Emergence of Semantic Knowledge: Experience-Driven Representational Change in a Simple Neural Network

James L. McClelland
Department of Psychology and
Center for Mind, Brain, and Computation
Stanford University

Some Phenomena in Conceptual Development

- Progressive differentiation of concepts
- Illusory correlations and U-shaped developmental trajectories
- Conceptual reorganization
- Domain- and property-specific constraints on generalization
- Acquired sensitivity to an object’s causal properties

- What underlies these phenomena?

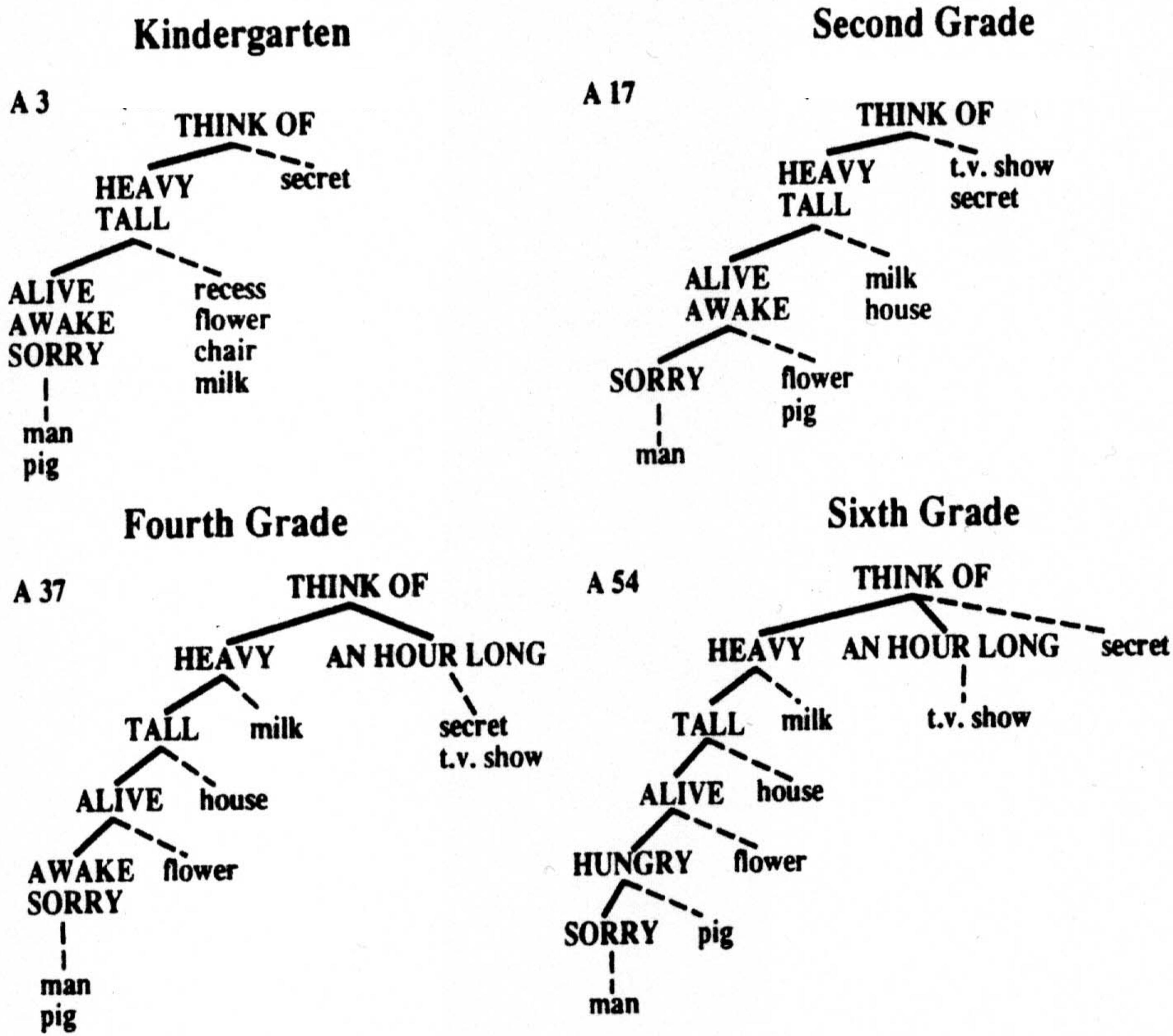
My Answer to this Question

- A domain general cognitive mechanism underlies the development and elaboration of conceptual knowledge.
- This mechanism can be embodied in a class of brain-like neural network models that adhere to three basic principles:
  - Representation is a pattern of activation over neuron-like processing units.
  - Knowledge is stored in the connections among the units.
  - Conceptual development and learning are the result of gradual changes in the strengths of the connections.
- Models built on these principles are sensitive to coherent covariation.
- This sensitivity is the main cause of all of the phenomena.

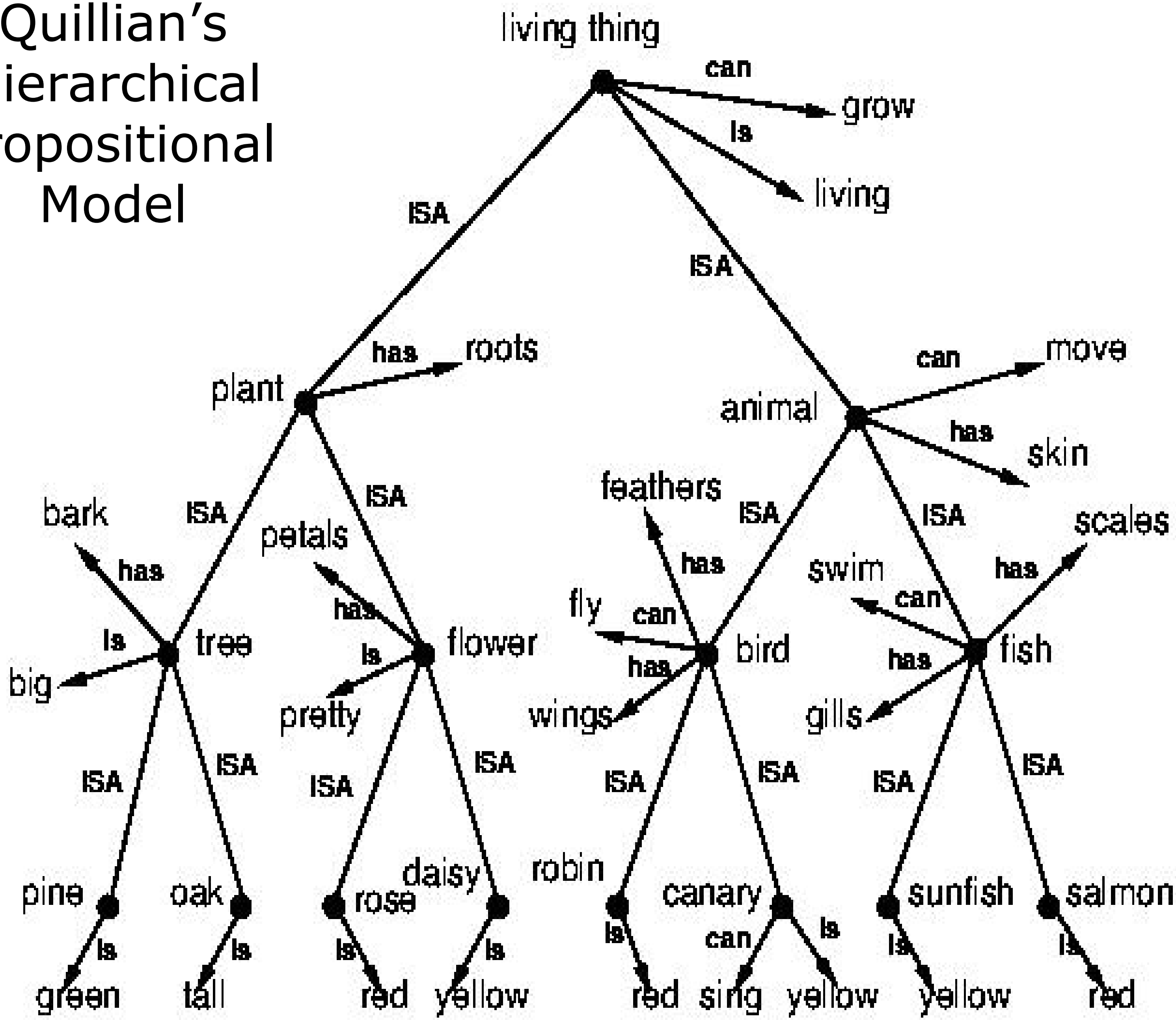
What is Coherent Covariation?

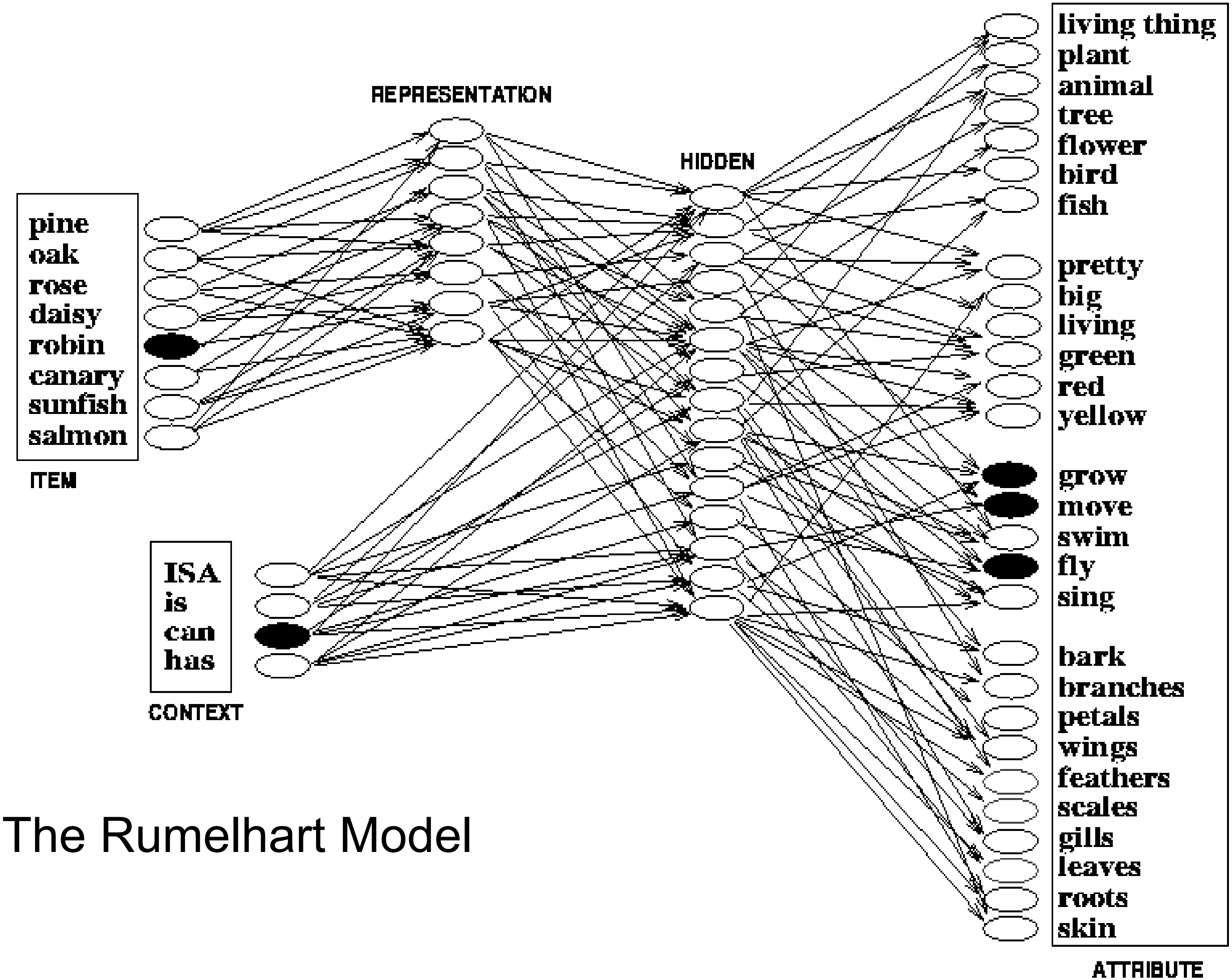
- The tendency of properties of objects to occur together in clusters.
  - One example:
    - Has wings
    - Can fly
    - Is light
  - And another:
    - Has roots
    - Has rigid cell walls
    - Can grow tall

