CS 461: Machine Learning Principles

Class 1: Sept. 5

Introduction & Course Logistics

Instructor: Diana Kim

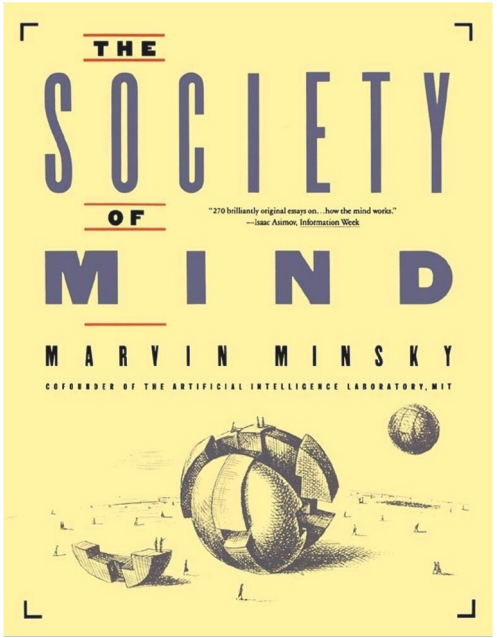
Q: What is learning?
How do we learn to swim?

+ change in mind and physical state
+ positive/ useful change
+ learning from repeats/ trials / following instructions/ practices/ and reduces errors gradually
+ continual process: memorize (can do it today, and can do it tomorrow too) ,
adapt to new situation (can swim in the indoor pool, in the lake, in the ocean)

Learning is
useful change in knowledge / behavior as a result of experience.

ENCYCLOPÆDIA Britannica

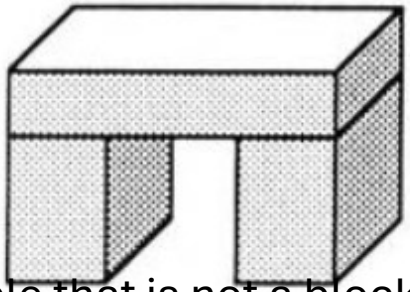
learning, the alteration of behaviour as a result of individual experience. When an organism can perceive and change its behaviour, it is said to learn.



Learning is making useful changes in the workings of our minds.

Learning Example:
(from observation to forming the mental descriptions for the concept of Arch)

“A Block-Arch Scenario”

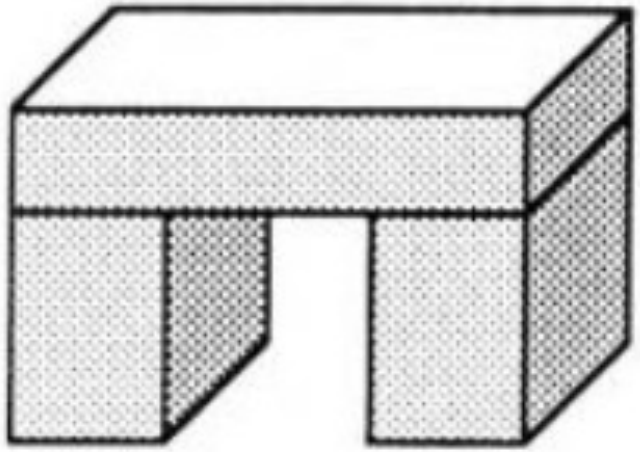


+ this example shows the learning process:
  building new mental descriptions → corresponding samples → check ”Arm-Change”

+ but what if “Arm-Change” is a necessary condition for block arch, so there exist a sample that is not a block arch?
  Think about the case classification is trained based on precision or recall objective.
  How it is important to set a right metric in learning?

from the Society of Mind by Marvin Minsky (1988)

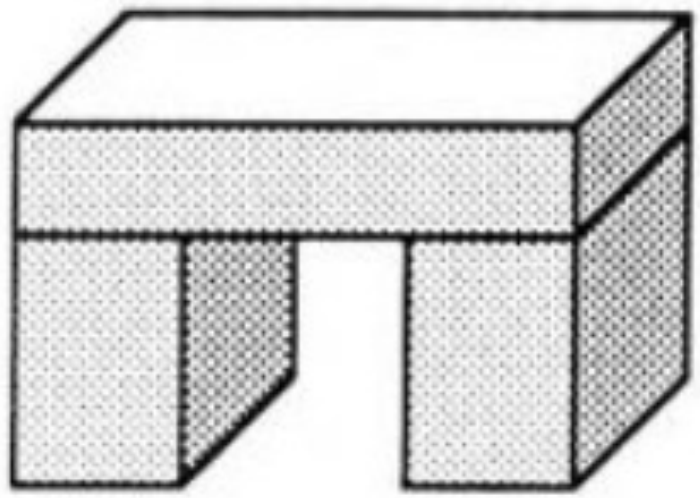
The Observations of “Hand-Change”:



Arch seems to cause a strange phenomenon:
when you push the car through it, your arm gets trapped.
You must release the car, and reach around to the other side of arch by changing hands.
Let’s call it “*Hand-Change*” phenomenon.

from the Society of Mind by Marvin Minsky (1988)

