6. Knowledge representation (KR)

6.1 What is knowledge?

- Knowledge is a theoretical or practical understanding of a subject or a domain.
- Knowledge is also the sum of what is currently known. (Negnevitsky, 2002)
- Apparently, knowledge is power.
- Structured information of real-world entities.
- Expert is one that possesses deep knowledge in some specific domain, which can be limited.

6.2 Two lines of research

1. Principles of correct reasoning
   - Find sound and complete inference rules
   - Truth-preserving operations on well-formed expressions
   - Philosophers and mathematicians
2. Nature of human understanding
   - How people acquire, associate and apply their knowledge of the world.
   - Common sense reasoning and natural language processing
   - Psychologists and linguists

6.3 What is KR?

- One of the most central issues in AI.
- Different representational technologies:
  - Rule-based systems
  - Semantic nets
  - Frames and scripts (object-oriented programming)
  - Formal languages, modal logic, and predicate calculus
  - Case-based reasoning
  - Concept maps
  - No representation
    - Brooks' intelligence without representation
    - Copycat (Mitchell, Hofstadter; 1993)
- Different roles it has in intelligent systems.
- Not a data structure

6.4 Roles of KR (I)

- A surrogate —a substitute for the real thing:
  - Allows one to reason about the world instead of acting in it.
  - Necessarily a simplification, imperfection, abstraction or include something not present in the world.
- A set of ontological commitments:
  - How should one think about the world?
  - Determines what can be represented and what not.
  - Defines what is important and relevant.
  - Provide a way of viewing the world, but not how to instantiate the view, e.g., how to choose prototypes for frames, or which inferences to make.

Ontology, in philosophy is the study of being and existence. It tries to describe the basic categories and relationships of being and existence in order to define entities.

Ontologia eli oppi olemisesta on filosofian osa-alue, joka tutkii todellisuuden luonnetta, ja sitä, mitkä ovat olemassa olevien asioiden perusyksiköt ja suhteet.

Freely adopted from Wikipedia, the free encyclopedia

6.5 Roles of KR (II)

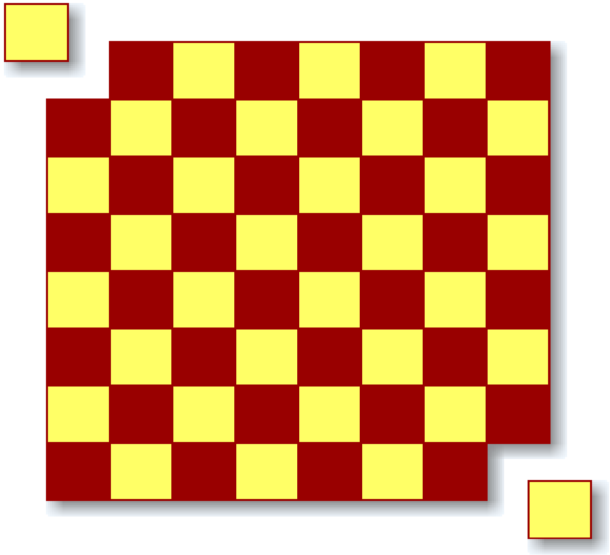
- Medium for human expression and communication:
  - How easy is it to think and talk? For instance, what kinds of things are so difficult that they are pragmatically impossible?
- Theory of intelligent reasoning:
  - What can be inferred, and what should be inferred.
  - What is intelligent reasoning?
    - Mathematics: formal calculus
    - Psychology: characteristic human behavior demonstrated in problem solving
    - Biology and behaviorism: Stimulus-response behavior
    - Probability theory: obeying axioms of probability theory
    - Economics: using values and preferences in maximizing utility
- Medium for efficient computation:
  - A way to organize information to facilitate reasoning

6.6 Important questions

- What is the representation representation of?
- Is it accurate and adequate? i.e., how close it is to the real thing.
- How well it serves problem solving?
  - How easy the problem is to solve, or can it be solved at all?
- How default, counterfactuals, untruths, and non-existence is represented if at all?
- How to deal with uncertainty and a dynamic environment?

6.7 Example: Mutilated chessboard

- A chess board can be covered with 32 dominoes that are size of two adjacent squares.
- If two diagonally opposite corners of the board are cut off, 62 squares remain. Can they be covered with 31 dominoes?



6.8 Rule-based systems

- One of the oldest and most used methods to represent domain knowledge in expert systems.
- Originally a model of human problem solving (Newell & Simon, 1972)
- Experts can express their knowledge in the form of rules.
- Knowledge represented as *if* ... *then* ... production rules:

IF <antecedent 1>
<antecedent 2>
...
<antecedent *n*>

THEN <consequent 1>
<consequent 2>
...
<consequent *m*>

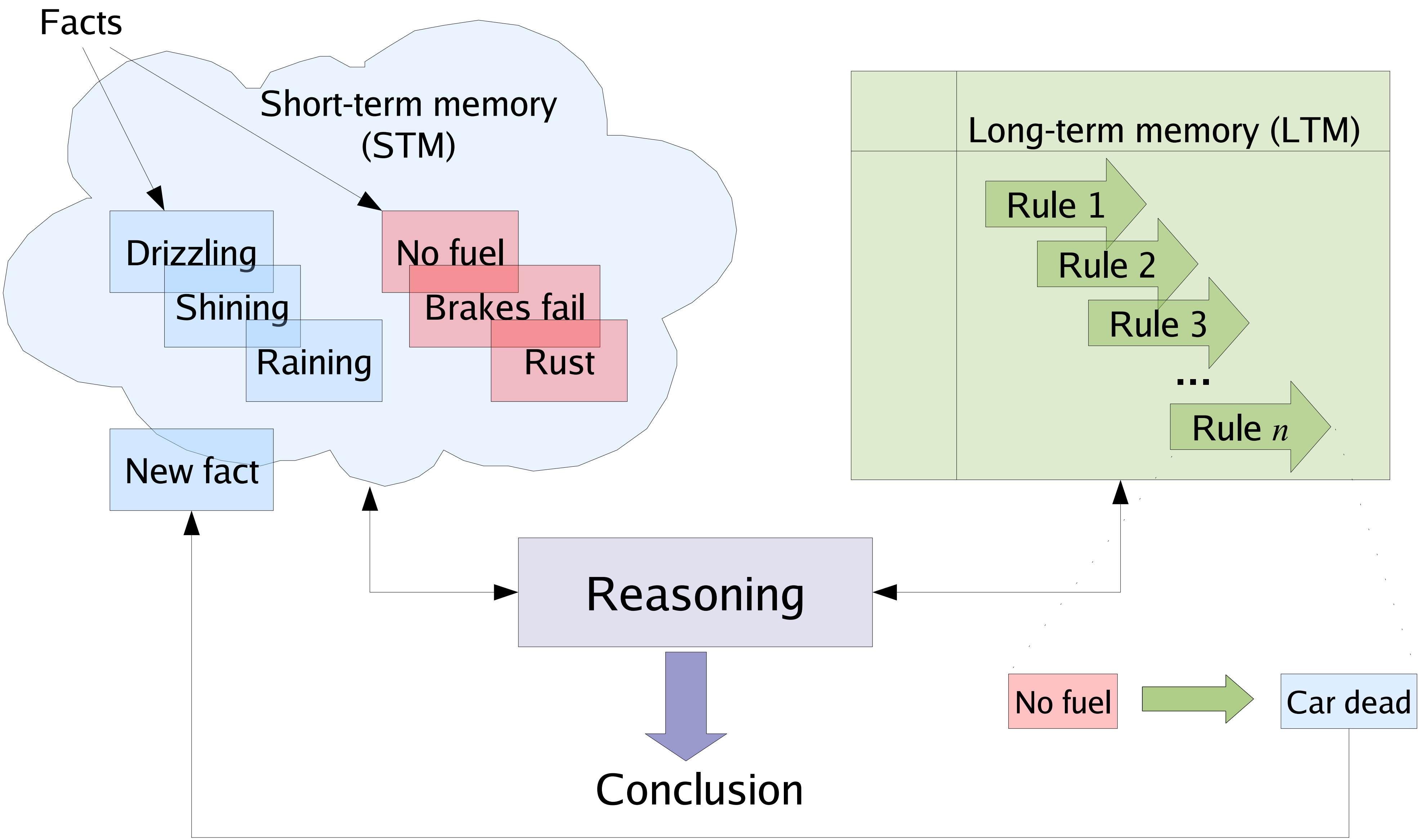
Recommendation:

