

General Examinations Proposal

Brandon Roy

March 2, 2009

Introduction

One of the central issues in studying intelligent systems is *meaning*. What is it, where does it come from, and how is it communicated? How might we build a system that can communicate meaningfully?

At the heart of this proposal is the study of language. For context, we begin with a treatment of child language acquisition. The goal is to provide a solid background in this and related areas. Children have the daunting task of learning language from the ground up, and are remarkably good at it. Attempts to do the same with machines have revealed just how challenging the task is. Technical work in the field of spoken language processing is reviewed, since we are interested in ways to build systems that can learn and communicate. This section concerns theoretical approaches to learning and inference as well as practical techniques. For example, what audio processing methods are useful for automatic speech recognition? What binds these sections together is a main area of interest, “computational semantics.” How do signals become meaningful abstractions? Which abstractions are worth learning? How do abstracted symbols link to one another, and to the external world?

A primary goal of working through this reading list is to emerge with a strong foundation in language acquisition, learning and communication in humans and machines. An additional, and perhaps more significant goal is to synthesize these areas and offer a new perspective to guide future research.

Main Area: Computational Semantics

Examiner

Deb Roy

Associate Professor of Media Arts and Sciences
Director of the Cognitive Machines Group
MIT Media Laboratory

Description

These readings approach ideas of semantics and meaning from multiple perspectives, including philosophy, artificial intelligence, and cognitive psychology. A central concern is how agents represent external and internal phenomena (eg. objects and events, intentions) and how they can communicate with one another about these phenomena. The perspective taken is a computational one: we'd like to think about how these ideas could be implemented with a computer, or understand what is missing if they can't.

Requirement

The completion requirement for this area will be a publishable-quality paper, evaluated by Deb Roy.

Examiner's signature: _____ Date: _____

Reading List

Foundations

- [1] Valentino Braitenberg. *Vehicles, Experiments in Synthetic Psychology*. Bradford Book, 1984.
- [2] D. Dennett. Evolution, Error and Intentionality. In *The Intentional Stance*, pages 287–321. The MIT Press, 1987.
- [3] Daniel C. Dennett. *Brainchildren: Essays on Designing Minds*, chapter 5: "Real Patterns", pages 95–120. The MIT Press, 1998.
- [4] Stevan Harnad. The symbol grounding problem. *Physica D*, 42:335–346, Jun 1990.
- [5] Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow. Behavior, purpose and teleology. *Philosophy of Science*, 10(1):18–24, 1943.

- [6] John R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.
- [7] Terry Winograd and Fernando Flores. *Understanding Computers and Cognition : A New Foundation for Design*. Addison-Wesley Professional, January 1987.
- [8] L. Wittgenstein. *The Blue and Brown Books*. Harper Perennial, 1965.

Cognitive Psychology

- [1] Lawrence W. Barsalou. Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(04):577–660, 1999.
- [2] Lawrence W. Barsalou. Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1435):1177–1187, July 2003.
- [3] J.L. Elman. Language as a dynamical system. In *Mind As Motion: Explorations in the Dynamics of Cognition*, chapter 8, pages 195–225. The MIT Press, 1995.
- [4] C.J. Fillmore. The Case for Case. In E.W. Bach and R.T. Harms, editors, *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1968.
- [5] J. Huttenlocher. Language and thought. In George A. Miller, editor, *Communication, language, and meaning: Psychological perspectives*, chapter 16, pages 172–184. Basic Books, 1973.
- [6] Jean Piaget. *The Child's Conception of the World*. Routledge & Kegan Paul, 1929.
- [7] Zenon W. Pylyshyn. Situating vision in the world. *Trends in Cognitive Sciences*, 4(5):197–207, May 2000.
- [8] Fei Xu and Joshua B. Tenenbaum. Word learning as bayesian inference. *Psychological Review*, 114(2):245–272, April 2007.