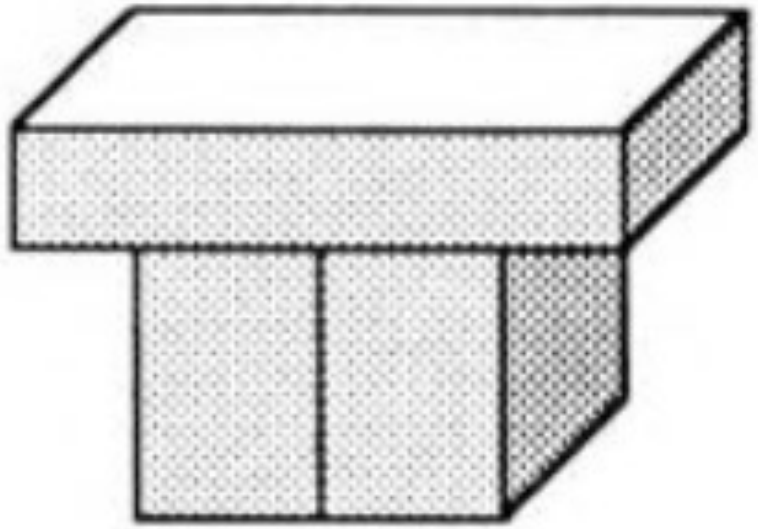
The block arch seems to cause the *Hand-Change* phenomenon.
Learning problem: Hmmm, what is the concept of Arch?



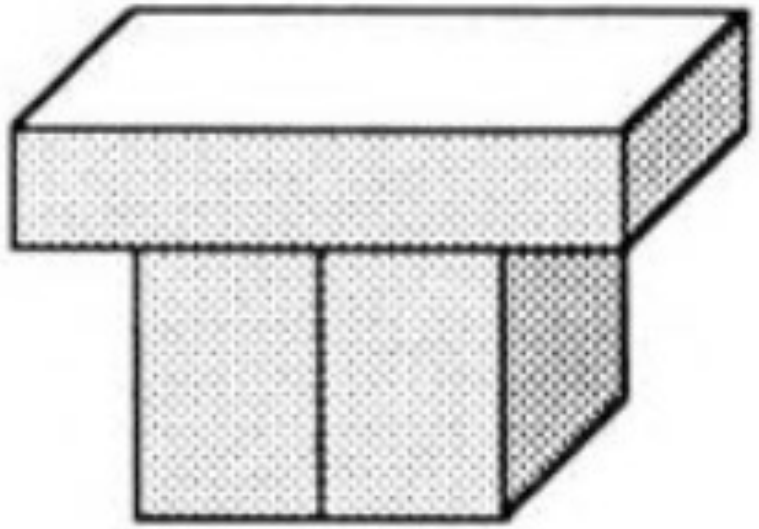
from the Society of Mind by Marvin Minsky (1988)

[Mental description 1]
 + Two standing blocks and a lying block.

[Mental description 1]
+ Two standing blocks and a lying block.



[Mental description 1]
+ Two standing blocks and a lying block.



hand-change disappears

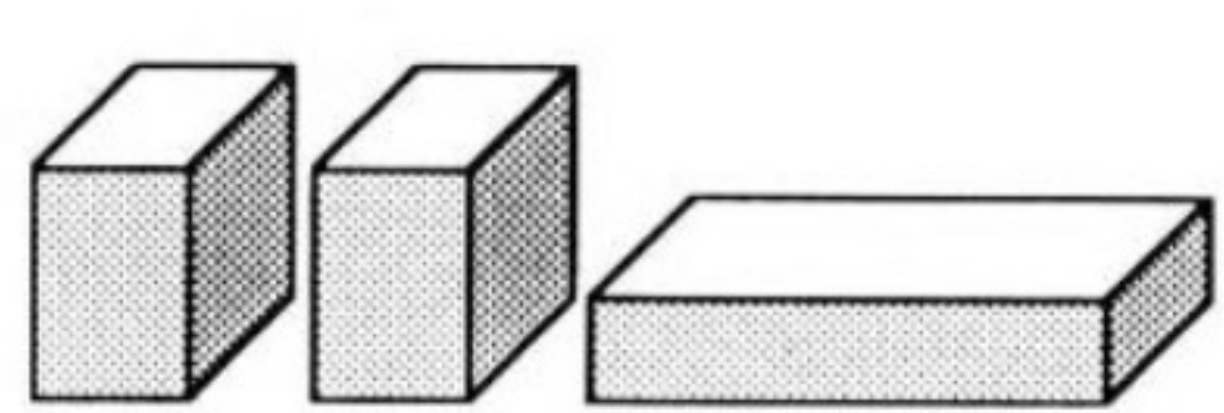
Update the mental description
to avoid the example failing in hand-change.

[Mental description 1]
+ Two standing blocks and a lying block.

[Mental description 2]
+ The standing blocks must not touch.

[Mental description 1]
+ Two standing blocks and a lying block.

[Mental description 2]
+ The standing blocks must not touch.



Hand-change disappears

Update the mental description
to avoid the examples failing in hand-change.

[Mental description 1]
+ Two standing blocks and a lying block.