IF is autumn
is cloudy
is drizzling
THEN take umbrella

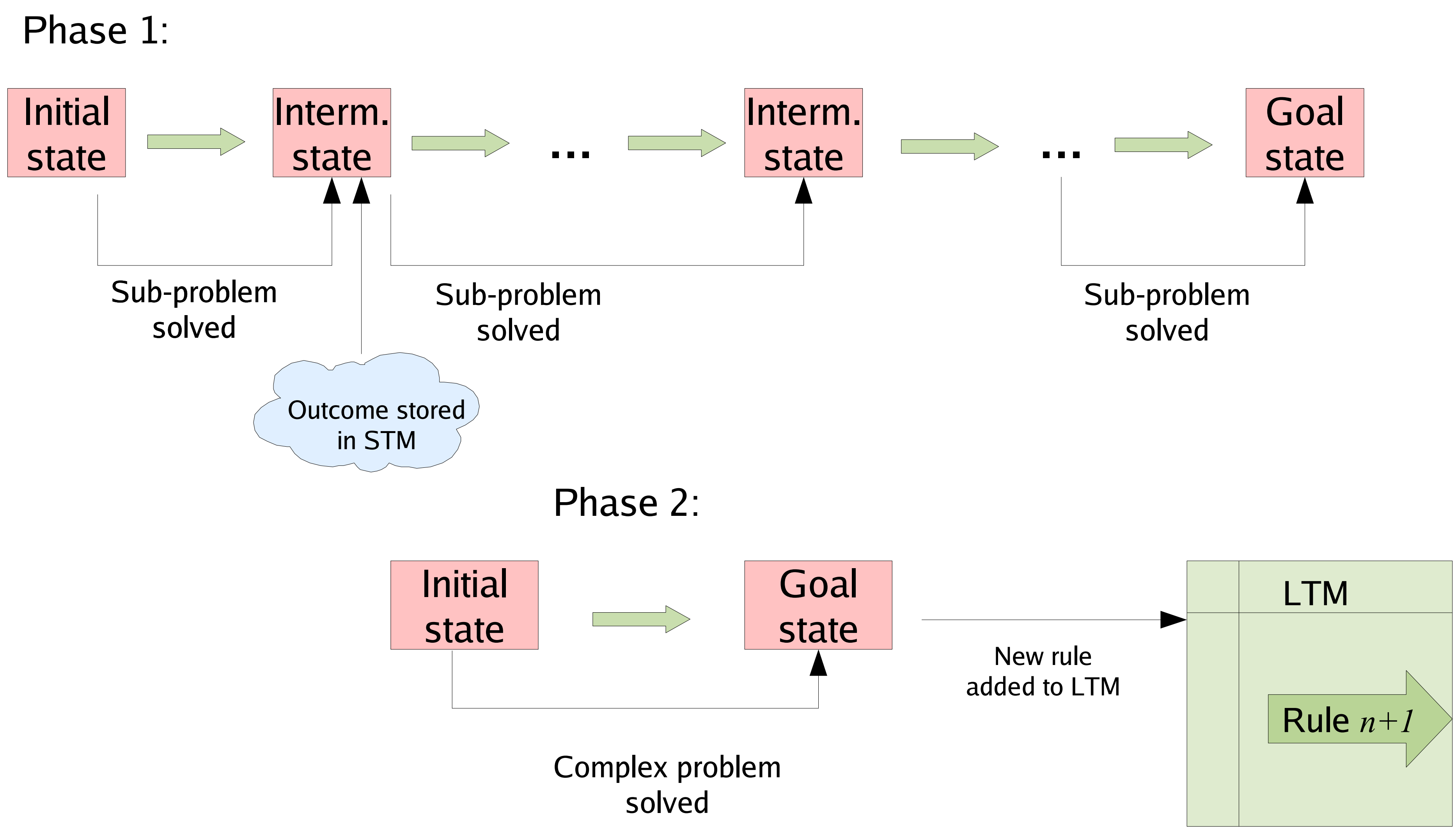
Relation:

IF fuel tank empty
THEN car is dead

6.9 Using production rules



6.10 Production compilation



6.11 Pros and cons of rule systems

- Advantages
  - “Natural”
  - Uniform structure
  - Transparent: a sequence of rules fired can be traced to come up with an explanation
  - Separation of knowledge and processing
- Shortcomings
  - No relation between rules; no hierarchical structure
  - Exhaustive search through all the rules until first fires
  - Information about a thing is stored in several (potentially unrelated) rules.
  - Order of rules in database affects the outcome (e.g., in conflict resolution)
  - Vulnerable to mistakes if data incomplete or fuzzy

