ARTIFICIAL INTELLIGENCE

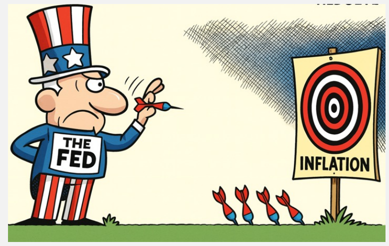
The Very Idea

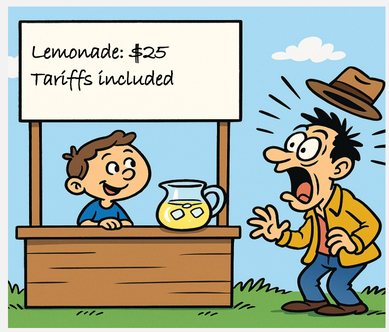
Vasant G. Honavar
Dorothy Foehr Huck and J. Lloyd Huck Chair in Biomedical Data Sciences and Artificial Intelligence
Professor of Data Sciences, Informatics, Computer Science, Bioinformatics & Genomics and Neuroscience
Director, Artificial Intelligence Research Laboratory
Director, Center for Artificial Intelligence Foundations and Scientific Applications
Associate Director, Institute for Computational and Data Sciences
Pennsylvania State University

vhonavar@psu.edu
http://faculty.ist.psu.edu/vhonavar
http://ailab.ist.psu.edu

Knowledge Representation

- One may blame inflation on labor shortage in key sectors of the economy
- Another may blame poor control of the money supply by the Federal Reserve
- Someone else may blame trade wars and the resultant increase in tariffs
- People use mental models or representations to make sense of their world





Knowledge Representation

- David Letterman's mention of Soup Nazi
  - You get the joke because you are an avid watcher of Seinfeld reruns
  - Your buddy does not because he is oblivious to pop culture
- People use mental models or representations to make sense of their world

Knowledge representation

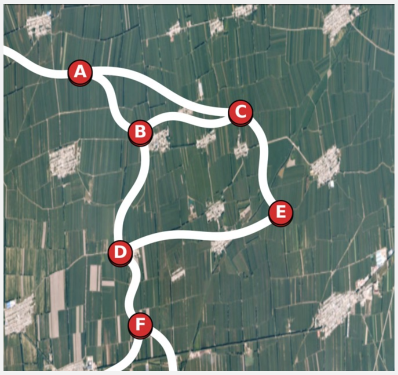
- Even the simplest of agents, to behave intelligently, need
  - Knowledge of their environments
- A smart thermostat needs to know
  - To turn the heat on if the room is too cold
    - Turning the heat on would make the room warmer
  - To turn the AC on if the room is too hot
    - Turning the AC on would make the room colder
  - How occupancy impacts the decisions

Knowledge representation

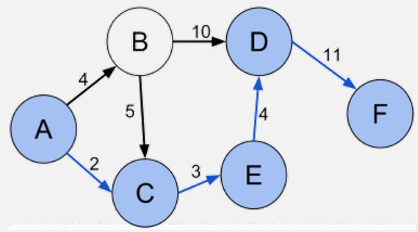
- What is knowledge?
- What kinds of knowledge are there?
- How can knowledge be encoded in a form that can be used by a machine?
- How does the form in which the knowledge is encoded impact, independently of content, the behavior of agents?
- How can an agent effectively use its knowledge to act in its environment?
- How can an agent communicate what it knows to other agents and humans?
- How can an agent acquire knowledge?

Example: Finding the shortest route from one city to another

- How can you find the shortest route from city A to city F?



