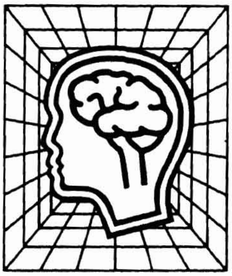
# 3

# The Representation of Knowledge



Intelligent behavior depends on being able to use stored knowledge about objects, processes, goals, causality, time, and action. Some of the questions that arise when trying to understand the nature and the representation of knowledge are:

- What is knowledge?
- Aside from the actual content, is the specific form in which knowledge is encoded an important factor in achieving intelligent behavior?
- How can knowledge be represented in the memory of a computer?
- Are there types of knowledge that cannot be described or discussed?

One of the remarkable attributes of human intelligence is the ability to convert a problem into a familiar form or *representation* that can be operated on using previously known techniques. In a simplified sense, this book puts forth the thesis that intelligence is largely the ability to create and manipulate descriptions. We are thus concerned with the nature of descriptions, what they are, their characteristics, and what their relation is to the things they describe.

This chapter will examine the concept of representation, and will discuss a variety of representations for different problem domains. We are particularly interested in representations that can be formally defined, and are thus suitable for computer mechanization.

## REPRESENTATION: CONCEPTS

Two people view a basketball game. One person sees ten players moving around in a random manner, and notices that a ball and two hoops seem to be the focuses of action. The other person knows the rules of the game, has followed the teams, and therefore sees the strategy being used by each team. Two people listen to the crash of thunder. One person interprets it as the gods expressing anger at the people's sins, and the other interprets the thunder in terms of electrical discharge from the clouds. Two people are discussing the economic recession. One feels that it is a worldwide situation caused by the oil crisis, while the other person views it as a natural result of poor tax policies.

In all of these cases, we have an observed phenomenon that is interpreted in accordance with a stored framework (model, metaphor, representation) that is used by the person to deal with the outside world. Different areas of human intellectual and emotional activities access different representations of the world with different attributes—they construct different "realities." For example, science uses representations that will only accommodate things that can be measured or observed.[16] Thus Newton's laws of motion deal with force, mass, and acceleration of objects, all of which can be measured by instruments. Religion, on the other hand, uses representations that deals with things that are not observable, such as *heaven, hell, angels,* as well as attributes that are not measurable, such as *goodness, evil,* and *holiness*. These different fields thus impose distinct requirements on their representations. For example, physics will require that a representation or model be able to predict the behavior of objects. Religion may expect the models or representations to affect the behavior and the mental state of its adherents.

A discussion of the role of representation in human thinking is given below. Later we will be mainly concerned with representations suitable for machine reasoning, i.e., representations that are formal or "rule-like." A general concern is how a machine might be given models or representations that would enable it to operate in the real world.

### Form vs. Content of Knowledge

Knowledge can be defined as the stored information or the models used by a person or machine to interpret, predict, and appropriately respond to the outside world. It is important to distinguish between the *form* and the *content* of knowledge. For example, addition of numbers can be performed by storing a *look-up table* containing the sums of all acceptable input pairs of integers. Alternatively, an electronic counter can be used which can successively be incremented by the two integers to obtain their sum. From a functional or content standpoint, the two approaches will produce identical answers, but from a representational standpoint, there are significant differences that influence how efficiently we can perform the given task. The look-up table would be very fast, but would require a very large memory store if we had to deal with large integers. The electronic counter would be much more efficient in its hardware requirements, but would be much slower in

---

[16] To be accurate, we note that many physical quantities cannot be measured directly, but must be inferred from other measurements.

producing the required answers. Thus, the specific structures by which knowledge is characterized, and its encoded form, can have a significant effect on its use in solving problems. Further, since no single representation or model can capture all aspects of a real object, an intelligent entity must employ a wide spectrum of representations to deal with the world.

## Representing Knowledge

A representation of a situation (or object, or problem) is a translation of the situation into a system consisting of a vocabulary that names things and relations, operations that can be performed on these things, and facts and constraints about these things. The primary distinguishing characteristics of a representation are (a) what information is made explicit, and (b) how the information is physically encoded. The purpose of a representation is to simplify the problem of answering a restricted class of questions about the given situation, and thus the selection of the representation must be goal-directed. At least two distinct representations are required to match the competence of some given computing mechanism to the requirements of solving a real world problem: the first representation provides efficient symbolic apparatus for answering questions about the given situation, and the second translates the solution techniques of the first into the operations and storage structures inherent in the machine.

The "15 game" is a good example of converting a problem to an alternative representation, using the new formulation to aid in solving the problem, and finally translating the result back into the original problem domain. In this game, two people take turns selecting numbers from 1 to 9. Once a number has been selected by one person, it is unavailable to the other. The person who first has exactly three numbers in his collection that add up to 15 wins the game. A sample game for two people, A and B, might be:

| | |
|---|---|
| A 5 ; B 3 | (A selects 5; B selects 3) |
| A 5,9; B 3,1 | (A selects 9; B chooses 1 to prevent A from achieving 15 on his next move) |
| A 5,9,4; B 3,1,2 | (A selects 4; B chooses 2 to block A) |
| A 5,9,4,6 win!! | (A selects 6 and wins with 4+5+6 = 15) |

Now suppose we analyze this game using the representation,

| | | |
|---|---|---|
| 2 | 9 | 4 |
| 7 | 5 | 3 |
| 6 | 1 | 8 |

Choosing a number in the 15 game corresponds to putting a marker in the tick-tack-toe board shown above. Thus, A choosing 5 is equivalent to putting A's marker in the center of the tick-tack-toe board; the game sequence can be shown in the new representation as:

| | | | |
|---|---|---|---|
| — — — | — A — | B A A | B A A |
| — A B | — A B | — A B | — A B |
| — — — | — B — | — B — | A B — |
| | | | (win for A) |
| 1 | 2 | 3 | 4 |

Note that in the third step A forced the win, since a move was chosen that provided A with two possibilities for three in a row, and B could only block one of these two possibilities. The tick-tack-toe representation and strategy were used to play the 15 game in an expert manner; we make our moves in the tick-tack-toe domain, and report back with the number corresponding to the tick-tack-toe location chosen.