6.12 Semantic nets

- Intuitive way of expressing knowledge as a directed graph
  - *Nodes* represent objects
  - *Links*, connecting the nodes, represent relationships between objects.
  - Links are labelled and directed.
  - Can represent specific individuals or classes of individuals.
- Meaning of an objects is defined by its associations with other objects.
- Characterize the nature of human understanding
  - Association of concepts
  - Knowledge organized hierarchically
- Rudimentary concept of inheritance

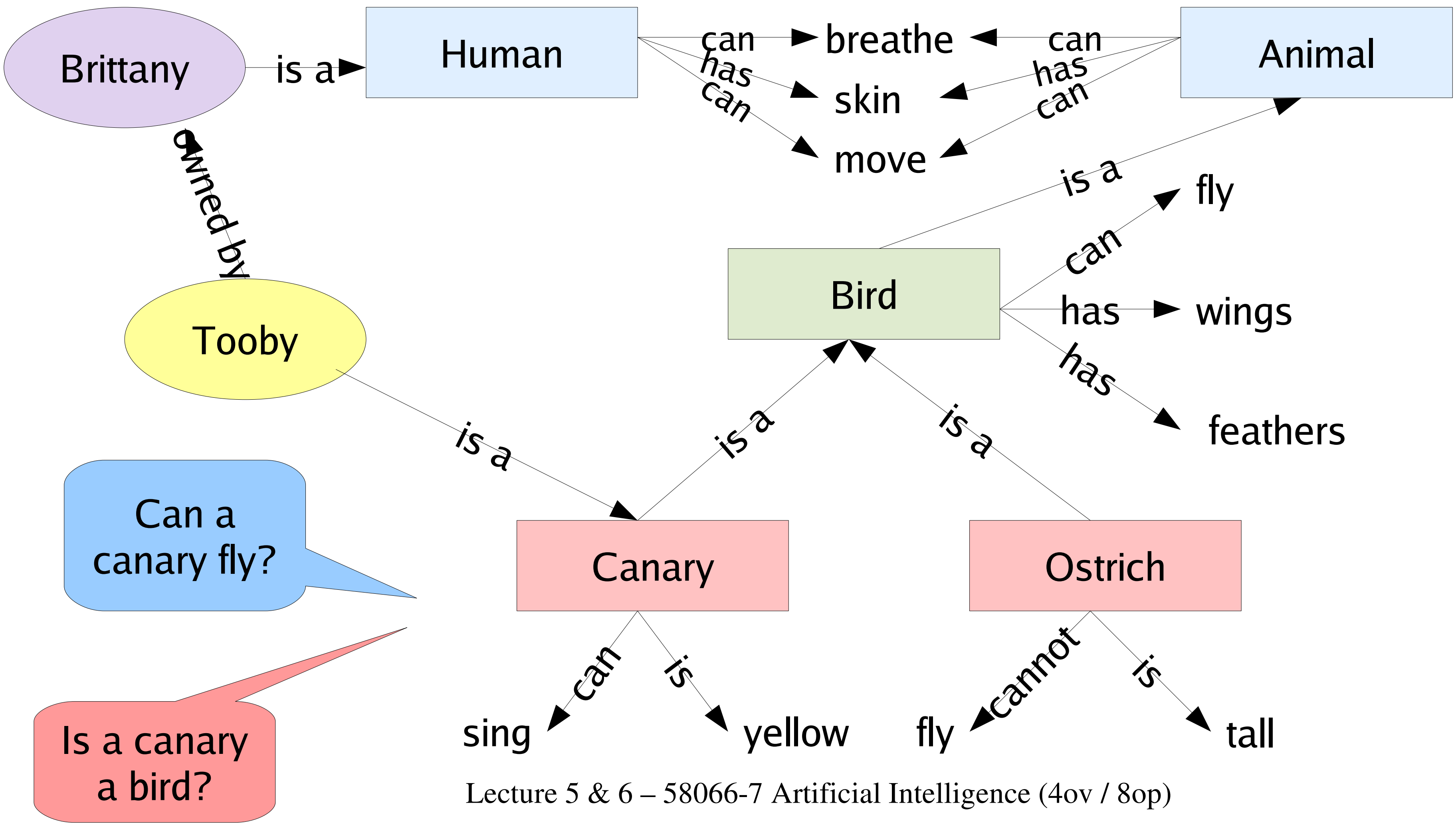
- First implementations in natural language understanding in 1960's; word meaning defined in term of other words.
- WordNet by George A. Miller at Princeton
  - Development started in 1985.
  - Semantic lexicon of English language.
  - Goals:
    - To provide a dictionary and thesaurus.
    - support automatic text analysis and AI applications.
  - Groups words in sets of synonyms (synsets).
  - Provides short and general definitions for words.
  - Records various semantic relations between synsets.
  - As of 2006 contains 150000 words in 115000 synsets.
  - Distinguishes nouns, verbs, adjectives and adverbs.
- http://wordnet.princeton.edu

6.13 Semantics of associations

- A is a part of B — B has A as a part (has-a hierarchy) *e.g., Fido has a tail*.
- A is the superordinate of B — B is subordinate of A (is-a hierarchy) *e.g., Fido is a dog.*
- A denotes the same as B — A denotes the opposite of B
- A is an instant of B ('a kind of' (ako) relation)

