

Overview of:
The Parallel Distributed Processing
Approach to Semantic Cognition
James L McClelland,
Timothy T Rogers

Nathan Feldman, Rick Shih

What is Semantic Cognition?

- Sets of relationships between objects and characteristics

- CAN

- IS

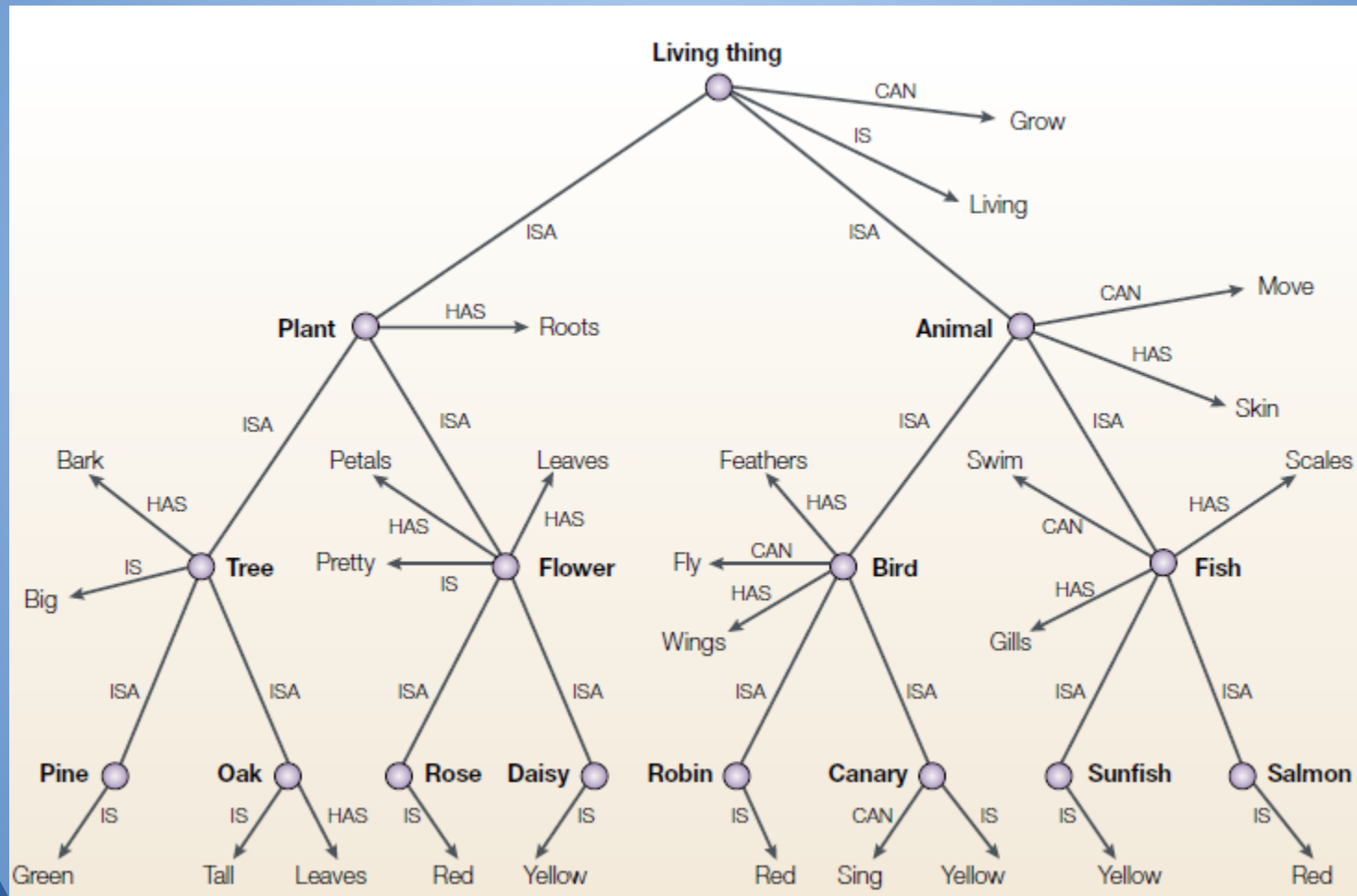
- ISA

- HAS

- Quillian's Model

- Rumelhart's Model

Quillian's Model



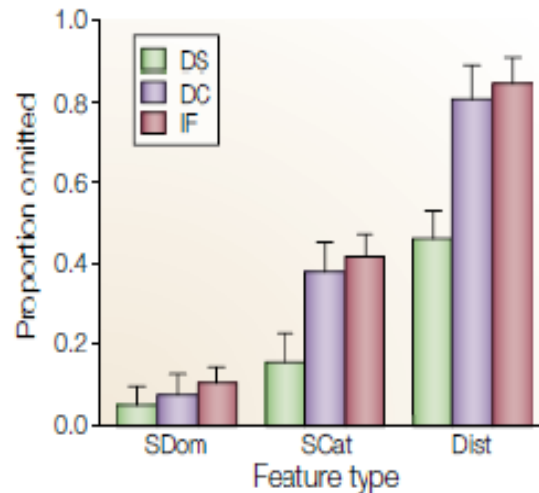
Quillian's Model

- Hierarchical
- Economic storage
- Immediate generalizations
- Represents semantic dementia well