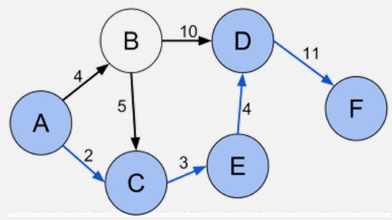
Arial view showing roads



A representation

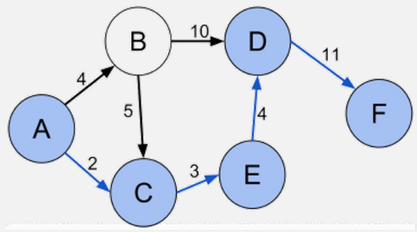
Example: Finding the shortest route from A to F

- How can you find the shortest route?
- Nodes represent cities
- Directed edges denote roads
- Numbers on edges denote distances
- Total distance along a path is the sum of the distances along edges



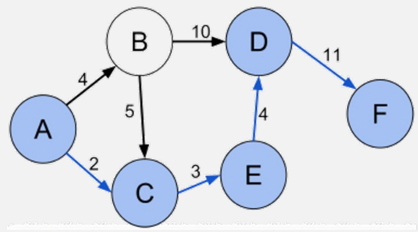
Example: Finding the shortest route from A to F

- Total distance along a path is the sum of the distances along edges
- What is the shortest path from A to F?



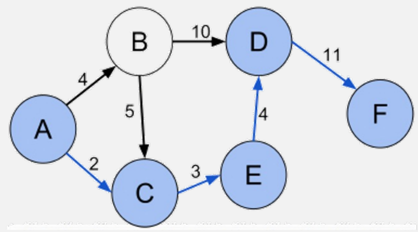
Example: Finding the shortest route from A to F

- What is the shortest path from A to F?
- A → C → E → D → F
- Why?



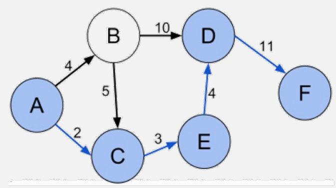
Example: Finding the shortest route from A to F

- Is there a systematic algorithm for solving shortest path problems?
- Yes!



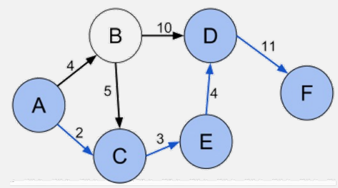
Example: Finding the shortest route from A to F

1. Generate 1-hop paths from the start location A to its neighbors and order them by their cost.
2. If the first path on the list ends in the destination, you have the solution.
3. If not, extend the first path on our list by one more step, and update the list, while maintaining the order.
4. Go back to step 2 and repeat.



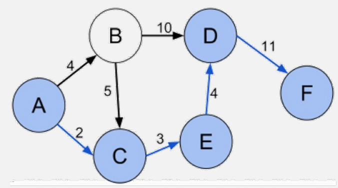
Example: Finding the shortest route from A to F

1. A (0)
2. AC (2), AB (4)
3. AB(4), ACE(5)
4. ACE(5), ABC(9), ABD(14)
5. ACED(9), ABC(9), ABD(14)
6. ABC(9), ABD(14), ACEDF(20)
7. ABCE(12), ABD(14), ACEDF(20)
8. ABD(14), ABCED(16), ACEDF(20)
9. ABCED(16), ACEDF(20), ABDF(25)
10. ACEDF(20), ABDF(25), ABCEDF(27)



- Now the first path on the list takes us from A to F.
- Because the list is sorted in increasing order of cost, the cost of ACEDF is $\leq$ the costs of all other paths on the list
- Hence, it must be the optimal (cheapest) path from A to F

Example: Finding the shortest route from A to F



1. Generate 1-hop paths from the start location A to its neighbors and order them by their cost.
2. If the first path on the list ends in the destination, you have the solution.
3. If not, extend the first path on our list by one more step, and update the list, while maintaining the order.
4. Go back to step 2 and repeat.