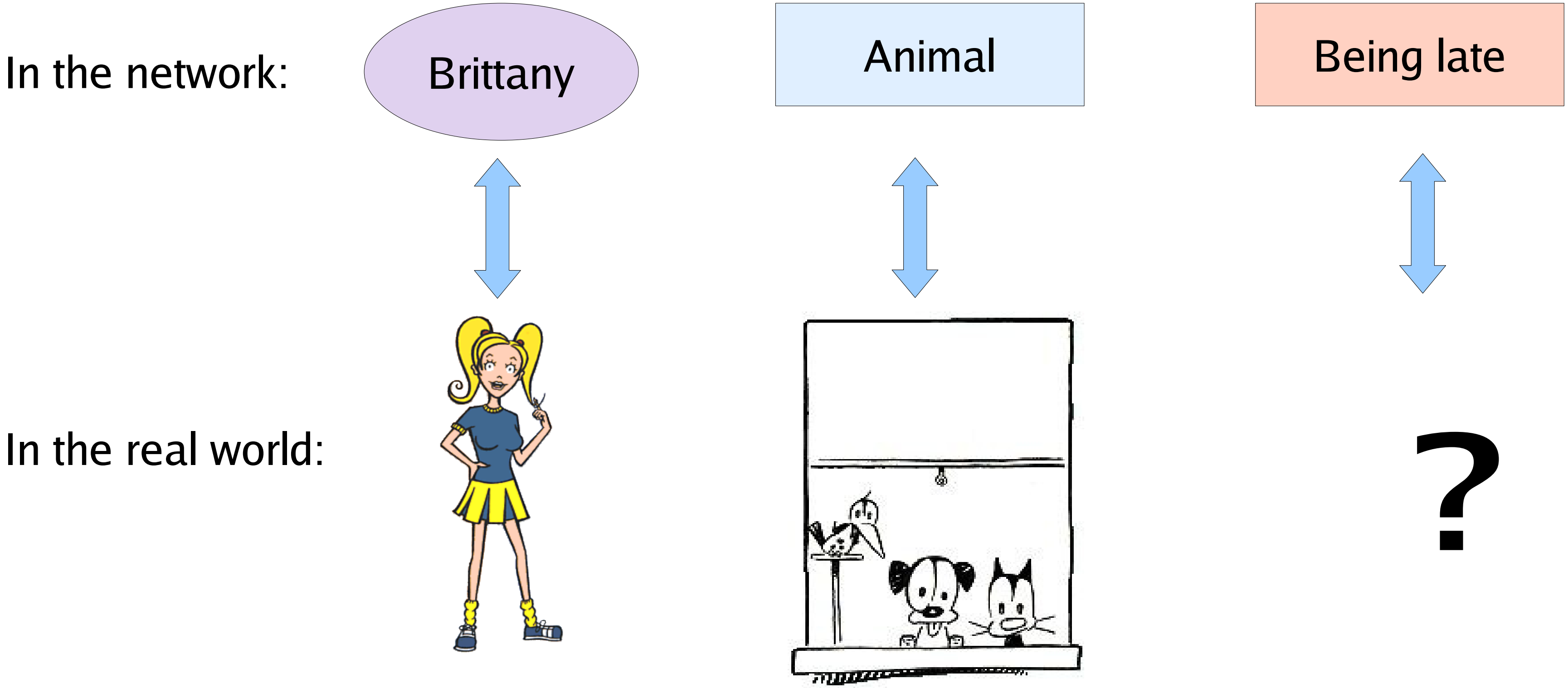
6.14 Example

Derived from laboratory experiments with human subjects (Collins and Quillian, 1969)



6.15 Limitations of semantic nets

- No formal semantics
  - An attempt to standardize link names
  - System is limited by the user's understanding of the meaning of the links.
- Negations cannot be represented.
- Symbol grounding problem:

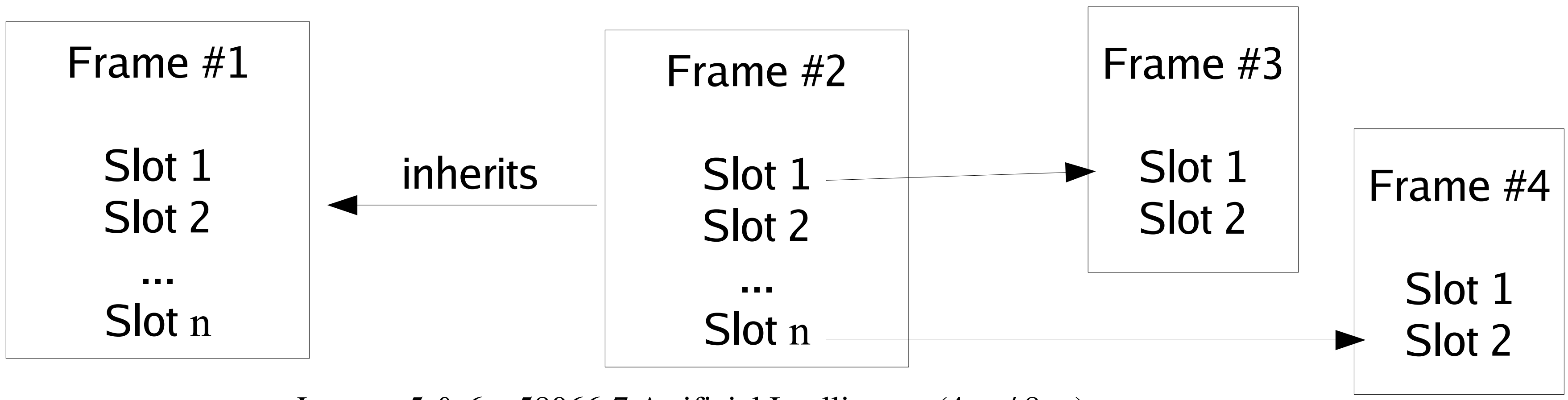


6.16 Frames

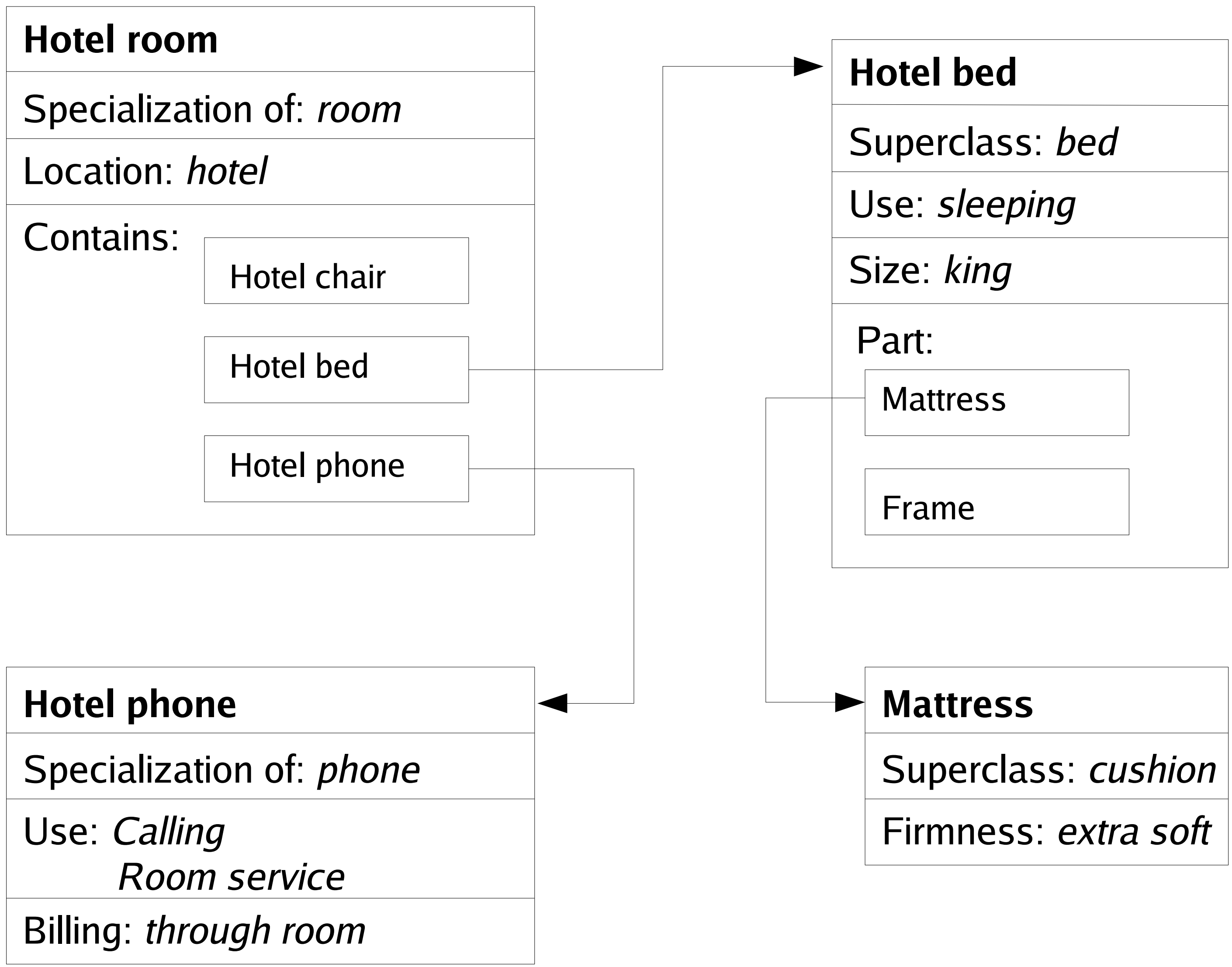
- Frame is a static structure that represents well-founded stereotyped situations (Marvin Minsky, 1975)
- Generalize what objects can be: physical objects, properties, place, situation, or feeling.
- Make inheritance more explicit
- Compared to other KR technologies
  - Compact representation; complex object in a single frame as opposed to in complex network structure
  - All information of an object stored in the same place as opposed to distributed among rules
  - Reflect the organizational structure of the domain; hierarchical organization as opposed everything is in nodes and links

