85-426/85-726: Learning in Humans and Machines
Spring 2018

Instructor: Charles Kemp (`ckemp@cmu.edu`)

Office Hour: Thursday 2:30–3:30 in BH 335A, or by appointment

TA: Qiong Zhang (`qiongz@andrew.cmu.edu`)

Office Hour: Tuesday 1:00 - 2:00 in BH 3rd floor lounge (BH 336D)

Classes: Monday and Wednesday, 1:30–2:50, BH 340A

Units: 9, which means that this course should occupy roughly 9 hours of your time every week (including 3 hours in class)

Web Page: `https://canvas.cmu.edu/courses/3493`

Readings: Readings for each class will be posted on Piazza. There is no required textbook but the following books are recommended for background reading on technical material:

- *Machine learning: A probabilistic perspective* by Kevin Murphy
- *Artificial Intelligence: A modern approach* by Russell and Norvig (2nd or 3rd edition)

Description:

This course explores how probabilistic methods can help to explain cognition and to develop intelligent machines. The applications discussed include perception, language, memory, categorization, reasoning, decision-making, and motor control.

Prerequisites: Basic programming in Python will be required for some problem sets. No technical background will be assumed beyond high school mathematics.

Objectives: After completing this course you should be able to:

1. Understand and apply Bayesian inference.
2. Describe and evaluate probabilistic approaches to core aspects of cognition, including reasoning, learning, and categorization
3. Characterize the relationship between probabilistic approaches to cognition and machine learning research.
4. Identify ways in which probabilistic models can be experimentally tested as models of cognition.
5. Develop and test simple models of cognition.

Assessment:

- Problem sets (40%: best 6 worth 6% each, remaining 2 worth 2% each)

- A project on probabilistic models of cognition. (30%)
- Presentation. (10%)
- Commentaries. (10%)
- Participation. (10%)

Piazza:

The Canvas site includes a link to Piazza. All readings and class announcements will be posted on Piazza rather than Canvas, so it's critical that you regularly check Piazza.

Please also use Piazza for asking questions about course material and logistics. In most cases you should post your questions to Piazza instead of emailing us directly — that way other students can help to answer your questions, and the whole class can see the answers. If you see a question on Piazza that you're able to answer, please jump in and do so.

Piazza allows you to post anonymously and to submit private posts that are visible to the instructor but not to other students. I have enabled these options just in case but am expecting that you will use them rarely if at all. If you have a question, chances are that someone else will have a similar question, and posting publicly allows everybody to benefit from your question.

Discussion topics on Piazza are organized using tags called folders. The `specific_classes` folder will include a post for each class. Each of these posts will include readings and a reading guide (if appropriate), and you should submit your commentaries as follow-ups to these posts. There are folders for for each homework assignment, and the project. The `course_content` folder is for questions about course material that aren't related to a specific reading (questions about a specific reading should be posted as followups to the reading guide for that reading). The `interesting_links` folder can be used to post information you come across that is relevant to the class and might interest other students. The `logistics` folder is for questions about course logistics—for example, about due dates, or about where to find a given resource. If your question doesn't fit into any of the folders, please post it anyway and use the `other` folder.

Class organization:

This class is designed to give you a deep understanding of the foundations of probabilistic models and a broad appreciation of how these models can be applied. Before Spring break we will focus mainly on foundations, and technical material will be presented in class and reinforced on the problem sets. Your goal should be to absorb *all* of the core technical material—if points come up in class that aren't absolutely clear, please ask.

After Spring break we will mostly focus on applications of probabilistic models. In the schedule below, any class marked with **[SP]** (for Student Presentation) is an applications class. Some of these applications rely on advanced techniques, and you should not expect to understand these applications at the same level at which you understand the core technical material. Instead, your goal should be to gain an overall sense of the problem that each application tackles and the general approach that it pursues.

Problem sets:

The problem sets are intended to help you learn the foundational technical material. Basic programming in Python will be required.

Presentations:

The applications will be mostly presented by students in the class. There are ten application classes (marked with **[SP]** in the schedule), and each of these classes will be led by a pair of students.