Semantic Dementia

Picture naming responses for JL

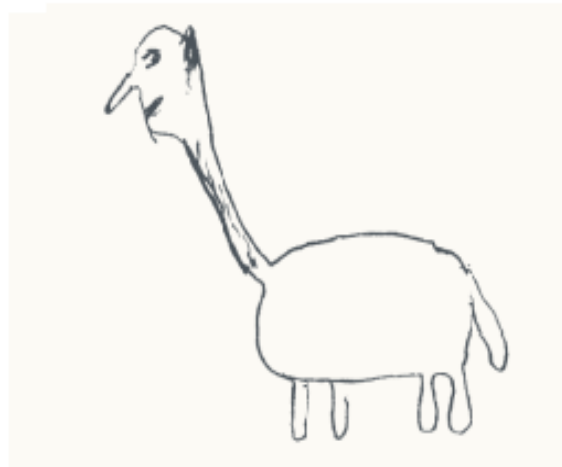
Item	Sept. 91	March 92	March 93
Bird	+	+	Animal
Chicken	+	+	Animal
Duck	+	Bird	Dog
Swan	+	Bird	Animal
Eagle	Duck	Bird	Horse
Ostrich	Swan	Bird	Animal
Peacock	Duck	Bird	Vehicle
Penguin	Duck	Bird	Part of animal
Rooster	Chicken	Chicken	Dog



IF's delayed copy of a camel



DC's delayed copy of a swan



JL,DS,DC,IF:
Patient Initials

SDom: General
Domain

SCat:
Superordinate
Category

Dist: Distinctive
Features

Quillian's Model

Weaknesses:

- Predicts that verifying specific properties is faster than general ones
- Doesn't specify when new superordinate categories should be introduced
- Sometimes general properties don't apply to the specific item (e.g., most plants have leaves, but pine trees have needles)

Graded Category Membership

- Humans perform semantic tasks more based on typicality and similarity
- Ex: subjects verify
 - 'robin is a bird', faster than 'chicken is a bird'.
 - 'chicken is an animal' faster than 'chicken is a bird'.
- This is because category verification occurs by comparing representations

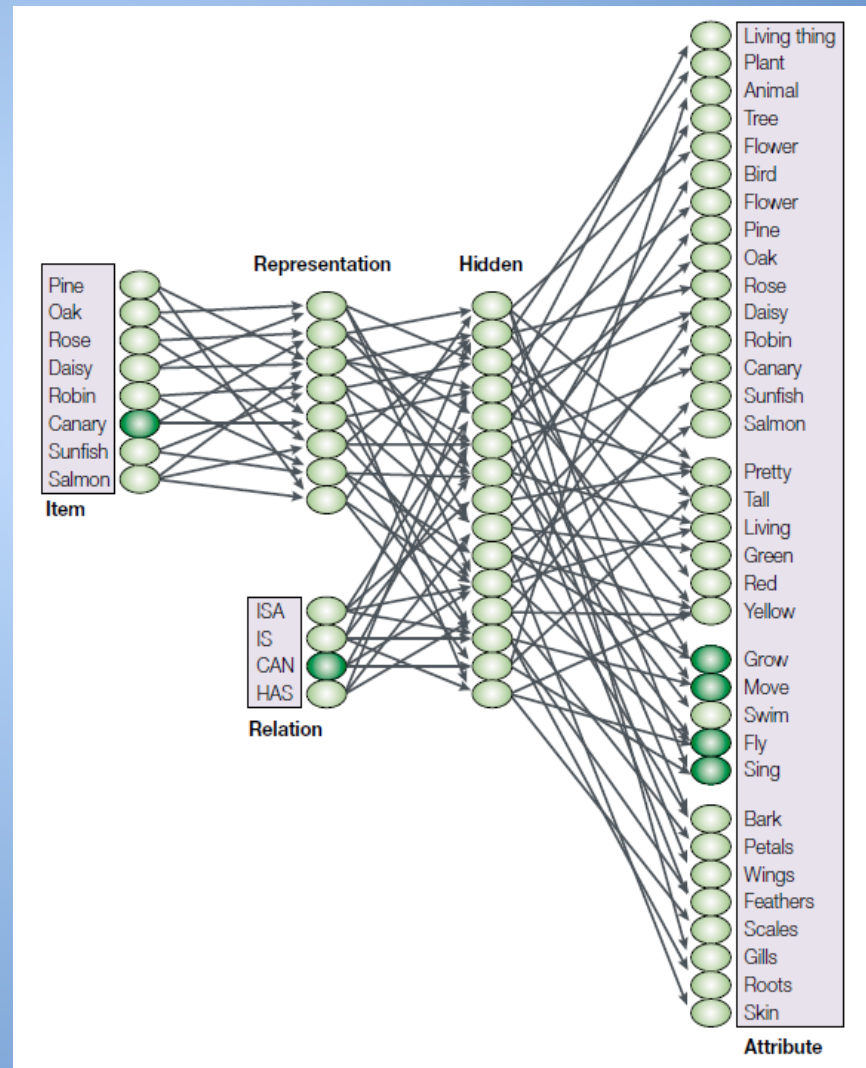
Graded Category Membership

Alternative models try to capture the notion that category membership is graded on feature values or proximity in multidimensional space. Examples:

- A Bayesian model (by J.R. Anderson) based on probabilistic relationship between categories and properties
- A model where category membership on summed similarity to known exemplars (D. L. Medin & M.M. Shaffer)

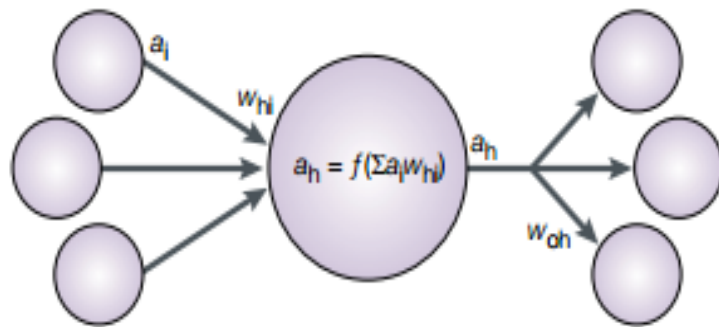
Rumelhart's Model: Parallel Distributed Processing (PDP)