### The Relation Between a Representation and Things Represented

As noted above, in its most general form a representation consists of a language for describing things in the world and a data structure or formalism for physically encoding the descriptions. A "model" is a specific description.

A natural language such as English or French is a representation in which the vocabulary (lexicon) has semantic content (meaning)—the words denote things in the world. In addition to the lexicon, the language includes constraints on what constitutes an acceptable structuring of words in a sentence, and rules for transforming words and sentences to account for singular, plural, tense, and sentence forms dealing with questions, commands, or statements. The written and spoken forms of a language are the data structures of the representation. Thus, a natural language is a representation with a "built-in" semantic content. The *meaning* of a description is fixed by the conventions of the language.

Logical and mathematical systems are also representations, but in such "formal languages" the vocabulary has no semantic content. To the extent that meaning can be ascribed to a logical expression, such meaning is not inherent in the expression but is an interpretation imposed on the expression by some outside agent. The same logical expression can be assigned completely different meanings by different agents.

Most of the representations we describe in this chapter have very simple vocabularies with no semantic content—they are largely data structures that can be used by a computer to store and transform information in a manner specified by a set of rules or algorithmic procedures. This lack of *inherent* meaning in the representations employed by a computer leads to the obvious question of whether it is possible for a computer to "understand" anything in a way a person might.

In his 1985 presidential address to the American Philosophical Association, Dretske asserts [Dretske 85]:

> . . . . all cognitive operations (whether by artifacts or natural biological systems) will necessarily be realized in some electrical, chemical, or mechanical operation over physical structures. . . . This fact alone doesn't tell us anything about the cognitive nature of the operation being performed—whether, for instance, it is an inference, a thought, or taking a square root. For what makes these operations into thoughts, inferences, or arithmetical calculations is, among other things, the meaning of, or the semantics of, those structures over which they are performed. . . . Unless the symbols have what we might call an intrinsic meaning, a meaning they possess which is independent of our communicative intentions and purposes, then this meaning must be irrelevant to assessing what the machine is doing when it manipulates them. The machine is processing meaningful (to us) symbols, but the way it processes them is quite independent of what they mean—hence, nothing the machine does is explicable in terms of the meaning of the

> symbols it manipulates or indeed, of their even having a meaning. . . . In order, therefore to approximate something of genuine cognitive significance, to give a machine something that bears a mark, if not all the marks, of the mental, the symbols a machine manipulates must be given a meaning of their own, a meaning that is independent of the user's purposes and intentions. Only by doing this will it become possible to make the meaning of these symbols relevant to what the machine does with them, possible in other words, to make the machine do something because of what the symbols mean, possible, therefore, to make these symbols mean something to the machine itself.

It should be noted that the arbitrary (even though fixed) conventions of a natural language are not enough to provide *intrinsic* meaning to a description of some aspect of the real world. This issue is further discussed later in this chapter and in Chapter 6.

## ROLE OF REPRESENTATION

In the case of the 15 game, an appropriately chosen representation served the purpose of converting a given problem into another problem that has a known solution. Some other roles of representation are:

- *Interpretation.* Sensory information can be interpreted by using internal representations (models) of real-world objects. For example, visual information can be interpreted by comparing the sensed visual data with stored descriptions of objects.
- *Organizing function.* A representation may allow us to organize information so that similarities and differences between objects and events are more readily identified. Plotting two sets of data on the same graph will visually show similarities and differences.
- *Questioning function.* Internal models lead us to ask questions about events. Why is a certain event occurring when our model predicts otherwise? We are thus guided to revisions in our models, the generation of a set of alternative models, or further attempts at data gathering from our surroundings.
- *Predictive function.* An internal model allows us to predict events that will result from actions. For example, a mathematical model of a rocket enables us to predict the motion of the rocket.
- *Deductive function.* Certain representations can be used to make new knowledge explicit by allowing deductions to be performed on the original knowledge. For example, given *All pit dogs are dangerous. This dog is a pit dog,* we can deduce *This dog is dangerous.*

## REPRESENTATIONS EMPLOYED IN HUMAN THINKING

The concept of representation as a way of selectively, and even creatively modeling the world, has proved to be one of the key ideas underlying our understanding of both human and machine intelligence. In this subsection we briefly describe the representations used by people to solve real-world problems. Our exposition here will be general and descriptive, in contrast to the more detailed and technical discussion of the rest of this chapter which deals with the formal representations suitable for use in computers.

## The Use of Models and Representations

Most people are unaware of their use of models in problem solving and in the way they view the world. As Robert North has stated [North 76]:

> Each of us carries around in his or her head a model of the world, of society, of the local community, of the family—even of oneself, and none of us can deal with any of these entities, even superficially, without reference to the appropriate mental construct or model. It is the only way we have of relating to other people and to our larger surroundings. We draw upon these models whenever we discuss affairs, whenever we vote, and whenever we plan for the future in any way.