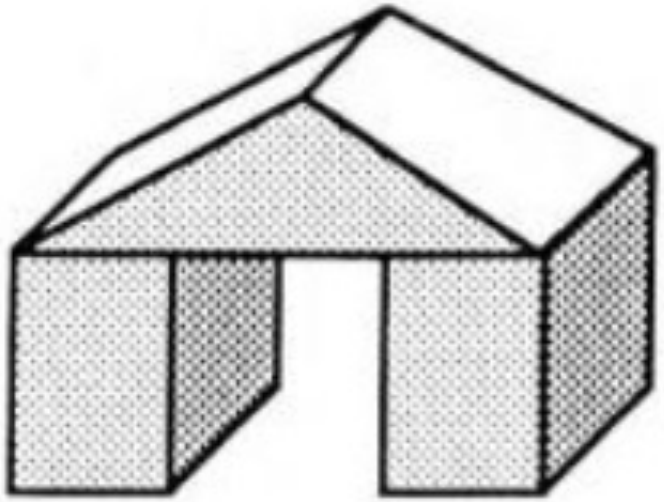
[Mental description 2]
+ The standing blocks must not touch.

[Mental description 3]
+ They must support the lying block.

[Mental description 1]
+ Two standing blocks and a lying block.

[Mental description 2]
+ The standing blocks must not touch.

[Mental description 3]
+ They must support the lying block.



✓

Another variation
But it produces hand-change.

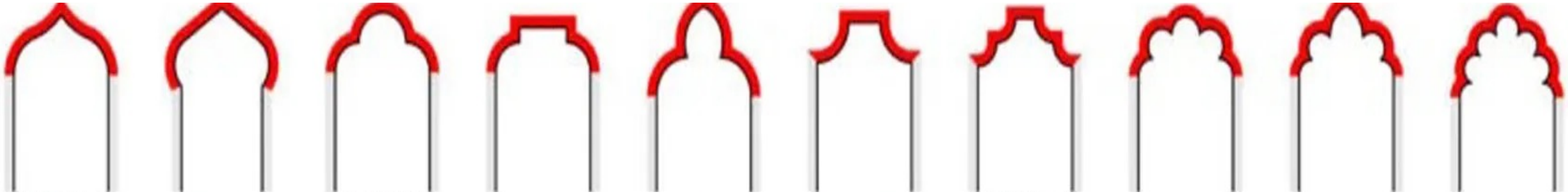
[Finalized Mental Descriptions about Block-Arch]

+ Two standing blocks and a lying block.
+ The standing blocks must not touch.
+ They must support the lying block.
+ The other things may be a wedge or a block.

Based on the descriptions,
we can recognize the various types of block-arches.

[Finalized Mental Descriptions about Block-Arch]

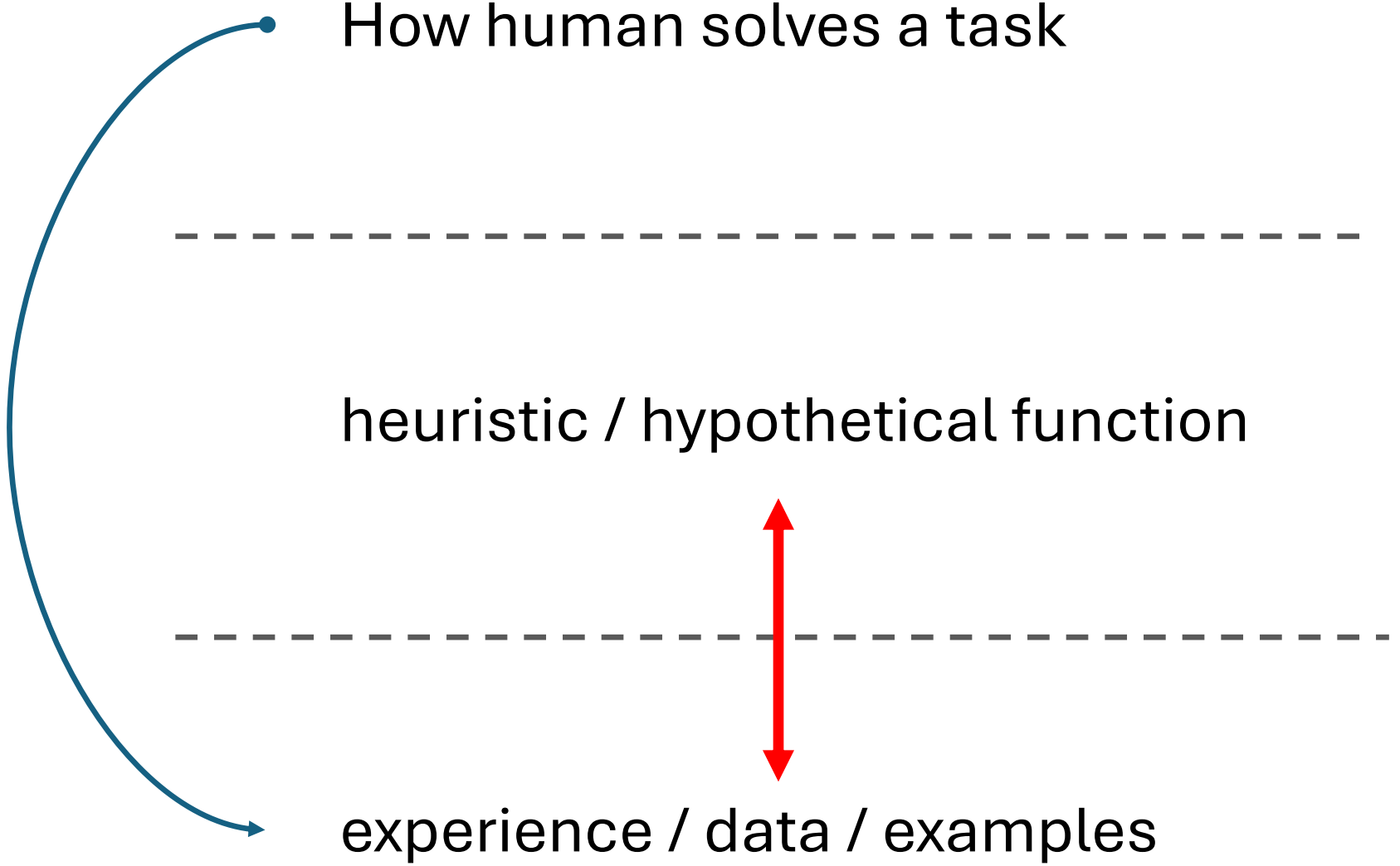
+ Two standing blocks and a lying block.
+ The standing blocks must not touch.
+ They must support the lying block.
+ The other thing may be a wedge or a block.