The PDP model is a type of neural network that distributes representations of things and processing across many units

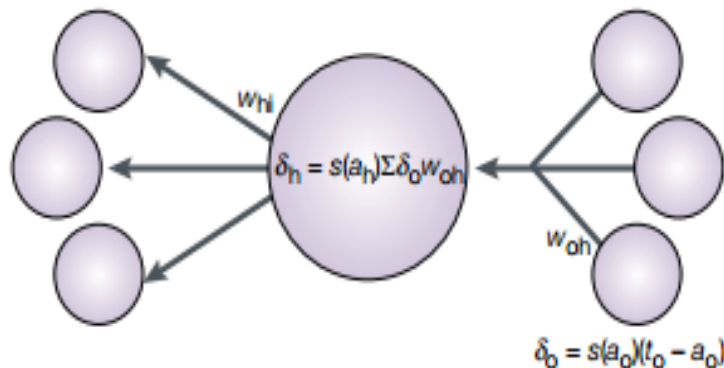


Rumelhart's Model: Parallel Distributed Processing (PDP)

Forward propagation of activation



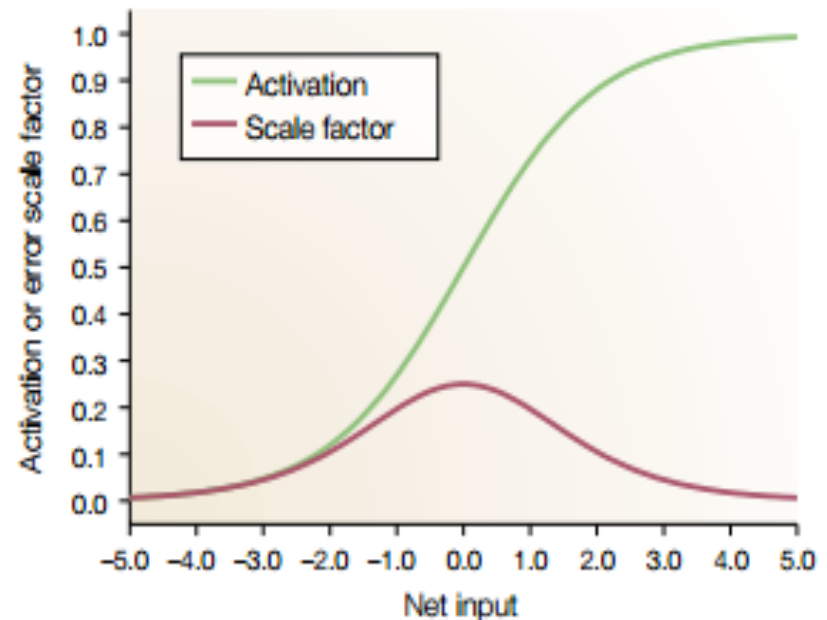
Back propagation of error



Connection weight changes

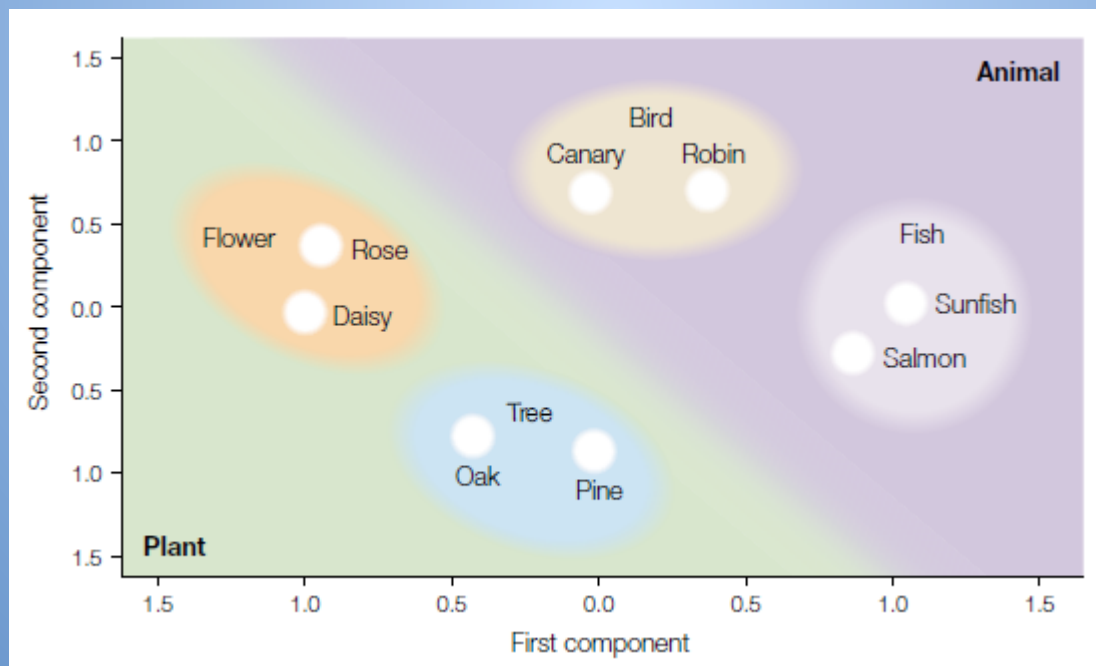
At the output layer: $\Delta w_{oh} = \epsilon \delta_o a_h$