- Note that this algorithm is kind of obvious once you have the graph representation

Role of representation in humans and machines

- How did we solve the shortest path problem?
- We used a representation that
  - Made the relevant aspects explicit
  - Discarded the irrelevant aspects
    - population of the cities
    - economy of the cities
    - Weather
    - what the driver had for breakfast…
  - If we can use such a representation to solve problems, so can a machine
- AI is about creating and manipulating representations of the world to solve problems, to exhibit intelligent behavior

Form versus content of representation

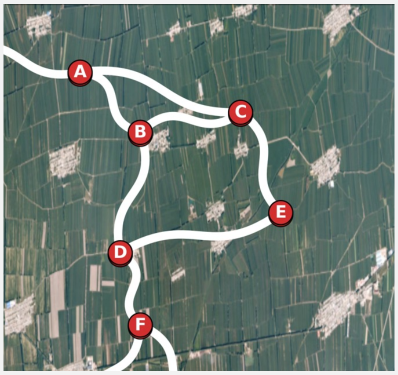
- Suppose you want to add two numbers between 0 and 10,000
  - Solution 1: Use a lookup table that stores sums of all possible pairs of numbers between 0 and 10,000
    - Adding the numbers 5 and 10 will take a single table lookup
  - Solution 2: Add numbers using a counter that starts at 0 and is successively incremented by the two numbers.
    - Adding the numbers 5 and 10 will take 15 increments of the counter starting at 0
  - Solution 3: Standard addition algorithm you learn in elementary school

Form versus content of representation

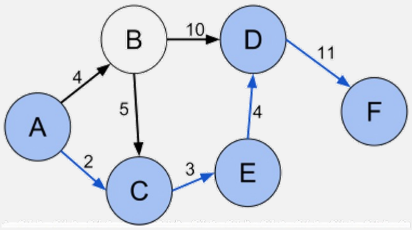
- Suppose you want to add two numbers between 0 and 10,000
  - Solution 1: Use a lookup table that stores sums of all possible pairs of numbers between 0 and 10,000
  - Solution 2: Add numbers using a counter that starts at 0 and is successively incremented by the two numbers.
  - Solution 3: Standard addition algorithm you learn in elementary school
- All three solutions are functionally equivalent
- But they differ greatly in their efficiency!
  - Solution 1 is faster, but requires a large memory
  - Solution 2 is slower, but requires little memory
  - Solution 3 is a little slower than 1, but significantly faster than 2, and efficient in terms of memory

Representation is a surrogate for the world

- Knowledge representation is like a cartoon – it emphasizes some aspects, ignores others



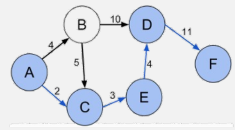
Arial view showing roads



A representation

Representation is a surrogate for the world

- Knowledge representation is like a cartoon – it emphasizes some aspects, ignores others



- What details did we ignore?
  - The terrain
  - The flora and fauna
  - The shape of the roads …
- What details did we keep?
  - Names of the cities
  - Road connectivity between cities
  - Distances traveled by road from each city to its direct neighbors
- The choice of what to ignore are as important as that of what to keep
- The resulting representation is all we needed to solve the shortest path problem!

Representation is a surrogate for the world

- A knowledge representation is fundamentally a surrogate for things we care about in the world
  - Individuals
  - Objects
  - Properties
  - Relationships
  - Processes
  - Actions

Representation is a surrogate for the world

- Knowledge representation enables an agent to substitute thinking for acting in the world
- Suppose an agent that knows that
  - the coffee that was just poured into a cup that it is holding is hot
  - if one were to drink hot coffee, it will burn one's tongue
- The agent can can infer the coffee that that was just poured into the cup it is holding, if it were to drink it, will burn its tongue
- Note that the agent figured out the effect of its action
  - By reasoning with what it knows
  - And not actually performing the action – drinking hot coffee and getting its tongue burned

Why substitute thinking for interacting with the world?