George Kelly [Kelly 55] focuses on the psychology of personal constructs, the creative capacity of living things to represent the environment, not merely to respond to it. The point of view that dominates this work is "constructive alternativism," the creation of alternative constructions to explain things in the universe. Some key ideas are:

- *Reality is subjective.* "Each person personally contemplates the stream of events upon which he or she is so swiftly borne," and builds a personal model of reality.
- *People as scientists.* Every person, in his or her own particular way, is a "scientist" whose ultimate aim is to predict and control. There are differences between the personal viewpoints of different people just as there can be differences between the theoretical points of view of different scientists.
- *Representing the universe.* Life in one part of the universe, the living creature, is able to represent another part, the environment. Because man can represent the environment, he can place alternative constructions on it. Man views the world using patterns or templates that he creates, and then attempts to "fit" these templates to the realities of which the world is composed. The fit is not always very good. Man creates his own way of seeing the world in which he lives, the world does not provide this for him. The same events can often be viewed in the light of two or more systems, yet the events do not belong to any system. A construct (specific pattern or model for interpreting some aspect of reality) is used to forecast events and is tested in terms of its predictive efficiency. In general, people improve their internal constructs by increasing the repertoire of constructs, by altering the existing ones to obtain an improved "match" with the world, or by combining constructs. Interpretations of the universe will never be perfect, and thus are always subject to revision or replacement.
- *Psychological relevance.* "Man, to the extent that he is able to construe his circumstances, can find freedom from their domination. The person who orders his life in terms of many special convictions makes himself a victim of circumstances. Every little prior conviction not open to review is a hostage he gives to fortune; it determines whether the events of tomorrow will bring happiness or misery. The person whose prior convictions encompass a broad perspective, and has cast these in terms of

principles rather than rules, has a much better chance of discovering regularities in 'world events' than someone with a limited and inflexible set of models."

## The Use of "Visual" Representations

There are many interesting examples of visual representations used by people to solve problems. For example, Koestler [Koestler 69] quotes Friedrich Kekule, the chemist who discovered the structure of the benzene ring in a dream:

> I turned my chair to the fire and dozed. Again the atoms were gamboling before my eyes. This time the smaller groups kept modestly in the background. My mental eye, rendered more acute by repeated visions of this kind, could now distinguish larger structures, of manifold conformation; long rows; all twining and twisting in snakelike motion. But look! What was that? One of the snakes had seized hold of its own tail, and the form whirled mockingly before my eyes. As if by a flash of lightning, I awoke.

The result of the dream was Kekule's insight that organic compounds such as benzene were closed rings rather than open structures.

The Nobel Prize–winning physicist Richard Feynman used a visual approach to the solution of particle physics problems which has become known as "Feynman diagrams." As described by Dyson [Dyson 79] this form of visual thinking was difficult to communicate to others:

> The reason Dick's physics was so hard for ordinary people to grasp is that he did not use equations. The usual way theoretical physics was done since the time of Newton was to begin by writing down some equations and then to work hard calculating solutions to the equations. . . . Dick just wrote down the solutions out of his head without ever writing down the equations. He had a physical picture of the way things happen, and the picture gave him the solutions, directly and with a minimum of calculations. It was no wonder that people who had spent their lives solving equations were baffled by him. Their minds were analytical; his was pictorial. My own training since the far-off days when I struggled with Piaggio's differential equations had been analytical. But as I listened to Dick and stared at the strange diagrams that he drew on the blackboard, I gradually absorbed some of his pictorial imagination and began to feel at home in his version of the universe.

The scientific literature has many additional examples of visual representations used to solve problems. The choice and application of such representations is very much a mystery. In our attempts to build intelligent machines we are currently limited to the formal representations described in the following sections.

## EFFECTIVENESS OF A REPRESENTATION

A good representation should allow all situations of interest to be easily described, and it should be stable, i.e., if the original situation changes slightly, its new representation should not be significantly different from the original representation. There should be little effect on the final form and content of the represented information even if there is a major change irrelevant to the class of questions that are of interest. The representation should also identify redundant information to allow compact storage in a physical (computer, brain) memory system.

The operations and data structures provided by the representation should result in simple computational procedures for answering questions relevant to the given situation. For example, in the matchstick puzzle of Chapter 1, the representation organized the information so as to reduce the size of the combinatorial search space. In the example of the 15 game, the representation converted a difficult game into a familiar one.

A representation should highlight the important information and thus simplify the relevance problem, as in the following "31 dominoes" problem. A domino is a rectangle with dimensions exactly equal to two adjacent grid squares (see Fig. 3-1).

Given an 8 × 8 grid (Fig. 3-1a), it is obvious that it is possible to find many arrangements in which 31 dominoes can be placed on the grid so as to cover 62 of the grid squares; Fig. 3-1(b) shows one such arrangement.

Are there always arrangements of the 31 dominoes that will leave any two selected squares uncovered? In particular, is there at least one such an arrangement that will leave two diagonally opposite squares uncovered, as shown in Fig. 3-1(c)? Try to solve this problem before looking at the answer we present below.

