Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How many features does it take to change a lightbulb?

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Author

Powers, David

Publication Date

2013

How many features does it take to change a lightbulb?

David M W Powers (powers@computer.org)
Beijing Municipal Lab for Multimedia & Intelligent Software
Beijing University of Technology, Beijing, China

Brain Signals Laboratory &
CSEM Centre for Knowledge & Interaction Technology
Flinders University, Adelaide, South Australia

Keywords: Computational Psycholinguistics, Sensor Fusion, Epistemic Robotics, Cognitive Linguistics.

Introduction

In the 1970s and 80s Cognitive Science and Cognitive Linguistics and Computational Psycholinguistics emerged as the boxes around our disciplines started to become straight-jackets, and research out of one discipline would start to make waves in others. The toy systems of Artificial Intelligence were reaching limits, and introspection by programmers and engineers was reinventing square wheels without any biological plausibility and in ignorance of relevant work across the cognitive sciences, while conversely, work in other fields often lacked the understanding of computability and complexity necessary to ensure that models were realistic and computationally plausible.

This is the starting point for the research program I have been undertaking for the last 35 years, seeking to build intelligent computer systems and computational cognitive models. The idea has been to try to build an intelligent system modelled on the way a baby learns about the world, culture, society and language. Conversely, the idea has been to explore theories from psychology, linguistics and neuroscience through the medium of computational models. The primary focus and agenda of our research program are summed up in Powers and Turk (1989): language and ontology are learned together through multimodal association.

Language & Ontology

Over the years, the breadth of the both the "Language" and "Ontology" learning aspect of the research has grown to include audio-visual speech, gesture and emotion recognition and synthesis, as well as robots both simulated and physical. The earliest models (Powers, 1983; 1984) self-organized with a clear dependency on closed class lexemes as the basic for syntactic structuring, and later work extended this to the levels of phonology and morphology (Powers, 1991;1997abc). In parallel, the same learning models, including both statistical and neurally based co-clustering models, were also used to learn noun, verb and preposition semantics in the context of a robot world simulation, and remain of major importance in our research (Pfitzner et al. 2009; Leibbrandt & Powers, 2010;2012).

The physical models ranged from a robot baby that turned and looked at you if your talked to it or touched it (Powers, 2002), whilst wheeled robots took on a life of their own (Powers et al., 2012) with simulated Teaching Head applications becoming a major focus (e.g. Milne et al., 2011-12) as an outcome of a major ARC/NHMRC Thinking Systems initiative that not only funded our "Thinking Head" project, but our colleagues' "Thinking Hand" and "Thinking Feet" projects.[1] Whereas we concentrated on hands and feet and wheels for locomotion, with fairly conventional path planners for navigation, and made use of conventional robotic grippers for grasping, or much safer simulated grasping for our Hybrid World (Newman et al., 2010), this Thinking Hand team concentrated on such matters as how to hold a glass or a light bulb without breaking it, whilst the Thinking Feet team looked at spiking models for navigation.

One of the core driving forces for our work at this point is the realization that our "five senses" actually hide a multitude of specific sensors and percepts *each*. For example the fingerprints on the hand distinguish the transverse motion of slip vs the normal force of pressure, in ensuring we neither drop nor crush the light bulb. Our two eyes and four types of visual transducer, and the different afferents and efferents involved in controling convergence and focus and aperture, combine with our inner and outer ears to direct our gaze and focus sound, with two different Nyquist tradeoffs of time vs frequency, with 3D balance and inertial sensing. Much of our focus is combining together different senses or subsenses, or discriminating out the different features from our combined senses that have particular value in the tasks we attempt.

This combination of multiple sensory or feedback inputs is called fusion (Lewis and Powers, 2000;2003) and complements processes of signal deconvolution (Li et al., 2003) and feature selection (Atyabi et al., 2012). Computationally it is not effective to learn by throwing all the mass of sensory input together into one big vector and trying to make sense of it ('early fusion'), but nor is it effective to try to deal with each sense or sensor on its own to do the task, and at the last minute vote to fuse sources or models ('late fusion'). Rather we need to look at the similarities and correlations (e.g. whose lips are moving to know who is talking to us) and dissimilarity and independence (viz. we don't want a committee of yes-men, but of independent thinkers, so we search a large space of potential solutions). The first step is often to figure out how many independent components, or clusters or features there are – or we can use algorithms that decide on the fly.

References

Ali, H. B. & Powers, D. M. W. (2011), Comparison of Region-based and Weighted Principal Component Analysis and Locally Salient ICA in terms of Facial Expression Recognition, *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, *368*: 81-89.

Atyabi, A., Luerssen, M.H., Fitzgibbon, S.P., & Powers, D.M., 2012. Evolutionary feature selection and electrode reduction for EEG classification. *IEEE Congress on Evolutionary Computation (CEC).*

Fitzgibbon, S. P., Powers, D. M. W., Pope, K., & Clark, C. R. (2007). Removal of EEG noise and artefact using blind source separation. *Journal of Clinical Neurophysiology 24*(3):232-243.

Fitzgibbon, SP, Lewis, T, Powers, DMW, Whitham, E, Willoughby, J & Pope, K. (2013), Surface Laplacian of Central Scalp *Electrical Signals is Insensitive to Muscle Contamination, IEEE Transactions On Biomedical Engineering, 60*:1, 4-9.

