Fundamentals of Computational Neuroscience

Chapter 9: Modular networks, motor control, and reinforcement learning

Mixture of experts

A. Absolute function

$f(x) = abs(x)$

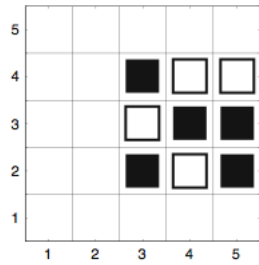B. Mixture of expert for absolute function
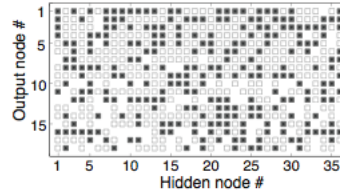
X — $\Sigma\Pi$ — $abs(x)$
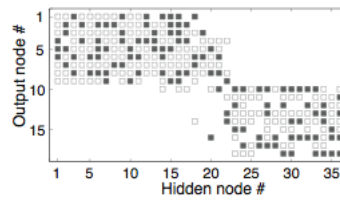
The What-and-Where task

A. Model retina with sample image

B. Without bias towards short connections
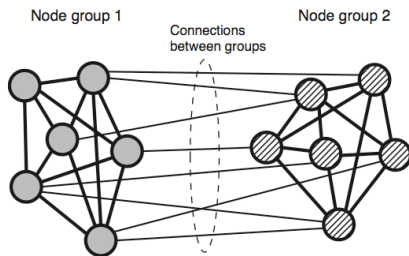
C. With bias towards short connections

Jacobs and Jordan, 1992

Coupled attractor networks
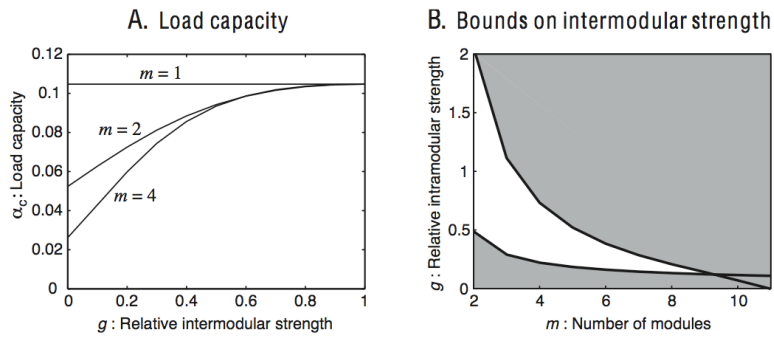
A. Coupled attractor networks