Differentiation of Conceptual Knowledge in Development


Kindergarten
A 3
THINK OF
HEAVY
TALL
secret
ALIVE
AWAKE
SORRY
recess
flower
chair
milk
man
pig
Second Grade
A 17
THINK OF
HEAVY
TALL
t.v. show
secret
ALIVE
AWAKE
milk
house
SORRY
flower
pig
man
Fourth Grade
A 37
THINK OF
HEAVY
AN HOUR LONG
TALL
milk
secret
t.v. show
ALIVE
house
AWAKE
SORRY
flower
man
pig
Sixth Grade
A 54
THINK OF
HEAVY
AN HOUR LONG
secret
TALL
milk
t.v. show
ALIVE
house
HUNGRY
flower
SORRY
pig
man

Quillian's Hierarchical Propositional Model



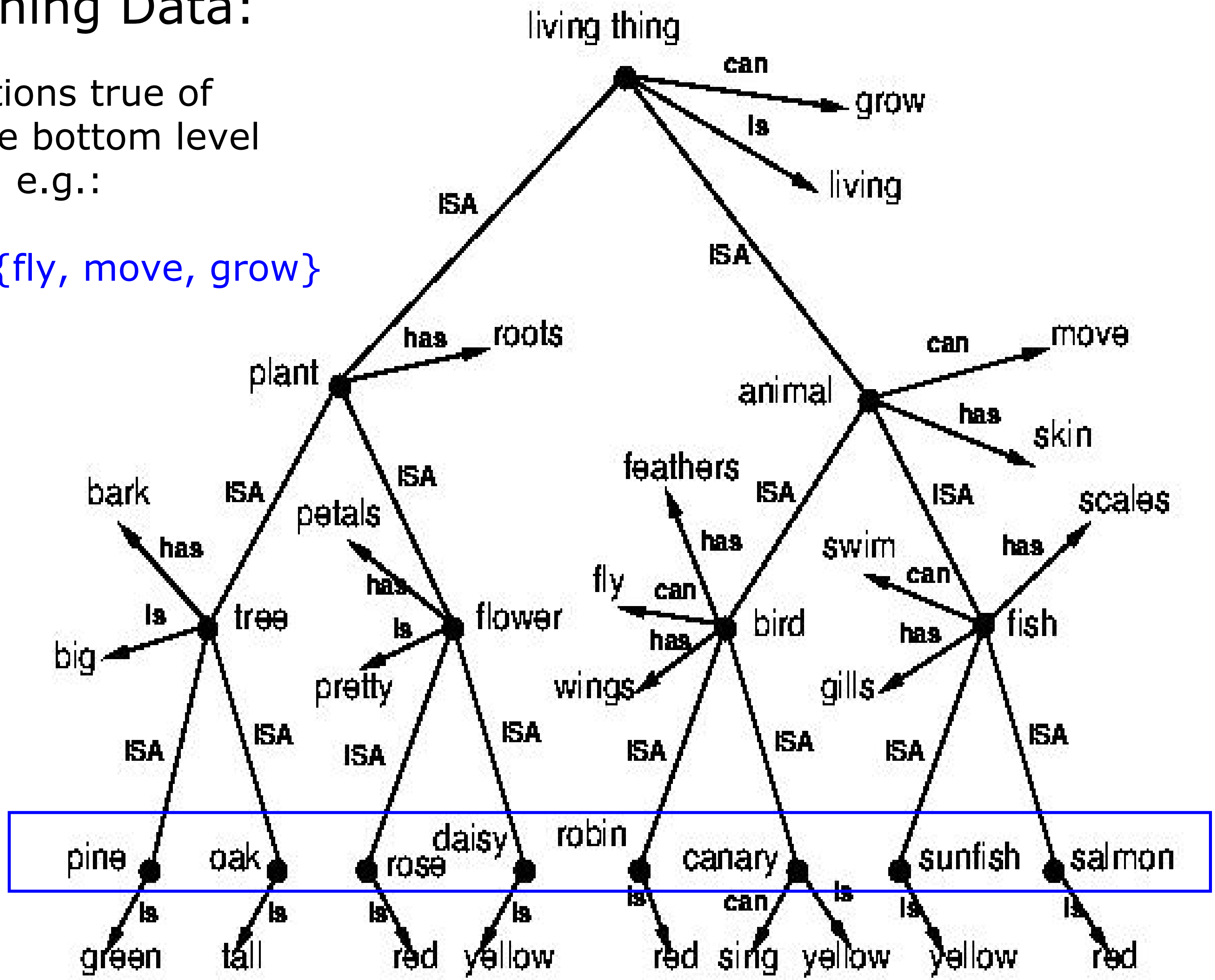


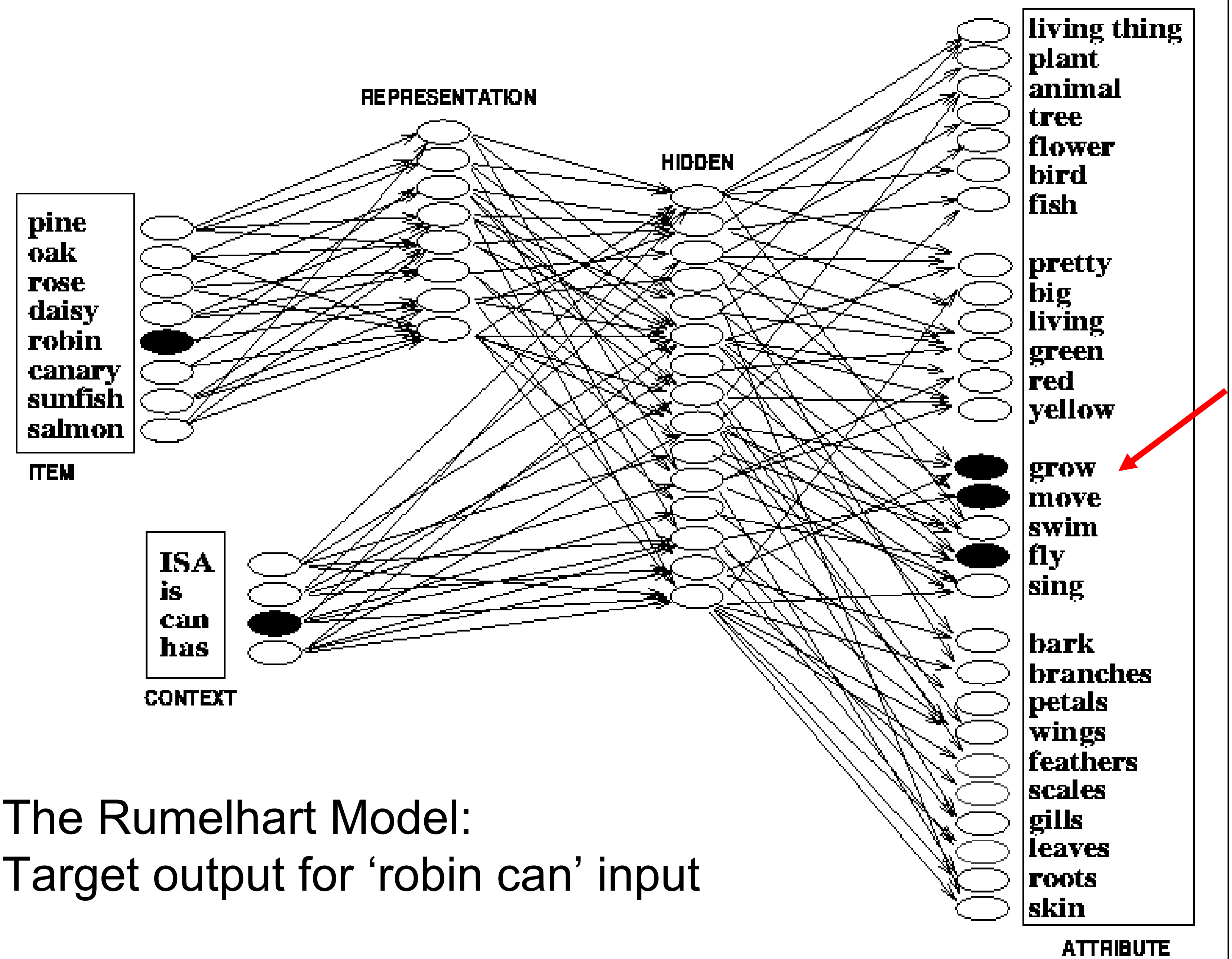
The Rumelhart Model

The Training Data:

All propositions true of items at the bottom level of the tree, e.g.:

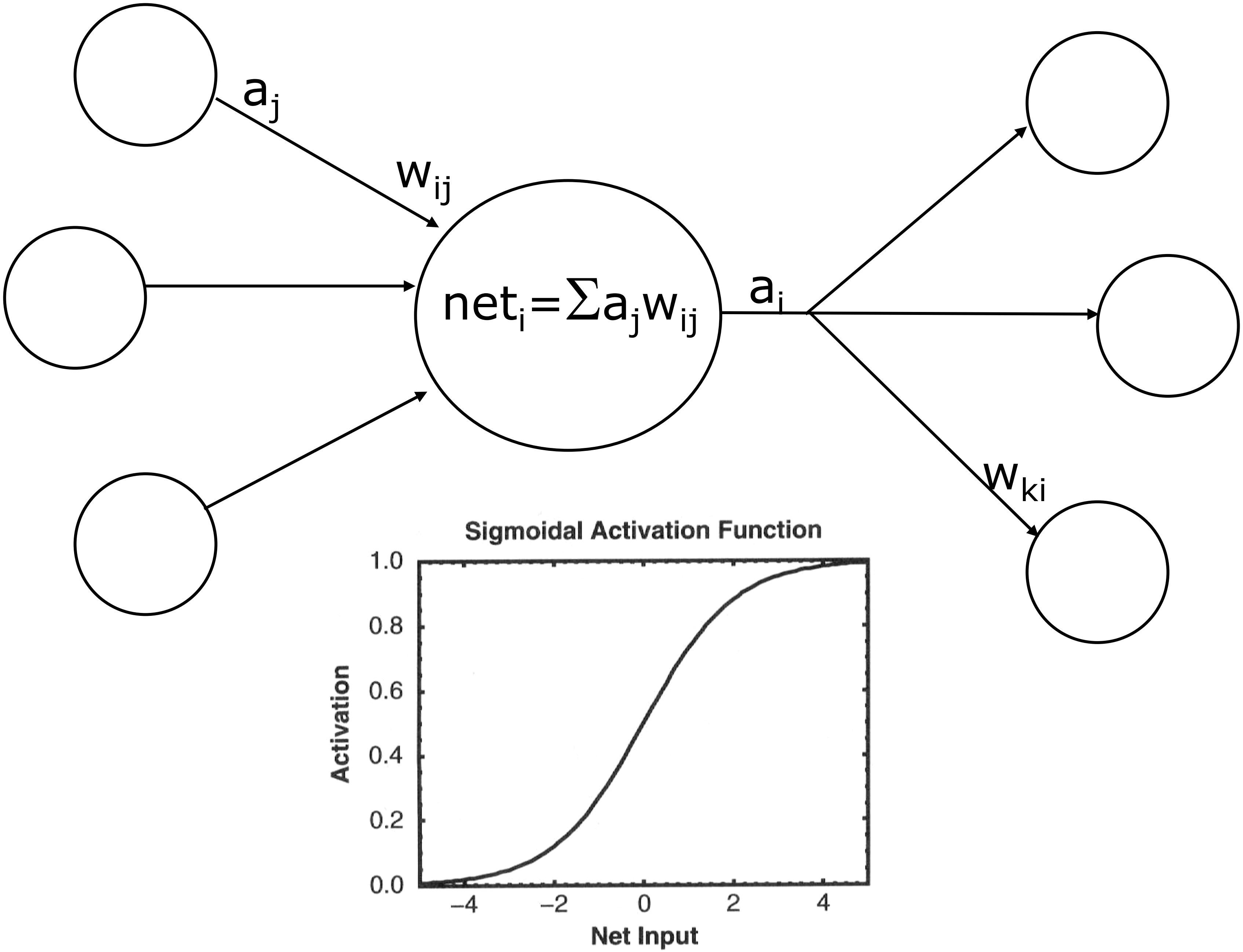
Robin can {fly, move, grow}



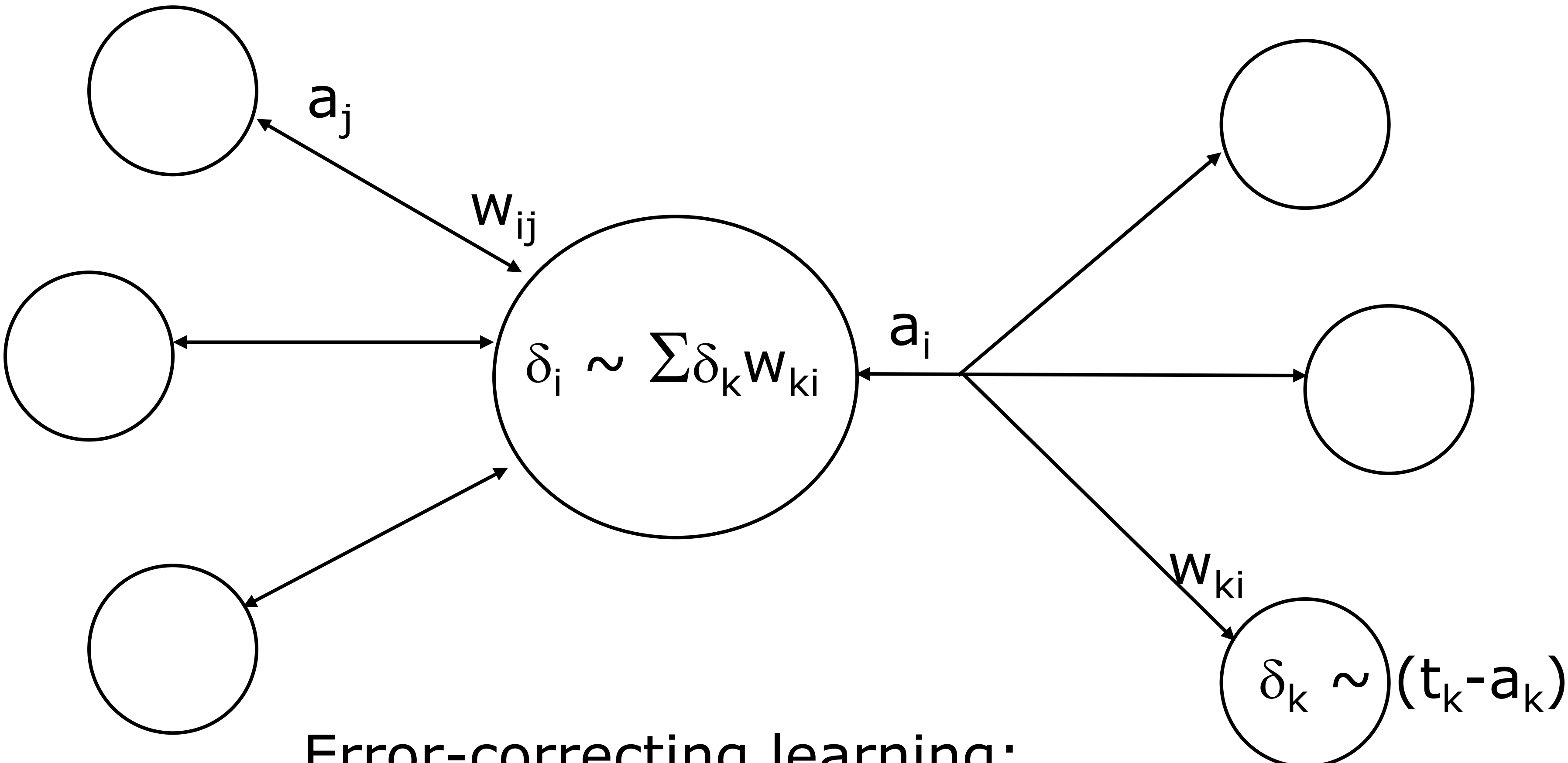


The Rumelhart Model:
Target output for 'robin can' input

Forward Propagation of Activation



Back Propagation of Error (δ)

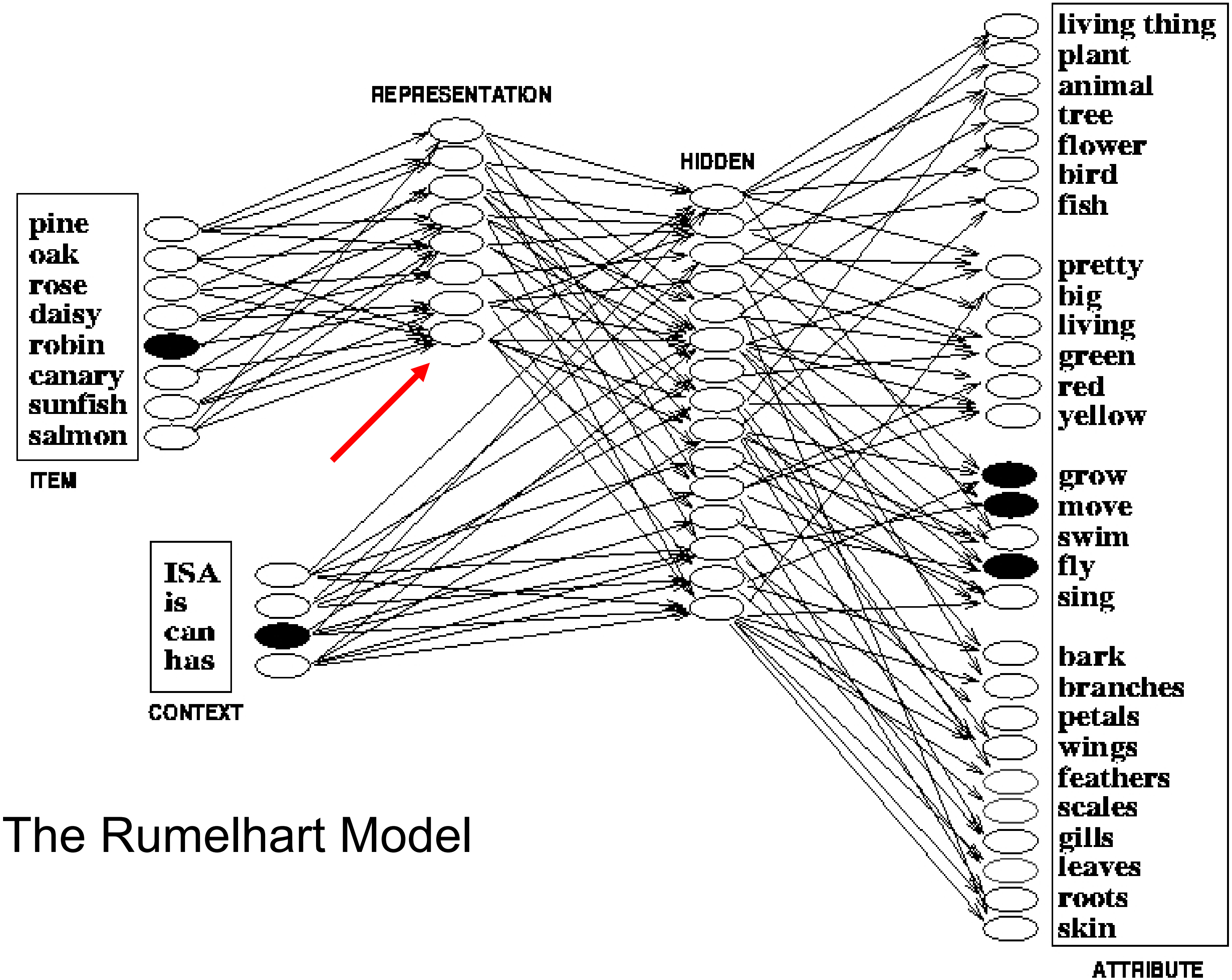


Error-correcting learning:

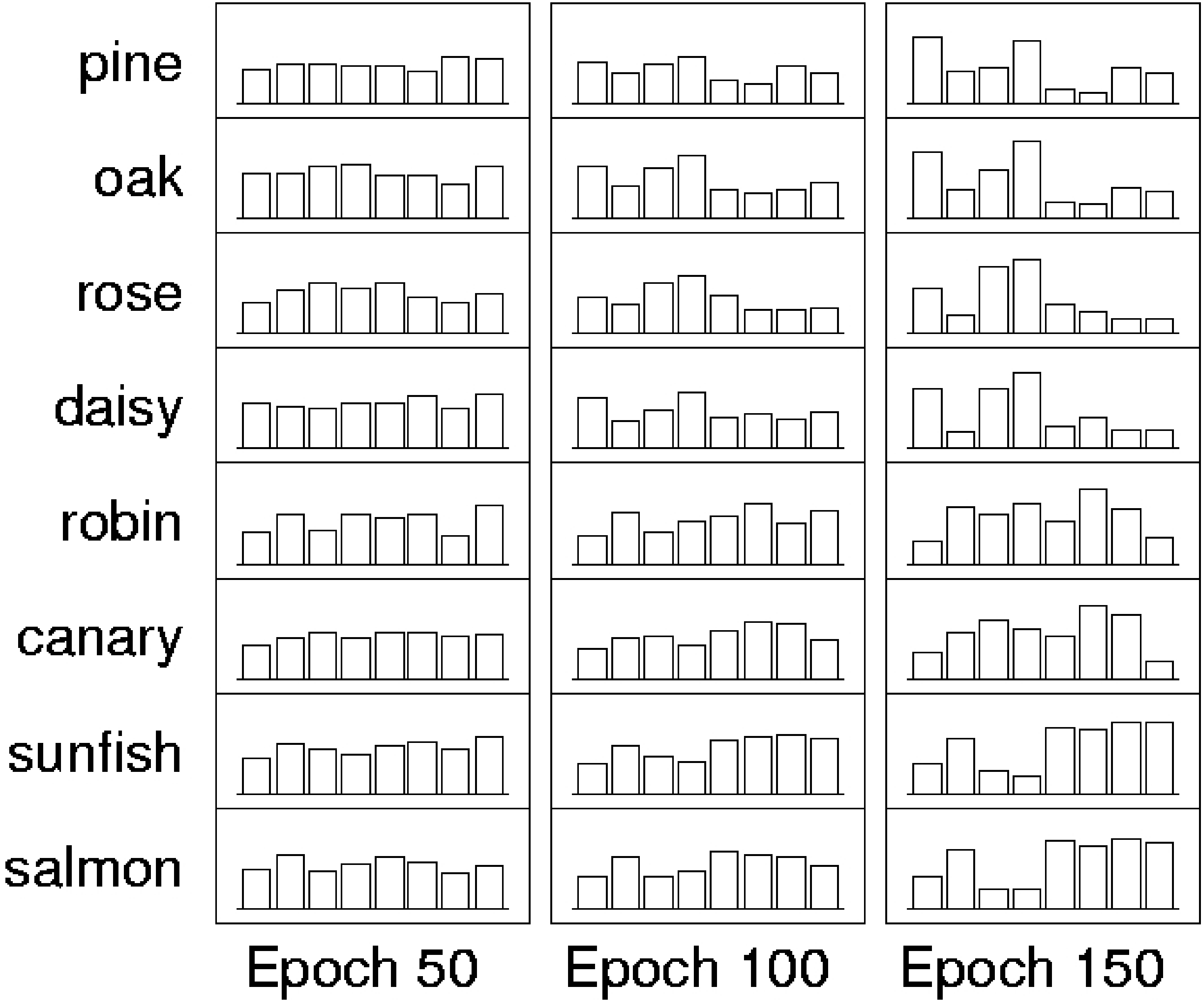
At the output layer: $\Delta w_{ki} = \varepsilon \delta_k a_i$

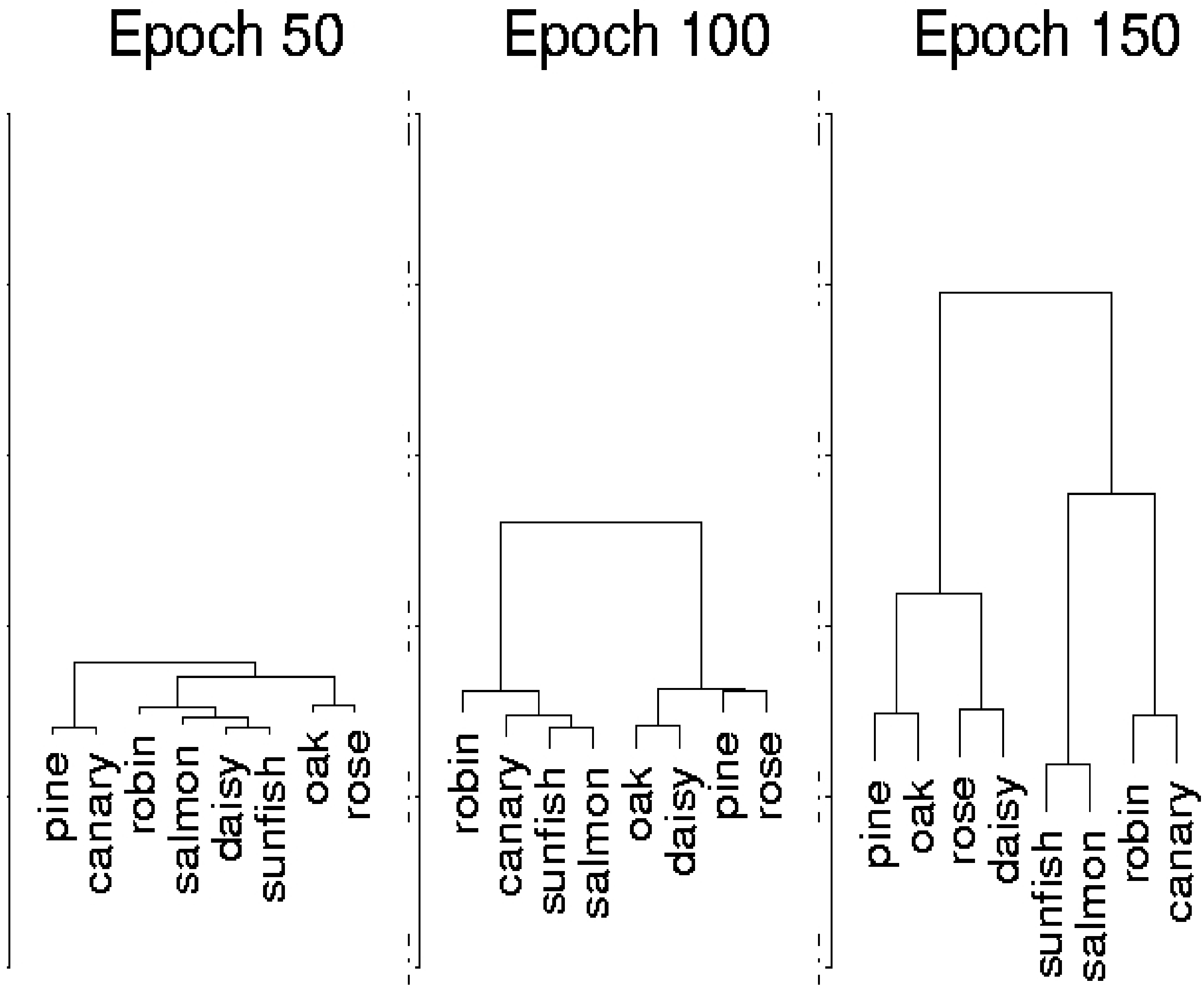
At the prior layer: $\Delta w_{ij} = \varepsilon \delta_j a_j$

…



The Rumelhart Model


pine
oak
rose
daisy
robin
canary
sunfish
salmon
Epoch 50
Epoch 100
Epoch 150


Epoch 50
Epoch 100
Epoch 150
pine
canary
robin
salmon
daisy
sunfish
oak
rose
robin
canary
sunfish
salmon
oak
daisy
pine
rose
pine
oak
rose
daisy
sunfish
salmon
robin
canary

What Drives Progressive Differentiation?

- Waves of differentiation reflect coherent covariation of properties across items.
- Patterns of coherent covariation are reflected in the principal components of the property covariance matrix.
- Figure shows attribute loadings on the first three principal components:
  - 1. Plants vs. animals
  - 2. Birds vs. fish
  - 3. Trees vs. flowers
- Same color = features covary in component
- Diff color = anti-covarying features

"Now wait just a minute…"

- Didn't you *tell the network* the taxonomic organization directly?
  - Pine ISA Tree, Plant
  - Robin ISA Bird, Animal
- Yes we did.
  - We do think names kids hear for things affect their conceptual representations.
- But labels aren't necessary as long as an item's properties exhibit coherent covariation.

Coherence Training Patterns

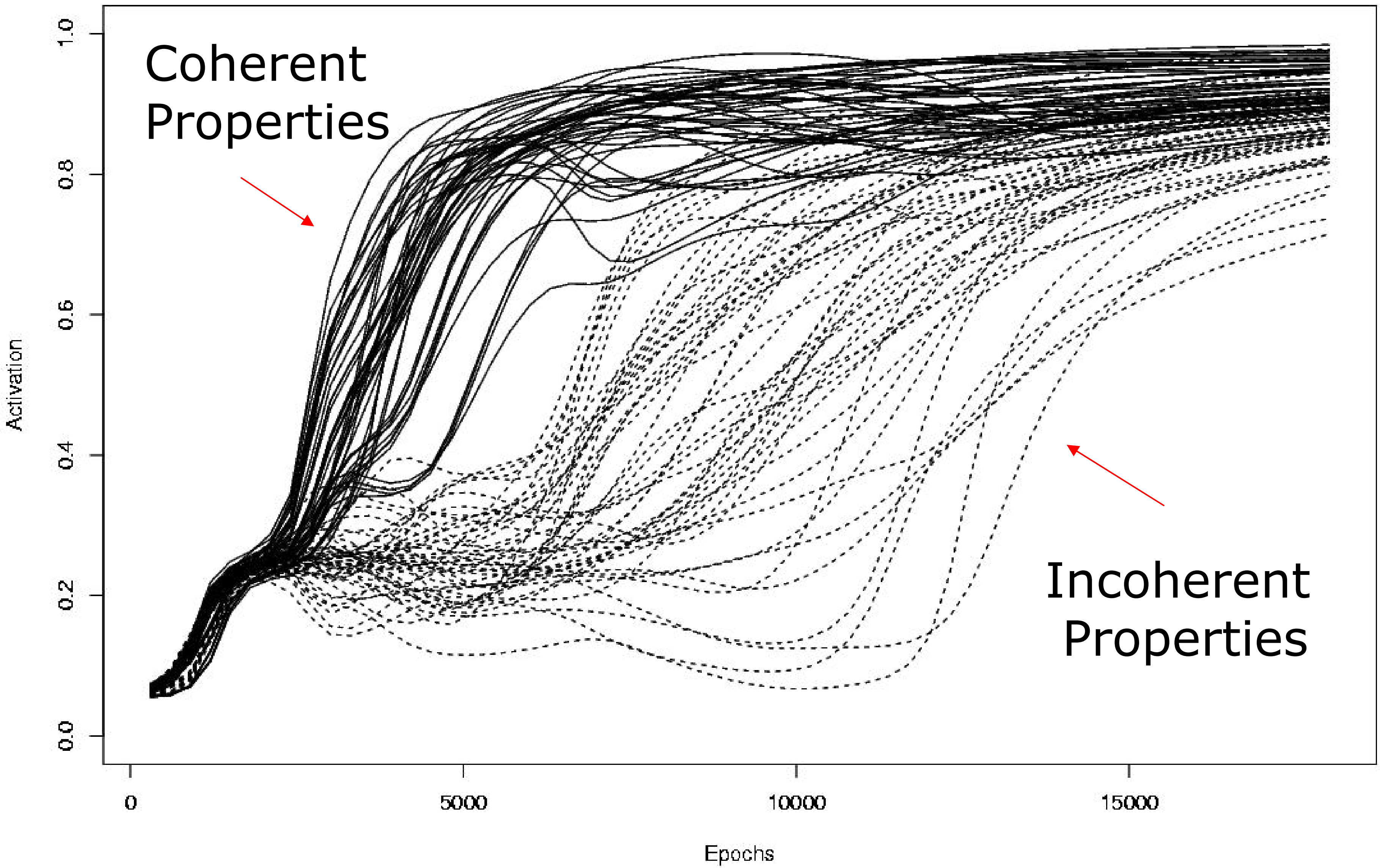
Properties

Items	Coherent												Incoherent											
1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0
2	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
4	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1
5	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0
6	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
7	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
8	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0
9	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0
10	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
11	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0
12	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
13	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0	0	1
14	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0	0	1	0	0	0	1	0
15	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0
16	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	1	0	0	0

