

EDITORIAL

Open Access



A critical time in computational cognitive science

David M W Powers 

Abstract

Understanding how people tick is an endeavour that has challenged us for millennia, both in informal settings and in increasingly formalized and scientific disciplines. Some are interested in the biology and others the behaviour. Some are focussed on language and others on culture or emotion. Some approach this as pure science worthy of understanding in its own right. Some approach it as applied science with value born of practical applications that improve our lifestyle and characterize our modern technological society.

Cognitive Science was born as Psychologists and Linguists found that the assumptions and predictions of their theories and models were spilling outside their discipline, as Neuroscience provided a neural substrate for models of perception and cognition that obviated the need for the postulation of hypothetical daemons, as Computer Science and Mathematics provided computational tools for testing models and theories that couldn't be tested empirically in the real world. Artificial Intelligence was also born in such an environment, with the early researchers exploring models of intelligence as much as developing intelligent programs. But Computational Intelligence and Cognitive Science increasingly reflect a contrast between an applied aim of building intelligent applications and entities, and a pure aim of understanding existing intelligent entities and functions, and we seem to lack a bridge between them.

Computational Cognitive Science aims to provide this bridge.

Introduction and scope

I am pleased to announce the launch of *Computational Cognitive Science*, a peer-reviewed open-access journal published by Springer. Our aim is to present work that is inherently interdisciplinary in nature and of multidisciplinary interest, that seeks to understand cognitive and behavioural processes at every level from the neuron to the meme, that seeks to combine the empirical rigours of the biomedical and behavioural sciences with the formal rigours of the computational and mathematical sciences.

Cognitive Science arose as new understanding about restrictions on the capability of machines was matched with corresponding syntactic restrictions on grammars and nesting characterizations of languages. Formal results in learnability of languages led to competing theories of language and language acquisition, with implications extending beyond Linguistics to Psychology, Neuroscience and Evolutionary Biology. Formal results about the learnability of concepts by neural networks led to another bifurcation,

with Artificial Neural Networks becoming a not particularly bioplausible tool of Computational Intelligence, and Biological Neural Networks continuing to be a source of novel architectures and techniques that spawned both useful insights in Neuroscience and specialized tools for Robotics.

Questions about the nature of thought, mind and intelligence were spawned by the rise of Artificial Intelligence, and led to deep philosophical questions that still challenge us, and have in turn spawned new understandings of our different disciplinary perspectives and terminologies, with different symbols and formalisms. Key dichotomies that emerged include: symbolic vs connectionist systems, and grounded vs ungrounded systems. We now talk about Symbolic AI, Strong AI, Weak AI, and even Good Old Fashioned AI (GOFAI)! Developmental Robotics now seeks to understand how language comes about as much as how it is grounded, in an engineering playground of simple robots. Cognitive Linguistics now seeks to understand how and why *metaphor* plays such a major role in language, in a paradigm that is clearly delineated from the nativist assumptions of Generative Linguistics. Machine Learning conversely exploits *similarity* and *diversity* as major characteristics that underlie the

Correspondence: David.Powers@flinders.edu.au
Centre for Knowledge & Interaction Technologies, Flinders University,
Adelaide, Australia

learning available for different problems or applications, but also presumably underlie the learning that takes place in our brains and bodies.

Computational Cognitive Science is in many ways an experimental journal. Indeed I would hope that it will always be prepared to experiment and innovate, to adopt and adapt the best new ideas, and to reject and discard the tired old ideas. *Computational Cognitive Science* is about paradigm breaking! Neither minor tweaks nor in depth analyses of flawed paradigms can help. We are interested in new data and new theories from the Cognitive Sciences that are mathematically consistent and computationally feasible, and thus empirically testable. We are interested in new systems and models from the Computational Sciences whether they are built on the naïve intuitions or insights of programmers or by the foremost experts using the latest data from Neurology or Psychology, but only to the extent that the work is explored in terms of its implications for Psychology or Linguistics or Sociology or Anthropology or Biology...

I am very pleased to be able to call on a broad editorial board who are distinguished for their multidisciplinary view of cognition, who are prepared on the one hand to engage in formal computational and mathematical modeling, and who are prepared on the other hand to engage with and develop the latest data and theories of the cognitive sciences. We are looking to expand the board further, so if you share the vision of *Computational Cognitive Science*, if you see computational modeling as a tool for testing cognitive theory, or you see cognitive theory as a tool for developing intelligent systems, then do get in touch and send me your CV and publications list.