Machine Learning?

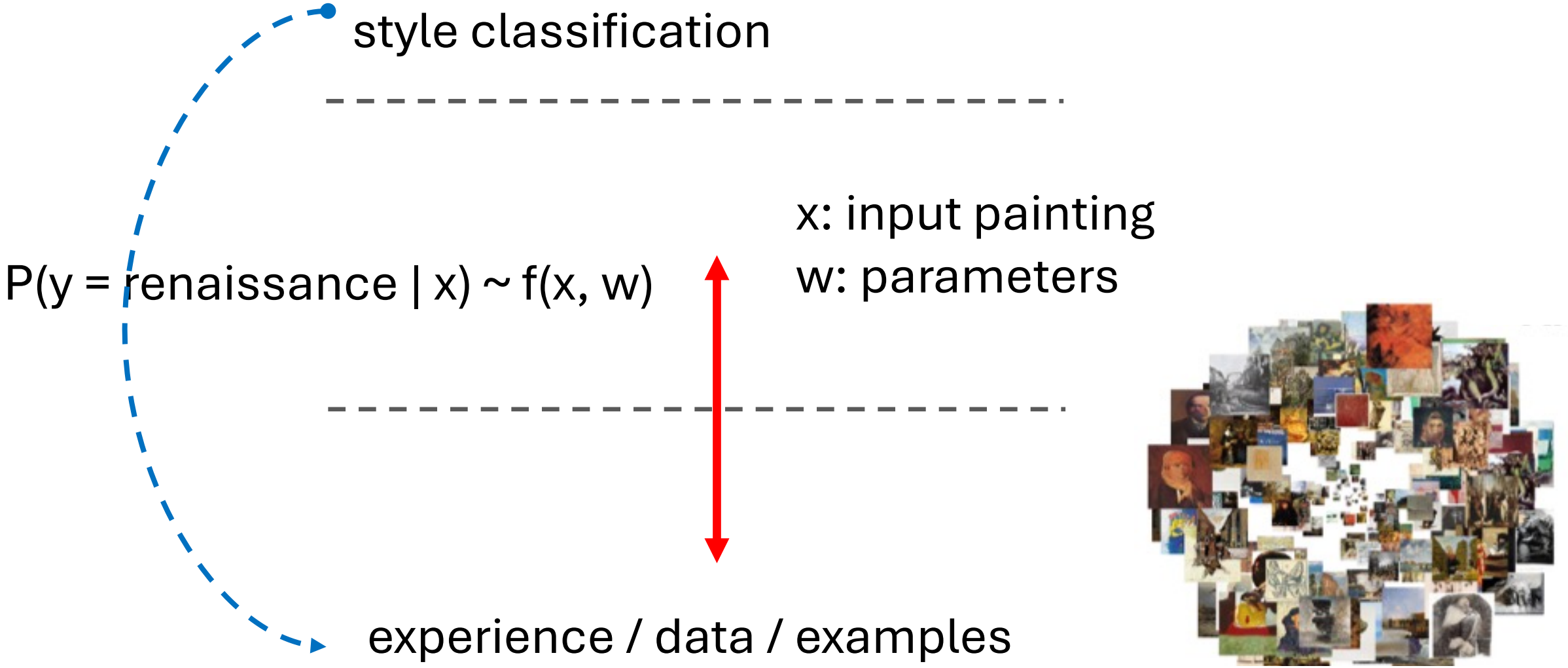
+ heuristic approach
+ a heuristic and mathematical function for a task
+ we don't know what function will be optimal, so we set a hypothetical model
+ the hypothetical model will be updated based on examples and performance metric (no direct instruction from programmers)
+ this course will be about: as you develop a ML system,
  (1) what tasks are aimed to
  (2) what mathematical models we can consider for the task
  (3) learning algorithms
  (4) evaluation: how can we improve generalization for unseen data?

Machine Learning?