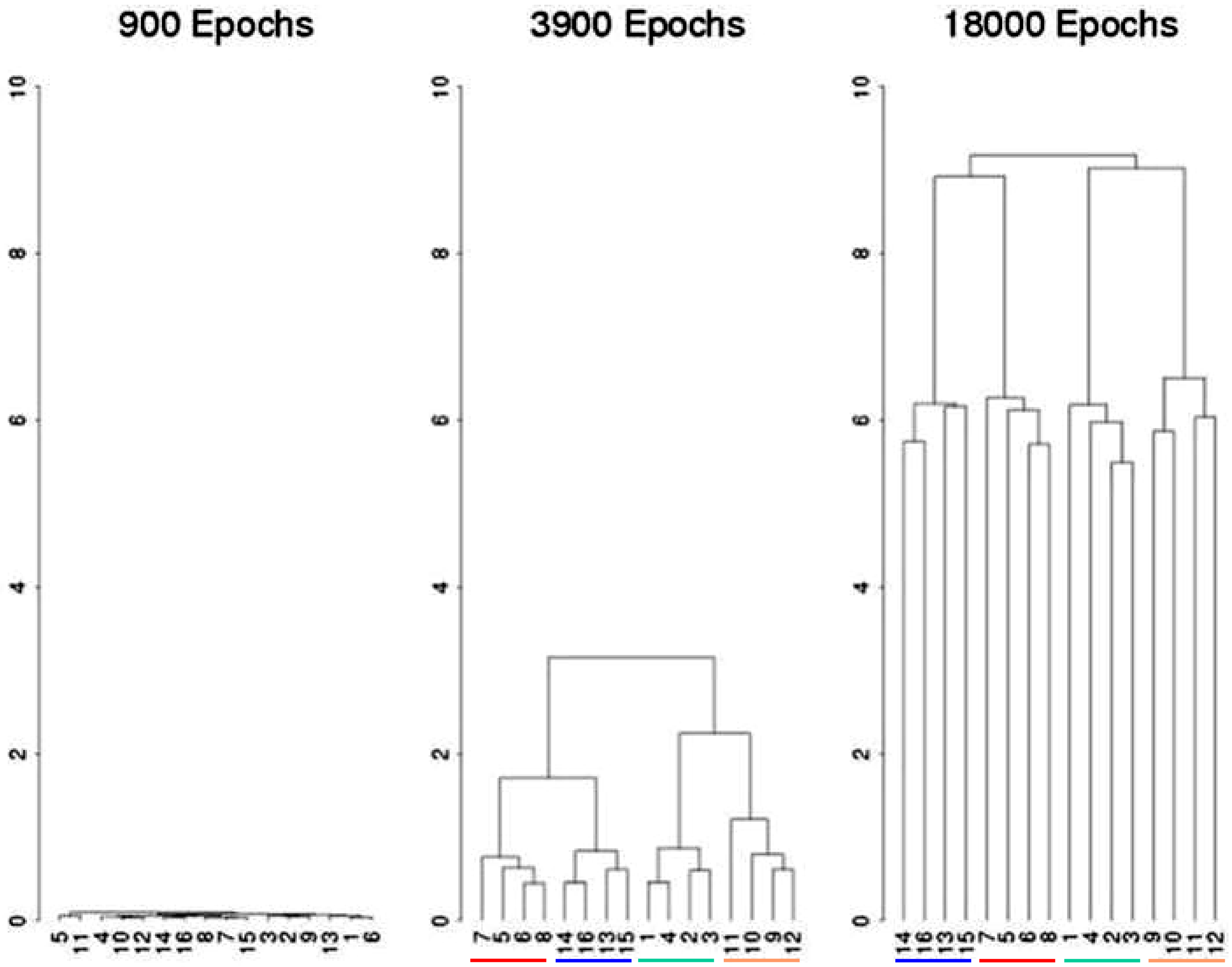
No labels are provided

Each item and each property occurs with equal frequency

Effects of Coherence on Learning



Effect of Coherence on Representation



Effects of Coherent Variation on Learning in Connectionist Models

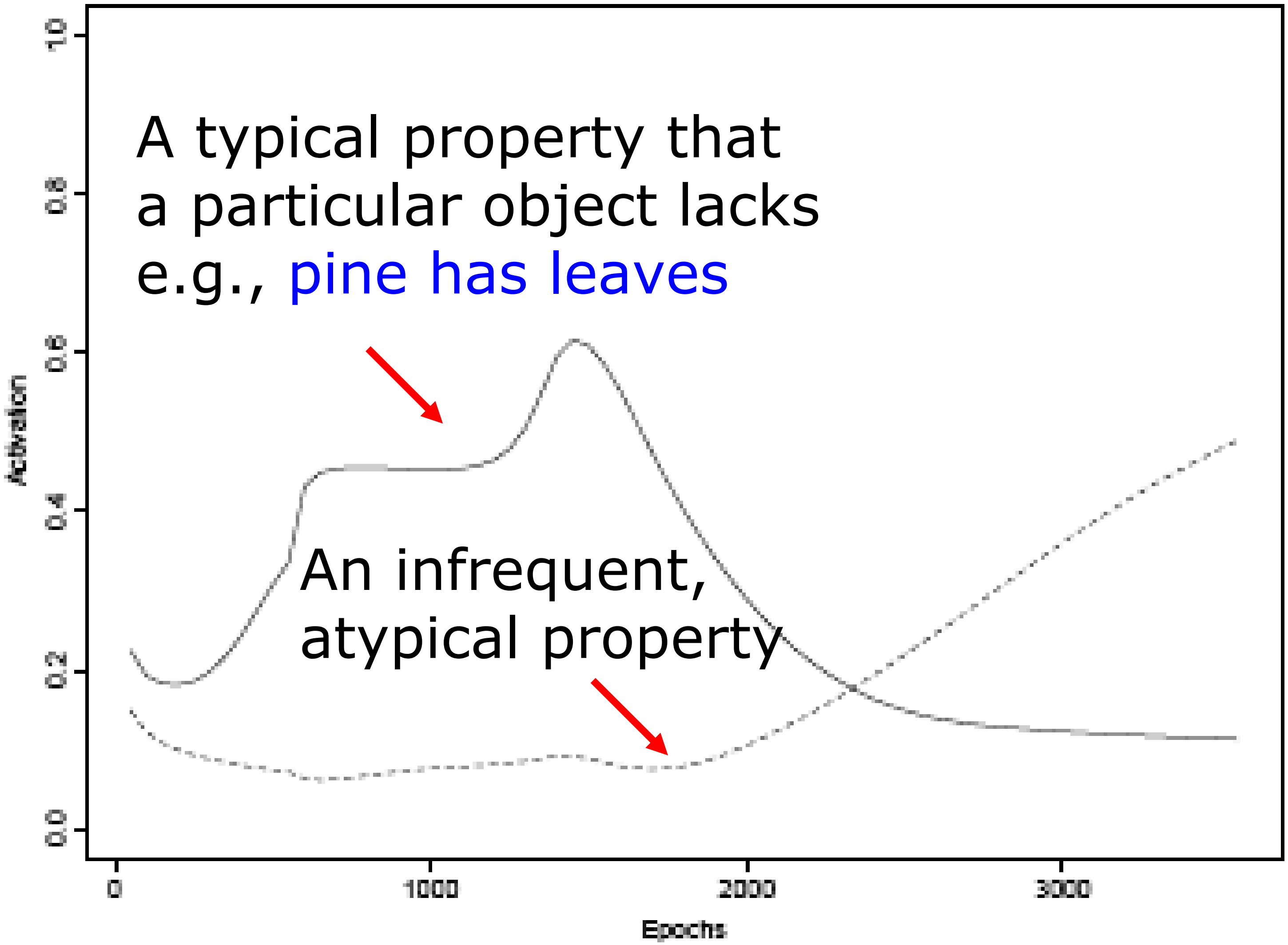
- *Attributes that vary together create the acquired concepts that populate our conceptual landscape, and determine which properties are central and which are incidental to a given concept.*
  - *Labeling of these concepts or their properties is in no way necessary.*
  - *But it is easy to learn names for such concepts.*
- *Arbitrary properties (those that do not co-vary with others) are very difficult to learn*.
  - *And it is harder to learn names for concepts that are only differentiated by such arbitrary properties.*

Where are we on that list of Phenomena?

- ✓ Progressive differentiation of concepts
- ➢ Illusory correlations and U-shaped developmental trajectories
- ➢ Conceptual reorganization
- ➢ Domain- and property-specific constraints on generalization
- Acquired sensitivity to an object’s causal properties

Illusory Correlations

- Rochel Gelman found that children think that all animals have feet.
  - Even animals that look like small furry balls and don't seem to have any feet at all.


A typical property that
a particular object lacks
e.g., pine has leaves
An infrequent,
atypical property
Activation
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
2000
3000
Epochs

Conceptual Reorganization (Carey, 1985)

- Carey demonstrates that young children ‘discover’ the unity of plants and animals as living things with many shared properties only around the age of 10.
- She suggests that the coalescence of the concept of living thing depends on learning about diverse aspects of plants and animals including
  - Nature of life sustaining processes
  - What it means to be dead vs. alive
  - Reproductive properties
- Can reorganization occur in a connectionist net?

Conceptual Reorganization in the Model

- Consider a domain in which superficial appearance information, which is not coherent with much else, is always available…

- And there is a pattern of coherent covariation that is contingently available in different contexts.