At the prior layer: $\Delta w_{hi} = \epsilon \delta_h a_i$

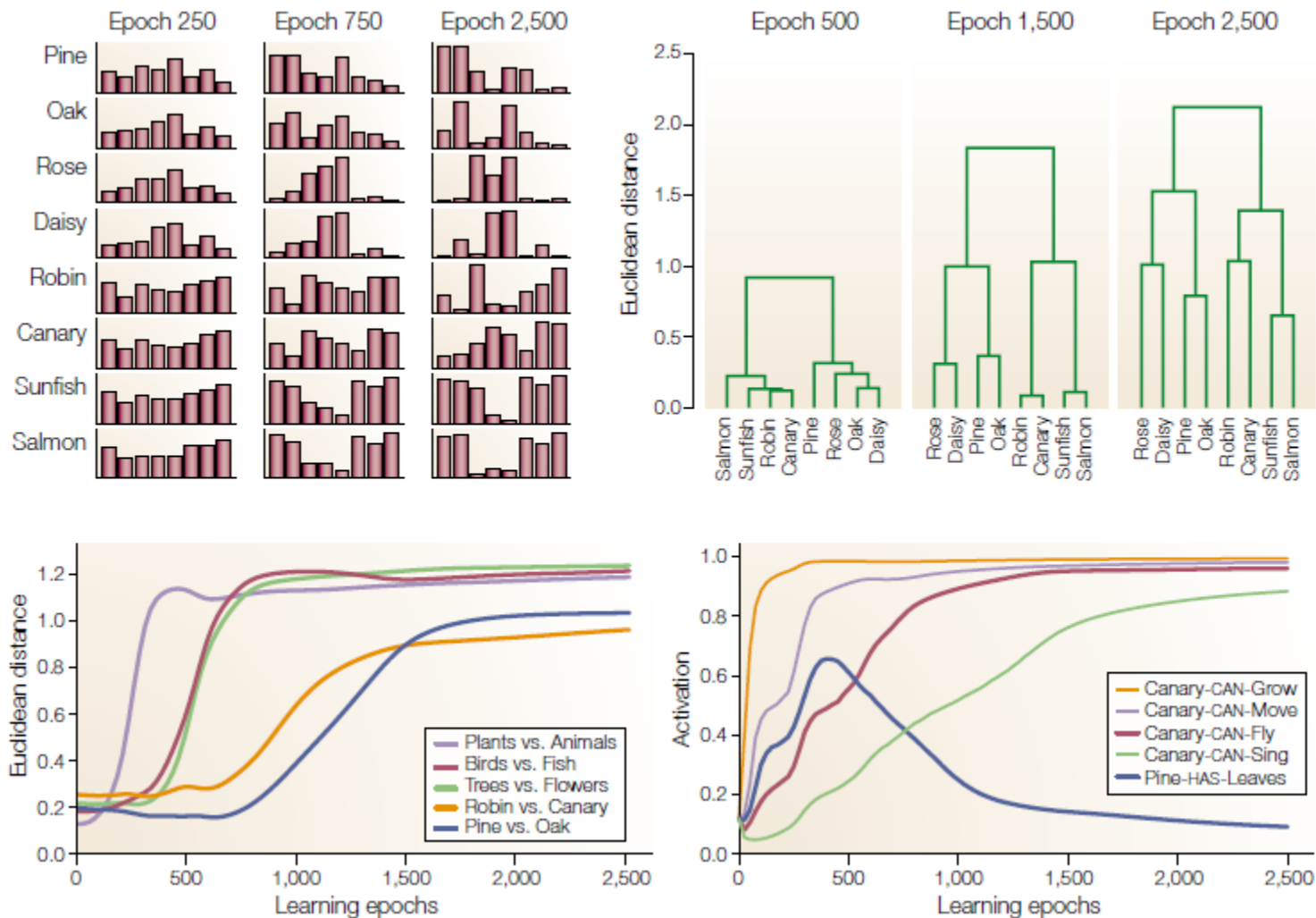


PDP

- Captures Quillian's hierarchical representation, while performing similar to human child development and performance