It is difficult for most people to prove the solution, namely, that the required configuration of dominoes cannot be found. However, if we color the grid with alternating black and white squares, as would be found on a checkerboard, we observe that the two diagonally opposite squares have the same color. Now we further observe that each domino covers exactly one black and one white square when placed on the board in a *legal* position. Thus, any possible legal covering with the 31 dominoes must leave one black and one white square uncovered. This condition is violated for the diagonally opposite squares. This simple solution to the posed problem was possible only after we changed the representation

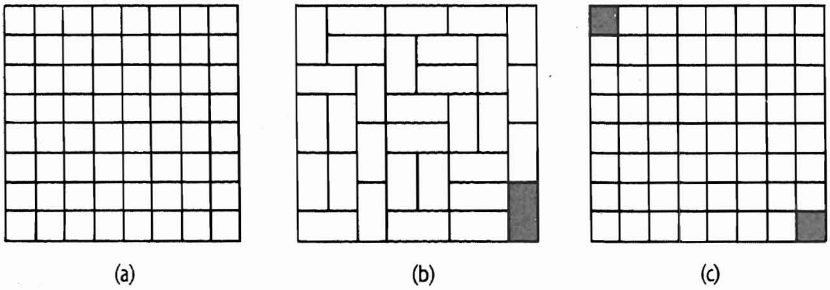

(a) (b) (c)

FIGURE 3-1 The 31 Dominoes Problem.

(a) An 8 × 8 grid. (b) One arrangement of 31 dominoes on 8 × 8 grid. (c) Desired coverage of 8 × 8 grid. Shaded areas on (b) and (c) are uncovered squares.

(by coloring the grid squares) to one that made explicit the critical constraint on any possible solution.

In addition to all the previous considerations, the utility of a representation often depends on its generality: there should be a number of reasonably distinct problem domains to which it can be applied.

## REPRESENTATIONS EMPLOYED IN ARTIFICIAL INTELLIGENCE

A significant portion of AI research is concerned with creating and studying the properties of symbolic representations, since such representations lie at the heart of planning, reasoning, and problem solving. It is surprising that there are only about ten distinct representational systems of broad generality currently employed in AI research. We will explore some of these in more detail in subsequent chapters. The major representations are indicated below, and several of particular interest to this book are discussed in following subsections.

*Feature space (or decision space).* A feature space is formed by assigning a problem-related measurement to each axis of a multidimensional space. Figure 3-2 shows a two-dimensional feature space with one axis representing weight and the other height. One of the points shown represents an individual 6 feet tall and 200 pounds in weight. Points that are close together in this representation represent persons or objects that have similar height-weight measurements.

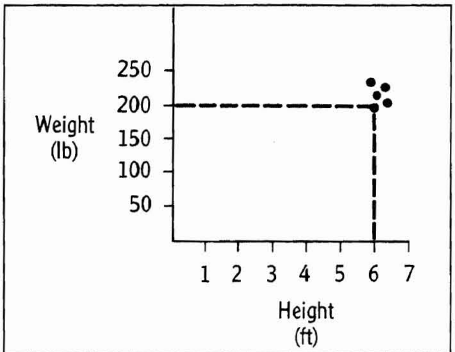


FIGURE 3-2
A Two-Dimensional Feature Space.

Point represents an individual who is 6 ft tall and weighs 200 lb. Cluster of points represents persons or objects with similar measurements.

*Relational graph/semantic net.* A tree or graph structure is typically used to describe relationships between objects (e.g., objects in a story), often for the purpose of general question answering. The net shown in Fig. 3-3 represents the facts: John is a man, John likes Mary, John sees object$_1$, and object$_1$ is a book. Using this net, it is possible to answer questions such as "Who does John like?," and "Does John see a book?" Note that to answer the second question it is necessary to trace through both the path "John sees object$_1$" and "object$_1$ is a book."

*Decision (or game) tree.* In a typical tree structure, each node, representing a state, is connected to one or more successor states. The goal is to traverse the tree from an initial state to a desired final state. In the example shown in Fig. 3-4 we begin with two sets, one with the two elements + + and the other with the single element \$. Two players take turns, and at each turn a player can choose any number of elements from one of the two

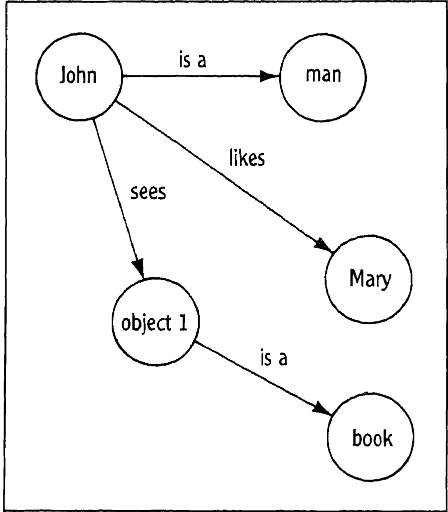


FIGURE 3-3 A Relational Net.

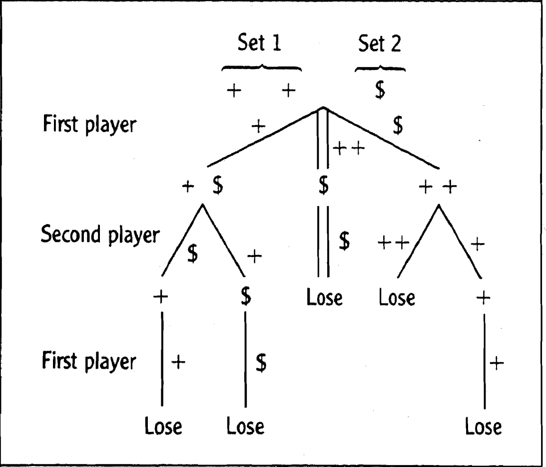


FIGURE 3-4 A Typical Game Tree.

Double lines mark the only path by which the first player can force a win.

sets. The person taking the last element loses. The tree shows that the first player has three choices +, + +, or $, and shows the configuration that results after each choice. Subsequent lines show the choices remaining to the second player, and then the first player, until no further alternatives are available for either. The exhaustive tree for this game shows that the first player should take both + + elements on the first move.