To help you prepare your presentation, you and your partner should arrange a meeting with us sometime during the week before your presentation. You and your partner should also prepare a reading guide that will be posted by midnight on the Friday preceding the presentation.

Commentaries:

There will be a reading assignment for every class, and you are expected to arrive in class with ideas and questions to discuss. To help you develop these ideas, you are required to submit short written commentaries before most classes.

A reading guide for each class will be posted on Piazza that identifies the sections of the readings you should focus on and the questions that should guide your reading. Feel free to use one of these questions as the basis for your commentaries. You can also use your commentaries as a chance to discuss other thoughts about the readings: for example a commentary might take one of the following forms:

1. Disagreement. Mention a claim that doesn't seem right to you.
2. Extension. Describe how the work could be usefully extended. Is there another way that the model could be tested? How might the model be revised to handle a broader range of cognitive phenomena?
3. Criticism. List at least one flaw or limitation with the approach presented.
4. Connection. Draw a connection between the reading and something else that you know about (e.g. something that we've discussed in a previous class, or that you learned about elsewhere).

Each commentary should not be long—one paragraph is typical. You can write up to half a page, but writing three thoughtful sentences is sufficient.

The presenters for the day will be reading your commentaries before class, so feel free to add a postscript with questions or thoughts about what you'd like to discuss during class. For example, if the reading made some point that was confusing but that you'd like to understand, include a note to this effect and we'll try to address it during class.

Commentaries can be submitted for every class except for the two "Project presentation" classes at the end of term. If you submit thoughtful commentaries for 20 or more classes, you will receive full credit for this component of the course. You don't need to submit a commentary on the day that you present one of the **[SP]** classes, but will nevertheless receive credit for a commentary on this day. Commentaries are graded pass/fail.

Turning in problem sets and commentaries:

Problem sets should be turned in by 1:30 on the due date. Submitting them in class is ideal, but you can also leave them in my mailbox (Baker Hall 336D). Late submissions will be accepted up to 24 hours after the deadline, and each student is granted one late submission without penalty.

Commentaries should be submitted by 8:30 am on the morning before class. Note that commentaries cannot be accepted after the class in which the paper is discussed. Commentaries should be posted on Piazza as followups to the reading guide for the relevant paper.

Participation:

This is an interactive course and its success will depend on your participation. 10% of your grade will be based on your participation in the class. You will receive full credit for this component if you engage with the material in class by asking questions and raising discussion points (including the points mentioned in your commentaries). I understand that different students have different personalities, and I'm not expecting that all students will be equally vocal in class. I do expect, however, that all students will speak up sometimes.

To make sure that the participation grade is based in part on something objective, we will be recording attendance at the start of each class. Please arrive on time. I will also take your level of activity on Piazza into account when computing the participation grade.

Project:

Your project can take one of several forms. You can develop a new model of some aspect of cognition, run a small experiment to test an existing computational model, implement an existing model and apply it to a new data set, or write a paper on some topic related to the class. Working on an existing research project is fine as long as this project matches the spirit of the class. You are encouraged to work in pairs, but working on your own is also fine.

The project includes three components:

1. Proposal. Submit no more than one page describing the question you plan to explore and the method you will use. Your proposal is due on February 28.
2. Check in. You will meet with me during the week of April 9 to discuss how your project is going.
3. Presentation. At the end of the semester you will briefly present your project. Your presentation should be about 8 minutes long, and you are encouraged to use visual aids (e.g. slides, handouts, or figures sketched on the board).
4. Writeup.

The proposal, check in and presentation will be graded pass/fail. You will receive full credit for handing in a complete proposal, meeting with me and giving a presentation.

