[Overview] [Goals and Assessment] [Readings] [Software] [Syllabus: 1 2 3 4 5]

85-419/719: Introduction to Parallel Distributed Processing

Spring 2017, Tue/Thu 10:30-11:50am, Porter 226C

Instructor: David Plaut
Baker 254N, x85145
Office hours: By appt. (arrange by email)
plaut@cmu.edu

TA: Ven Popov
Baker 455, x8112
Office hours: TBD Baker Hall
popov@andrew.cmu.edu

Announcements

Assignments

Overview

The goal of the course is to introduce the basic principles of parallel distributed processing (also known as connectionist or neural network modeling) and to illustrate how these principles provide insight into human cognitive processing. In addition, the course will cover some issues in neural and cognitive development, cognitive impairments due to brain damage, and some basic computational issues. The course also attempts to introduce the general practice of studying cognition through computational modeling and analysis. There will be computer simulation exercises in addition to readings. Homework assignments will generally require you to report the results of simulations you have carried out, to analyze these results, and to think critically about some issues raised in the readings. There will also be a final project that will typically involve simulation modeling.

The course is divided into five sections. The first three cover basic topics in parallel distributed processing. For each of these, a **homework assignment** is handed out at the beginning of the section and is due at the end of the section (n.b. the third section's homework is split into two parts). At the end of the third section, you will also be required to submit a one-page **proposal** outlining the final project you intend to carry out. This will be returned with feedback at the beginning of the fourth section (right after Spring Break), and you will be expected to get started on your project immediately thereafter. You should be on the lookout throughout the earlier sections of the course for topics or issues that you find particularly interesting and would like to pursue in more detail in a project. The fourth section focuses on applications from a range of perceptual, linguistic and cognitive domains, and will be followed by a **take-home essay** (1000-1200 words) based on class lectures and readings. The final section will be devoted to a brief **oral report** from each student on the topic of their project. A 12-15 page **final paper** (5000-7000 words) based on the project is due one week after the end of classes. There is **no final exam** for the course.

In general, there are assigned readings for each lecture that are intended to prepare you to participate in the class discussion for that day. In addition, there may be optional background readings (marked with "*opt:*" and in parentheses in the Syllabus) that serve either as the basis for the lecture, to present an alternative point of view, or simply to make available to you relevant material that we won't have time to cover in class. Optional readings are also a good source of ideas for projects. There are no required readings on days when something is due, but you are still expected to attend class, hand in your homework, and draw on the material you have already learned in order to participate in the discussion.

Course Goals and Assessment

Below are the broad **goals** of the course and how each is assessed (listed in brackets).

- Extend breadth of knowledge of cognitive psychology, including theoretical perspectives, research findings, and applications [through assigned readings, homeworks, and in-class discussions]
- Foster familiarity with diverse experimental paradigms used in psychology [through hands-on experience with computational modeling, assessed via homework assignments and the final project].
- Engender the ability to read and critique psychological articles [assessed through homework assignments, take-home exam, and in-class discussions of assigned readings].
- Improve skill in oral and written presentation [through the oral project presentation and written final project report].
- Increase facility in designing psychological studies to address research questions [through the design of a final modeling project].
- Foster critical thinking and creativity [through in-class discussions and the formulation and execution of a final project].

The **grading** in the class will be divided up as follows:

Homework 1:	10%
Homework 2:	15%
Homework 3:	15%
Homework 4:	10%
Project proposal:	5%
Take-home essay:	10%
Oral presentation/class participation:	5%
Final project	30%

Assignments should be uploaded within Canvas and are due at the beginning of class on the date listed in the Syllabus (usually a Tuesday). **Late penalties** will be assessed as follows: Homeworks handed in late but before 5pm of the next day (usually a Wednesday) will be penalized by 5% of the total possible points (i.e., the graded score will be multiplied by 0.95); those handed in before 5pm of the following weekday (usually a Thursday, but a Monday if the homework was due on a Thursday) will be penalized by 10%; those handed in later than that but before graded papers are returned will be penalized by 15%. Papers may not be handed in for credit after other students' graded homeworks are returned and feedback is posted, except with explicit permission from the instructor. Late homeworks may be submitted to the instructor by email (pdf file). The 5% for class participation will be based on contributions to class discussions throughout the semester, and on the quality of the oral project report.