6.17 Frame architecture

- Like records or objects in programming languages.
- Frames can represent classes or instances.
- Consists of slots, that can be
  - Frame identification
  - Relation to other frames
  - Attribute values (particularly default values)



6.18 Example


Hotel room
Specialization of: *room*
Location: *hotel*
Contains:
Hotel chair
Hotel bed
Hotel phone
Hotel bed
Superclass: *bed*
Use: *sleeping*
Size: *king*
Part:
Mattress
Frame
Hotel phone
Specialization of: *phone*
Use: *Calling*
Room service
Billing: *through room*
Mattress
Superclass: *cushion*
Firmness: *extra soft*

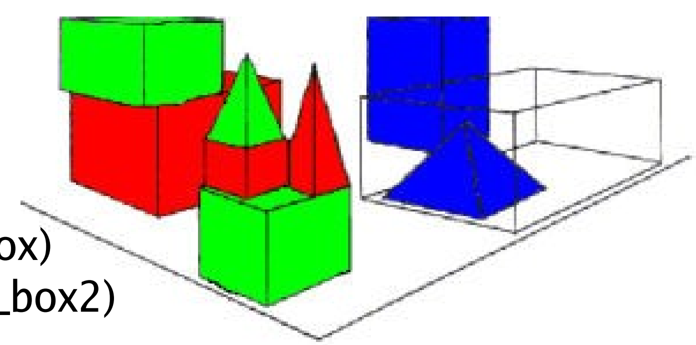
6.19 Frame problem

- Introduced by McCarthy and Hayes in 1969
- Not related to KR technology called frames!
- Related to reasoning about a dynamic (changing) world
- *Frame problem is philosophy:* how rational agent bounds the set of beliefs that needs to be changed once the action is executed.
- *Frame problem in AI:* how to avoid the specification (in formal logic) of all the aspects that do not change as a result of an action.

6.20 Frame problem in AI

- After action is executed, the state of the world changes according to the action's effect.
- Most actions do not change most of the world.
- Non-effects of an action need to be specified for an artificial system → *frame axioms*.
- As the problem size increases keeping track of the things that do not change gets unbearable.
- The problem: how to formulate the effect of an action without having to write down obvious non-effects of the action? → computational issue

on_top(*green_box1, red_box1*)
on_top(*red_pyramid*, *green_box2*)
...
on_floor(*blue_box*)



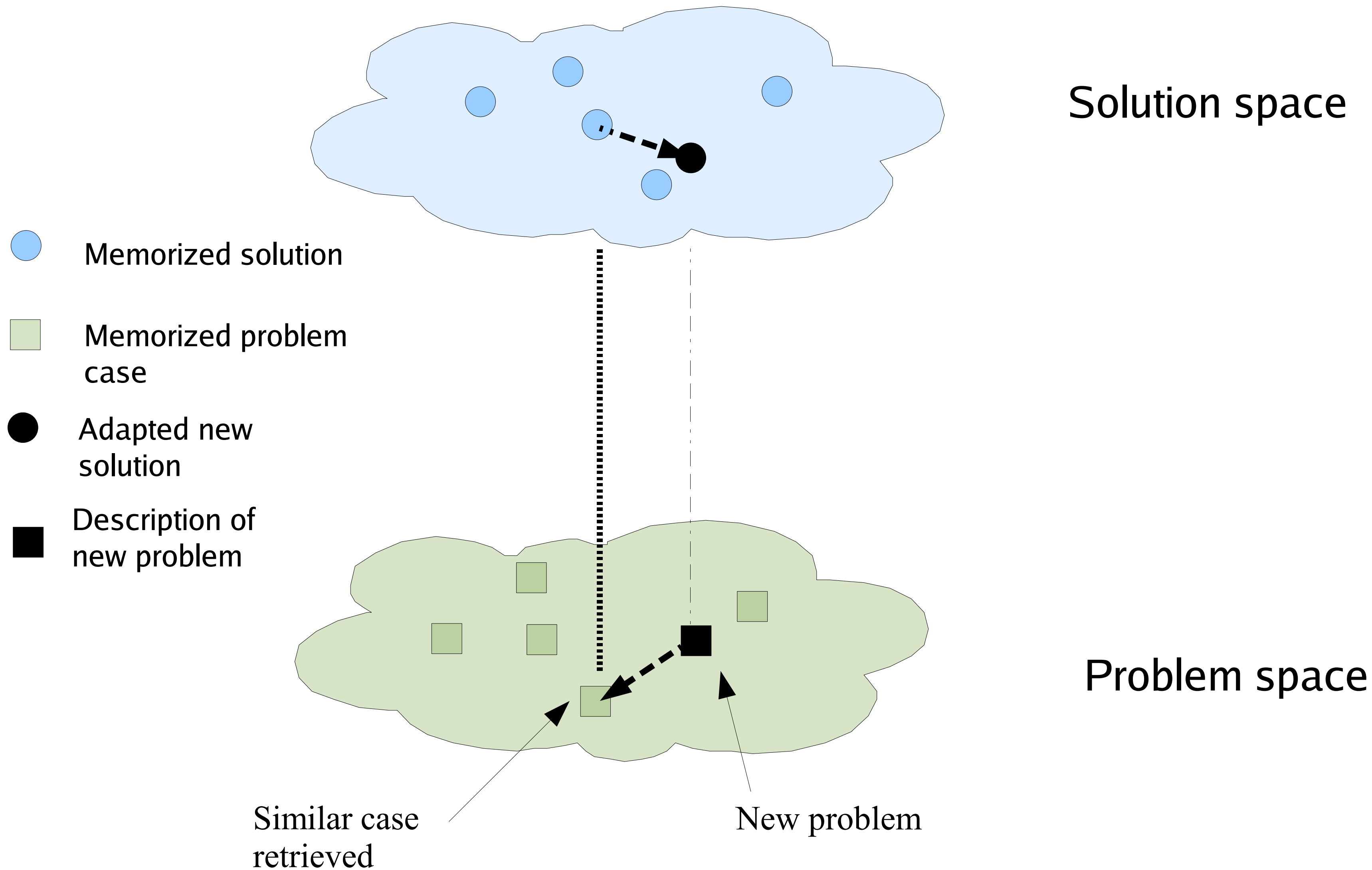
move(green_box1, blue_box)