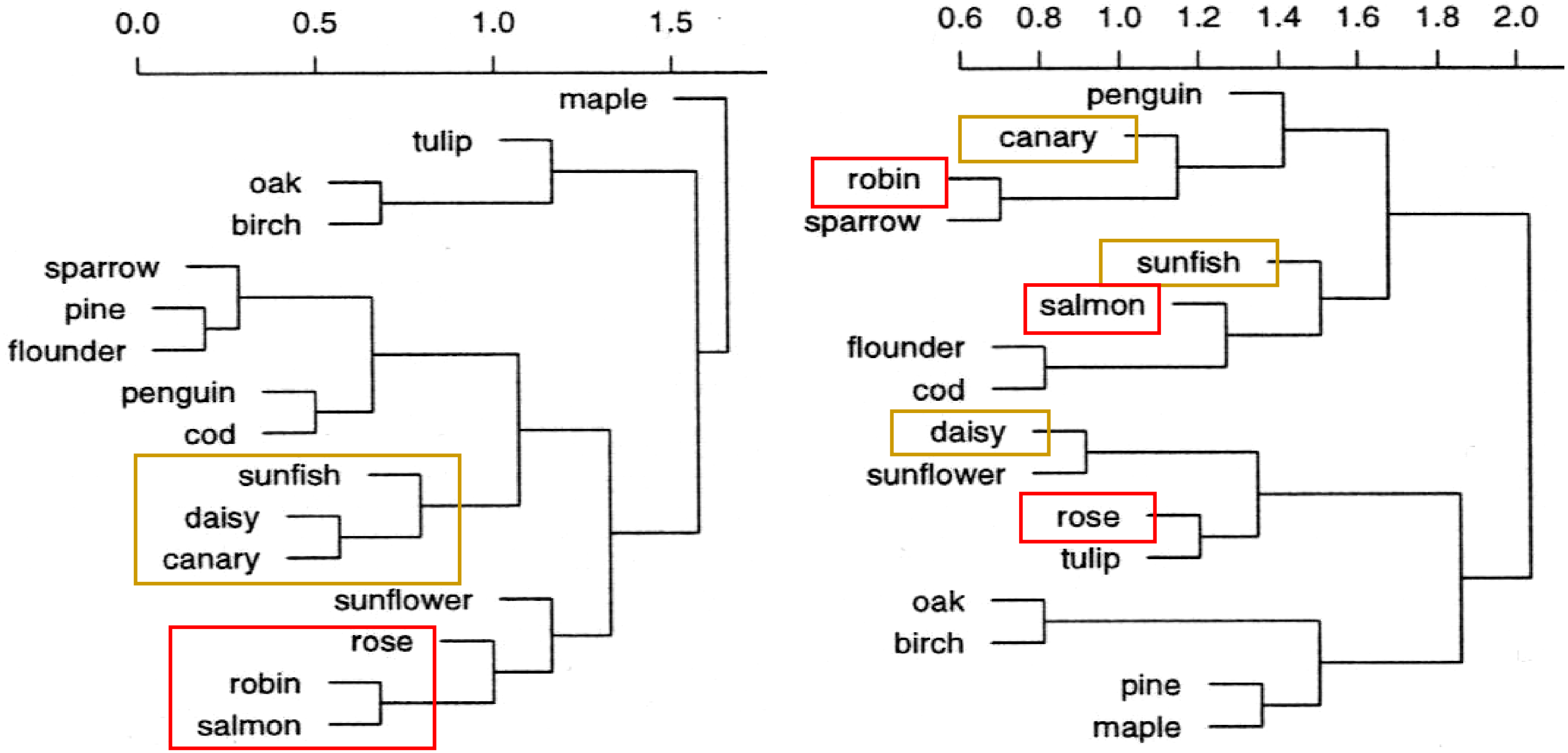
- The model forms initial representations based on superficial appearances.

- … and gradually discovers the shared structure that cuts across the different contexts, reorganizing its representations.

Reorganization Simulation

- Data from 21 plant and animal concepts is used in training.
- In each training pattern, the input is an item and one of the three relations: ISA, HAS, or CAN.
- The target includes all of the superficial appearance properties (IS properties) plus the properties appropriate for the relation.

Organization of Conceptual Knowledge Early and Late in Development

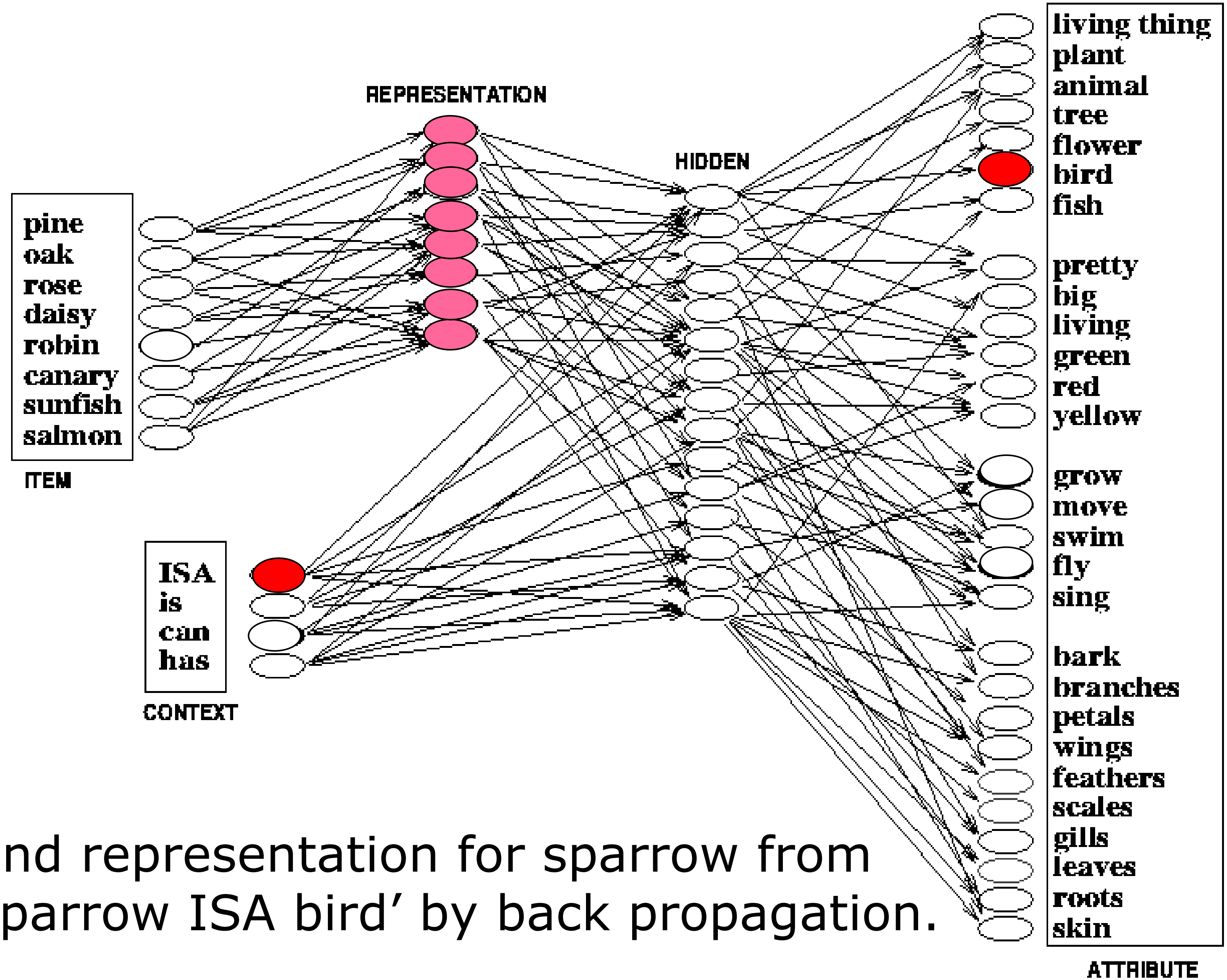


Domain Specificity

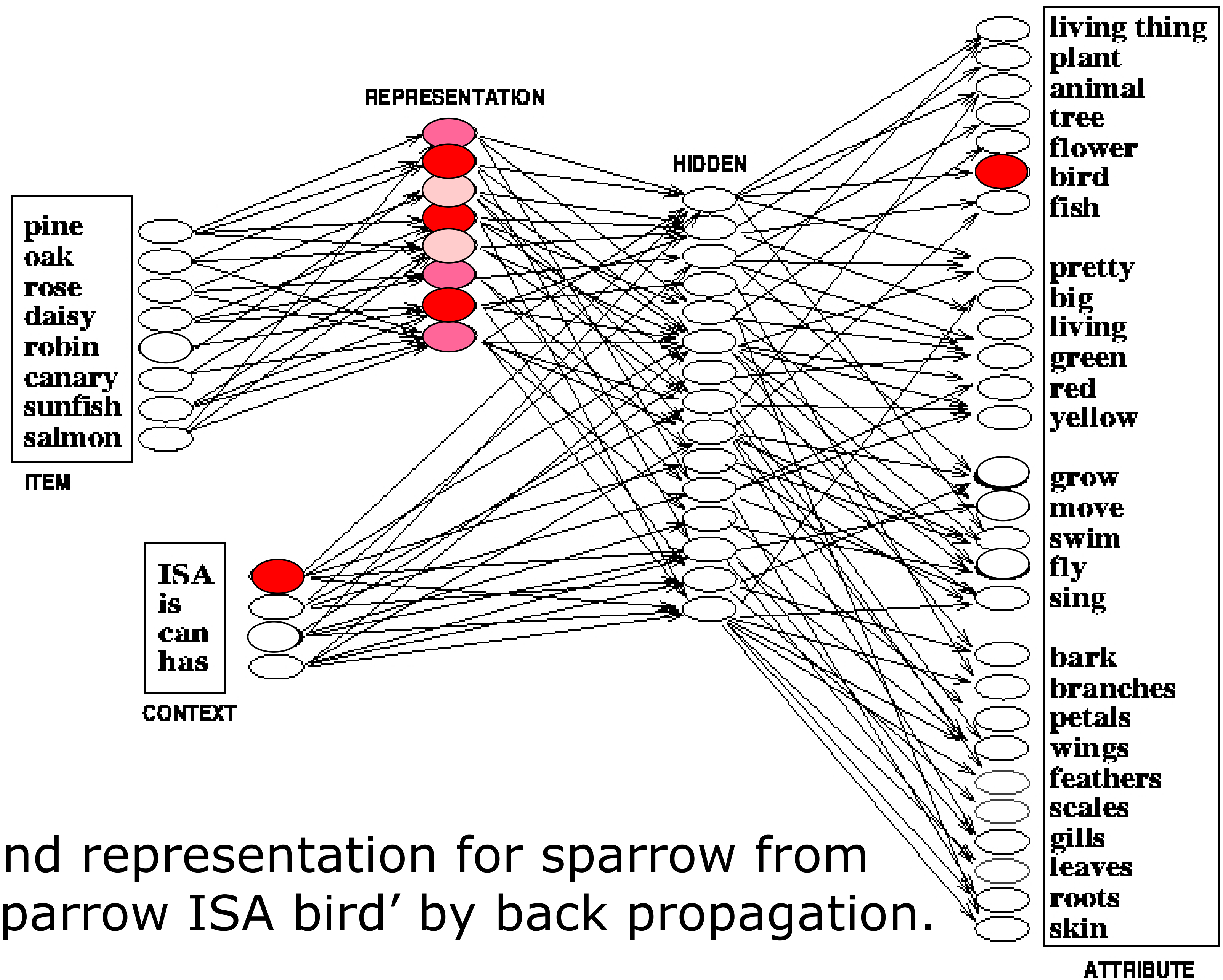
- What constraints are required for development and elaboration of domain-specific knowledge?
  - Are domain specific constraints required?
  - Or are there general principles that allow for acquisition of conceptual knowledge of all different types?