Node group 1   Node group 2

Connections between groups

B. The left–right universe with letters

```
0000011100000    1111111110000
0000011100000    1111111111100
0000111110000    1110000001110
0000110110000    1110000011100
0001100011000    1110001110000
0011100011100    1111111000000
0011111111100    1111111000000
0111111111110    1110001110000
0111000001110    1110000011100
0111000001110    1110000001110
0111000001110    1111111111100
0111000001110    1111111110000
```
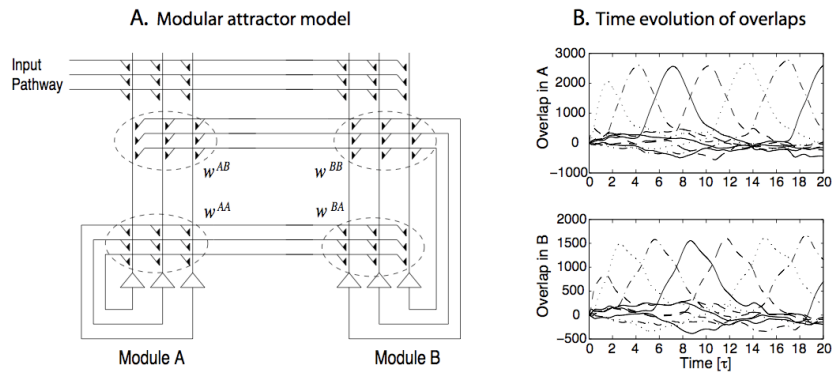