on_top(green_box1, blue_box)
on_top(red_pyramid, green_box2)
...
on_floor(*blue_box*)

6.21 Case-based reasoning (CBR)

- Retrieve stored relevant *cases* from memory, and adjust them to the current situation:
  - Case = problem + solution + derivation of the solution
  - Case-based reasoning = retrieval + analogy + adaptation + learning
- Why?
  - To model human behavior
  - Design technology to make AI systems more effective
- Interpretative CBR
  - Uses old cases to classify or characterize new situations.
  - Medical education is not just study of anatomy, physiology and disease, but also case histories of patients.
  - Lawyers use past court cases: how law is interpreted now depends on how it has been interpreted in past.
- Problem-solving CBR
  - Uses prior cases to suggest solutions to new problems.

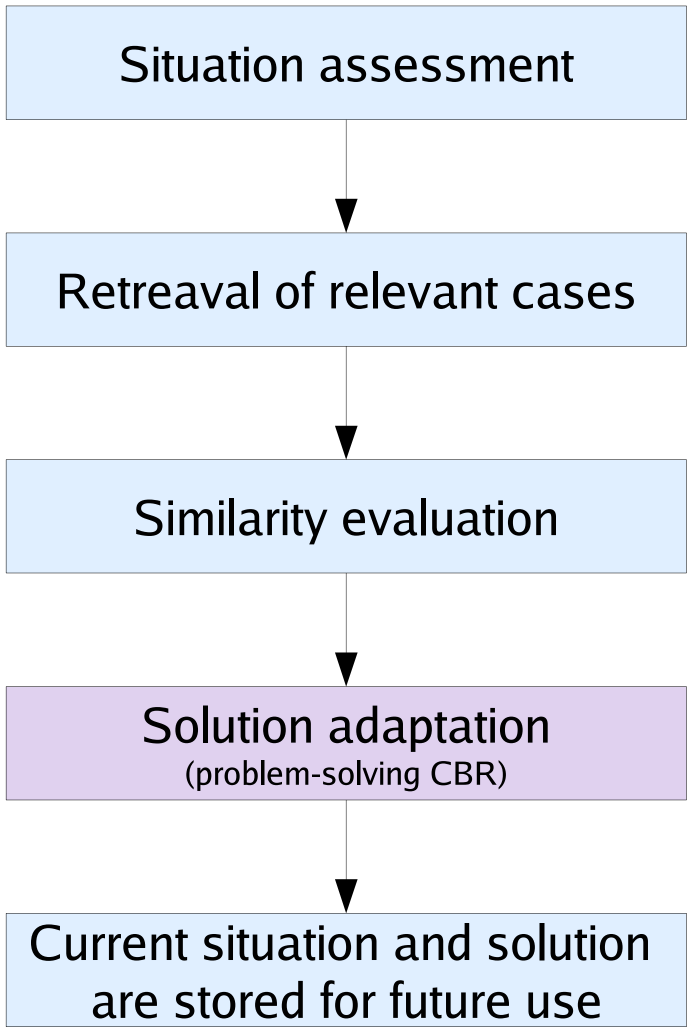
6.22 CBR idea


Solution space
Memorized solution
Memorized problem case
Adapted new solution
Description of new problem
Problem space
Similar case retrieved
New problem

6.23 CBR ideology

- Based on two tenets:
  - World is regular; similar problems have similar solutions.
  - Types of problems an individual encounters tend to recur.
- Can be applied both in routine and novel problems, but can handle only cases that (can) happen in practice.
- No domain theory required.
- Learns from failures and successes:
  - Task failures — learn better solutions
  - Expectation failures — learn to anticipate similar problems in the future (e.g., failing plan)
- Like in EBL allows learning from single examples.
- Unlike EBL does not generalize at storage, but in retrieval.

6.24 CBR learning



To determine relevant features of the current situation

Top-down: Use problem description to find relevant cases
Bottom-up: Use features that help to discriminate between cases

Which features to compare?

Central open challenge in CBR:
- usually done by a rule-based system
- problem: considerable need of domain knowledge
- Adapt the context, not the case

Store both successes and failures
→ speed-up
→ acquisition of knowledge

6.25 Concept maps

- Graphical tools for organizing and representing knowledge
- Developed by Joseph D. Novak in 1972 for the study of children's understanding of science.
- Consist of nodes (ovals or boxes) representing concepts, and relationships between concepts represented by linking lines.
- Links are labelled by *linking words* or *linking phrases* specifying the nature of relationship.

