

This sort of reasoning shifts the burden of learning from the learning process to the problem solving process. Generalization is done at “run time”, when solving a specific problem.

Case-Based Reasoning

This is a *heuristic* form of reasoning. It generates solutions that are plausible, but probably not solutions that are optimal.

In order to reason by analogy, an agent must be able to determine the **degree of similarity** between particular observations.

Then, after “deriving” a hypothesis and testing it, the situation is stored, along with its hypothesis and result, as another case for future use.

Thus, learning is (merely) the result of storing a new example in memory.