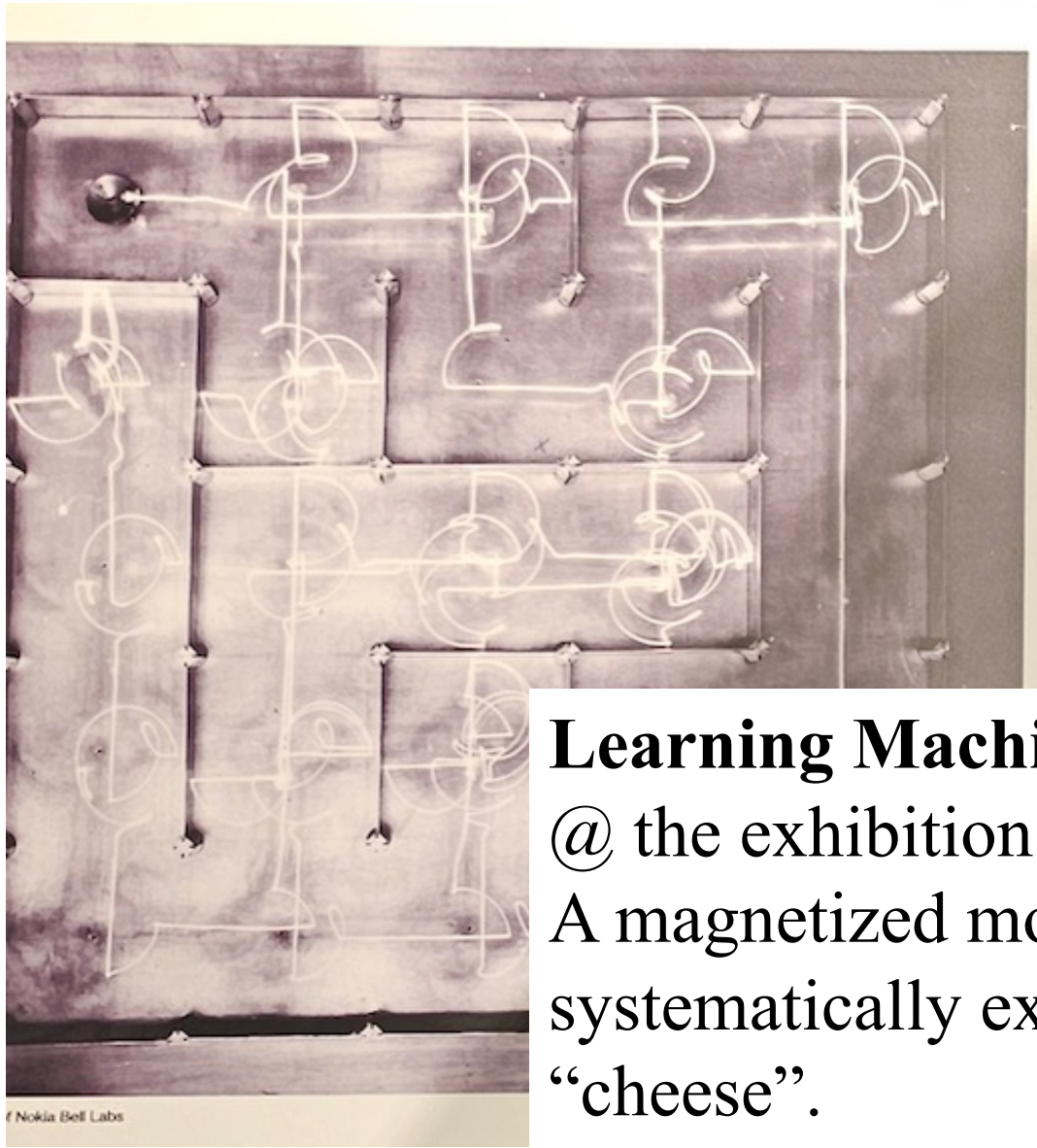


Machine Learning?

The heuristic function is kept updated based on the past examples.



Early Example: The Learning Machine by Claude Shannon (1950)