Academic integrity

All submitted work must be solely the product of your own original work for this course. You must not work with other students on homeworks, and you must not look at solutions to problems from previous semesters of this course (from past students or the instructor), even if you have access such solutions. Your final paper must appropriately cite the sources on which it is based, particularly when text from a source is included verbatim or paraphrased. Out of class, you are encouraged to discuss issues and content related to the course with other students, as well as possible final paper topics. However, you must not discuss a specific homework assignment until after the submissions for it are graded and returned. The **minimum** penalty for a violation of academic integrity is to receive zero points on the relevant assignment, and all violations will be reported to the Office of Student Affairs for possible further disciplinary action.

Readings

There is **no required text** for the course. All assigned and optional readings, as well as lecture slides, are available as downloadable pdf files from links in the Syllabus below. Other course materials (e.g., handouts, assignments, etc.) will be made available via links at the top of this web page. The following texts contain some of the course readings and may be useful as general references:

- **PDP1:** Rumelhart, D.E., McClelland, J.L., & the PDP research group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* Cambridge, MA: MIT Press.
- **PDP2:** McClelland, J.L., Rumelhart, D.E., & the PDP research group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models.* Cambridge, MA: MIT Press.
- **PDP Handbook:** McClelland, J.L. and Rumelhart, D.E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises.* Cambridge, MA: MIT Press.
- **MPR:** McLeod, P., Plunkett, K. and Rolls, E.T. (1998). *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford: Oxford University Press.
- **CCN:** O'Reilly, R.C., Munakata, Y., Frank, M.J., and Hazy, T. (2014). *Computational cognitive neuroscience.* http://grey.colorado.edu/CompCogNeuro/index.php/CCNBook/Main

---

Software

We will be using a software package called "Lens" (for Light Efficient Network Simulator), developed by former CMU CS graduate student Doug Rohde. Lens runs under Windows, Mac OSX, and Linux. The **main website** for Lens is http://tedlab.mit.edu/~dr/Lens/ but **don't** install Lens from that site. You can download a file containing a precompiled version of Lens here:

- **Windows:**
  Download the file Lens-windows.zip and use WinZip or a similar program to unzip the file. Inside the resulting folder will be a folder called "Lens". Read the file "README.rtf" in this directory for further instructions. Because of differences in the way Windows and Unix-based systems handled spaces, it will simplify things if you put the Lens directory at the top level of the drive (i.e., C:\Lens).
- **Mac OS X:**
  Download the file Lens-OSX.zip and double-click it to create an unzipped version. Inside will be a folder called "Lens". Read the text file README.txt in this folder for further instructions.

- **Linux:**
  Download the file Lens-linux.tar.gz. Open a terminal and untar it with the command "tar xzf Lens-linux.tar.gz". This will create a directory called "Lens". Read the text file README in this directory for further instructions.

If you have any problems getting Lens running, contact the TA or instructor. After installing Lens, you should look at the online manual at http://tedlab.mit.edu/~dr/Lens/, particularly the instructions under "Running Lens" and the Tutorial Network under "Example Networks". The precompiled versions of Lens come with a offline (local) copy of the manual that can be accessed by pointing your web browser at `Manual/index.html` in the Lens directory.

Syllabus

This syllabus is **subject to change** throughout the course, so be sure to revisit this web page frequently.