*State transition graph (or sequential machine).* This representation uses nodes and labeled links. The traversal of any particular link requires that the input conditions specified by its label are satisfied. This compact representation can be used to represent any algorithmic procedure, as discussed extensively in Chapter 2. The state diagram shown in Fig. 3-5 shows a "parity machine" that receives as input a string of 1's or 0's, and determines whether the number of 1's in the string is odd or even.

The machine goes into the ODD state when an odd number of 1's has been input, and into the EVEN state for an even number of 1's. The machine must be initialized to start in the EVEN state. The arrows indicate the transition to the same or the other state, and the 0 or 1 indicates the input symbol that causes the transition.

*Frames.* A frame is a way of representing knowledge about the objects and events common to a particular situation. The elements of a given situation are stored as entries in the "slots" of the frame (see Chapter 6). In Fig. 3-6 we show an open frame for DOG, and a filled-in

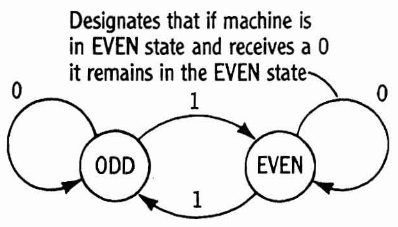


FIGURE 3-5
State Transition Graph for a Parity Machine.

frame for a specific dog, DOG-1. Note that there are slots that have default values, i.e., values that will apply unless otherwise specified. If we are told that DOG-1 is a three-legged, white and black-spotted Dalmatian named Penny , we can record this information in the DOG-1 frame on the right.

*Logic.* The propositional and predicate calculus are formalizations of the process of inferring new information from existing facts (see Chapter 4). The notation given below indicates that for all objects, if the object is a dog, then that object is an animal.

(ALL X)[DOG(X) → (ANIMAL(X)) ]

*Mathematics.* Mathematical representations are pervasive in all areas of AI; e.g., representations such as power series, Fourier transforms, the matrix form, and spatial coordinate systems play an important role in many areas of machine perception. The mathematical expression below specifies the relationship between $R$, $x$, $y$, and $A$.

$$R = x\cos A + y\sin A$$

*Procedural representations.* Knowledge about the world can be formulated in terms of procedures that allow specific tasks to be performed, as in

PROCEDURE: BOIL WATER
1-Obtain pot, and put water in it
2-Put pot over range burner, and turn on burner
3-Turn off burner when steam rises

This is an example of knowing by "knowing how." Thus, given the procedure we might know how to boil water;

```
(?                              (DOG-1
 A DOG                           A DOG
  WITH [BREED =?]                 WITH [BREED =DALMATIAN]
       [FEET  =4]default               [FEET  =3]
       [EARS  =2]default               [EARS  = 2] (default value)
       [NAME  =?]                      [NAME  =PENNY]
       [SIZE  =?]                      [SIZE  = ?]
       [COLOR =?])                     [COLOR =WHITE WITH BLACK SPOTS])
```

FIGURE 3-6 A Frame for the Concept "A Dog."

(a) Concept of a dog using frame representation. (b) Instantiation of the frame for a specific dog.

however, we might not understand the concept "boiling water."

*Production systems.* Production systems use rules of the form, IF *condition A* is satisfied THEN *consequence B* follows. The production rule approach is discussed in Chapter 7.

*Isomorphic/iconic/analogical representations.* These are representations for which there is a direct structural relation to some of the properties of the domain being represented (see below and Chapter 9). In Fig. 3-7, we show that a house plan is an isomorphic representation of the actual physical house.

## Feature Space (or Decision Space)

As indicated in the previous subsection, a feature space is formed by assigning a problem-related measurement to each axis of a multidimensional space. This representation can be used for many purposes, but is especially relevant for decision making and classification tasks. For example, consider the problem of classifying a person into the category *man* or *woman* given the person's height and weight measurements. In Fig. 3-8 we again show a two-dimensional feature space that uses height as one dimension and weight as the other. A (*height, weight*) measurement set, such as (5′10″, 175 lb) is then repre-

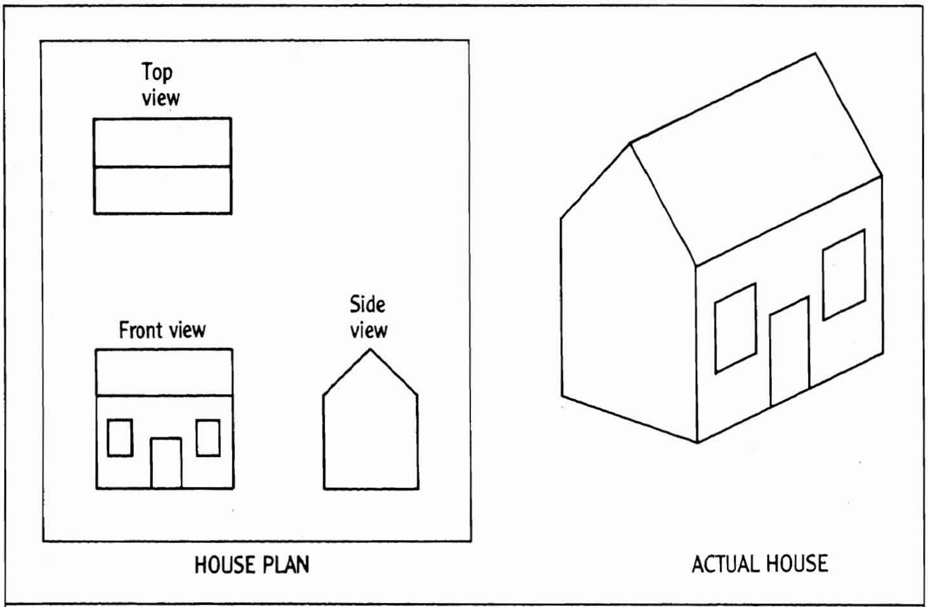


FIGURE 3-7 Isomorphism Between a Drawing and the Physical Object.