Our aim is that any undergraduate student or research student or early career academic or traditional professor that comes to *CCS* will be able to see broader perspectives on their work, will find introductory material, reviews and pointers to help them understand the research they read, as well as find an editorial board who are a resource for navigating the interdisciplinary mire rather than an arbitrary jury that decides whether your work is publishable or not. Our aim is to have people from multiple disciplines review your work and help frame it or introduce it appropriately for the diverse disciplinary groups.

CCS is interdisciplinary, and thus will publish work that falls between the cracks - research that may take a while to be understood in terms of its relationship with traditional disciplines and paradigms. *CCS* is also multidisciplinary, and thus authors and referees will normally be familiar with and operating in multiple disciplines at once - but we can't be masters of everything, and our aim is not just to provide reviewers for your work, but collaborators who will help introduce it to the rest of

this multidisciplinary nexus we call *Computational Cognitive Science*.

Today we are only at the beginning of a daunting challenge - but with *your* help we can do it!

Format

Computational Cognitive Science has a very flexible format that is designed to guide authors from many different disciplines with many different levels of knowledge. We thus publish not only research papers but review papers, introductions, critiques, commentaries, thematic series and special sessions. Each of these has guidelines in relation to length, but we don't expect you to follow these blindly. In particular, we assume that each paper will include both a computational focus and a cognitive focus, but where many different disciplines are involved, we are happy to provide the space to ensure that the work is accessible to each of these audiences and reflects its impact in each of the relevant fields. We can facilitate this with boxed text that sits outside the main text and provides additional resources that will help your audiences get the most out of your paper.

In some cases you may want to contact us, or we may contact you, to expand your contribution into a *treatment*. A *treatment* will be larger with boxed contributions containing both introductions and critiques or commentaries from the perspective of the different disciplines. The target paper, like some books, will effectively have chapters in it by other authors, and is citeable as a whole with the name of the target authors, although individual boxed contributions will also be citeable.

In some cases we agree simply that an introductory or a tutorial paper is appropriate, rather than overloading a single paper. The two papers by Jim Peterson in our first volume (Peterson, J.K. (2015a). *Computation in networks*. *Computational Cognitive Science*, 1. in press; Peterson, J.K. (2015b). *Nodal computation approximations in asynchronous cognitive models*. *Computational Cognitive Science*, 1. in press) are in this category, putting a more mathematical spin on what we mean by 'computational' - the original paper on nodal computation in asynchronous cognitive models seemed to need a separate introduction to the mathematical basis of computation in networks.

Most commonly, a paper and its authors will sit primarily in one discipline, using a particular kind of methodology in a particular application. However, the paper should explicitly connect with work carried out in other disciplines, and in particular compare the results of a model with the results of empirical research on human or animal subjects.

As we launch, the paper by Lindh-Knuutila and Honkela (2015) is in this category, employing standard topic modelling techniques (ICA and LDA) with clustering

methods represented notably by the biologically inspired Self-Organizing Maps (SOMs). This connects with both Cognitive Linguistics and Psycholinguistics and the paper addresses the challenge that our current theories of grammatical and semantic categories are not well motivated, particularly across languages and cultures. On the other hand these semantic intuitions, and the categories, models and theories from Linguistics, have become the basis for the grammars and taxonomies of Artificial Intelligence, and so the paper compares hand crafted theories and resources with those that are generated by different computational models with their contrasting assumptions and biases.

The launch paper by Komosinski and Kups (Komosinski, M., & Kups, A. (2015). Time-order error and scalar variance in a computational model of human timing: simulations and predictions. *Computational Cognitive Science*, 1, in press) by contrast introduces a new computational model, the Clock-Counter Timing Network, to explore the human sense of time and in particular issues of temporal discrimination and pacemaking. In this work, the fit of the model with empirical data is tested as the influence of its parameters is explored. It represents one step along a path that will include lower level modeling and potential emergence of properties that are explicitly specified in the present model, and also poses the ongoing challenge of quantifying and explaining the values of its many parameters.