- Acting in the world can be costly
  - Imagine assembling furniture you bought from Ikea
  - You better have a plan for assembling the furniture before you start
  - Mistakes can be costly
- Acting in the world can be dangerous
  - Imagine the scenario with hot coffee
  - You better think about the consequences before gulping down a hot cup of coffee

Representation as a surrogate for the world

Action Description
$Pickup(A, B)$
$Putdown(A, Table)$

State Description
$On(A, B)$
$On(B, Table)$

State description
$On(A, Table)$
$On(B, Table)$

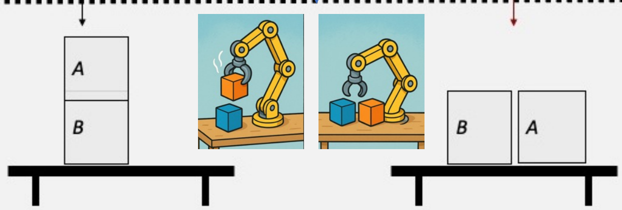
Representation

represent

represent

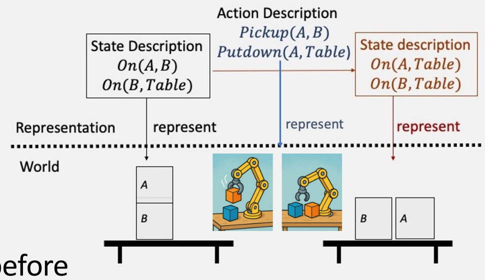
represent

World



Representation is a cartoon

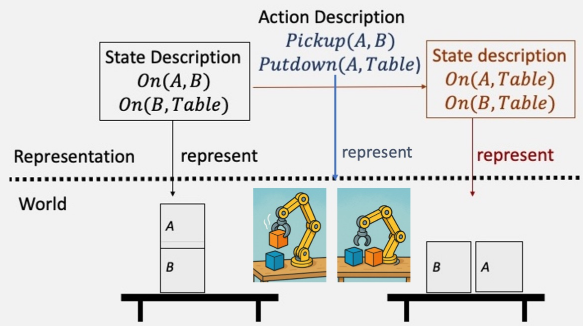
- What details did we ignore?
  - Weight of the blocks
  - Size of the blocks
  - Color of the blocks
  - The force to be applied
  - ....
- What details did we keep?
  - What is true of the world before the actions are executed
  - What would be true after the actions are executed



- The choice of what to ignore is as important as the choice of what to keep

How is a representation useful?

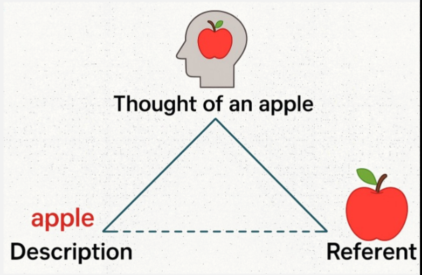
- Representation offers descriptions of the world states and actions
- The descriptions can be manipulated by an algorithm
- The results of running the algorithm give us the consequences of actions in the world



- For a representation to be useful, there must be a coupling or correspondence between descriptions and their physical counterparts

What makes a representation meaningful?

- What is the meaning of ``apple''?
  - ``apple'' has no intrinsic meaning
  - The use of ``apple'' to refer to apples is social convention
- The meaning of ``apple'' is learned from experience
  - What apple looks like
  - What it smells like
  - What it feels like
  - What it tastes like



- The correspondence between descriptions of objects, actions, individuals .. to their referents in the world is what makes the representation meaningful to the agent

Discuss

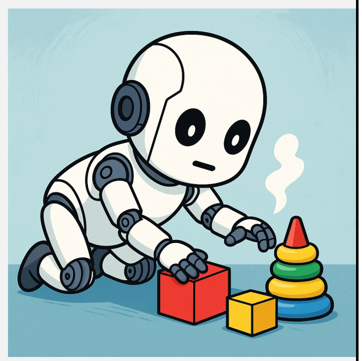
- Can a robot assign to `apple' the same meaning as you?
  - Why or why not?
- Do you assign the same meaning to `apple' as your human friends?
  - Why or why not?