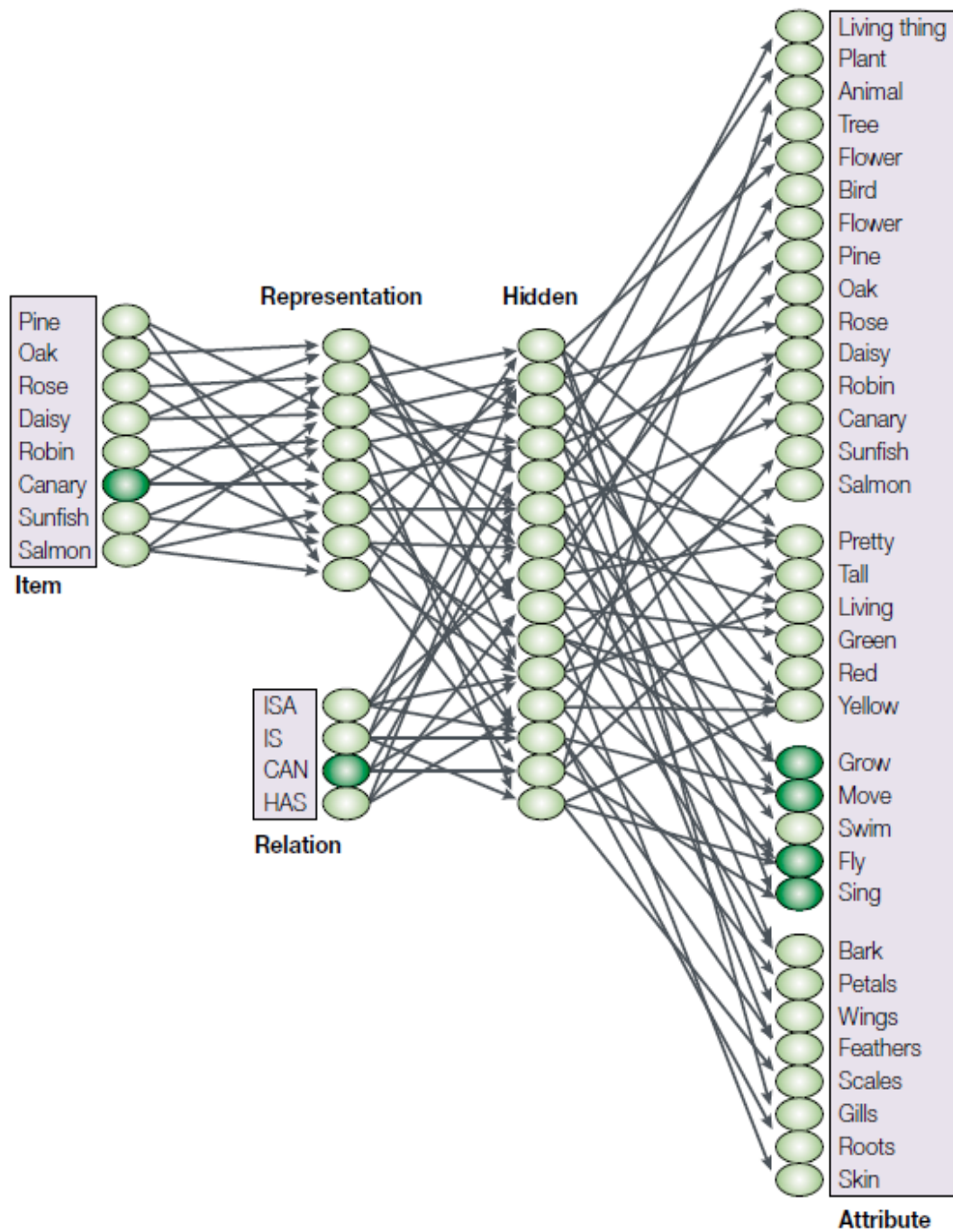


PDP: System Learning



PDP: Human Learning

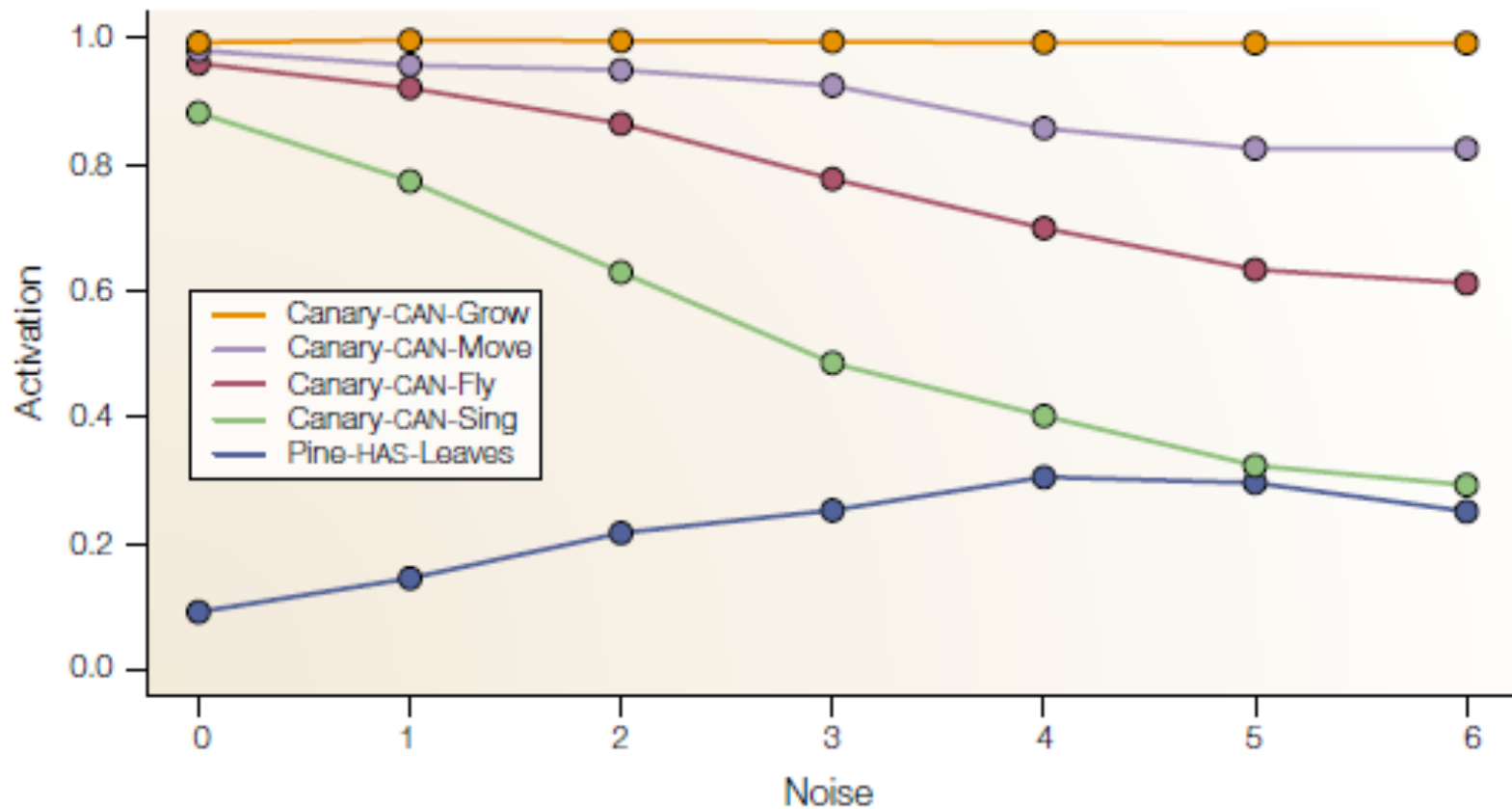
- Humans are able to learn new to categorize new information quickly
- The PDP is able to learn accurate semantics for a new entity "Sparrow" by only allowing weights into the representation layer to change. Entities like "Canary" and "Robin" shared characteristics aide in this process.
- The PDP model does not allow quick learning for new entities that are dissimilar to any previous entities, as it results in CATASTROPHIC INTERFERENCE
- McClelland, McNaughton, & O'Reilly propose a complementary learning system theory to address this issue.



