Knowledge-based systems and the need for learning

The implementation of a knowledge-based system can be quite difficult. Furthermore, the process of reasoning with that knowledge can be quite slow.

This raises two questions:

1. Is there an easier way to gather all the information that needs to be encoded in a knowledge base?
2. Is it possible to "speed up" the process of reasoning?

Can Computers Learn?

Learning a new set of facts.

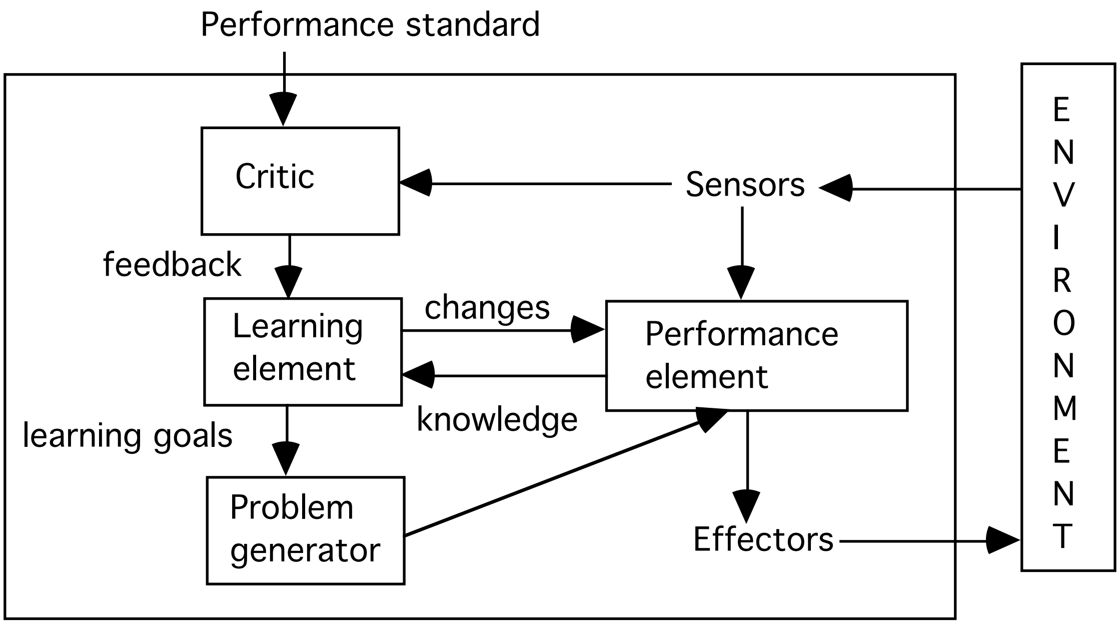
Learning how to do something.

Improving on something that was already learned.

Simon's Definition of Machine Learning

"Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time." [*Machine Learning I*, 1983, Chapter 2.]

Below is a diagram that shows one way that a learning element can fit into an AI system. Not every component or connection shown below is necessarily present in every system.



[Figure from Russell and Norvig]

Design of a Learning Element

There are many issues that impact the design of a learning element, including:

- Which components of the performance element are to be improved.
- What representation is used for those components.
- What feedback is available.
    - Supervised learning
    - Unsupervised learning
    - Reinforcement learning
- What prior information is available.

Inductive Concept Learning

Of the general learning tasks we saw above, this falls into the category of "learning new facts"

Consider the following example: Say that we're physicians who are treating patients with a particular type of cancer. We'd like to predict, among other things, whether a patient will have a recurrence of the disease. Let's assume that we have at our disposal the medical records of other patients. If we've tracked the health of these patients for many years, we know which patients have had a recurrence and which have not. Perhaps we can analyze the records to determine the attributes that might be predictive of recurrence.

Inductive – algorithm tries to induce a "general rule" from a set of observed instances; does not make use of explicit background knowledge about the domain of application.

Supervised – the learning algorithm is given the correct concept label for each input example -- in the example above, the label "recurrence" or "not" for each patient.

A Simple Example: Work or Play?

	Outlook	Temperature	Humidity	Windy?	Plan
Example 1	Sunny	Mild	High	False	Play
Example 2	Sunny	Hot	High	True	Play
Example 3	Overcast	Hot	High	False	Work
Example 4	Rainy	Mild	High	False	Work
Example 5	Rainy	Cool	Normal	False	Work
Example 6	Rainy	Cool	Normal	True	Work
Example 7	Sunny	Hot	Normal	False	Play
Example 8	Sunny	Cold	Normal	True	Work

Each example (instance) is a possible day, described by the features Outlook, Temperature, Humidity, Windy

The target concepts to be learned are "a good day to work" and "a good day to play"

Examples of Application Domains

Medical diagnosis:
- patient symptoms and tests → has disease / does not have disease
- patient symptoms and tests → disease

Other diagnostic applications, such as diagnosis of customer-reported telephone problems.

Stock market prediction:
- closing price of last few days → market will go up or down tomorrow

Cell phone fraud detection:
- characteristics of the call (destination, location of caller, duration, etc) → fraudulent / not fraudulent

---

Face recognition for security:
- facial characteristics → identification of individual

Vehicle steering:
- image viewed from camera → direction and degree to turn the wheel

---

Inductive Concept Learning

Induction task:

Given: a collection of examples described by attributes, each labeled with a class name

Return: a classification "rule" that can determine the class of any object from its attributes' values.

Inductive Bias

Recall the claim that these algorithms make no use of explicit background knowledge of the domain of application. This is true. However, it would be inappropriate to claim that these algorithms are completely "knowledge-free."

The attributes describing the examples are generally fixed.
The representation of the knowledge to be learned is set by the algorithm.
Examples may have been selected in a biased manner.
Etc.

Implicit knowledge can play an important role in many other ways as well.

Def. Any preference for one hypothesis over another, beyond mere consistency with the examples, is called a **bias**.

Inductive Learning Hypothesis

Any hypothesis found to approximate the target concept well over a sufficiently large set of training examples will also approximate the target concept well over any other unobserved examples.

Assumptions of classifier learning algorithms:
- The sample represents the population.
- The features permit discrimination.