Human-designed representations in AI systems

- In AI systems with representations built by human designers
  - The correspondence between the surrogate or description and the aspects of the world they refer to is established by the human designers of the AI system
  - The meaning or semantics of the representation are provided by the human designers
  - In AI systems with human-supplied semantics, there is no necessary correspondence between the string of letters ``apple'' and what it refers to, except in the mind of the designer
- Such machines may not ``understand'' anything the way we do
- The result is competence without comprehension

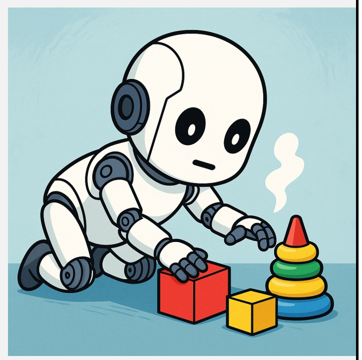
Self-acquired representations in AI systems

- In robots with representations that are learned from interaction with the world
  - The correspondence between the surrogate or description and the aspects of the world they refer to arises from coupling of the representations with sensory percepts and actions
  - The meaning or semantics of the representation are acquired by the robot
  - The descriptions used by the robot are meaningful to it just as descriptions used by us are meaningful to us



Self-acquired representations in AI systems

- Robots that learn the meaning of their representations by interacting with the world ``understand'' the world in ways a human might
  - although the robot's understanding might differ from ours because of differences in descriptions and their semantics
- Such robots can exhibit competence with comprehension



All representations are wrong, but some are useful
A lesson from Alice in Wonderland

- “What a useful thing a pocket-map is!” I remarked.
- “Map making. That's another thing we've learned from your Nation”, said Mein Herr. “But we've carried it much further than you.
- “What do you consider the largest map that would be really useful?''
- About six inches to the mile.
- Only six inches! exclaimed Mein Herr.
- “We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all ! We actually made a map of the country, on the scale of a mile to the mile!''

Sylvie and Bruno Concluded, Lewis Carroll, 1893

All representations are wrong, but some are useful
A lesson from Alice in Wonderland

- "Have you used it much?" I enquired.
- "It has never been spread out, yet," said Mein Herr
- "The farmers objected: they said it would cover the whole country, and shut out the sunlight! So we now use the country itself, as its own map, and I assure you it does nearly as well."
- Lesson: A representation must be a cartoon of reality to be useful
  - It must retain details relevant to the problem at hand
  - It must discard details that have little bearing on the problem at hand

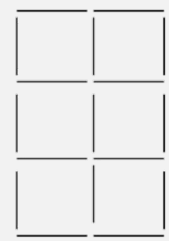
Sylvie and Bruno Concluded, Lewis Carroll, 1893

All representations are wrong, but some are useful

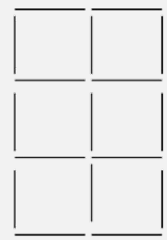
- The most accurate representation of the world is the world itself
- Representations are cartoons
  - They keep some details, throw out other details
  - The choice of what to discard is as important as that of what to keep
  - More detail is not necessarily better
- Because all representations are cartoons of reality, it is inevitable that they are in some respects wrong
- All reasoning about the natural world relies on representations
- All reasoning about the natural world, by humans or machines, will at some point yield conclusions that are incorrect
  - Example: Newton's laws break down at subatomic scale
- We can assess the usefulness of representations only relative to the tasks we want to perform

Sticks and Squares Problem

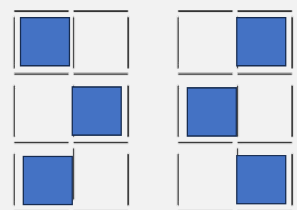
- We have 17 sticks arranged to form 6 squares
- Your task is to remove exactly 5 sticks so that we are left with exactly 3 squares and no orphan sticks