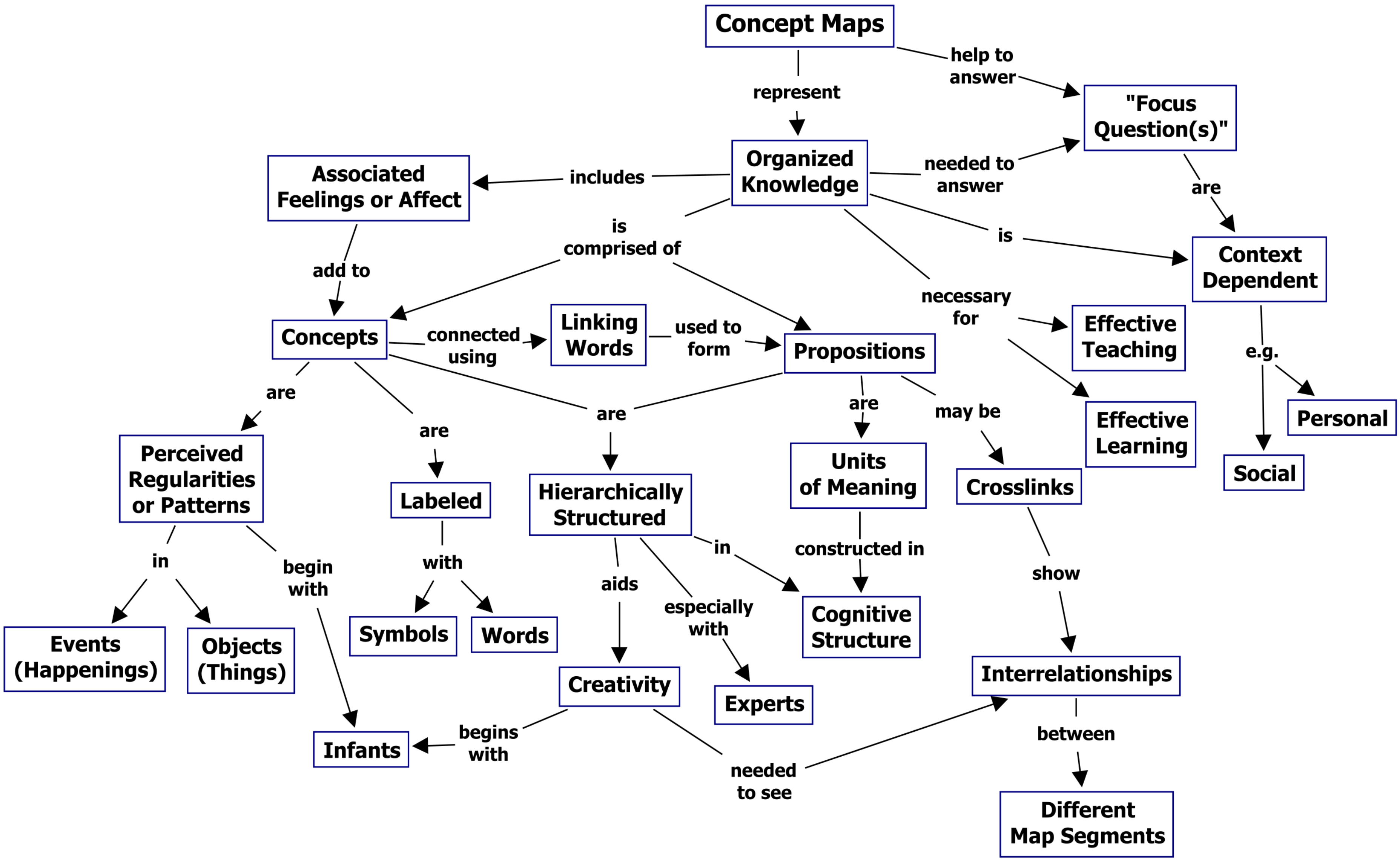
6.26 Some definitions

- *Concept* =
  - Perceived regularity in event or objects
  - Records of events or objects
- Concept is designated by a *label*:
  - One word
  - Symbol (e.g., %)
  - Several words
- *Proposition* is a unit of meaning that contains two or more connected concepts forming a meaningful statement.

6.27 Properties of concept maps

- Usually downward branching tree structure, but can have upward links.
- Hierarchical: more general concepts at the top
- Cross-links
  - Relationships or links across concepts maps form different knowledge domains.
  - Clarifies how concept in one domain of knowledge is related to concept in another domain.
- May contain specific examples (not concepts)

6.28 Cmap of concept maps



6.29 Use of concept maps

- To answer a specific question: quest to understand some specific situation or event which gives the map the context.
- In education for
  - Learning concepts (e.g., note taking)
  - Teaching specific topics
  - Evaluation/assessment of understanding
  - Curriculum planning
- Capture tacit expert knowledge, i.e., knowledge that is hard to articulate in words.
- Sharing and communication of knowledge
- Brain-storming

6.30 The CMap Tools

- http://cmap.ihmc.us/
- Developed at the Institute for Human and Machine Cognition (IHMC), Florida.
- Software tool that can be used to construct, navigate, share and criticize knowledge models represented as concept maps.
- The Cmap Tools client can be downloaded free from the above site. Supported platforms are Windows, MacOS, Linux and Solaris (Sparc).

6.31 Case study: Subsumption architecture

- Rodney Brooks, 1986
- Designed to implement physical robots
- Complicated intelligent behavior is decomposed into many “simple” horizontally layered behavioral modules.
  - Each layer implements one goal of the system: exploration, wandering, and obstacle avoidance.
  - Higher layers are more abstract.
  - Each layer's goal subsumes the goals from layers below. For instance, 'explore the world' subsumes 'wander around', which subsumes 'avoid an object.'
  - Each layer accesses all sensory data, and generate actions to actuators.
- Simple reactive behaviors constitute coherent, goal-directed behavior. No global state.
- Ideally, complex hierarchy of behaviors can sufficient to produce increasingly intelligent behavior.

6.32 Subsumption philosophy

- World is too complicated, unpredictable and dynamic, to be reliably represented.
- Uncertainty in sensing → accurate world model impossible ("World is its own best model.")
- No learning, no planning, no language, prewired patterns of behavior.
- Advantages:
  - Modularity
  - Emphasis on iterative development and testing of real-time systems.
- Disadvantage
  - Conflicting goals → cannot have many layers.
  - Low flexibility at runtime.

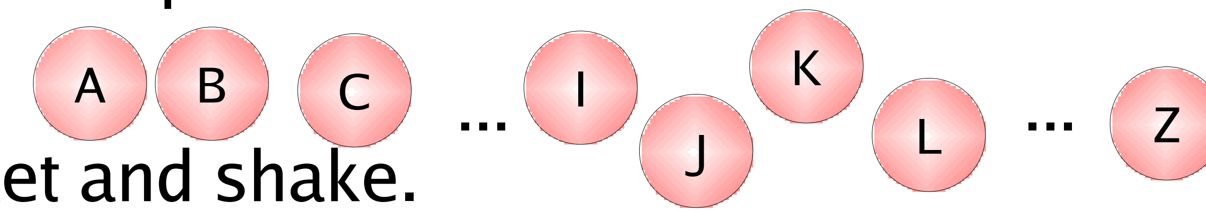
6.33 Case study: Copycat

- Architecture built by Melanie Mitchell 1993 (as a student of Douglas Hofstadter at Indiana University)
- Models high-level perception in domain of solving problems of analogy: “Hat is to head as glove is to what?”
- Works on an idealized textual domain:
  - *abc* is to *abd* as *ijk* is to *?*
  - *abc* is to *abd* as *iijjkk* is to ?
- Driving idea is so called “fluid concepts.”
- Semantic network like architecture that constantly evolves as a response to the experience within the environment.