Section 1: Processing and Constraint Satisfaction

Jan 17 (Tue): Overview and basic principles (slides) **[HOMEWORK 0: Install Lens] [HOMEWORK 1 POSTED]**

- Rogers, T.T. (2009). Connectionist models. In Squire, L. (Ed.), Encyclopedia of Neuroscience, Volume 3, pp. 75-82. Oxford: Academic Press.
- (*opt:* Rogers, T.T. and McClelland, J.L. (2014). Parallel distributed processing at 25: Further explorations in the microstructure of cognition. *Cognitive Science, 38,* 1024-1077. **Up to the top of p. 1041 only.** This goes into a bit more depth than the Rogers chapter.)
- (*opt:* O'Reilly et al. (2014). Introduction, CCN, Chapter 1.)
- (*opt:* O'Reilly et al. (2014). Neuron, CCN, Chapter 2.)

Jan 19 (Thu): Lens tutorial **[BRING LAPTOP]**

Jan 24 (Tue): Constraint satisfaction

- McLeod, P., Plunkett, K. and Rolls, E.T. (1998). The attraction of parallel distributed processing for modelling cognition. MPR, Chapter 2.
- (*opt:* O'Reilly et al. (2014). Networks, CCN, Chapter 3.)

Jan 26 (Thu): Schema theory (slides)

- Rumelhart, D.E., Smolensky, P., McClelland, J.L., & Hinton, G.E. (1986). Schemata and sequential thought processes in PDP models. PDP2, Chapter 14. **Up to page 38 (top) only.**

Jan 31 (Tue): Psychological implications (slides) **[HOMEWORK 1 DUE] [HOMEWORK 2 POSTED]**

- (*opt:* McClelland, J.L. & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375-407.)
- (*opt:* McClelland, J.L. & Elman, J.L. (1986). The TRACE model of speech perception. PDP2, Chapter 15. The speech analog of the IA model.)
- (*opt:* Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in normal and aphasic speakers. *Psychological Review, 104*, 801-838. A model of speech production based on similar principles as the IA model.)

Section 2: Simple Learning and Distributed Representations

Feb 2 (Thu): Correlation-based learning (Hebb rule) (slides)

- McLeod, P., Plunkett, K. and Rolls, E.T. (1998). Pattern association. MPR, Chapter 3.
- (*opt:* O'Reilly et al. (2014). Learning, CCN, Chapter 4.)

Feb 7 (Tue): Error-correcting learning (Delta rule)

- (*opt:* Hinton, G.E. (1989). Connectionist learning procedures. *Artificial Intelligence, 40*, 185-234. **Up to page 198 only.** An alternative presentation of the delta rule.)

Feb 9 (Thu): Distributed representations (slides)

- Hinton, G.E., McClelland, J.L., & Rumelhart, D.E. (1986). Distributed representations. PDP1, Chapter 3.
- (*opt:* Plaut, D.C. and McClelland, J.L. (2010). Locating object knowledge in the brain: A critique of Bowers's (2009) attempt to revive the grandmother cell hypothesis. *Psychological Review, 117*, 284-290. A position piece that argues against localist representations in the brain.)

Feb 14 (Tue): Psychological implications (slides) **[HOMEWORK 2 DUE] [HOMEWORK 3 POSTED]**

- (*opt:* McClelland, J.L. & Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159-188.)
- (*opt:* Farah, M.J. and McClelland, J.L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General, 120*, 339-357. A pattern-associator model of selective deficits for living vs. nonliving things.)

Section 3: Learning Internal Representations

Feb 16 (Thu): Back-propagation (slides)

- Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning internal representations by error propagation. PDP1, Chapter 8.
- (*opt:* Hinton, G.E. (1989). Connectionist learning procedures. *Artificial Intelligence, 40*, 185-234. **From page 198 to end only.** More detailed treatment of back-propagation with example simulations.)
- (*opt:* Rumelhart, D.E., Durbin, R., Golden, R., and Chauvin, Y. (1995). Backpropagation: The basic theory. In Y. Chauvin & D.E. Rumelhart (Eds.), *Back-propagation: Theory, architectures, and applications* (pp. 1-34). Hillsdale, NJ: Erlbaum. An advanced treatment of back-propagation cast in terms of probabalistic inference.)