Inference and Generalization in the PDP Model

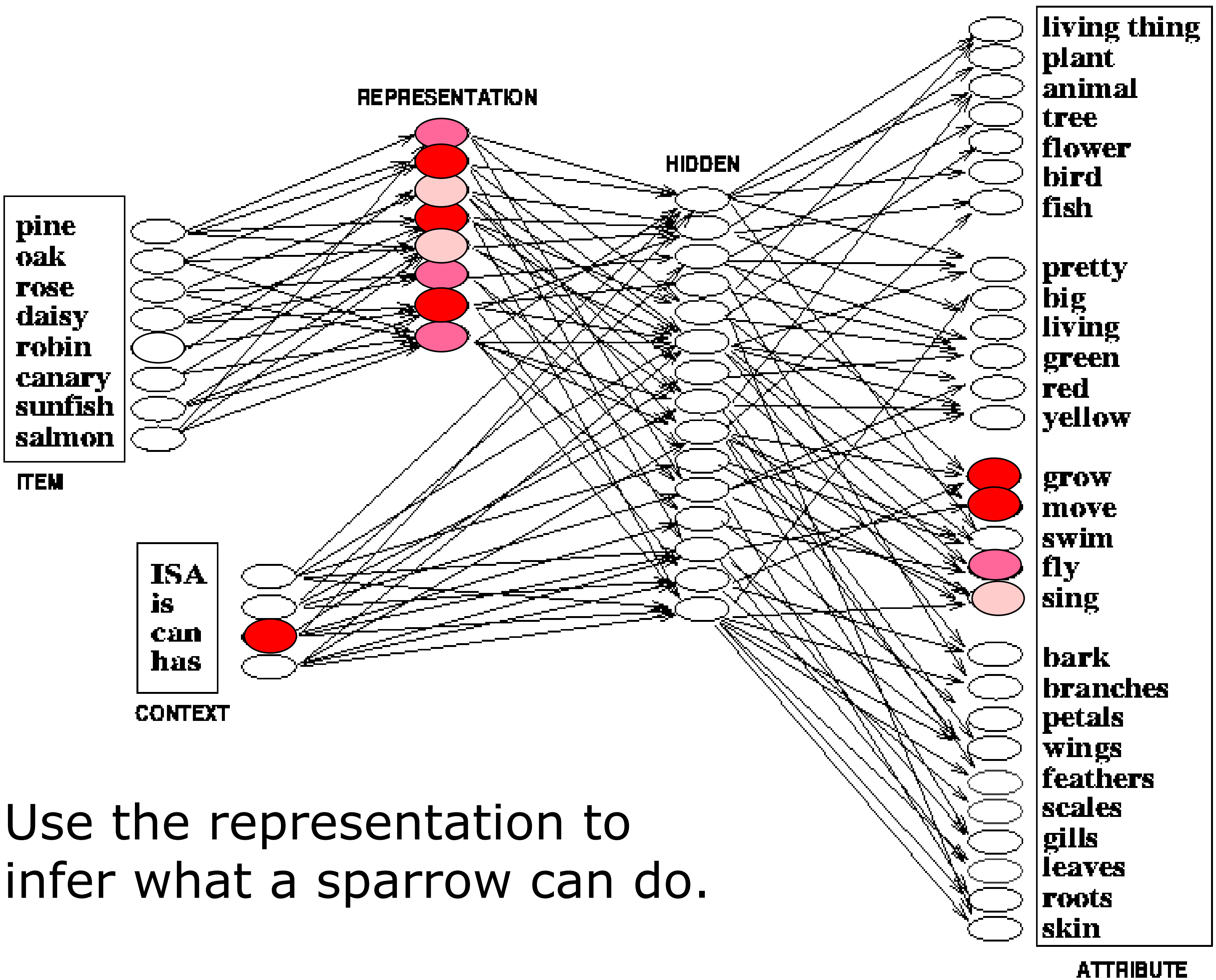
- A semantic representation for a new item can be derived by error propagation from given information, using knowledge already stored in the weights.



Find representation for sparrow from ‘sparrow ISA bird’ by back propagation.



Find representation for sparrow from ‘sparrow ISA bird’ by back propagation.



Use the representation to infer what a sparrow can do.

Inference and Generalization in the PDP Model

- A semantic representation for a new item can be derived by error propagation from given information, using knowledge already stored in the weights.

- Crucially:

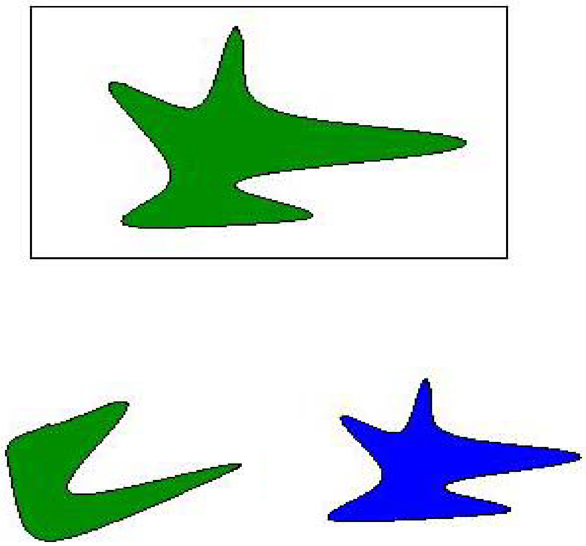
  - The similarity structure, and hence the pattern of generalization, depends on the knowledge already stored in the weights.

Domain Specificity

- What constraints are required for development and elaboration of domain-specific knowledge?
  - Are domain specific constraints required?
  - Or are there general principles that allow for acquisition of conceptual knowledge of all different types?

Differential Importance (Marcario, 1991)

- 3-4 yr old children see a puppet and are told he likes to eat, or play with, a certain object (e.g., top object at right)
  - Children then must choose another one that will “be the same kind of thing to eat” or that will be “the same kind of thing to play with”.
  - In the first case they tend to choose the object with the same color.
  - In the second case they will tend to choose the object with the same shape.



- Can the knowledge that one kind of property is important for one type of thing while another is important for a different type of thing be learned?

- They can in the PDP model, since it is sensitive to domain-specific patterns of coherent covariation.

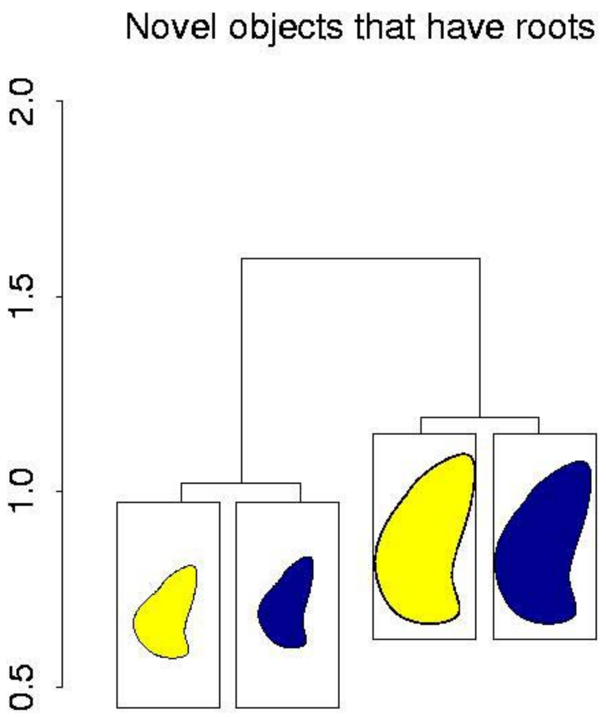
Adjustments to Training Environment

- Among the plants:
  - All trees are large
  - All flowers are small
  - Either can be bright or dull
- Among the animals:
  - All birds are bright
  - All fish are dull
  - Either can be small or large
- In other words:
  - Size covaries with properties that differentiate different types of plants
  - Brightness covaries with properties that differentiate different types of animals

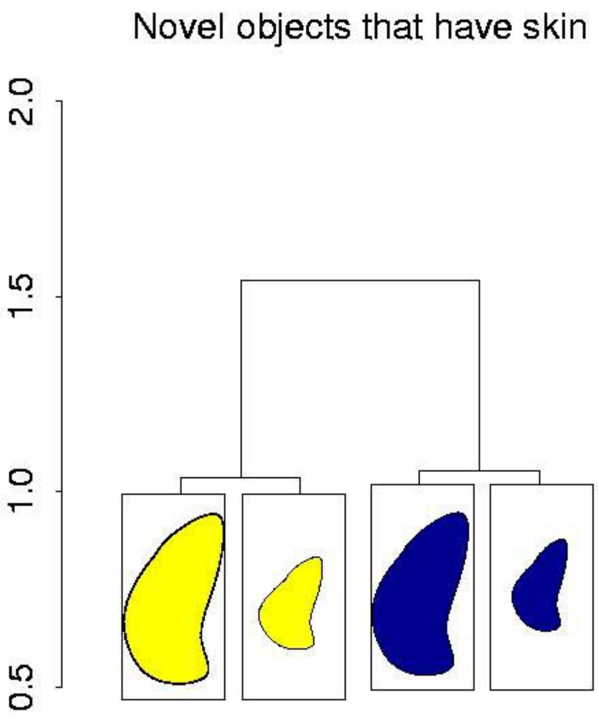
Testing Feature Importance

- After partial learning, model is shown eight test objects:
  - Four “Animals”:
    - All have skin
    - All combinations of bright/dull and large/small
  - Four “Plants”:
    - All have roots
    - All combinations of bright/dull and large/small
- Representations are generated by using back-propagation to representation.
- Representations are then compared to see which animals are treated as most similar, and which plants are treated as most similar.