AI / Computational Models

- [1] Gary L. Drescher. The schema mechanism. In Stephen Jose Hanson, Werner Remmele, and Ronald L. Rivest, editors, *Machine Learning: From Theory to Applications*, volume 661 of *Lecture Notes in Computer Science*, pages 125–138. Springer, 1993.
- [2] N.D. Goodman, V.K. Mansinghka, and J.B. Tenenbaum. Learning grounded causal models. In *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society*, 2007.

- [3] Barbara J. Grosz and Candace L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.
- [4] Daniel Jurafsky and James H. Martin. *Speech and Language Processing*, chapter 17–21. Prentice Hall, 2nd edition, 2008.
- [5] Walter Kintsch. Symbol systems and perceptual representations. In Manuel de Vega, Arthur M. Glenberg, and Arthur C. Graesser, editors, *Symbols and Embodiment: Debates on Meaning and Cognition*, pages 145–163. Oxford University Press, 2008.
- [6] M. Minsky. A Framework for Representing Knowledge. MIT AI Laboratory Memo 306, Massachusetts Institute of Technology, Cambridge, MA, USA, June 1974.
- [7] Allen Newell. Physical symbol systems. *Cognitive Science*, 4:135–183, 1980.
- [8] Deb Roy. Grounding words in perception and action: computational insights. *Trends in Cognitive Sciences*, 9(8):389–396, August 2005.
- [9] Deb Roy. A mechanistic model of three facets of meaning. In Manuel de Vega, Arthur M. Glenberg, and Arthur C. Graesser, editors, *Symbols and Embodiment: Debates on Meaning and Cognition*, chapter 11. Oxford University Press, 2008.
- [10] Deb Roy and Niloy Mukherjee. Towards situated speech understanding: visual context priming of language models. *Computer Speech & Language*, 19(2):227–248, April 2005.
- [11] Deb K. Roy and Alex P. Pentland. Learning words from sights and sounds: a computational model. *Cognitive Science*, 26(1):113–146, 2002.
- [12] R.C. Schank and R.P. Abelson. *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. Lawrence Erlbaum Associates, 1977.
- [13] J.M. Siskind. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61(1-2):39–91, 1996.
- [14] C. Yu and D.H. Ballard. A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, 70(13-15):2149–2165, 2007.

Technical Area: Spoken Language Processing

Examiner

Dr. Allen Gorin

Director, Human Language Technology Research

U.S. Department of Defense

Fort Meade, MD

Description

To take the idea of *computational* semantics seriously requires a solid technical foundation. The technical section of this proposal draws from the literature on automatic speech and language processing as well as machine learning. One motivation for covering this area is inherently practical: if we want to build systems that learn from sensory input, what tools are the most effective? How do we obtain useful structures for computation from perceptual input? This reading list attempts to cover work in speech and language processing, along with both established and new work in machine learning and pattern recognition.

Requirement

The written requirement for this area will be a 24 hour take home examination to be administered and evaluated by Allen Gorin.

Examiner's signature: _____ Date: _____

Reading List

Speech and Audio Processing

- [1] Albert S. Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press, September 1994.
- [2] J. P. Campbell. Speaker recognition: a tutorial. *Proceedings of the IEEE*, 85(9):1437–1462, 1997.
- [3] Y. Chow, M. Dunham, O. Kimball, M. Krasner, G. Kubala, J. Makhoul, P. Price, S. Roucos, and R. Schwartz. BYBLOS: The BBN continuous speech recognition system. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 12, pages 89–92, 1987.
- [4] S. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word

recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing* [see also *IEEE Transactions on Signal Processing*], 28(4):357–366, 1980.