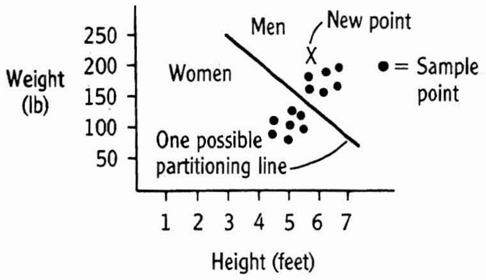


FIGURE 3-8
Example of Partitioning a Two-Dimensional Feature Space as the Basis for Making Classification Decisions.

sented as a point in the feature space. It is possible to partition the space so that points that are typical examples of their class lie in a particular partition. In Fig. 3-8 the space has been partitioned into the classes *men* and *women*, based on a set of typical members of each class. Note that a new point, (6′, 225 lb) falls in the male class. If the measurements are good indicators of the classes being represented, then the data for each class "cluster" into a compact region, and regions for distinct classes are well separated.

A well-chosen feature space can be partitioned into regions such that the points in each region belong to a single class of objects or events. One problem is to partition the feature space, based on a given set of labeled samples, called the *training set*. The partitioning boundaries can be generated using techniques from statistical decision theory, and linear and nonlinear programming. Basically, a *cost* of making a classification error is defined and is used to determine boundaries producing *expected-least-cost* classifications. For example, in the one-dimensional case shown in Fig. 3-9, we plot a height distribution for the classes men and women. We can select a decision line for separating the two classes, and if it is more "expensive" to make the mistake of classifying a man incorrectly as a woman, we will position the line toward the lower height values.

## Decision Tree/Game Tree

We noted that a decision tree is often used to describe exhaustively all the consequences that can arise from some initial situation (state), assuming each state can

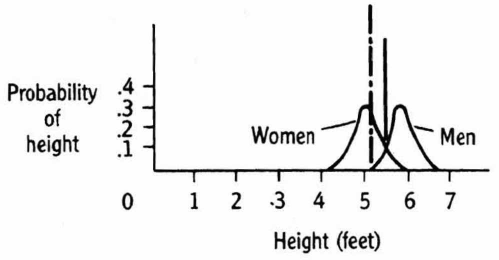


FIGURE 3-9
One-Dimensional Classification: Distinguishing Men from Women on the Basis of Height.

The decision line on the left (interrupted) is selected if classifying a man incorrectly is more "expensive" than misclassifying a woman. The solid line on the right should be used if both types of misclassification errors are equally costly.

only give rise to a specified number of successor states as a result of the application of a given set of operations or actions. For example, a tree could provide an explicit representation of all possible moves in a game of checkers. Many techniques are available to make an efficient traversal of the tree, and to limit the number of paths that must be examined.

One problem that often arises in the decision tree/graph representation is finding the shortest or longest path between two nodes in the graph, or, alternatively, determining which paths have a high "payoff" with respect to achieving some given goal. This type of problem is often handled using *relaxation,* and an example is shown in Fig. 3-10. Suppose we want to go from an origin node to a destination node taking the shortest path. In Fig. 3-10(a), we show the basic computation employed. In this example there are

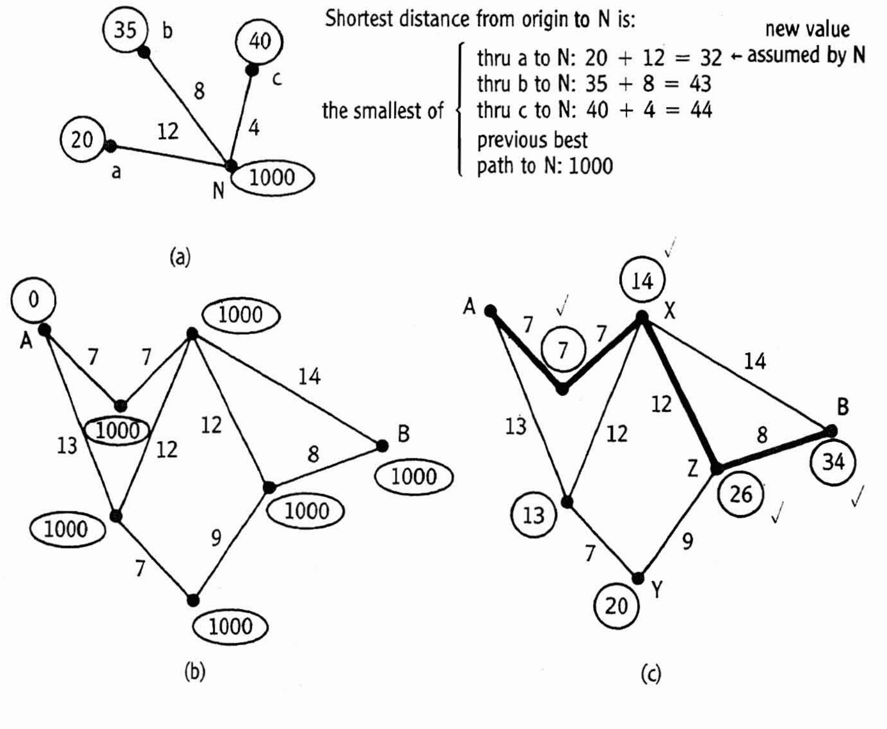


FIGURE 3-10 Finding the Minimum Path in a Graph using Relaxation.

