

## **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.

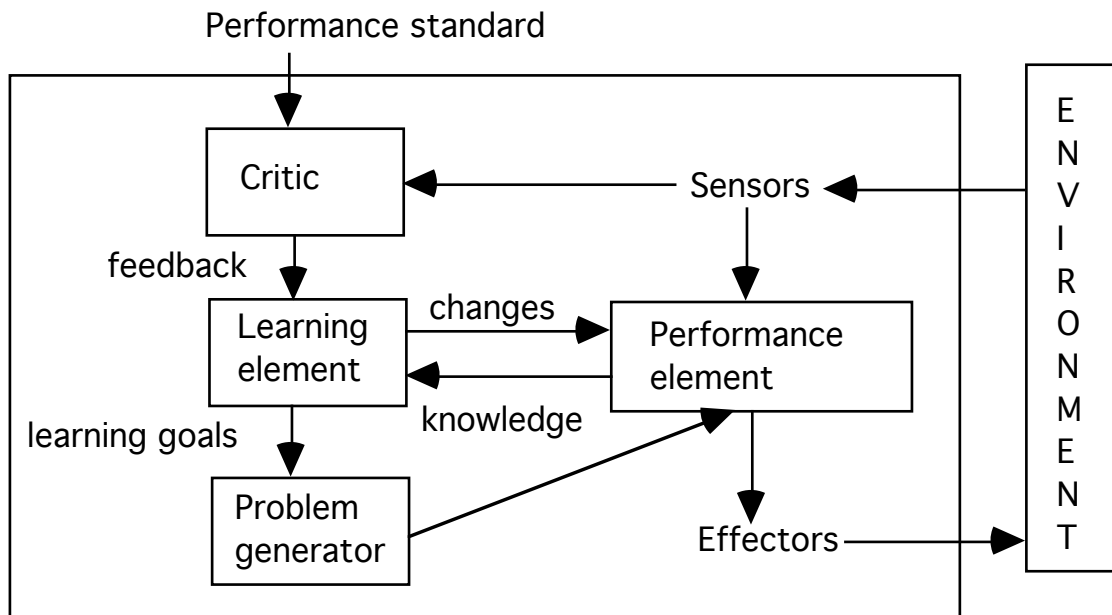
Learning how to do something.

Improving the ability of something 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.]




---

### 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.

(First three from Russell and Norvig)

---

### Inductive Concept Learning

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

Learning here is viewed as fitting a function to observed data.

**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 value of the function for particular inputs, and changes its representation of the function to try to match the information provided.

An example is a pair  $(x, f(x))$ , where  $x$  is the input and  $f(x)$  is the output of the function applied to  $x$ .

---

**Example: Send patient home from hospital post-op?**

Major Operation?	Family at Home?	Old?	Send Home?
Yes	No	Yes	No
Yes	No	No	No
No	No	Yes	No
No	Yes	Yes	Yes
No	No	No	Yes

Each example (instance) is a patient case described by the attributes Major Operation?, Family at Home?, and Old?

The target concept,  $f$ : patient case  $\rightarrow$  {Yes, No}

---

**Examples of Application Domains**

**Medical diagnosis:**

- patient symptoms and tests  $\rightarrow$  has disease / does not have disease
- patient symptoms and tests  $\rightarrow$  disease

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

**Stock market prediction:**

- closing price of last few days  $\rightarrow$  market will go up or down tomorrow

**Cell phone fraud detection:**

- characteristics of the call (destination, location of caller, duration, etc)  $\rightarrow$  fraudulent / not fraudulent

**Face recognition for security:**

- facial characteristics  $\rightarrow$  identification of individual

**Vehicle steering:**

- image viewed from camera  $\rightarrow$  direction and degree to turn the wheel
-

---

## Building Classifiers

Learn about the domain. Write a program that maps inputs to outputs.

Example. Telephone trouble diagnostic system used by Nynex, Bell Atlantic, Verizon...

Automate the process using data in the form of observations  $(x, f(x))$

---

### (Supervised) Inductive Concept Learning revisited

#### Induction task:

**Given:** a collection of examples

**Return:** a function  $h$  (hypothesis) that approximates  $f$  (the target concept)

OR

**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 that earlier we said 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 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 function well over a sufficiently large set of training examples will also approximate the target function well over any other unobserved examples.

Assumptions of classifier learning algorithms:

- The sample represents the population.
- The features permit discrimination.