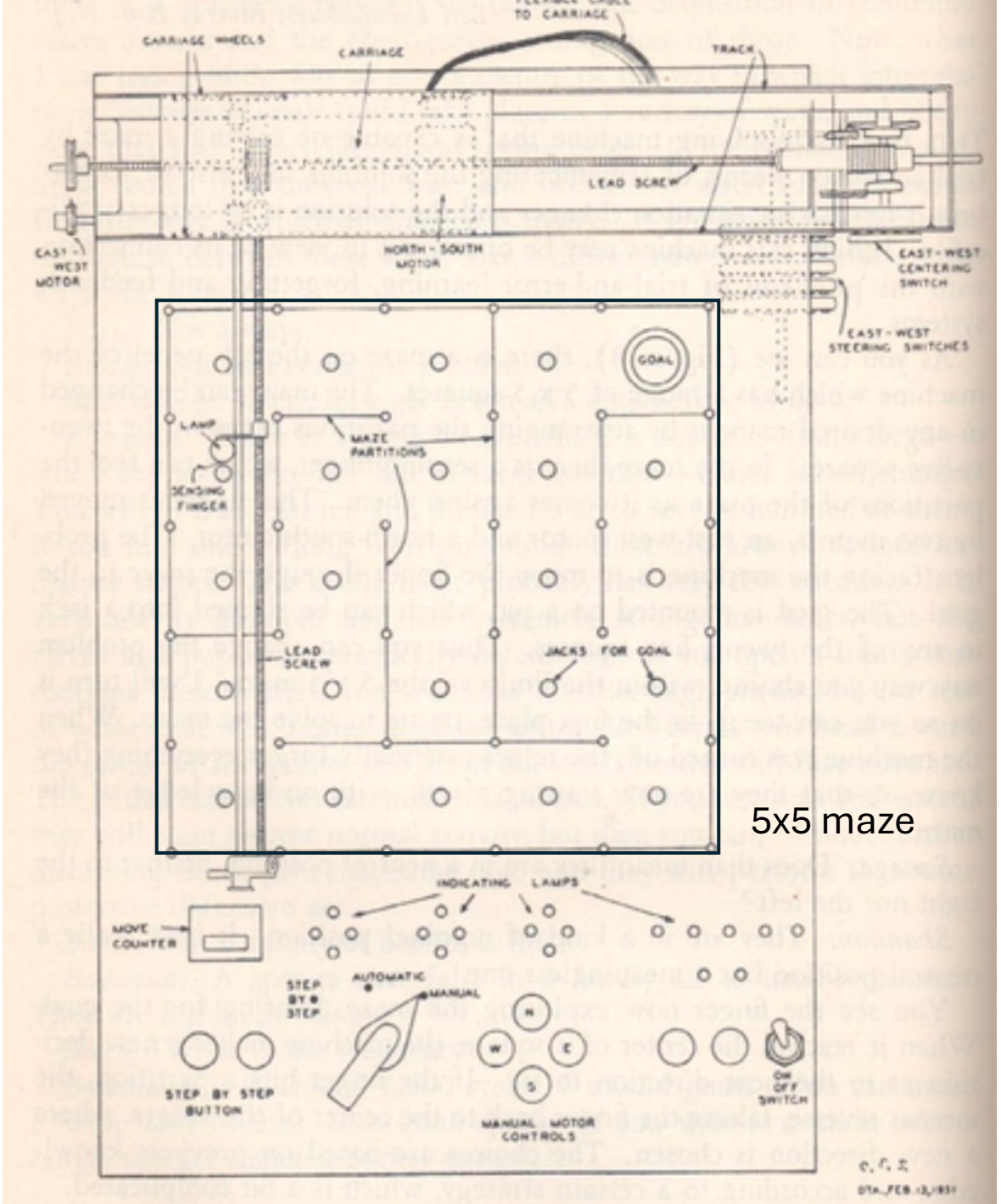


LEARNING MACHINE

Betty Moore was a mathematician at Bell Labs when she met Claude Shannon in 1948. They married a few months later in March 1949. For their first Christmas, Betty bought Claude the largest Erector set she could find. Within hours, they were immersed in a remarkable maze-building project.

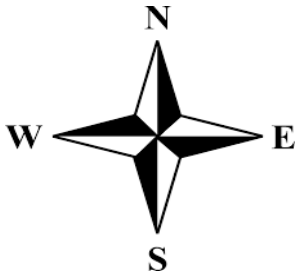
Claude Shannon wanted to build a mechanical device with an electronic "brain" made of relays used for telephone circuits. A magnetized "mouse" named Theseus would systematically explore the maze until it found the "cheese." Placed in the maze a second time, *Theseus* would go straight to the cheese. Change the maze partitions and *Theseus* would "forget" the old solution and start the learning process anew. Demonstrated at state fairs and the National Academy of Science, Shannon's device is considered one of the earliest —and most vivid—examples of machine learning.

Learning Machine
@ the exhibition "MIT Objects" at MIT museum.
A magnetized mouse named "Theseus" would systematically explore the **maze** until it found the "cheese".



5x5 maze

A Maze Solving Machine



- A <u>finger</u>
  to move Theseus E/W/N/S

- There is a <u>memory</u> for each square
  to remember the direction by which the
  the figure left the square on its last visit.