Sticks and Squares Problem



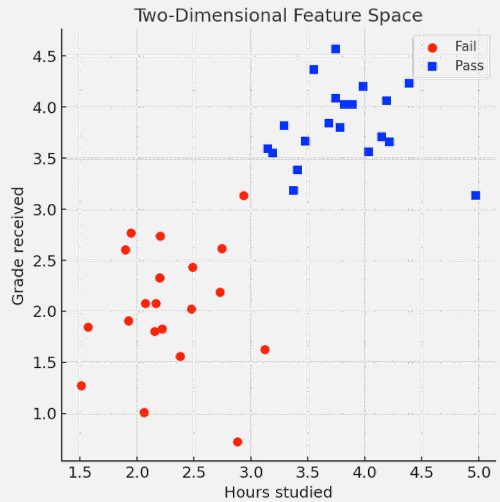
- We have 17 sticks arranged to form 6 squares
- Your task is to remove exactly 5 sticks so that we are left with exactly 3 squares and no orphan sticks
- Level of abstraction – sticks
  - $\binom{17}{5} = \frac{17!}{(17-5)!5!} = 6188$ configurations



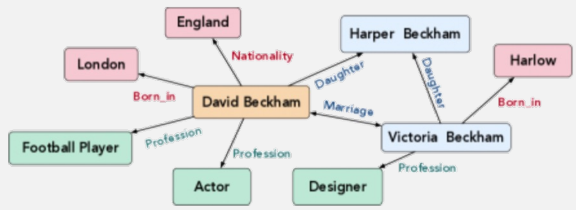
- Level of abstraction – squares
  - $\binom{6}{3} = \frac{6!}{(6-3)!3!} = 20$ configurations
- Add some knowledge of geometry
  - Solution pops out

Feature spaces

- A feature space is constructed by assigning a problem-related attribute or feature or measurement to a distinct dimension of a multidimensional space.
- Points that are close together in feature space represent objects that have similar feature values
- Feature spaces are used by machine learning systems



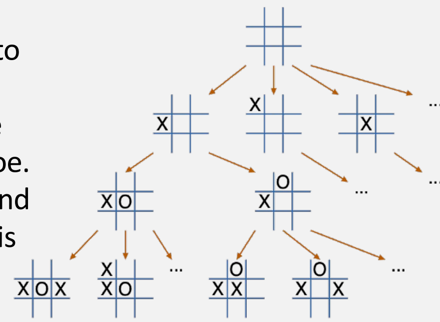
Relational networks



- Relational networks are graphs where the nodes represent individuals or other entities and links denote relationships between entities
- Such a representation can be used to answer queries about individuals and their relationships to other individuals or entities
  - Who is David Beckham married to?
  - Who is the daughter of David Beckham and Victoria Beckham?
  - Who is both an actor and a Football player? etc.

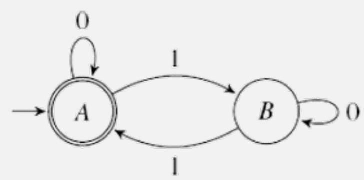
Game trees

- In a game tree, each node, representing a state, is connected to one or more successor states
- Figure shows the game tree for the familiar 2-person game of tic-tac-toe.
- Two players (X and O) take turns, and at each turn the player can place his or her mark (X or O) in an empty square.
- The goal is to traverse the tree from an initial state to a desired goal state.
- The player who gets three X's or three O's along a row, column, or diagonal wins.



State transition graphs

- This representation is a generalization of finite state automaton. The nodes denote states and labeled links denote conditions that the input must satisfy for the corresponding state transition to occur.
- The state diagram in Figure shows a ``even parity machine''



Logic

- The propositional and predicate logic, and their extensions, e.g., logics of knowledge and belief formalize of the process of inferring assertions about the world that are true from existing facts about the world.
- Given All men are mortal, that is, Socrates is a man, we can deduce using predicate logic that Socrates is mortal