PDP: Semantic Dementia

- Perturb the inputs into the representation layer using random noise.
- Distinctive characteristics are lost faster than shared characteristics
- More general characteristics replace distinctive ones, even when the general characteristics shouldn't apply

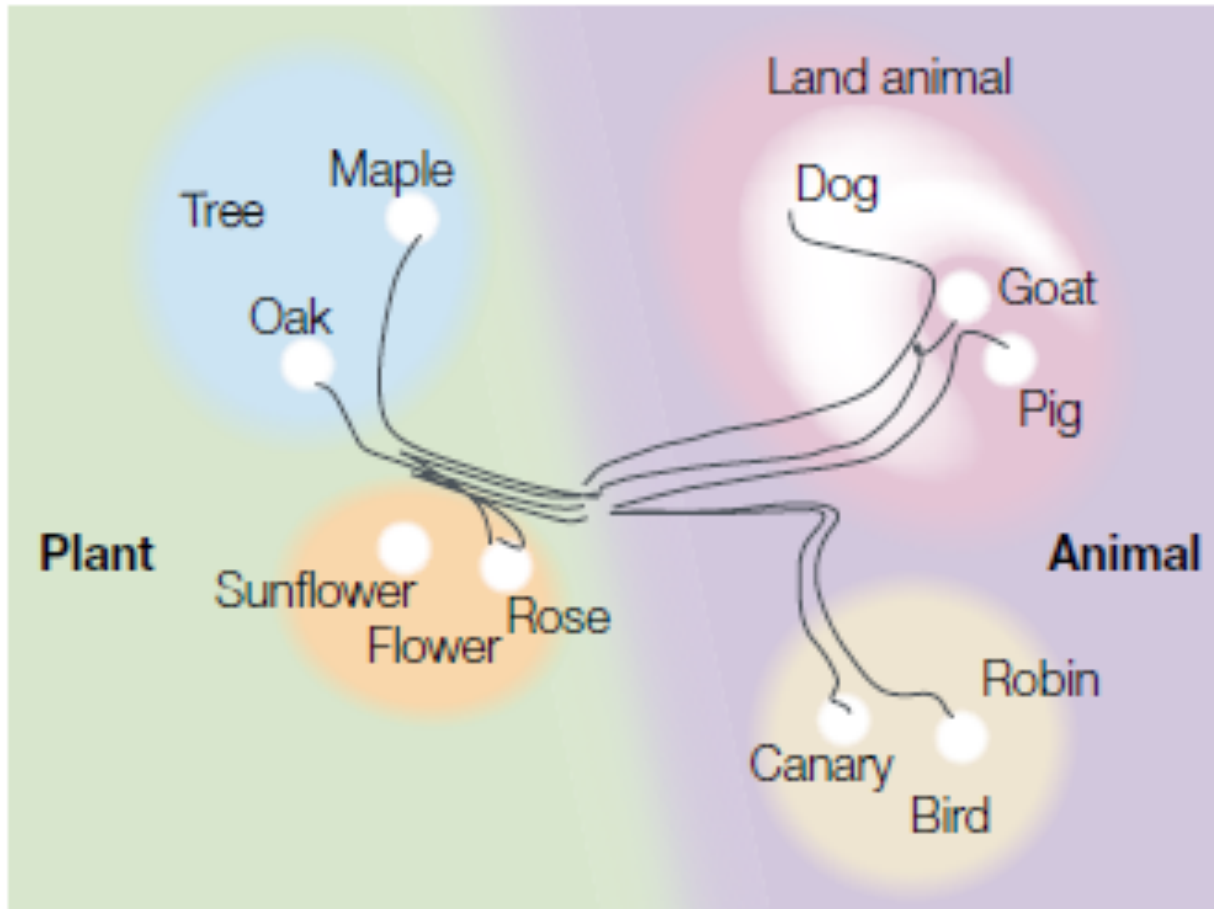
PDP: Semantic Dementia



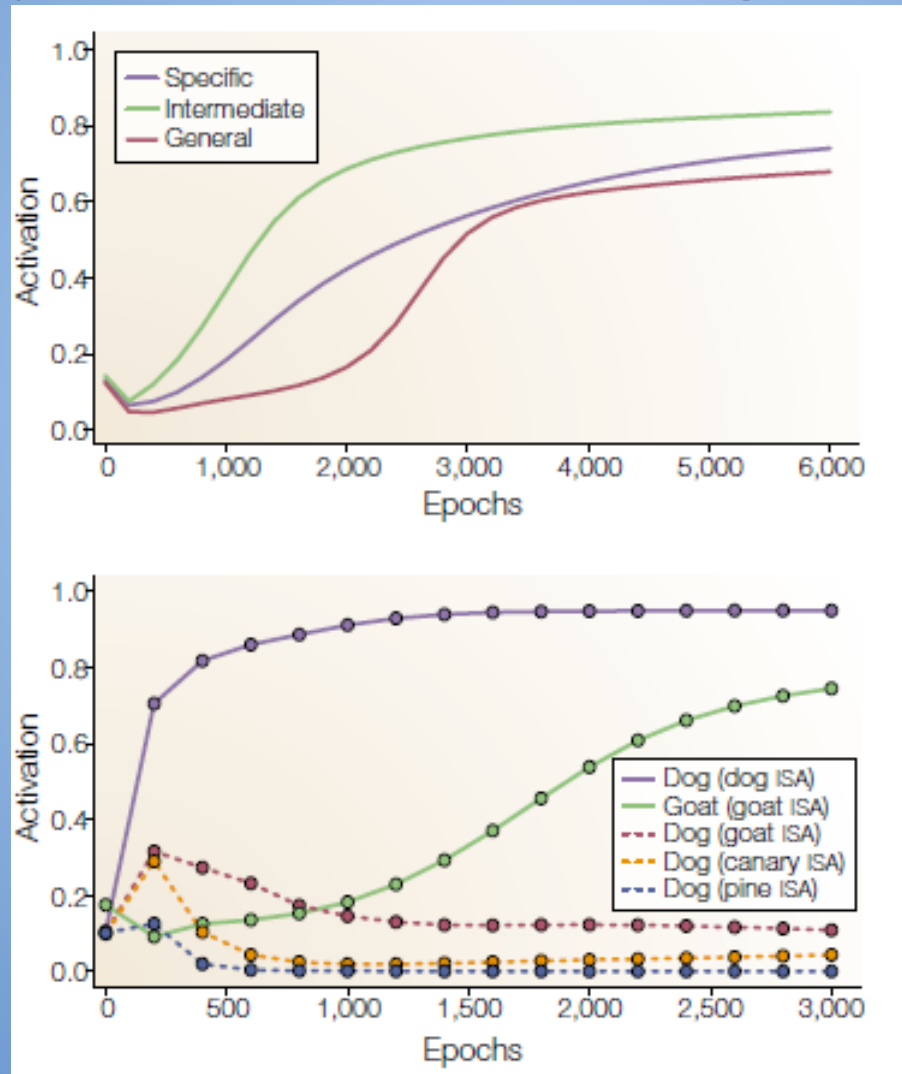
Basic Level

- We tend to identify objects primarily at the basic, intermediate level
- We learn words like 'tree', 'bird', and 'dog' earlier than 'plant' or 'animal', or 'canary', 'pine', or 'poodle', partly because we hear these more
- To address this, make basic level objects more frequent in training data