Class policies:

To help everybody stay focused on the course material, please do not use devices (including phones, tablets, and computers) in class unless specifically invited to do so. Some of the reasons for this policy are described at

- `http://educationnext.org/should-professors-ban-laptops-classroom-computer-use-affects-student-learning-study/`
- `http://www.psychologicalscience.org/index.php/news/releases/take-notes-by-hand-for-better-long-term-comprehension.html`
- `https://medium.com/@cshirky/why-i-just-asked-my-students-to-put-their-laptops-away-7f5f7c50f368`

You are encouraged to discuss the assigned readings with others, but the commentaries that you submit should be your own. You may discuss the problem sets with others to some extent. High-level discussions about how to approach the problems are permissible and encouraged. Implementation details (including fine-grained details about specific algorithms) are best figured out on your own, and I'm expecting that your code will not be identical to the code of any other student in the class. If you choose to carry out a project in collaboration with another student, the two of you should submit a single written report.

Accommodation of disabilities:

If you need accommodation for a disability, please talk to me within the first two weeks of semester.

Schedule and readings

The submission dates below are firm, but the list of readings is provisional—please refer to the website for the most current information. Some readings are very long and the website will indicate which sections are important, which can be skimmed, and which can be skipped entirely. The reading guide for each week will be posted by midnight on the Friday preceding that week.

January 17: Introduction and administration

R. A. Jacobs and J. K. Kruschke. Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 8–21, 2011.

(OPTIONAL) Z. Ghahramani, Probabilistic machine learning and artificial intelligence. *Nature*. 521, 2015, p 452 – 459.

January 22: Basic Bayesian inference

Problem Set 1 out

T. L. Griffiths, C. Kemp, and J. B. Tenenbaum. Bayesian models of cognition. In R. Sun, editor, *The Cambridge handbook of computational cognitive modeling.* Cambridge University Press, 2008. Pages 1–7 of the pdf.

E. T. Jaynes. *Probability Theory: The Logic Of Science.* Cambridge University Press, Cambridge, 2003. Chapter 1: Plausible Reasoning. Pages 1 – 6, 13–15.

January 24: Bayesian concept learning

K. P. Murphy. Bayesian concept learning. Pages 1 - 11 of the pdf.

January 29: Levels of analysis

Problem Set 1 due, Problem Set 2 out

David Marr. *Vision.* W. H. Freeman, San Francisco, CA, 1982. Chapter 1: The philosophy and the approach.

January 31: Bayesian networks (Directed Graphical models)

T. L. Griffiths, C. Kemp, and J. B. Tenenbaum. Bayesian models of cognition. In R. Sun, editor, *The Cambridge handbook of computational cognitive modeling.* Cambridge University Press, 2008. Pages 12–18 of the pdf.

February 5: Causal graphical models

Problem Set 2 due, Problem Set 3 out

Y. Hagmayer, S. A. Sloman, D. A. Lagnado, and M. R. Waldmann. Causal reasoning through intervention. In A. Gopnik and L. Schultz, editors, *Causal learning: Psychology, philosophy and computation.* Oxford University Press, Oxford, 2007.

February 7: Functional causal models

C. G Lucas and C. Kemp An improved probabilistic account of counterfactual reasoning. *Psychological Review*, 122(4):700–734, 2015.

February 12: Coins and Dice

Problem Set 3 due, Problem Set 4 out

T. L. Griffiths, C. Kemp, and J. B. Tenenbaum. Bayesian models of cognition. In R. Sun, editor, *The Cambridge handbook of computational cognitive modeling.* Cambridge University Press, 2008. Pages 7–11 of the pdf.

(OPTIONAL) D. Navarro and A. Perfors. An introduction to the beta binomial model.

February 14: Gaussian distributions

K. P. Murphy *Machine Learning: A Probabilistic Perspective.* MIT Press, Cambridge MA, 2012. Pages 32 and 38.

M. S. Landy, M. S. Banks and D. C. Knill. Ideal-Observer models of cue integration. In J. Trommershäuser, K. Kording and M. S. Landy, editors, *Sensory Cue Integration*, pages 5 – 29. Oxford University Press, 2011.

(OPTIONAL) R. Jacobs. Optimal linear cue combination.

February 19: Inference Algorithms

Problem set 4 due, Problem set 5 out

T. L. Griffiths, C. Kemp, and J. B. Tenenbaum. Bayesian models of cognition. In R. Sun, editor, *The Cambridge handbook of computational cognitive modeling.* Cambridge University Press, 2008. Pages 33–37 of the pdf.

