- **Two most fundamental concerns of AI researchers**
  - ***Knowledge Representation*** (capture ***K*** in a language)
    - **capturing knowledge in a way suitable for computer manipulation**
  - ***Search*** (problem-solving techniques)
    - **systematically explores a space of problem states**

Symbolic AI

- ***Problem states*: initial knowledge and its inferences, e.g.,**
  - **intermediate steps in reasoning**
  - **different board configurations in a board game, such as chess and go**
- ***State space*: the collection of problem states**
- ***Inference rules*: inferring new knowledge from initial knowledge**

Solving Problems with Search Algorithms

Input: a problem *P*

Preprocessing:

Define *states* and a *state space*
Define *Operators*
Define a *start* state and *goal* set of states.

Processing:

Activate a Search algorithm to find a *path* from **start** to one of the **goal** states.

Two Key Components of Heuristic Search

Each Heuristic Search consists of two parts

- A heuristic **measure**
- An **algorithm** that uses the heuristic measure to **search** the state space

Data, Information, Knowledge

Data: are any facts (words), numbers, or text that can be processed by a computer.
Information: The patterns, associations, or relationships among all this *data* can provide *information*.
Knowledge: Information can be converted into *knowledge* about historical patterns and future trends.

What Is Association Rule Mining?

- Association Rule Mining

= Searches for **relationships** between **items** in a dataset.

- ❖ Finding frequent patterns, associations, correlations, or causal structures among sets of items or objects in information repositories.
- ❖ Frequent pattern: pattern (set of items, sequence, etc.) that occurs frequently in a database.

- Rule form: $X \rightarrow Y$

Two Tasks in ARM

- To reduce the computational complexity of mining association rules, we can divide the problem into two subtasks:
  1. Frequent Itemset Generation: Find all the itemsets that satisfy the *minsup* threshold. These itemsets are called frequent itemsets.
  2. Rule Generation: Extract all the high-confidence rules from the frequent itemsets. These rules are called strong rules.
- The computational requirements for frequent itemset generation are generally more expensive than those for rule generation.

Association Rule Generation

Given a **frequent itemset** Y, find all non-empty subsets $X \subset Y$ such that $X \rightarrow Y - X$ satisfies minimum confidence requirement.

If $|Y| = k$, then there are $2^k - 2$ candidate association rules (ignoring: $Y \rightarrow \varnothing$ and $\varnothing \rightarrow Y$)

For each frequent itemset, Y, generate all non-empty subsets $X \subset Y$
  For every non-empty subset X of Y do
    if ($\sigma(Y)/\sigma(X)$) >= *minconf*
      output rule $X \rightarrow (Y - X)$
end

If Y is not a frequent pattern, we do not consider the rules derived from its subsets.

注意topic4里 association rule和association candidate的区别， 还有一个$3^n - 2^{n+1} + 1$