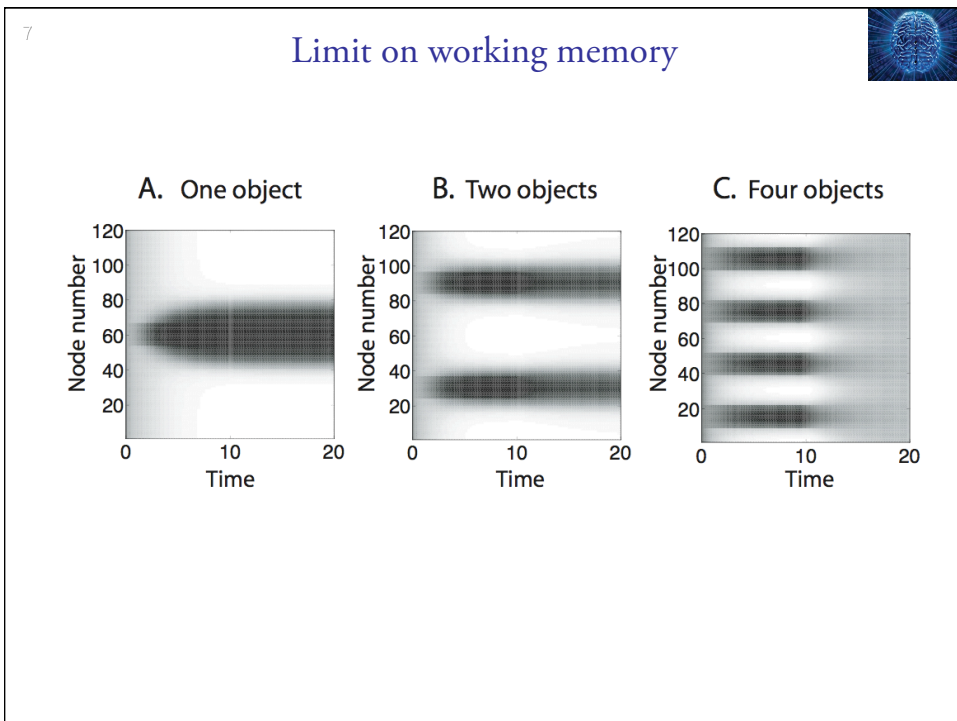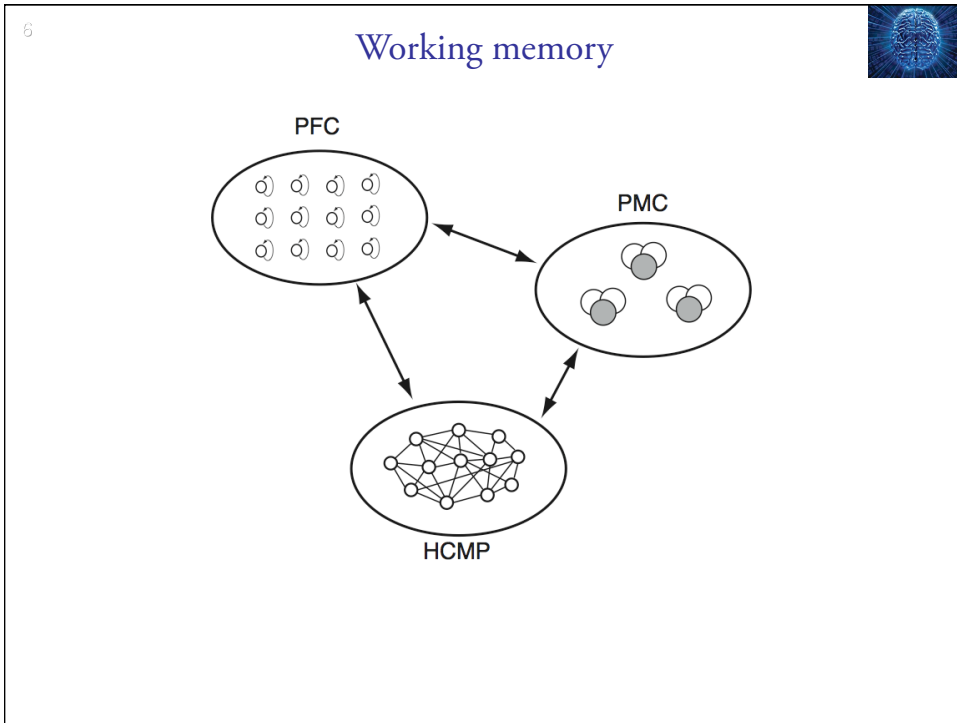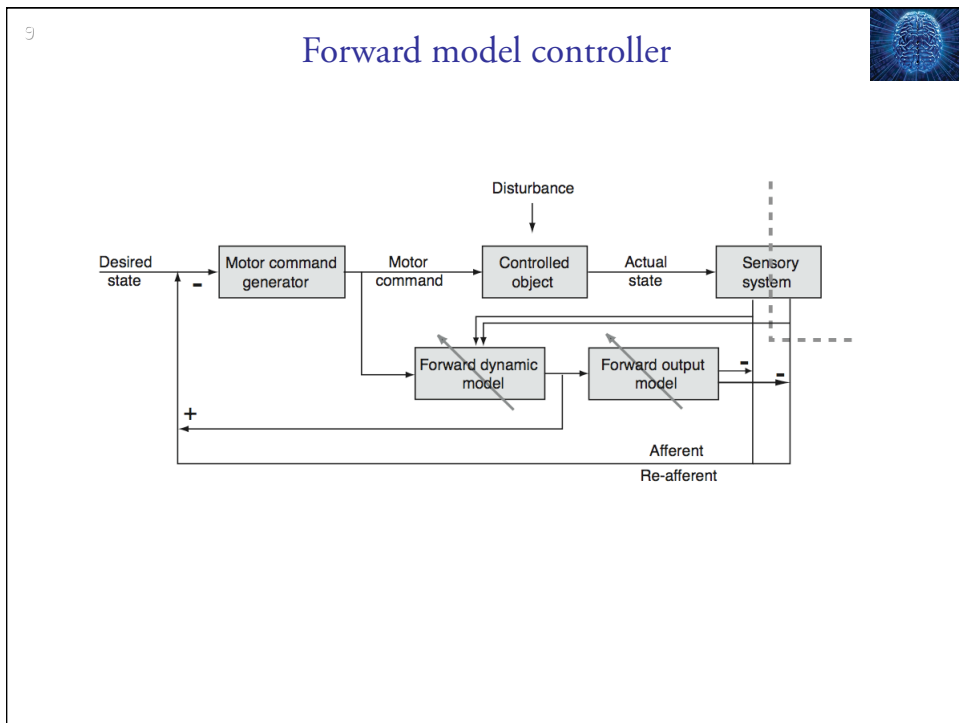
Limit on modularity
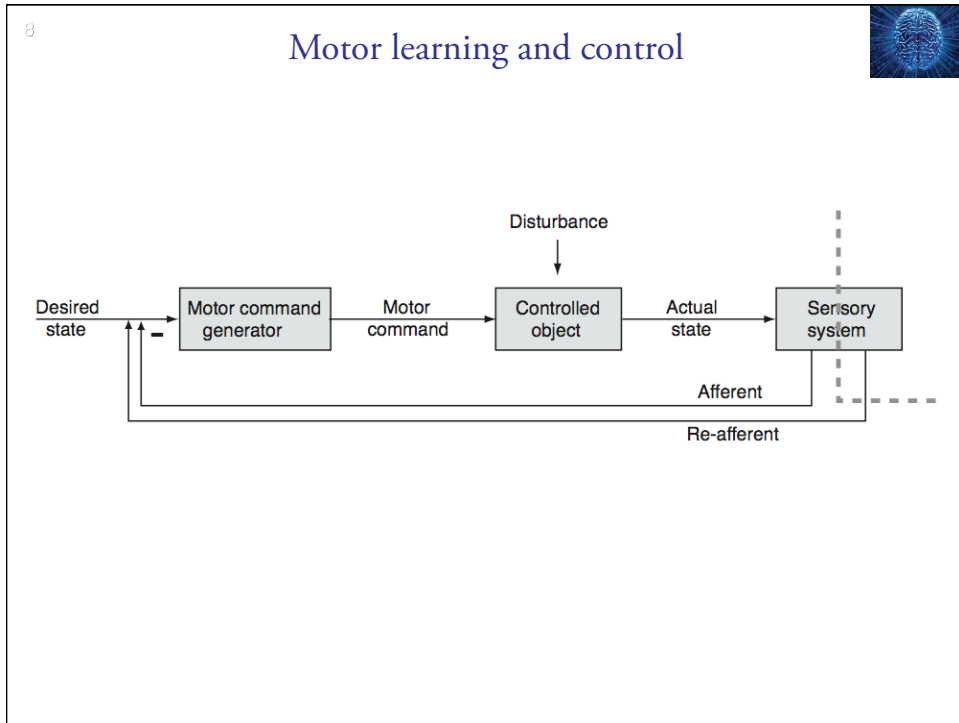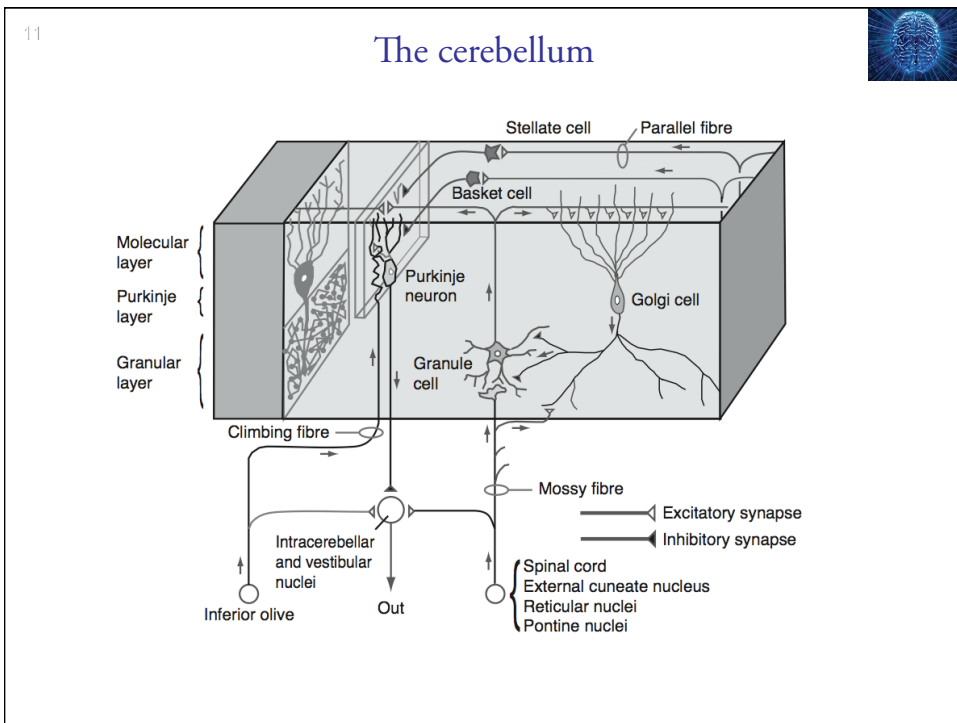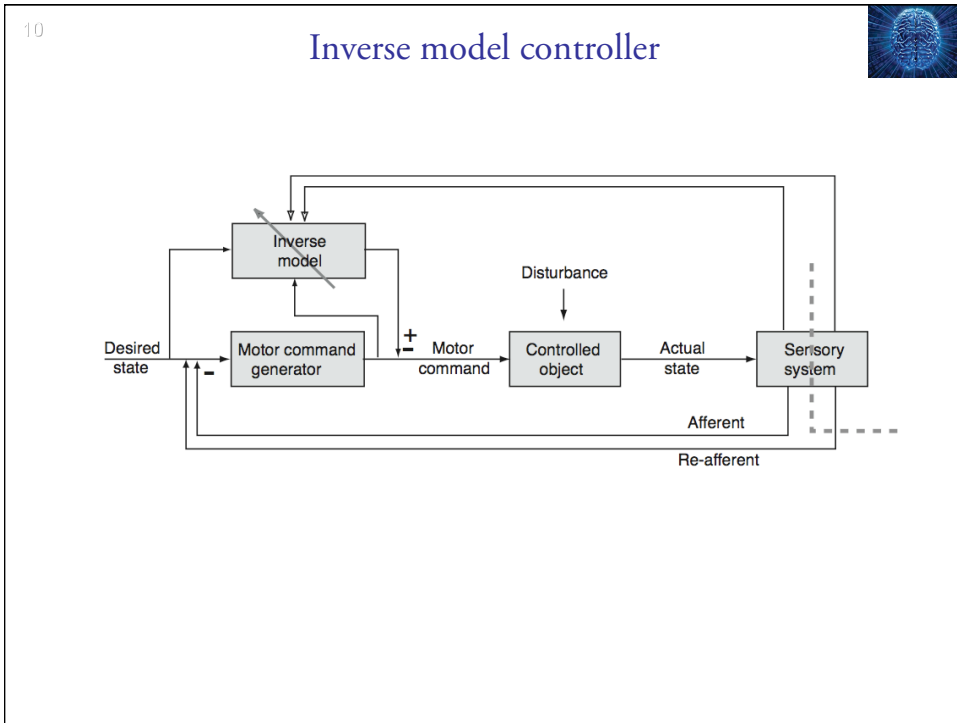
A. Load capacity