- [5] Sabine Deligne and Frédéric Bimbot. Inference of variable-length linguistic and acoustic units by multigrams. *Speech Communication*, 23(3):223 – 241, 1997.
- [6] Raul Fernandez and Rosalind W. Picard. Classical and novel discriminant features for affect recognition from speech. In *Interspeech*, pages 473–476, September 2005.
- [7] A. L. Gorin, G. Riccardi, and J. H. Wright. How may i help you? *Speech Communication*, 23(1/2):113–127, 1997.
- [8] T. Hain, S. Johnson, A. Tuerk, P. Woodland, and S. Young. Segment Generation and Clustering in the HTK Broadcast News Transcription System, 1998.
- [9] Greg Kochanski. Prosody beyond fundamental frequency. In S. Sudhoff, D. Lenertov’a, R. Meyer, S. Pappert, P. Augurzy, I. Mleinek, N. Richter, and J. Schließer, editors, *Methods in Empirical Prosody Research*, pages 89–122. Walter de Gruyter, June 2006.
- [10] L. Lamel, L. Rabiner, A. Rosenberg, and J. Wilpon. An improved endpoint detector for isolated word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing* [see also *IEEE Transactions on Signal Processing*], 29(4):777–785, 1981.
- [11] C.M. Lee, S.S. Narayanan, and R. Pieraccini. Combining acoustic and language information for emotion recognition. In *Seventh International Conference on Spoken Language Processing*. ISCA, 2002.
- [12] Alex Park and James R. Glass. Unsupervised word acquisition from speech using pattern discovery. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 409–412, 2006.

Language processing

- [1] Eugene Charniak. *Statistical Language Learning*. Bradford Books, 1993.
- [2] Carl de Marcken. The unsupervised acquisition of a lexicon from continuous speech. Technical report, Cambridge, MA, USA, 1995.
- [3] Carl de Marcken. *Unsupervised Language Acquisition*. PhD thesis, Massachusetts Institute of Technology, September 1996.
- [4] A. L. Gorin, S. E. Levinson, L. G. Miller, A. N. Gertner, A. Ljolje, and E. R. Goldman. On adaptive acquisition of language. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 601–604 vol.1, 1990.

- [5] Allen Gorin. On automated language acquisition. *The Journal of the Acoustical Society of America*, 97(6):3441–3461, 1995.
- [6] Pat Langley and Jaime G. Carbonell. Language acquisition and machine learning. In Brian MacWhinney, editor, *Mechanisms of Language Acquisition*, chapter 5, pages 115–155. Lawrence Erlbaum Associates, 1987.
- [7] Ronald Rosenfeld. A maximum entropy approach to adaptive statistical language modelling. *Computer Speech & Language*, 10(3):187 – 228, 1996.
- [8] N. Tishby and A. Gorin. Algebraic learning of statistical associations for language acquisition. *Computer Speech & Language*, 8(1):51–78, 1994.

Techniques

- [1] Francis Bach and Michael Jordan. Learning spectral clustering, with application to speech separation. *Journal of Machine Learning Research*, 7:1963–2001, October 2006.
- [2] David M. Blei, Thomas L. Griffiths, Michael I. Jordan, and Joshua B. Tenenbaum. Hierarchical topic models and the nested chinese restaurant process. In *Advances in Neural Information Processing Systems 16: Proceedings of the 2003 Conference*. The MIT Press, 2004.
- [3] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, January 2003.
- [4] C.J.C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [5] S. Della Pietra, V. Della Pietra, J. Lafferty, R. Technol, and S. Brook. Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):380–393, 1997.
- [6] M.D. Escobar and M. West. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90(430):577–588, 1995.
- [7] I.J. Good. The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika*, 40(3-4):237–264, 1953.
- [8] Daniel Jurafsky and James H. Martin. *Speech and Language Processing*. Prentice Hall, 2nd edition, 2008.
- [9] T. Krishnan and S. C. Nandy. Discriminant analysis with a stochastic supervisor. *Pattern Recognition*, 20(4):379–384, 1987.
- [10] J.B. Kruskal and D. Sankoff. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*. Addison-Wesley Reading, Mass., 1983.

