Introduction to Parallel Distributed Processing

Basic Principles

- Basic unit functions
- Constraint satisfaction
- Schema theory
- Correlation-based learning (Hebb)
- Error-correcting learning (Delta)
- Localist vs. distributed representations
- Distributed memory model
- Back-propagation
- Generalization and overfitting
- Temporal learning and recurrent networks
- Bolzmann machines/Contrastive Hebbian learning
- Unsupervised learning/generative models
- Deep learning
- Reinforcement learning and forward models
- Tricks of the trade / designing representations

Applications

- Cognitive development
- High-level vision and attention
- Semantics
- Memory and the hippocampus
- Statistical learning
- Language: Morphology (past-tense)
- Language: Sentence processing
- Cognitive control and executive functions

Understanding complex information-processing systems

Marr (1982)

Computational theory
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?

Representation and algorithm
How can this computational theory be implemented? What is the representation for the input and output, and what is the algorithm for the transformation?

Hardware implementation
How can the representation and algorithm be realized physically?

Some objections to PDP (Rumelhart & McClelland, PDPI:4, 1986)

- PDP models are too weak
  - Limitations of one-layer learning
  - Stimulus equivalence
  - Recursion
- PDP models are not cognitive
  - Eliminates rules and structured representations
- PDP models are the wrong level of analysis (cf. Bayesian approaches)
  - At implementational rather than algorithmic level
  - Quantum vs. Newtonian physics analogy
- PDP modeling is merely reductionism to neuroscience
  - Emergent properties
- Neuroscience cannot yet constrain cognitive theories
- PDP models lack neural realism
- PDP and nativism vs. empiricism
- Why are people smarter than rats?
- Conscious knowledge and explicit reasoning

Weaknesses of PDP (Norman, PDPII:26, 1986)

- The type-token problem
  - Simultaneous multiple instances of the same concept
- Variables
  - Symbol manipulation
- The need for extra, evaluative structure
  - A "teacher" in supervised learning
- More than one system at a time is required
  - Cognition is not just a single settling
  - Different part of the network have to do different things
- Learning and consciousness
  - Procedures for learning (controlled $\Rightarrow$ automatic)

"Future" directions (Rumelhart & McClelland, PDPII, 1986)

- Higher-level processes
  - Sequential symbol processing
  - Language processing
- Learning and architecture
- Neuroscience
  - Neuropsychology
  - Physiology and anatomy

Impact of PDP (Rogers & McClelland, 2014)

- Representation and processing: No rules required
- Origins of knowledge and developmental change: Rethinking innateness
- Cognitive neuropsychology: Explanation without the transparency assumption
- Machine learning: Deep learning
- Impact on theories in specific domains
  - Interactive processes in perception, language, and cognition
  - Reading and language processing
  - Optimality theory (linguistics)
  - Long-term memory
  - Semantics/conceptual knowledge
  - Cognitive control
  - Sequential processing

Limitations and controversies (Rogers & McClelland, 2014)

- Limitations and criticisms of back propagation
  - Catastrophic interference
- Lack of transparency
- PDP models cannot capture abstract relational structure
- PDP models are too flexible

Current issues and future directions (Rogers & McClelland, 2014)

- Probabilistic models of cognition
- Role of "statistical learning" widely appreciated in development
- Acceptance of sensitivity to both specific and general information
- Resurgence of neural networks in machine learning
- The advent of computational cognitive neuroscience
- Distributed representations are being taken seriously by cognitive neuroscience

"Models are important, not as expressions of belief, but as vehicles for exploring the implications of ideas"