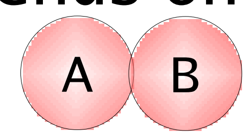
This idea shared with Brooks' subsumption architecture.

6.34 Copycat idea

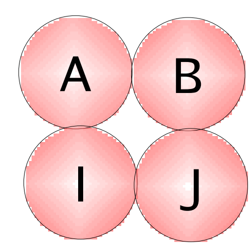
- Imagine balls with letters painted on them with Velcro strips attached:



- Put balls in the bucket and shake.
- As the balls collide they may stick together and bond.
- Strength of the bond depends on how strongly related the letters on the balls are:



- Pairs may also bond together if the relationships between the associated letters are the same:



- Shaking will also break bonds.
- Eventually the system stabilizes, and the bonding ceases.

6.35 Copycat architecture

- Three main components:
  - The Workspace
  - The Coderack
  - The Slipnet
- Temperature
  - Measures the degree of perceptual organization in the system.
  - Higher the temperature, less information is available, and decisions are made more in random.
  - Lower the temperature, the more confident the system is that it has a solution.

6.36 Workspace and coderack

- Workspace
  - Global working (short-term) memory that other components of the system can inspect.
  - Initially contains the input and eventually the answer.
- Coderack
  - Contains *codelets*, pieces of code that operate on the objects in the workspace.
  - Build perceptual structures in workspace and activate concepts in Slipnet.

6.37 Slipnet

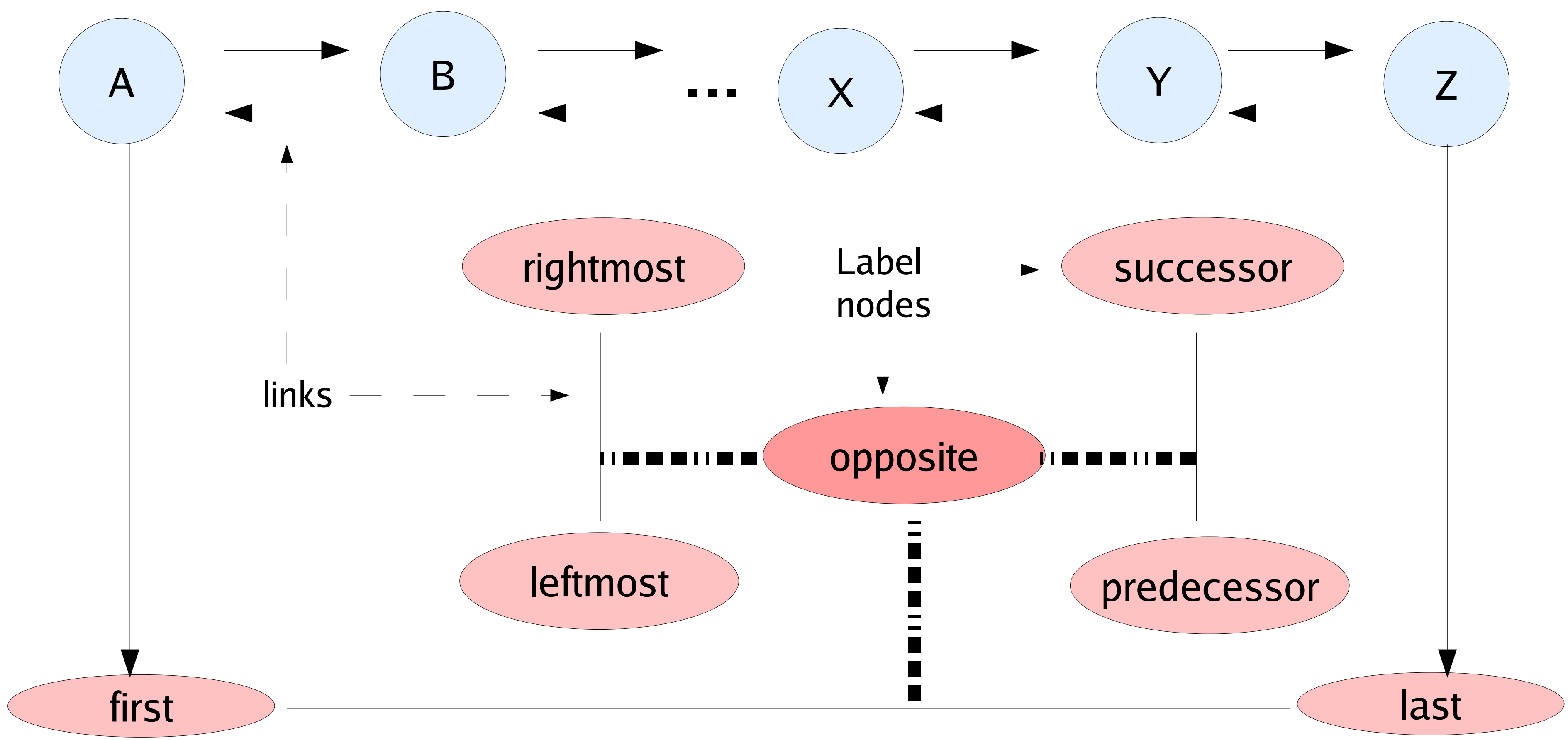
- System's long-term memory
- Nodes are the center of concepts:
  - letters 'a', 'b', ...
  - relations such as 'opposite', 'last', 'successor' etc.
- Concepts have activation level and pre-assigned 'depth'; 'successor' is deeper than 'last'.
- Concept's scope includes the nodes within certain distance.
- Nodes are connected by links that are labeled by (labelling) nodes.

The activation level of the labelling nodes determines the distance between connected nodes.

The distance determines the strength of association between concepts.

The spread of activation from node to node is controlled by the distance of the nodes.

Activation decay is determined by the depth of the concept.



6.38 Limitations of Copycat

- The domain is semantically too weak to be impressive.
- No learning; it starts a new in every task.
- It cannot deal with sequences that are not successors or predecessors (e.g., strings with every other letter skipped)
- It cannot handle mappings with more than one change.
- It cannot create analogies of its own.
- A large collection of hand-coded parameter values that can be (or must be) tweaked to every task.