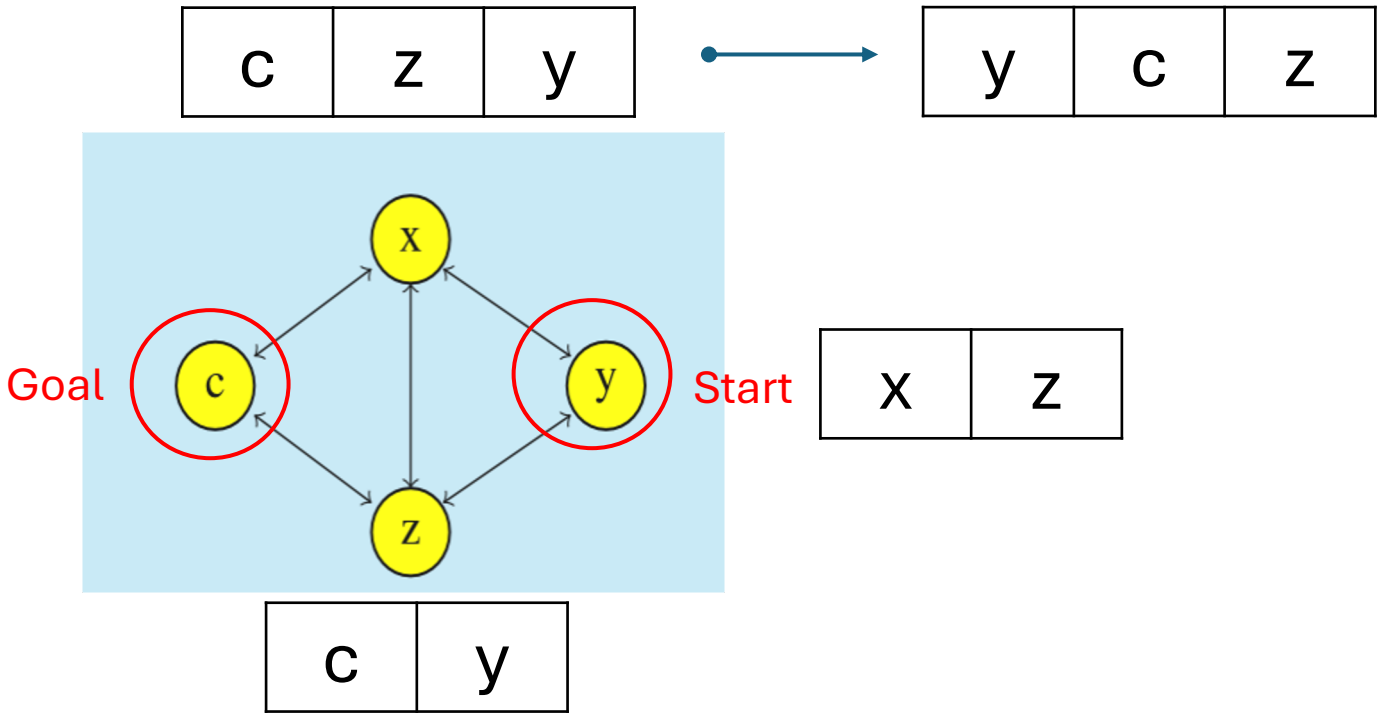
  As the figure comes to the square again,
  the memory is <u>updated counterclockwise</u>.
  ex) left easterly and updated to northern

- Two modes of Machine Operations:
  Exploration and Goal strategy

From Cybernetics, Transaction of the Eighth Conference, 1951
"Presentation of a Maze-Solving Machine" by Claude Shannon

Rotating-Edge Algorithm

Q: search for the node **c** starting from **y?**



The memory updated counterclockwise implements Rotating Edge Algorithm.

From IEEE BITS the Information Theory Magazine "Shannon, Euler, and Mazes" by Robert G. Gallagher (2021)

Theseus is a learning machine

- Theseus finds a path in the maze by trial & error method.
- Theseus memorizes the found path.
- It also automatically explore the new path when the maze is changed.

Machine Learning Definition from Textbook (Murphy2022)

1.1 What is machine learning?

A popular definition of **machine learning** or **ML**, due to Tom Mitchell [Mit97], is as follows:

> A computer program is said to learn from experience E with respect to some class of tasks T, and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

Machine Learning & AI

Many of ML schemes already existed before AI appearance, like decision, estimation theory, and optimization theory. But they are reframed as ML in the context of AI to emphasize its ability of learning: self adjustment based on examples.

Machine Learning as a subfield for Artificial Intelligence (AI)

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

[Summer, 1956]

The Birth of AI in 1956

The study is to proceed on the basis of conjecture that every aspect of learning or any other feature of intelligence can in principle be precisely described that a machine can be made to simulate it.

from "*A Proposal for The Dartmouth Summer Research Project on AI*"