PDP: System Learning



PDP: System Learning



PDP: Coherent Covariation and Causality

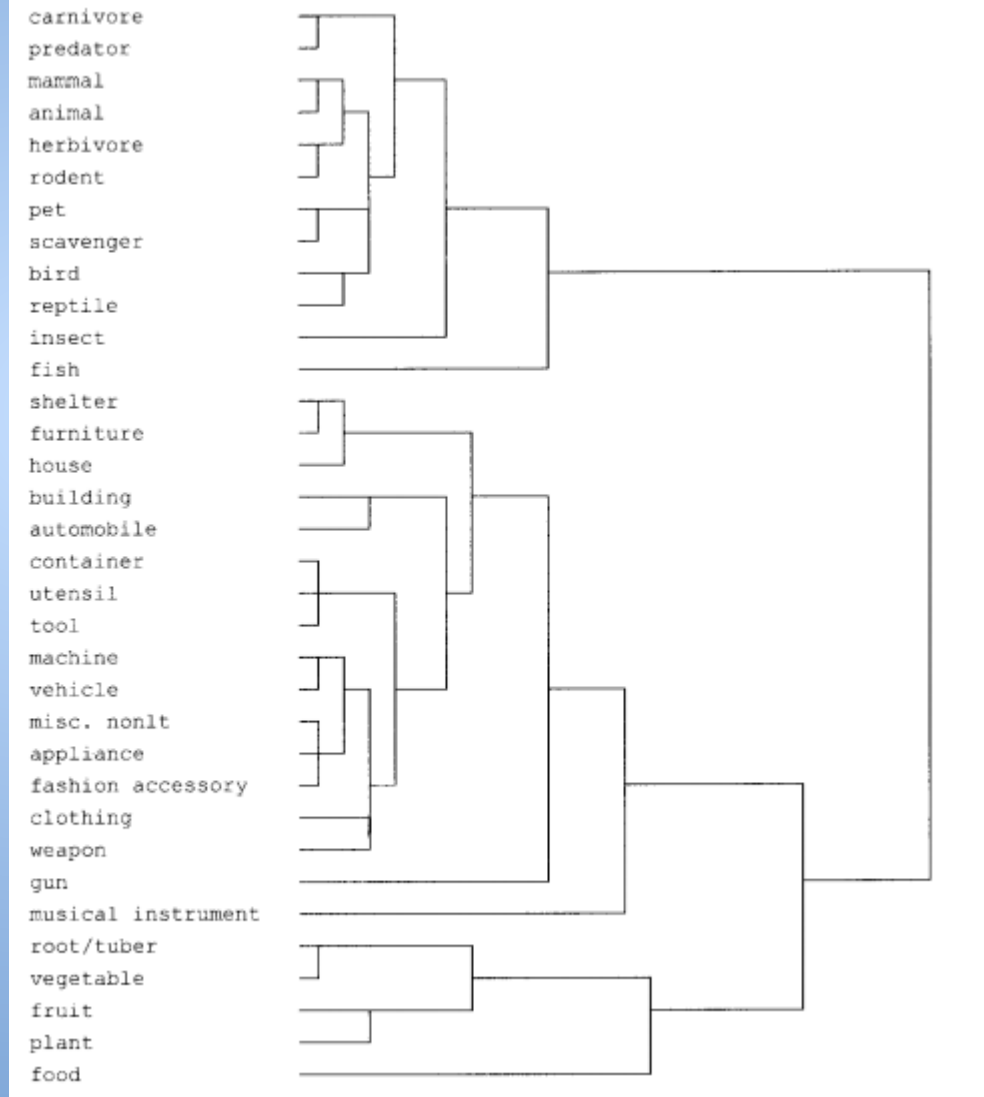
- Different properties help distinguish between objects in different domains
 - (e.g. size plays bigger role when distinguishing between flowers and trees than when distinguishing between birds and fish)
- The PDP network shows similar learning ability by the covariance within the representation layer.
 - Size (IS large vs IS small) affects the representation more in plants than it does in animals
- This coherent covariation helps capture knowledge about causality (e.g. hollow bones and having wings cause an animal to be able to fly)