- [11] Andrew Y. Ng, Michael I. Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, volume 14, 2002.
- [12] A. Poritz. Linear predictive hidden markov models and the speech signal. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 7, pages 1291–1294, May 1982.
- [13] A. Poritz. Hidden Markov models: a guided tour. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 7–13, 1988.
- [14] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [15] R.E. Schapire. The Boosting Approach to Machine Learning: An Overview. *Nonlinear Estimation and Classification*, pages 149–172, 2003.
- [16] Matthias Seeger. Gaussian processes for machine learning. *International Journal of Neural Systems*, 14(2):69–106, 2004.
- [17] Naftali Tishby, Fernando C. Pereira, and William Bialek. The information bottleneck method. In *Annual Allerton Conference on Communication, Control and Computing*, volume 37, pages 368–377, 1999.
- [18] Geoffrey I. Webb. Multiboosting: A technique for combining boosting and wagging. *Machine Learning*, 40(2):159–196, 2000.

Theory

- [1] Nelson Blachman. The Amount of Information that y Gives about X . *IEEE Transactions on Information Theory*, 14(1):27–31, January 1968.
- [2] Rudi Cilibiasi and Paul Vitanyi. Clustering by compression. *IEEE Transactions on Information Theory*, 51:1523–1545, 2005.
- [3] K S Fu and T L Booth. Grammatical inference: introduction and survey—part i. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(3):343–359, 1986.
- [4] K S Fu and T L Booth. Grammatical inference: introduction and survey—part ii. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(3):360–375, 1986.
- [5] M. Mohri, F. Pereira, and M. Riley. The design principles of a weighted finite-state transducer library. *Theoretical Computer Science*, 231(1):17–32, 2000.
- [6] John R. Pierce. *An Introduction to Information Theory: Symbols, Signals & Noise*. Dover Publications, 1980.

- [7] Ray J. Solomonoff. A formal theory of inductive inference, part 1. *Information and Control*, 7(1):1–22, March 1964.
- [8] Ray J. Solomonoff. A formal theory of inductive inference. part 2. *Information and Control*, 7(2):224–254, June 1964.
- [9] George Kingsley Zipf. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley Press, 1949.

Contextual Area: Human Communicative Development

Examiner

Brian MacWhinney

Professor of Psychology

Department of Psychology

Carnegie Mellon University

Description

This section focuses on language development, emphasizing early language acquisition in children. From identifying the words in a stream of speech, to learning the syntax of a language, to learning the meanings of words and how to use them, the child language-learner has an impressive set of linked problems to solve. The readings in this section cover these and other aspects of human communicative development.

Requirement

The written requirement for this area will be a 24 hour take home examination to be administered and evaluated by Brian MacWhinney.

Examiner's signature: _____ Date: _____

Reading List

- [1] E. Bates and J.C. Goodman. On the Emergence of Grammar from the Lexicon. In Brian MacWhinney, editor, *The Emergence of Language*, chapter 2, pages 29–79. Lawrence Erlbaum Associates, 1999.
- [2] Heike Behrens. The input-output relationship in first language acquisition. *Language and Cognitive Processes*, 21(1-3):2–24, April 2006.
- [3] P. Bloom. *How Children Learn the Meanings of Words*. The MIT Press, 2000.
- [4] Michael R. Brent and Jeffrey Mark Siskind. The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2):33–44, 2001.
- [5] Roger Brown. *A First Language*. Harvard University Press, 1973.
- [6] Jerome S. Bruner. *Child's talk: Learning to Use Language*. WW Norton, 1983.
- [7] Herbert H. Clark. *Using Language*, volume 23. Cambridge University Press, 1996.