Feb 21 (Tue): Temporal learning and recurrent networks (slides)

- Elman, J.L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.
- (*opt:* Williams, R.J. & Zipser, D. (1995). Gradient-based learning algorithms for recurrent networks and their computational complexity. In Y. Chauvin & D.E. Rumelhart (Eds.), *Back-propagation: Theory, architectures, and applications* (pp. 433-486). Hillsdale, NJ: Erlbaum. A more advanced treatment of recurrent back-propagation, including back-propagation through time and alternatives.)

Feb 23 (Thu): Generalization and overfitting (slides)

- (*opt:* Morgan, N. & Bourlard, H. (1990). Generalization and parameter estimation in feedforward nets: Some experiments. In D.S. Touretzky (Ed.), *Advances in Neural Information Processing Systems 2*. San Mateo: CA: Morgan Kaufmann, 630-637. Discussion of cross-validation.)
- (*opt:* Weigand, A.S., Rumelhart, D.E., & Huberman, B.A. (1991). Generalization by weight-elimination with application to forcasting. In R.P. Lippmann, J.E. Moody, & D.S. Touretzky (Eds.), *Advances in Neural Information Processing Systems 3*. San Mateo: CA: Morgan Kaufmann, 875-882. More sophisticated form of weight decay.)

Feb 28 (Tue): Boltzmann machines / Contrastive Hebbian learning **[HOMEWORK 3 DUE] [HOMEWORK 4 POSTED]** (slides)

- (opt: Hinton, G.E. & Sejnowski, T.J. (1986). Learning and relearning in Boltzmann Machines. PDP1, Chapter 7.)
- (opt: Hinton, G.E. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences, 11,* 428-434. Multi-layer generative models.)

Mar 2 (Thu): Unsupervised learning / generative models (slides)

- (opt: Becker, S. (1991). Unsupervised learning procedures for neural networks. *International Journal of Neural Systems, 2,* 17-33. General overivew of unsupervised learning.)
- (*opt:* Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE, 78*, 1465-1480. Topographically constrained competitive learning.)

Mar 7 (Tue): Deep learning (slides)

- (*opt:* LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature, 521,* 436-444.)
- (*opt:* Hinton, G.E. (2014) Where do features come from? *Cognitive Science, 38,* 1078–1101.)

Mar 9 (Thu): Reinforcement learning / forward models (slides) **[HOMEWORK 4 DUE] [PROJECT PROPOSAL DUE]**

- (*opt:* Barto, A.G. (1995). Reinforcement learning: Reinforcement learning in motor control. In M.A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 804-813). Cambridge, MA: MIT Press.)
- (*opt:* Tesauro, G. (1995). Temporal difference learning and TD-Gammon. *Communications of the ACM, 38*, 58-68.)
- (*opt:* Jordan, M.I., and Rumelhart, D.E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16*, 307-354.)
- (*opt:* Pickering, M.J. and Clark, A. (2014). Getting ahead: Forward models and their place in cognitive architecture. *Trends in Cognitive Sciences, 18,* 451-456.)

Mar 14 (Tue): NO CLASS (Spring Break)

Mar 16 (Thu): NO CLASS (Spring Break)

Section 4: Applications

Mar 21 (Tue): Cognitive development (slides)

- Munakata, Y. and McClelland, J.L. (2003). Connectionist models of development. *Developmental Science, 6*, 413-429.
- (*opt:* Munakata, Y., McClelland, J.L., Johnson, M.H. & Siegler, R. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104*, 686-713.)
- (*opt:* Mareschal, D. and Quinn, P.C. (2001). Categorization in infancy. *Trends in Cognitive Sciences, 5,* 443-450.)
- (*opt:* Yermolayeva, Y. and Rakison, D.H. (2013). Connectionist modeling of developmental changes in Infancy: Approaches, challenges, and contributions. *Psychological Bulletin, 140,* 224-255.)