PDP: Coherent Covariation and Causality

- Still does not utilize the same mechanism as causality.
- Coherent Covariance can overgeneralize
 - (e.g. canaries and daisies are both yellow, but that doesn't imply daisies have wings)
- Further research into the model in order to incorporate logical relationships between characteristics
 - building representations from sequences of events

Analysis of Data: Cree and McRae

- Cree and McRae (Extended Model from Wu and Barsalou's 2002 Knowledge Type Taxonomy)
- 34 Feature classes: entity, taxonomic, made-of, systemic
- Covariance Clustering of domains



Gathering Data

- Participants were given 30 entities with the task of identifying all words that come to mind associated to it
- 541 Concepts

Clustering

- Trends using Hierarchical Clustering Analysis
 - Creature categories pattern together
 - Fruits and vegetables pattern together
 - Fruits and vegetables can pattern with either creatures or nonliving things
 - Nonliving foods (like cake) can be patterned with living things
 - Musical instruments can be patterned with living things

Sensory/Functional Dichotomy

- The brain separates processing of sensory and functional characteristics
- Sensory inputs can be broken down into the visual cortex and auditory cortex
- Rumelhart's model is a single system

Semantic Dementia

- Deterioration of the living things domain
 - General trend in case studies that the domain of living things is affected over nonliving things
 - Statistical probability supports this because of a multitude of shared features
- Does not address the distinct feature deterioration patterns of semantic dementia

Domain	Category	distinguishing features
Nonliving	Appliance	49.1 (17.5)
Nonliving	Miscellaneous nonliving thing	47.8 (17.7)
	Foods	45.4 (12.3)
Nonliving	Building	45.3 (12.7)
Nonliving	Machine	43.3 (14.2)
Nonliving	Tool	40.4 (18.2)
Nonliving	Fashion accessory	36.8 (17.1)
Nonliving	Shelter	35.8 (19.7)
Nonliving	Clothing	35.6 (17.4)
Nonliving	House	35.2 (19.8)
Nonliving	Container	34.5 (18.4)
Nonliving	Vehicle	33.7 (16.3)
Nonliving	Automobile	33.7 (17.1)
Nonliving	Furniture	33.1 (15.4)
Nonliving	Utensil	30.4 (17.4)
Nonliving	Weapon	29.3 (15.4)
Fruit/vegetable	Root/tuber	28.6 (22.7)
Fruit/vegetable	Vegetable	28.3 (18.0)
Fruit/vegetable	Plant	28.0 (17.0)
Creature	Insect	26.4 (12.9)
	Musical instruments	25.3 (15.8)
Creature	Mammal	25.1 (16.0)
Creature	Rodent	21.6 (15.2)
Creature	Animal	21.0 (14.4)
Fruit/vegetable	Fruit	20.7 (12.5)
Creature	Pet	20.7 (15.2)
Creature	Carnivore	20.1 (11.9)
Nonliving	Gun	19.7 (12.5)
Creature	Scavenger	19.7 (12.8)
Creature	Predator	19.4 (10.0)
Creature	Reptile	17.8 (7.8)
Creature	Fish	16.0 (15.7)
Creature	Bird	16.0 (11.9)
Creature	Herbivore	15.7 (12.7)

Note. Standard deviations are in parentheses.

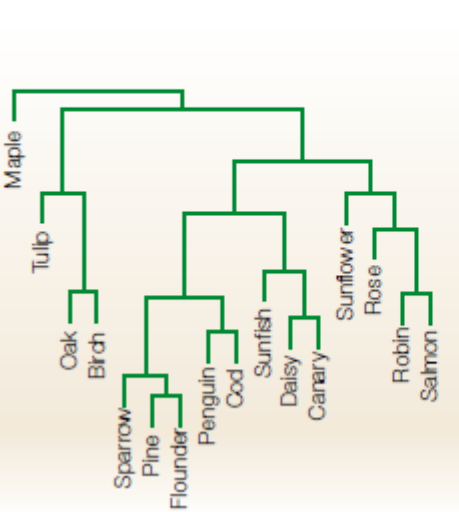
References

McClelland, J.L. & Rogers, T.T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4, 310-322

Cree, G.S. & McRae, K. (2003). Analyzing the Factors Underlying the Structure and Computation of the Meaning of Chipmunk, Cherry, Chisel, Cheese, and Cello (and Many Other Such Concrete Nouns). *Journal of Experimental Psychology: General*, Vol. 132, No. 2, 163-201

Questions???

Epoch 500



Epoch 2,500