Similarities of Obtained Representations

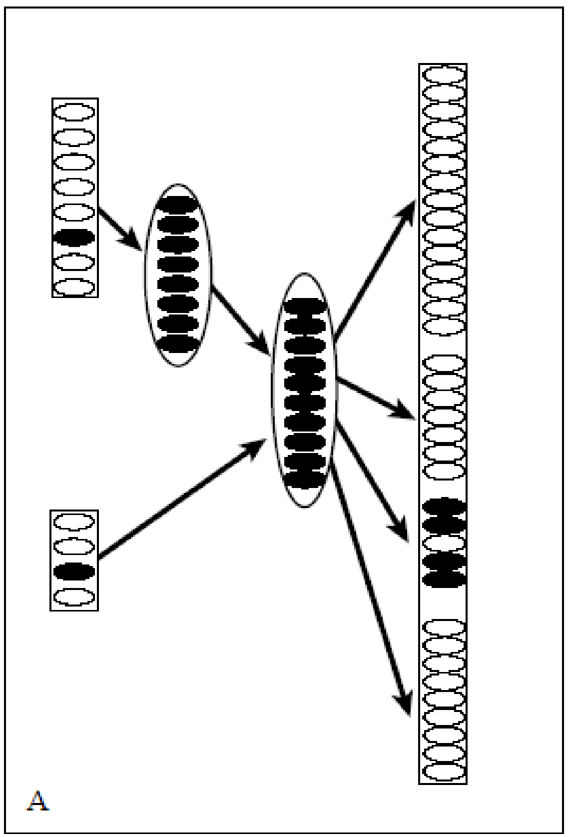


Size is relevant for Plants



Brightness is relevant for Animals

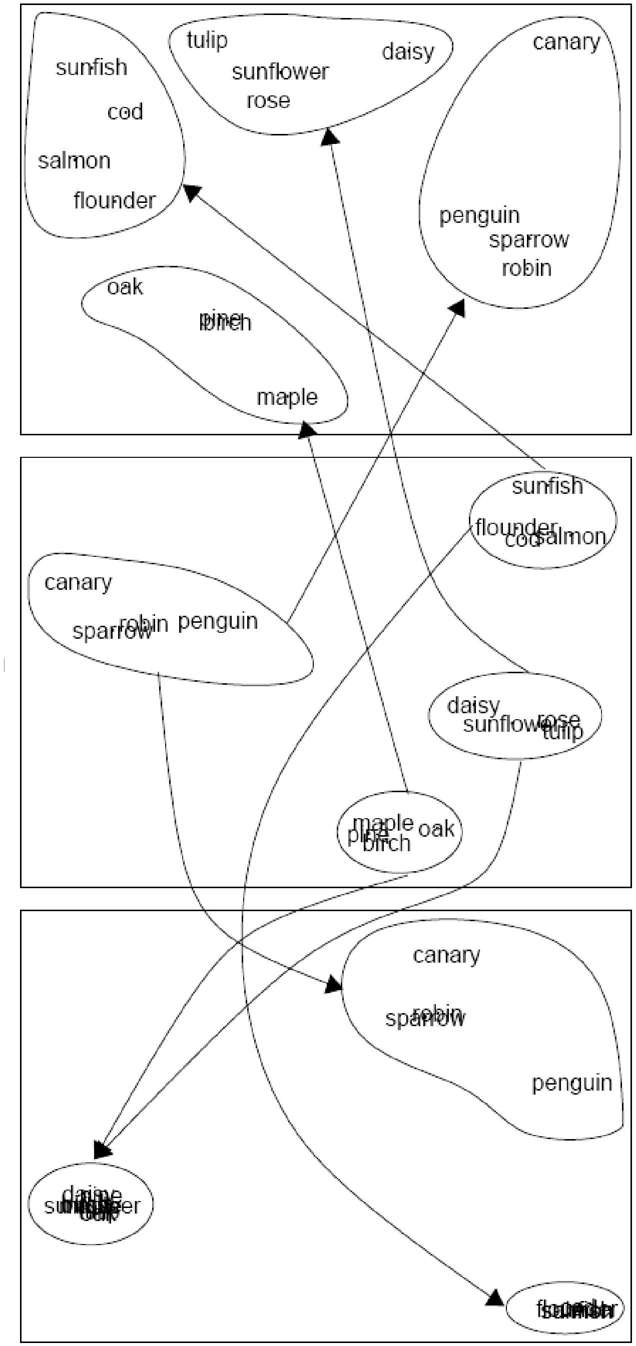
Differential Generalization of Different Types of Properties



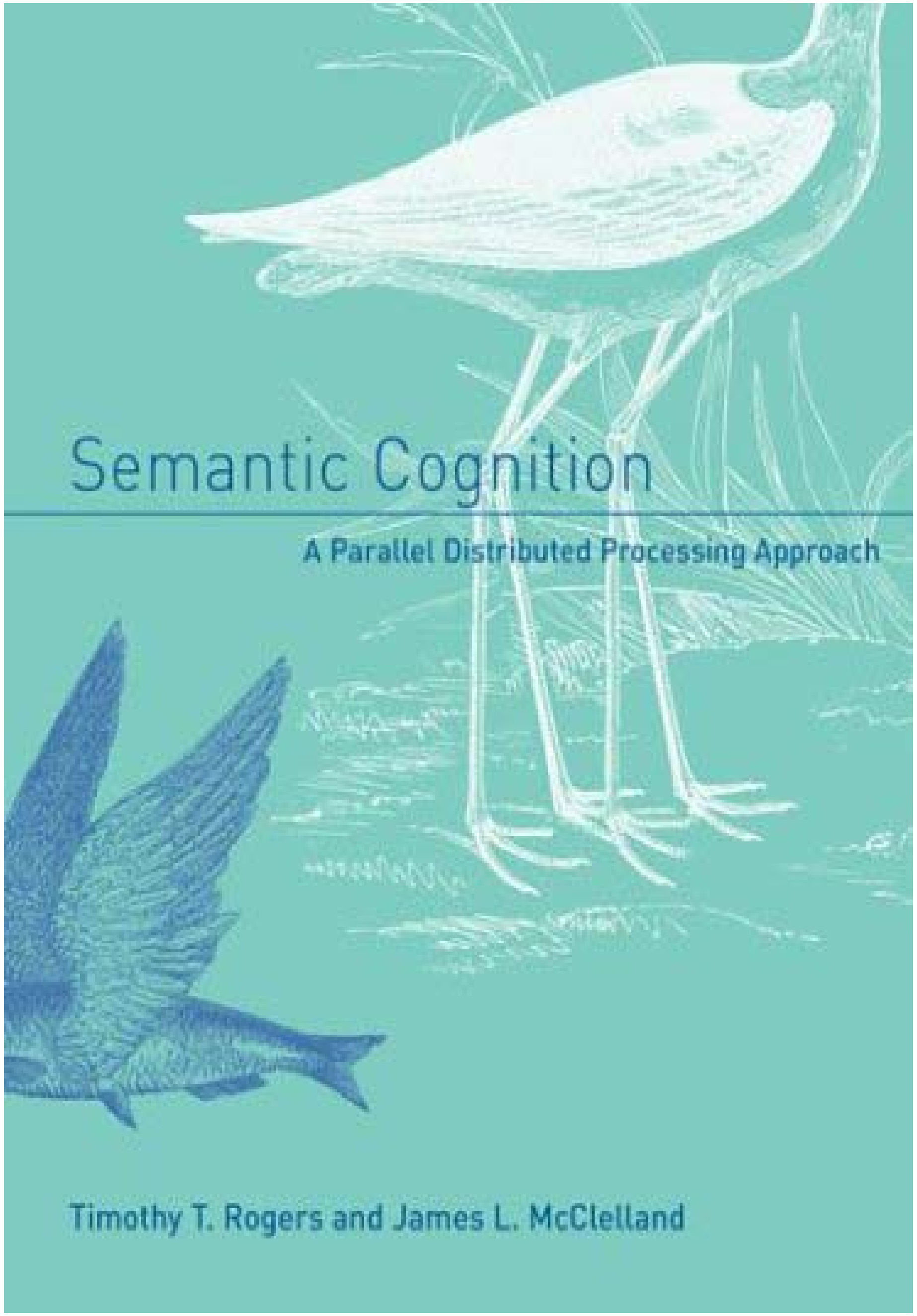
IS context

Shared Representation and HAS context

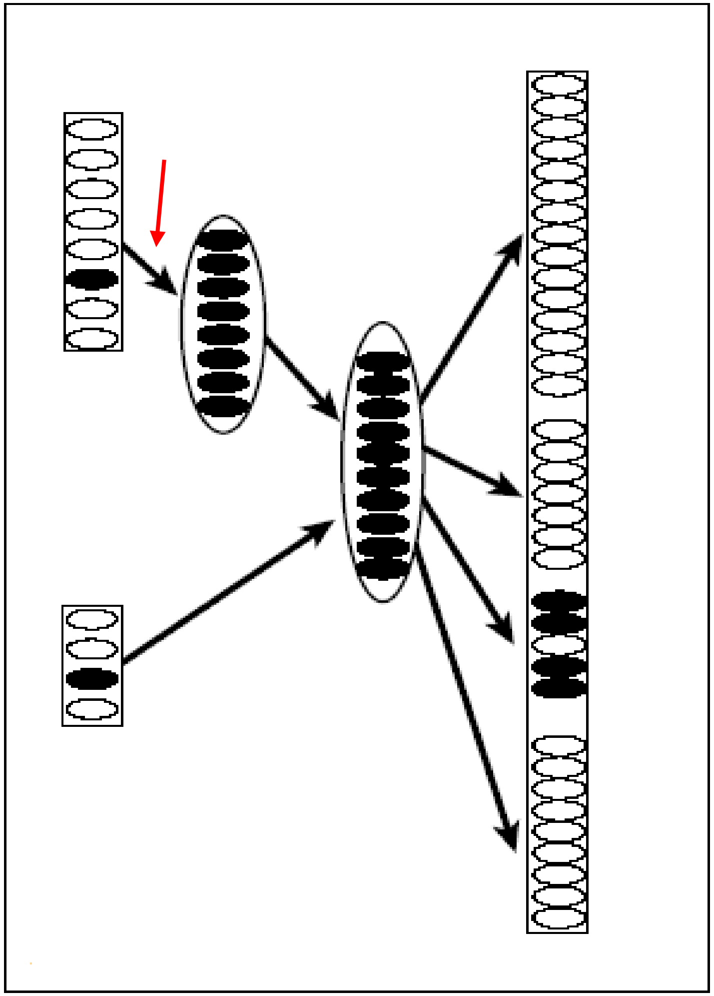
CAN context

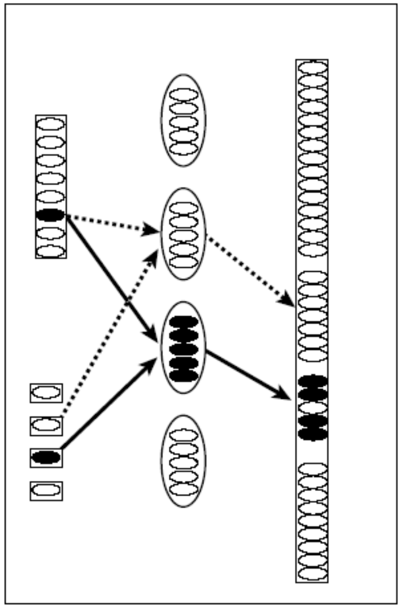


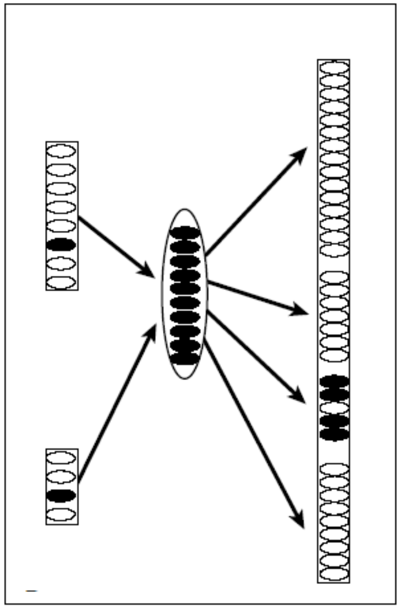
- In Rogers and McClelland (2004) we also address:
  - Conceptual differentiation in prelinguistic infants.
  - Many of the phenomena addressed by classic work on semantic knowledge from the 1970's:
    - Basic level
    - Typicality
    - Frequency
    - Expertise
  - Disintegration of conceptual knowledge in semantic dementia
  - How the model can be extended to capture causal properties of objects and explanations.
  - What properties a network must have to be sensitive to coherent covariation.



Coherence Requires Convergence







A Cognitive Dynamical System in the Brain?

- The human brain appears to be a cognitive dynamical system, judging from the changes we see over the course of development.
- Neural networks can provide a idealized model of this system.
- The model we have explored to date is far simpler than the brain.
- But new algorithms and ever-faster computers allow increasing sophistication.
- Perhaps someday they will allow us to capture the cognitive dynamical system in the brain somewhat better than we do today.