B. Bounds on intermodular strength

Sequence learning

A. Modular attractor model

B. Time evolution of overlaps

Working memory

PFC

PMC

HCMP

Limit on working memory

A. One object   B. Two objects   C. Four objects

Motor learning and control

Forward model controller

Inverse model controller

The cerebellum

Reinforcement learning

The basal ganglia

A. Outline of basic BG anatomy

C. Recordings of SNc neurons and simulations

14

Temporal difference learning

A. Linear predictor node

B. Temporal delta rule

C. Temporal difference rule

15

Actor-critique and Q-learning

B. Actor-critic model of BG

D. Q-learning model of BG

Actor-critique controller

Further readings

Robert A. Jacobs, Michael I. Jordan, and Andrew G. Barto (1991), **Task decomposition through competition in a modular connectionist architecture: the what and where tasks**, in **Cognitive Science** 15: 219–250.

Geoffrey Hinton (1999), **Products of experts**, in **Proceedings of the Ninth International Conference on Artificial Neural Networks**, ICANN '99, 1:1–6.

Yaneer Bar-Yam (1997), **Dynamics of complex systems**, Addison-Wesley.

Edmund T. Rolls and Simon M. Stringer (1999), **A model of the interaction between mood and memory**, in **Networks: Comptutation in neural systems** 12: 89–109.

N. J. Nilsson (1965), **Learning machines: foundations of trainable pattern-classifying systems**, McGraw-Hill.

O. G. Selfridge (1958), **Pandemonium: a paradigm of learning, in the mechanization of thought processes**, in **Proceedings of a Symposium Held at the National Physical Laboratory**, November 1958, 511–27, London HMSO.

Marvin Minsky (1986), **The society of mind**, Simon & Schuster.

Akira Miyake and Priti Shah (eds.) (1999), **Models of working memory**, Cambridge University Press.

Daniel M. Wolpert, R. Chris Miall, and Mitsuo Kawato (1998), **Internal models in the cerebellum**, in **Trends Cognitive Science** 2: 338–47.

Edmund T. Rolls and Alessandro Treves (1998), **Neural networks and brain function**, Oxford University Press.

James C. Houk, Joel L. Davis, and David G. Beiser (eds.) (1995), **Models of information processing in the basal ganglia**, MIT Press.

Richard S. Sutton and Andrew G. Barto (1998), **Reinforcement learning: an introduction**, MIT Press.

Peter Dayan and Laurence F. Abbott (2001), **Theoretical Neuroscience**, MIT Press.