(a) Basic approach. (b) Road network example. (c) Shortest path shown by thick line; e.g., shortest way from Z back to origin is: via $X = 14 + 12 = 26$, since via $Y = 20 + 9 = 29$.

paths to node N coming from nodes a,b, and c. Attached to each node is a circled number specifying the shortest distance from the origin node (not shown) to a,b, and c respectively. The shortest distance from the origin to N is then obtained as follows: For each node a, b, and c, form the sum of its circled number and the distance from that node to N. Select the smallest sum or the current value as the shortest distance from the origin to N. In Fig. 3-10(b) we show a graph, and in Fig. 3-10(c) the final results of the computation. A is the origin node and B is the destination node. The circled values in Fig. 3-10(b) are used to initialize the computation; the value 1000 could have been replaced by any number larger than any reasonable final answer. The computation described in Figure 3-10(a) can be performed at any node in any order. When the computation can no longer produce a change, (i.e., a reduction of an existing circled value, anywhere in the network), the circled value at the destination is the desired shortest path length. Note that, in general, we may have to perform the "update" computation of Fig. 3-10(a) many times at each node before no further change is possible. When we have completed the computation, we trace back through the circled nodes that led to this lowest number. This is shown by the check marks in Fig. 3-10(c).

Another approach to this problem uses various *pruning* techniques that ignore unpromising paths. One of the best known pruning techniques, the *alpha-beta heuristic* shown in Box 3-1, reduces the number of branches that must be analyzed by ignoring obvious *loser* branches.

Since the decision tree and other graph structure representations require an explicit description of the complete problem domain, they do not seem capable of dealing with the potentially infinite problems that are the core problems of AI. For example, the decision tree appears to require an exhaustive listing of all alternatives. However, there are three methods for removing this difficulty. (It should be noted that these methods do not guarantee that the "best solution" will always be found.) First, we can provide a single number estimate, an *heuristic evaluation function*, that indicates the value of exploring the (potentially infinite) remainder of the tree extending beyond some given node. Second, we can throw away information that does not appear to be important in finding a best solution, based on the assumption that only a few aspects of the problem need be considered to obtain an acceptable answer *most of the time*. Finally, if a mechanism exists for generating portions of the tree as needed, then only those parts of a tree that we wish to examine need to be generated; typically, only a finite portion of a potentially infinite tree need be searched to find a desired solution.

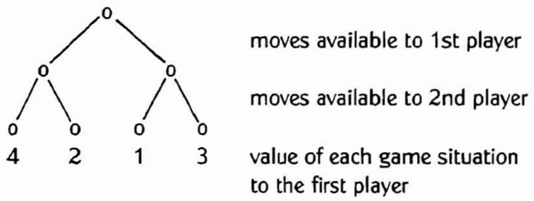
## Isomorphic/Iconic/Analogical Representations

We usually do not appreciate the remarkable "isomorphic" representation known as a road map (Fig. 3-11). The road map can be used to answer an unbounded set of very complex questions. For a current location on the map, we might ask, What is the nearest major town, and how far away is it? What is the closest highway intersection where at least three roads come together? Note the complexity of trying to answer such questions by using tabular data, or other similar symbolic representations. For the first question, one would have to store the locations of all towns and be able to determine the road distance to each town from every road point. The second question would require that we either prestore or compute the intersection of all roads, and be able to determine the road distance of these intersections for all locations in the map. One can see some of the problems of storing the map information in a symbolic format that is still capable of providing the answer to any question that could have been answered by looking at the map itself.

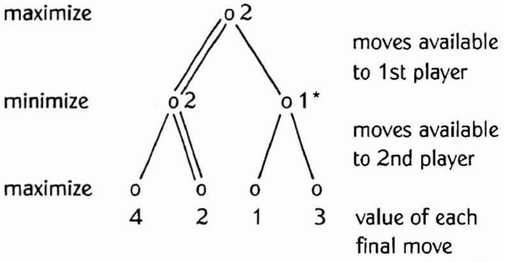
The term "isomorphic", "iconic", or "analogical" representation is used to denote representations for which there is a direct structural and metric relation to some of the properties of the domain being represented. Technically, this type of a relationship is called an *isomorphism*, and we can say that an isomorphic representation is able to represent implicitly those properties of the domain preserved by the isomorphism. An interesting example of an isomorphic representation is the

## BOX 3-1 Game Trees and the Alpha-beta Heuristic

A game can be represented as a tree, where alternating levels indicate the moves available to each opponent. In the tree below, we show the value of the game situation to the first player as the lowest level of the tree:



The first player tries to select a move that will result in the best game situation, i.e., the highest value after the second player has made his selection, while the second tries to minimize this value. Using a mini-max analysis, we can project values up to all nodes of the tree. In the example below, if we look at the bottom row, we know that the minimizing player will choose the branch corresponding to a value of 2, rather than the branch corresponding to 4. Similarly, given the choice of a 1 or a 3 on the bottom row, the minimizing player will choose the 1. The maximizing player, given the choice of 2 or 1 on the second row will choose the 2. Thus, the best that the first player can achieve is the value 2, obtained by selecting the left-hand branch.



Complete evaluation of a game tree is usually impractical, and is indeed unnecessary. In the above tree, if the analysis is carried out from left to right, the maximizing player would eliminate the choice marked with * as soon as the value of 1 is projected up from the lowest level, since this is less than the other node (whose value, 2, is assumed to have already been determined). Thus, it is not necessary to evaluate the lowest level node whose value is 3.