Jia, X., Han, Y., Powers, David M.W., Zheng, Z., Sha, l., Bao, X. and Fei, J., (2012). Spatial and temporal visual speech feature for Chinese phonemes, *Journal of Information and Computational Science 9*:14, 4177-4185.

Leibbrandt, R. E. & Powers, D. M. W. (2010), Frequent frames as cues to part-of-speech in Dutch: why filler frequency matters. *32nd Annual Meeting of the Cognitive Science Society (CogSci2010)*, 2680-2685.

Leibbrandt, R. E. & Powers, D. M. W. (2012). Robust Induction of Parts-of-Speech in Child-Directed Language by Co-Clustering of Words and Contexts, *Joint Workshop of ROBUS and UNSUP*, *Conference of the European Chapter of the Association for Computational Linguistics*, 44-54.

Lewis, T. W. & Powers, D. M. W. (2000). Lip Feature Extraction using Red Exclusion. In P. Eades and J. Jin (Eds), *Visualisation 2000*, *CRPIT 2*:61-70.

Lewis, T. W. and D. M. W. Powers (2003). Audio-Visual Speech Recognition using Red Exclusion and Neural Networks. *Journal of Research and Practice in Information Technology 35,* 1:41-64.

Li, Y., Powers, D. M. W. & Pope, K. (2003). A new approach to blind signal deconvolution using recurrent neural networks. *International Journal of Knowledge-Based Intelligent Engineering Systems 7*:2, 62-69.

Milne, M., Leibbrandt, R. E., Lewis, T. W., Luerssen, M. H. & Powers, D. M. W. (2010). Development of a virtual agent based social tutor for children with autism spectrum disorders. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN2010)*, 1555-1563.

Milne, M., Luerssen, M. H., Lewis, T. W., Leibbrandt, R. E. & Powers. D. M. W. (2011). Designing and Evaluating Interactive Agents as Social Tutors for Children with Autism Spectrum Disorder. In Diana Perez-Marin and Ismael Pascual-Nieto (Eds.) *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*, Madrid, Spain: IGI Global.

Newman, W., Franzel, D., Matsumoto, T., Leibbrandt, R. E., Lewis, T. W., Luerssen, M. H., and Powers, D. M. W. (2010). Hybrid world object tracking for a virtual teaching agent. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN2010)*, 2244-2252.

Pfitzner, D. M., Leibbrandt, R. E. & Powers, D. M. W. (2009) Characterization and Evaluation of Similarity Measures for Pairs of Clusterings, *Knowledge and Information Systems 19*:361-394.

Powers, D. M. W. (1983). Neurolinguistics and Psycholinguistics as a Basis for Computer Acquisition of Natural Language," *SIGART, 84*, 29-34.

Powers, D. M. W. (1984) Natural Language the Natural Way, *Computer Compacts*, 100-109.

Powers, D. M. W. & Turk, C. C. R. (1989). *Machine Learning of Natural Language*, Berlin: Springer-Verlag.

Powers, D. M. W. (1991) Goals, Issues and Directions in Machine Learning of Natural Language and Ontology, *SIGART Bulletin 2:1*, 101-114.

Powers, D. M. W., Clark, C. R., Dixon, S. E. & Weber, D. L. (1996). Cocktails and Brainwaves: Experiments with Complex and Subliminal Auditory Stimuli, *Proceedings of the Australian and New Zealand Conference on Intelligent Information Systems*, 68-71.

Powers, D. M. W. (1997). Learning and Application of Differential Grammars, *CoNLL97: Computational Natural Language Learning*, 88-96.

Powers, D. M. W. (1997). Perceptual Foundations for Cognitive Linguistics, *International Conference on Cognitive Linguistics*, 173.

Powers, D. M. W. (1997). Metaphor and Learning: The Phonology of Syntax, *International Conference on Cognitive Linguistics*, 260-265.

Powers, D. M. W. (1997). Unsupervised learning of linguistic structure: an empirical evaluation, *International Journal of Corpus Linguistics 2*#1, 91-131.

Powers, D. M. W. (2002). Robot babies: what can they teach us about language acquisition?" In J. Leather and J. Van Dam, (Eds) *The Ecology of Language Acquisition*, Kluwer Academic

Powers. D. M. W. (2003). Recall and Precision versus the Bookmaker. *International Conference on Cognitive Science*, 529-534.

Powers, D. M. W. (2011), Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation, *Journal of Machine Learning Technologies 2*:1, 37-63.

Powers, D. M. W. (2012). The Problem of Kappa", *13th Conference of the European Chapter of the Association for Computational Linguistics*, 345-355.

Powers, D. M. W., Atyabi, A. & Anderson, T. (2012). MAGIC Robots - Resilient Communications, *Spring World Congress on Engineering & Technology, IEEE*:591-594

Yang, D. Q. and Powers, D. M. W. (2010), Using Grammatical Relations to Automate Thesaurus Construction, *Journal of Research and Practice in Information Technology 42*#2:129-146

---

[1] http://www.arc.gov.au/ncgp/sri/TS_sumapps_06.htm