(OPTIONAL) W. R. Gilks, S. Richardson and D. Spiegelhalter, Introducing Markov chain Monte Carlo. In W. R. Gilks, S. Richardson and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 1995.

February 21: Topic models [SP]

T. L. Griffiths, C. Kemp, and J. B. Tenenbaum. Bayesian models of cognition. In R. Sun, editor, *The Cambridge handbook of computational cognitive modeling.* Cambridge University Press, 2008. Pages 37–43 of the pdf.

D. Blei. Probabilistic topic models. *Communications of the ACM.* 55:4, 2012.

February 26: Model Selection

Problem set 5 due, Problem set 6 out

D. J. C. MacKay. *Information theory, inference, and learning algorithms.* Chapter 23: Model Comparison and Occam's razor. Cambridge University Press, Cambridge, 2003.

February 28: The automated statistician [SP]

Project proposal due

J. R. Lloyd and D. Duvenaud and R. Grosse and J. B. Tenenbaum and Z. Ghahramani. Automatic construction and natural-language description of Nonparametric Regression models. *Association for the Advancement of Artificial Intelligence (AAAI)*, 2014.

March 5: Memory I

Problem set 6 due, Problem set 7 out

P. Hemmer and M. Steyvers. Integrating episodic and semantic information in memory for natural sciences. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, 2009.

March 7: Memory II

J. R. Anderson and L. J. Schooler. The adaptive nature of memory. In E. Tulving and F. I. M. Craik, editors, *Handbook of memory.* Oxford University Press, New York, 2000.

March 12 and 14: Spring Break

March 19: Categorization

Problem set 7 due, Problem set 8 out

R. L. Goldstone and A. Kersten. Concepts and Categorization. *Handbook of Psychology*, 591–621, 2003.

March 21: Rational process models

T. L. Griffiths, E. Vul and A. N. Sanborn. Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, 21:263–268, 2012.

March 26: Neural networks

J. Pearl, Belief networks revisited. *Artificial Intelligence*, 59, pp 49 – 56, 1993.

J. L. McClelland, Integrating probabilistic models of perception and interactive neural networks: a historical and tutorial review. *Frontiers in Psychology*, 4, 2013.

March 28: Probabilistic programming [SP]

Extracts from N. D. Goodman and J. B. Tenenbaum. Probabilistic Models of Cognition. `http://probmods.org/v2`, 2016.

April 2: Reasoning [SP]

Problem set 8 due

T. L. Griffiths and J. B. Tenenbaum. Optimal predictions in everyday cognition. *Psychological Science*, 17(9):767–773, 2006.

April 4: Decision making [SP]

S. Sher and C. R. M. McKenzie Framing effects and rationality. In N. Chater and M. Oaksford, editors, *The probabilistic mind: Prospects for Bayesian cognitive science*, pages 79–96. 2009.

April 9: Motor control [SP]

Project check in week

J. Trommershäuser, L. T. Maloney and M. S. Landy. Decision making, movement planning and statistical decision theory. *Trends in Cognitive Science*, 12(8):291–297, 2008

April 11: Vision [SP]

R. Lafer-Sousa, K. L. Hermann, and B. R. Conway. Striking individual differences in color perception uncovered by 'the dress' photograph. *Current Biology*, 25(13):R545–R546, 2015.

April 16: Language [SP]

D. Jurafsky. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In J. Hay R. Bod and S. Jannedy, editors, *Probabilistic linguistics.* MIT Press Cambridge, MA, 2002.

April 18: Social reasoning [SP]

P. Shafto, N. D. Goodman, and M. C. Frank. Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, 7(4):341–351, 2012.

April 23: Philosophy of Science [SP]

Howson and Urbach, Bayesian versus non-Bayesian approaches to confirmation. In A. Eagle (Ed), *Philosophy of Probability: Contemporary Readings*, Routledge, 2010.

April 25: Criticisms of Bayesian cognitive models

J. S. Bowers and C. J. Davis. Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138(3):389–414, 2012.

April 30: Project presentations

May 2: Project presentations