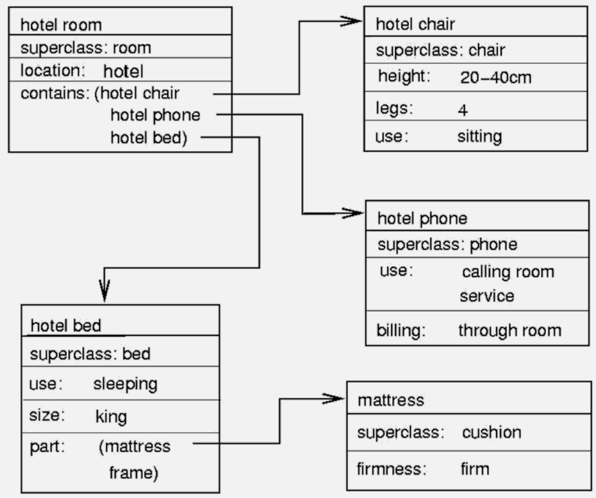
Procedural representation

- A procedural representation describes knowledge about the world in terms of a procedure that instructs an agent to perform a specific task.

Procedure Boil Water
(1) Obtain pot, and put water in it
(2) Put pot over the stove, and light the stove
(3) Turn off the stove when steam rises

Frame representation

- A frame provides a means of representing knowledge about prototypical objects, relationships, and events.
- The aspects of a given scenario are stored as entries in the slots of the frame.
- A partial frame representation of a hotel room is shown in Figure



Production(rule) systems

- Production systems use rules of the form IF condition THEN consequence
- Such rules form the basis of the so-called expert systems
- Applications range from configuring computer systems to medical diagnosis

Rules of the Road

- IF the stoplight is red AND there is no "No turn on Red" sign AND you have stopped, THEN a right turn is ok.
- IF the stoplight is green THEN a right turn is ok.
- IF right turn is ok THEN turn right.

Isomorphic representations

- Isomorphic representations enforce a direct structural relation between the representation and some of the properties of the world being represented.
- Road maps offer an example of a widely used isomorphic representation
- For any 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 meet?

KR is a set of ontological commitments

- What does an agent care about?
  - ✓Entities
    - coffee, liquid, tongue
  - ✓Properties
    - being hot, being able to burn
  - ✓Relationships
    - Coffee is a liquid
- KR involves abstraction, simplification
  - A representation is
    - a model of the world
    - like a cartoon
  - All representations are wrong, but some are useful

KR involves a set of epistemological commitments

- What can we know?
  - Boolean logic
    - Is a proposition True or False?
  - Probability theory
    - What is the probability that a given proposition true?
  - Decision theory
    - Which choice among a set of candidate choices is the most rational?
  - State-transition system
    - What is the next state for given current state and action?
  - Relational graphs
    - Relations between individuals
  - Feature spaces
    - Values of specific features for an individual

Knowledge representation is a theory of reasoning

- How can knowledge be encoded ?
  - Syntax
- What does the encoded knowledge mean?
  - Inferences that are sanctioned by the semantics
- What can we infer from what we know?
  - Inferences that can be performed by algorithms
- How can we manage inference?
  - What should we infer from among the things we can infer?

KR formalisms

- Logical
  - e.g., Propositional Logic, First order logic, Description logic
- Probabilistic
  - e.g., Bayesian networks
- Grammatical
  - e.g., Context free grammars
- Structured
  - e.g., frames – as in object-oriented programming
- Decision theoretic
- …

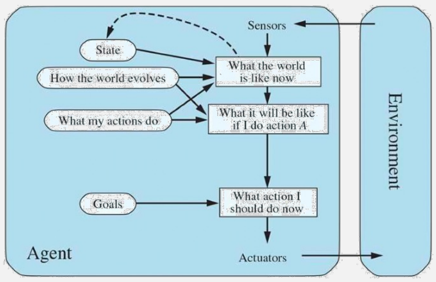
KR is a medium for efficient computation

- Reasoning = computation
- Anticipated by Leibnitz, Hilbert
  - Can all truths be reduced to calculation?
  - Is there an effective procedure for determining whether or not a conclusion is a logical consequence of a set of facts?
- KR involves tradeoffs between
  - Expressivity and tractability (decidability, efficiency) tradeoff
    - The more you can say (using a representational formalism), the less you can effectively do (within the formalism)
  - General purpose reasoning versus special-purpose, domain-specific inference
  - Declarative versus procedural knowledge