For AI problems, there are two poles
heuristic approach vs. symbolic reasoning

- Heuristic

+ focus on end-effect

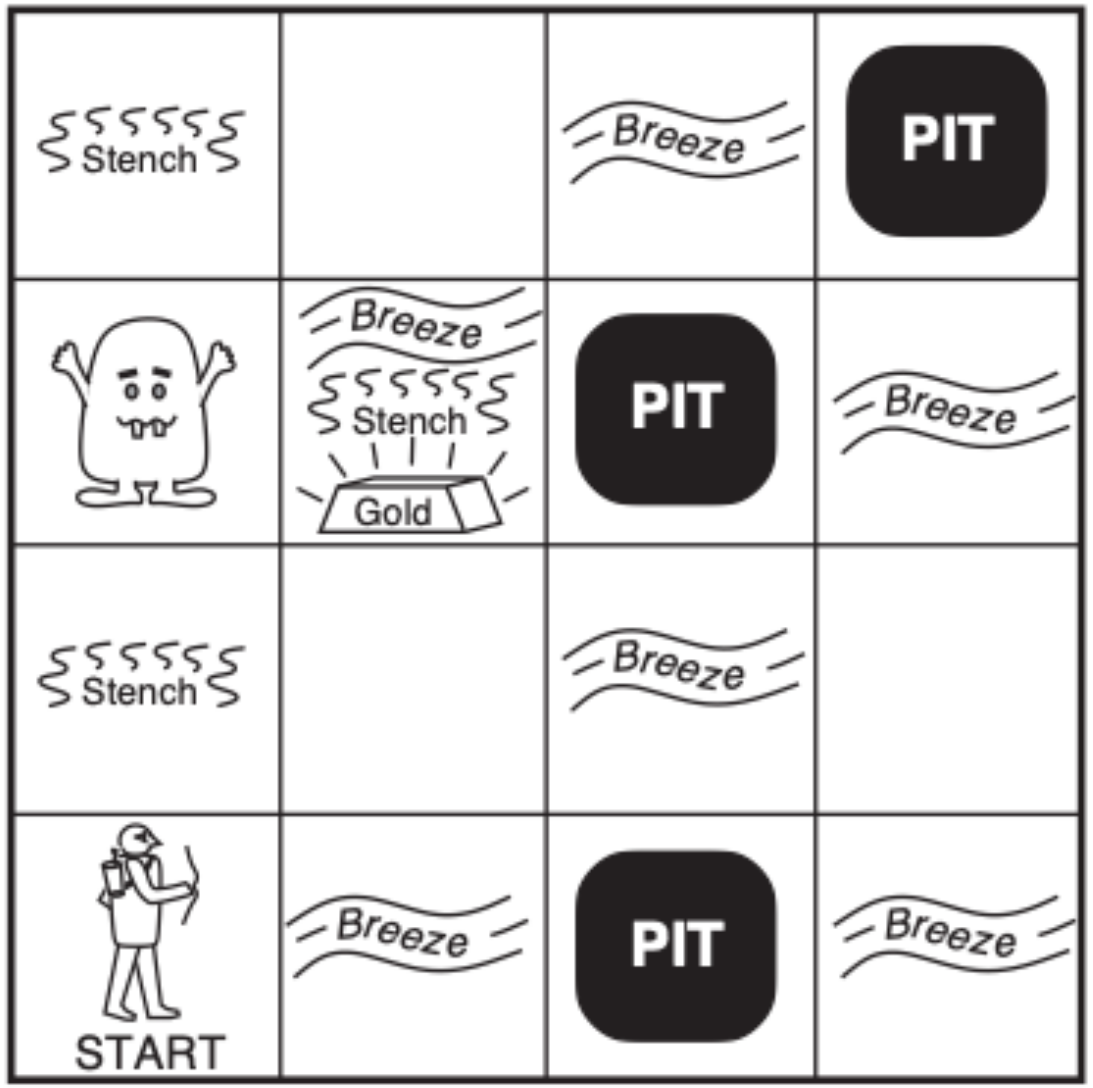
+ make it work

- Symbolic

+ focus on
the representation of true knowledge

+ serial & logical processing of symbolic
expression to drive new knowledge

Example of Symbolic Reasoning



1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

1

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 A B OK	3,1 P?	4,1

2

1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 A S OK	2,2 OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

3

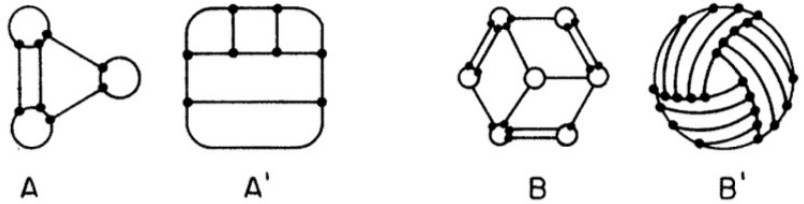
From Chap. 7 Artificial Intelligence: A Modern Approach, Stuart J. Russell

Heuristic Programming Approaches to solve AI problems

- Make a machine to Search
  Q: how effectively test all possible solutions and find an optimal?

- Make a machine to do Pattern Recognition:
  Q: what patterns/features would be useful for AI problem?

  A≡A' and B≡B'
  (topological equivalent
  for connectivity relations)



- Make a machine to do Learning from past and predict future:
  Q: how could we find a model
  to generalize the experience from past so enable to predict future?

From "Steps Toward Artificial Intelligence" by Marvin Minsky [1961]

Heuristic Programming Approaches to solve AI problem

- Make the machine to Search
  Q: how effectively search all possible solutions?

- Make the machine to do Pattern Recognition:
  Q: what patterns/features would be useful for AI problem?

- Make the machine to do Learning from past and predict future:
  Q: how could we find a model
  to generalize the experience from past so enable to predict future?

From “Steps Toward Artificial Intelligence” by Marvin Minsky [1961]

Learning cannot be pursued
without considering Pattern Recognition.

Machine Learning Principles

Machine Learning:
learning a heuristic function from data and empirical measurement

$$y = f(x) + \varepsilon$$

it is important to have

- right performance metric
- data samples reflecting the nature of the target task well
- (why style classification accuracy is around 60% for 20 styles while Image Net achieves around 90% for 1000 objects. https://paperswithcode.com/sota/image-classification-on-imagenet
- having a right hypothetical space according to the number of data samples
- (why it is not a good idea to just pursue a complex modeling? Why do we need to consider the number of data points we have?)

Machine Learning Principles:

1. Define the target task:
   classification, regression, density estimation, learning latent information

2. Functional modeling:
   discriminative (generative), parametric (non-parametric), complexity decision

3. Data collection and feature extraction

4. Learning algorithms:
   empirical performance metric, batch or online, optimization methods

5. Evaluation

CS 461 Course Logistics
(syllabus in Canvas)