Mar 23 (Thu): High-level vision and attention (slides)

- Mozer, M.C. and Sitton, M. (1998). Computational modeling of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341-393). Hove, England: Psychology Press/Erlbaum.
- (*opt:* Mozer, M.C. and Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, *2*, 96-123.)
- (*opt:* Behrmann, M., Zemel, R.S. and Mozer, M.C. (1998). Object-based attention and occlusion: Evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1101-1036.)
- (*opt:* O'Reilly et al. (2014). Perception, CCN, Chapter 6.)

Mar 28 (Tue): Semantics (slides)

- McClelland, J.L. and Rogers, T.T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Neuroscience Reviews, 4*, 310-322.
- (*opt:* Plaut, D.C. (2002). Graded modality-specific specialization in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology, 19*, 603-639.)
- (opt: Rogers, T.T. and McClelland, J.L. (2008). Precis of *Semantic cognition: A Parallel Distributed Processing approach.* *Behavioral and Brain Sciences, 31,* 689-749.)

Mar 30 (Thu): Memory and the hippocampus (slides)

- McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419-457.
- (*opt:* O'Reilly et al. (2014). Memory, CCN, Chapter 8.)

Apr 4 (Tue): Language: Morphology (slides)

- The past-tense debate. *Trends in Cognitive Sciences, 6*, 456-474. [Pinker, S. & Ullman, M T. The past and future of the past tense, 456-463; McClelland, J.M. & Patterson, K. 'Words *or* Rules' cannot exploit the regularity in exceptions: Reply to Pinker and Ullman, 464-465; McClelland, J.M. & Patterson, K. Rules or connections in past-tense inflections: What does the evidence rule out?, 465-472; Pinker, S. & Ullman, M T. Combination and structure, notgradedness, is the issue: Reply to McClelland and Patterson, 472-474.]
- (*opt:* Rumelhart, D.E. & McClelland, J.L. (1986). On learning the past tenses of English verbs. PDP2, Chapter 18. Original model.)
- (*opt:* Seidenberg, M.S., and Plaut, D.C. (2014). Quasiregularity and its discontents: The legacy of the past tense debate. *Cognitive Science, 38,* 1190-1228. Discussion of more general theoretical issues.)

- (*opt:* Plaut, D.C. & Gonnerman, L.M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*, 445-485. Modeling of derivational morphology.)

Apr 6 (Thu): Language: Sentence processing (slides)

- McClelland, J.L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4*, 287-335.
- (*opt:* Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.)
- (*opt:* Rohde, D.L.T., and Plaut, D.C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition, 72*, 67-109. Follow-up modeling casting doubts on Elman's (1993) conclusions.)
- (*opt:* Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559-586). Hillsdale, NJ: Erlbaum. Traditional treatment of language.)
- (*opt:* O'Reilly et al. (2014). Language, CCN, Chapter 9.)

Apr 11 (Tue): NO CLASS (Passover)

Apr 13 (Thu): Cognitive control and executive function (slides) **[TAKE-HOME ESSAY POSTED]**

- Cohen, J.D., Aston-Jones, G., and Gilzenrat, M.S. (2004). A system-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation versus exploration. In M.I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 71-90). New York: Guilford Press.
- (*opt:* Rumelhart, D.E., Smolensky, P., McClelland, J.L., & Hinton, G.E. (1986). Schemata and sequential thought processes in PDP models. PDP2, Chapter 14. **From page 38 to end only.**)
- (opt: Botvinick, M. and Plaut, D.C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review, 111,* 395-429.)
- (*opt:* O'Reilly et al. (2014). Executive functions, CCN, Chapter 10.)

Apr 18 (Tue): Future directions **[TAKE-HOME ESSAY DUE]**

- (*opt:* Rogers, T.T. and McClelland, J.L. (2014). Parallel distributed processing at 25: Further explorations in the microstructure of cognition. *Cognitive Science, 38,* 1024-1077. **From page 1043 to end only.**)

Apr 20 (Thu): NO CLASS (Spring Carnival)

Section 5: Project Progress Reports

Apr 25 (Tue):
Apr 27 (Thu):
May 2 (Tue):
May 4 (Thu):
May 12 (Fri): FINAL PROJECT PAPER DUE