KR is a medium of expression and communication

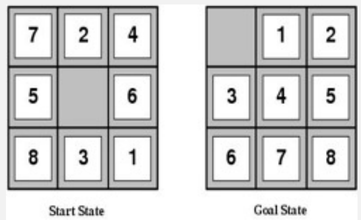
- If we assume shared
  - Ontological and epistemological commitments
  - KR formalism (syntax, semantics, reasoning)
- Then KR is a medium for
  - Expression
    - How general is it?
    - How precise is it?
    - Is it expressive enough?
  - Communication
    - Can we talk or think in the language?
    - Can we communicate the things we want to communicate?
    - What things are difficult to communicate?

Problem solving agents



- The agent is given a goal
- Agent is not told precisely how to achieve its goal
- Achieving a goal may require a long sequence of actions
- Agent needs to figure out how to achieve its goal
- Agent needs a model (representation) of the world and how its actions affect the world

Example: 8-puzzle



- States?
  - Position of each tile on the board
- Initial state?
  - Any state can be initial
- Actions?
  - {*Left*, *Right*, *Up*, *Down*}
- Goal test? Check whether goal configuration is reached

Problem Formulation

Simplifying assumptions

- Discrete, fully observable states
  - 'in class', 'at home'
- Discrete actions
  - Mary executes action 'Go home' in state 'in class' to reach the 'at home' state
  - In this setup, we can't speak of Mary being on her way home
- Passive environment
  - All state changes are due to the agent's own actions
  - Mary can't end up at home because her mom picked her up

Representation

A representation

- Maps each (physical) state of the external environment into the corresponding abstract state via sensors
- Maps each (physical) action on an environmental state into an abstract action on the corresponding abstract state
- Maps effects of an abstract action on an abstract state into a corresponding effect on the corresponding physical state via effectors

The mapping from

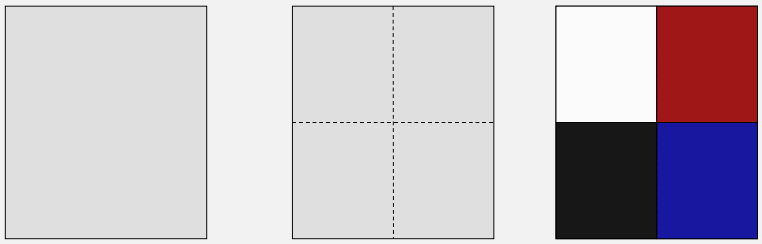
- environmental states and abstract states is many to one
- abstract state to an environmental state is one to many

Representation

The mapping from

- environmental states and abstract states is many to one
- abstract state to an environmental state is one to many

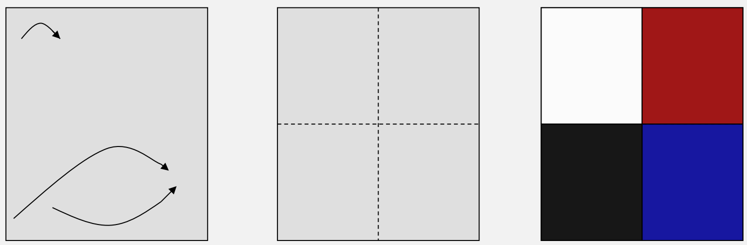
A representation induces a partition over environmental states



- 4 abstract states, 12 abstract actions (lateral, vertical, or diagonal moves)
- 4 abstract states, 8 abstract actions (lateral, vertical but not diagonal moves

Representation

- Effects of abstract actions in the abstract state space is fully deterministic and predictable.. But..
- The corresponding effects of the physical actions on the world are predictable to the extent allowed by
  - the resolution of the representation and
  - the fidelity of sensors and effectors



Representation

A representation

- Is a surrogate inside an agent's 'brain' for entities that exist in the external world
- Derives its semantics through grounding (sensors, effectors)
- Embodies a set of ontological commitments – assumptions about the entities, properties, relationships, and actions that we care about

Properties of representations

- expressive power
- complexity, ..