- [8] Beverly A. Goldfield and Steven J. Reznick. Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17(1):171–183, February 1990.
- [9] Alison Gopnik. Three types of early word: the emergence of social words, names and cognitive-relational words in the one-word stage and their relation to cognitive development. *First Language*, 8(22):49–69, February 1988.
- [10] J. Huttenlocher, W. Haight, A. Bryk, M. Seltzer, and T. Lyons. Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(236–248), 1991.
- [11] Natasha Z. Kirkham, Jonathan A. Slemmer, and Scott P. Johnson. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83(2):35–42, 2002.
- [12] Patricia K. Kuhl. Early language acquisition: Cracking the speech code. *Nat Rev Neurosci*, 5(11):831–843, November 2004.
- [13] M. L. Laakso, A. M. Poikkeus, J. Katajamaki, and P. Lyytinen. Early intentional communication as a predictor of language development in young toddlers. *First Language*, 19(56):207–231, June 1999.
- [14] P. Li, X. Zhao, and B. MacWhinney. Dynamic Self-Organization and Early Lexical Development in Children. *Cognitive Science*, 31(4):581–612, 2007.
- [15] B. MacWhinney. The competition model. In B. MacWhinney, editor, *Mechanisms of Language Acquisition*, chapter 8, pages 249–308. Lawrence Erlbaum Associates, 1987.
- [16] B. MacWhinney. A unified model of language acquisition. In J. Kroll and A. de Groot, editors, *Handbook of bilingualism: Psycholinguistic approaches*, pages 49–67. Oxford University Press, 2004.
- [17] B. MacWhinney. How Mental Models Encode Embodied Linguistic Perspectives. In R.L. Klatzky, B. MacWhinney, and M. Behrmann, editors, *Embodiment, Ego-Space, and Action*, Carnegie Mellon Symposia on Cognition Series, chapter 11, pages 369–410. Psychology Press, 2008.
- [18] G.F. Marcus. *The algebraic mind*. The MIT Press, 2001.
- [19] E.L. Newport and R.N. Aslin. Innately constrained learning: Blending old and new approaches to language acquisition. In *Proceedings of the 24th Annual Boston University Conference on Language Development*, volume 1, 2000.
- [20] W. O’Grady. *Syntactic carpentry: An Emergentist Approach to Syntax*. Lawrence Erlbaum, 2005.
- [21] Ann M. Peters. *The Units of Language Acquisition*. Cambridge University Press, 1983.
- [22] T. Regier. Emergent constraints on word-learning: a computational perspective. *Trends in Cognitive Sciences*, 7(6):263–268, June 2003.

- [23] T. Regier. The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, 29:819–865, 2005.
- [24] J.R. Saffran. Statistical language learning: mechanisms and constraints. *Current Directions in Psychological Science*, 12(4):110–114, 2003.
- [25] J.R. Saffran, R.N. Aslin, and E.L. Newport. Statistical Learning by 8-Month-Old Infants. *Science*, 274(5294):1926–1928, 1996.
- [26] J.R. Saffran, E.L. Newport, and R.N. Aslin. Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, 35(4):606–621, 1996.
- [27] Mark S. Seidenberg. Language Acquisition and Use: Learning and Applying Probabilistic Constraints. *Science*, 275(5306):1599–1603, 1997.
- [28] Mark S. Seidenberg, Maryellen C. MacDonald, and Jenny R. Saffran. Does Grammar Start Where Statistics Stop? *Science*, 298(5593):553–554, 2002.
- [29] LB Smith. Learning how to learn words: An associative crane. *Becoming a word learner: A debate on lexical acquisition*, pages 51–80, 2000.
- [30] C.E. Snow. Conversations with children. *Language acquisition*, 2:69–76, 1986.
- [31] E.D. Thiessen. Statistical learning. In E. Bavin, editor, *Cambridge Handbook of Child Language*. Cambridge University Press, In press.
- [32] E.D. Thiessen, E.A. Hill, and J.R. Saffran. Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, 7(1):53–71, 2005.
- [33] S.P. Thompson and E.L. Newport. Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning and Development*, 3(1):1–42, 2007.
- [34] M. Tomasello. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, 2003.
- [35] Lev Vygotsky. *Thought and Language*. The MIT Press, 1986.
- [36] Heinz Werner and Bernard Kaplan. *Symbol Formation: An Organism Developmental Approach to the Psychology of Language*. Wiley New York, 1963.