The alpha-beta heuristic is a tree-pruning algorithm that formalizes the following concept: Whenever we project a value to a parent node from a lower node that is better than the existing value of the parent node, check how that parent node now compares with other nodes on its level. It may be that no further exploration is needed below that parent node.

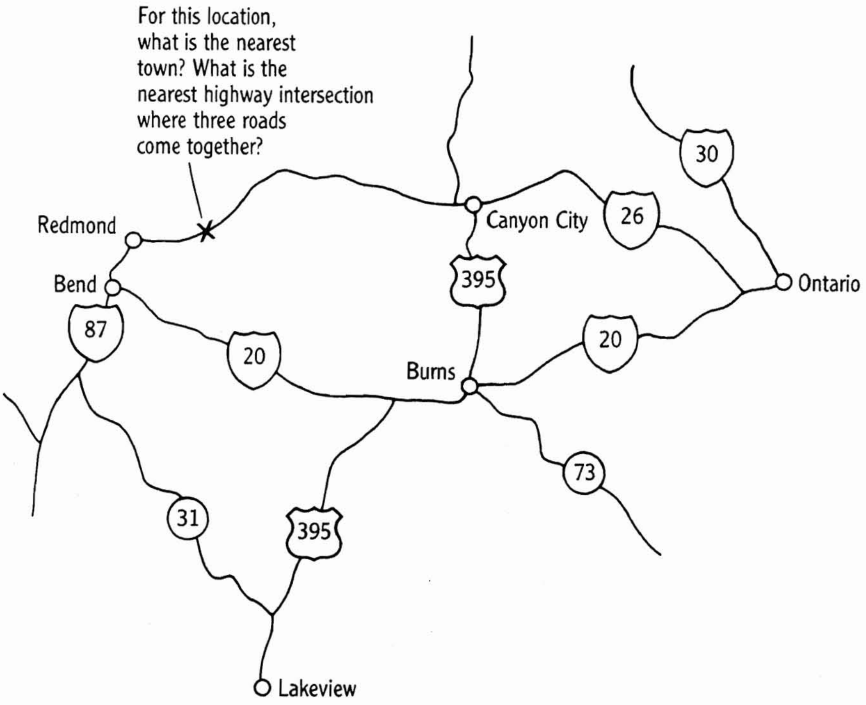


FIGURE 3-11 Example of an Iconic/Isomorphic Representation.

string model for determining the shortest path in a graph (see Box 2-8). Consider the problem of finding the shortest route from city A to city B over some given network of roads. Construct a "map" made out of pieces of string to represent the roads connecting the cities. Let the length of each piece of string be proportional to the length of the road segment it represents, and knot together the strings at places where the roads intersect. Now with both hands, grasp the points in the string map corresponding to cities A and B and pull these points in opposite directions. The shortest path will correspond to the road segments represented by the strings supporting the rest of the dangling road network.

Solving a problem using an isomorphic representation is often similar to performing a physical experiment on a "real-world" situation, as opposed to obtaining the solution by an algorithmic technique applied to a symbolic description. A physical experiment, unlike a symbolic solution, can proceed without complete specification or understanding of the problem domain. Thus, at least in part, the power of an isomorphic representation resides in the fact that there is no need to make explicit the problem domain constraints and relationships, since they are captured by the structure of the representation. Even if understood, attempting to make such knowledge explicit is often impractical because of the enormous amount of detail needed to capture the many aspects of the natural world. Isomorphic (iconic) representations are discussed in Chapter 9.

## DISCUSSION

It is generally acknowledged that most elements comprising an AI system cannot function without knowledge of the application environment. For a computer system, this knowledge must be represented in some formal notation that can be manipulated for the purposes of storage, retrieval, and inference making. A basic philosophical question concerns the extent to which the complexities of the world can be reduced to a manageable set of symbolic relations susceptible to logical analysis (see [Nilsson 83]). There are those who feel that many subtle concepts cannot be captured using a formal representation [Pentland 83]. Some examples of things that are not readily represented by a symbolic description are a person's face, a taste, the sound of a musical instrument, and a smell.

The assumption that we can capture people's knowledge, actions, and experiences in a computer program by using formal representations has also been challenged by the phenomenologists.[17] Martin Heidegger, a leading phenomenologist, believes that our implicit beliefs and assumptions cannot be made explicit. As Winograd [Winograd 86] says, "Heidegger rejects both the simple objective stance (the objective physical world is the primary reality) and the simple subjective stance (my thoughts and feelings are the primary reality), arguing instead that it is impossible for one to exist without the other. The interpreted and the interpreter do not exist independently: existence is interpretation and interpretation is existence"[p.31]. If the phenomenologists are correct, we can never capture the subtleties of interpretation required to function in the world until we find some way of capturing in the machine the interactive nature of interpretation.

We sidestepped a basic problem in representation: Suppose an intelligent entity has a wide spectrum of representations available. How can it determine which representation or model of the world is applicable for a given situation? People seem to select appropriate representations for real-world problems without difficulty. This problem of knowing which representation to use at any given time arises in many contexts, e.g., in the frame selection problem discussed in Chapter 6.

Finally, a question that still plagues our attempts to achieve machine intelligence: If a suitable model is not currently available, how can one systematically obtain a new and efficient model for the given situation? People are very adept at developing new representations when their existing ones are inadequate, but we have no idea how this is accomplished.

---

[17] Phenomenology is a philosophical examination of the foundations of experience and action.