Conferences, workshops, projects and themes

We are additionally inviting the proposal of Special Themes and Monographs. Our initial Special Themes will be on *Embodied Conversational Agents*, *Small-brained Animal Neural Models*, and *Computational Models of Language Development*. As we publish continuously rather than in issues, we end up with a Thematic Series rather than a single Special Issue. However, we do propose a timeline that allows us to issue an initial set of papers collectively, but we also have the facility to link in later submissions, including comments, commentaries and critiques. We are also keen to see to special sessions or subseries at conferences becoming Special Themes, and will also consider turning collected editions of Special Theme papers into *Cognitive Science and Technology* books.

Computational Cognitive Science thus invites conference and workshop organizers to contact us about expanded versions of papers from conferences and workshops, and particularly welcomes people to propose special sessions at appropriate venues in parallel with setting up a Special Theme within *CCS*. The Special Theme on *Computational Modeling of Language Development* is a case in point, cooperating with the Conference on Computational Natural Language Learning and the Workshop on

Cognitive Aspects of Computational Language Learning. Note that the focus must be in the intersection of computational models and biological cognition.

Another model is where a large community or inter-mural project is looking to publish its output and maintain the momentum with an ongoing Thematic Series. The Thematic Series on *Embodied Conversational Agents* arose in this way. Similarly societies and centres may like to adopt *CCS* as a primary medium or an official organ. The initial sponsor of *Computational Cognitive Science* as a free Open Access journal is the Centre for Knowledge and Interaction Technologies at Flinders University.

Critique, referencing and quotation

Computational Cognitive Science has an interdisciplinary, transdisciplinary and multidisciplinary nature that means we draw many threads together from diverse places, and that many of our audience will be encountering some of these for the first time. For this reason we use the Harvard Author (Date) approach to citation in order to provide a resource that helps researcher learn their way around the disciplines, and the remainder of this editorial aims to provide an understanding of the conventions that *CCS* is adopting as we draw together many disciplines as well as a skirt some areas where the are radically different views and methodological controversies.

In referencing research, it is important to cite the earliest, seminal, ideas that lead to the approach, the earliest seeds that provide the deepest roots. It is also important to cite the pivotal points where major advances were reported, along with the latest, culminal, work where the newest work is reported and we have the best view of the state of the field. All authors should be named (or at least the first three) for individual works, in the order listed in the papers. For groups of works where relatively large numbers of papers are listed, it is sufficient to indicate the group of authors clearly, along with the relevant dates, although the precise membership and order may vary between papers.

The face of the researchers that your readers see is their formal citation details, unless you flesh it out with actual quotations – and if you are critiquing someone's views or claims, it is important to quote their actual words rather than argue against an unsubstantiated strawman.

It is important to dig back to the original papers to understand how they were thinking, where the insights came from, what the assumptions were. It is fatal to get into 'Chinese Whispers' where we see a 27th hand version of a standard theory or mathematical formalism that has lost all its humanity - we no longer understand where it came from, or what motivated the developer, or what assumptions were explicitly or implicitly made, but

neglected or forgotten over time. Some people use models and formalisms in isolation, without explaining the insights and intuitions, and without spelling out the axioms and assumptions that underlie them. Formalisms are intended to provide a clear and unambiguous language that is easy to understand, but it is important to clearly explain the model and the notation, and identify the assumptions and insights that lie at their origin – with appropriate citations back to the seminal work.

Almost all models, every abstraction, is based on decisions about which details can be modeled, which details are important, which details can be neglected, which details can be thrown away. This means models are in general based on assumptions that are known to be false, but are sufficiently closely approximated individually or in common combinations as to work well. The models of particular interest to *CCS* are those that are based on assumptions that come from the biology. They may not be true as models of the physical world, but if they seem to represent the same assumptions that our neural substrate effectively makes, biologically and behaviourally, as they enable an organism to understand its physical world, then these are appropriate assumptions for us to explore.

Evaluation and statistics

A major issue in both Computational Science and Cognitive Science is evaluation. Statistics was largely developed to provide a way of dealing quantitatively with the kind of qualitative data we traditionally have in the Behavioural Sciences. However, different disciplines have different traditions and conventions when it comes to designing, evaluating and writing up work. This issue is one that is of great concern to many journal editors, and has become a significant area of research for me personally, and so the rest of this editorial will shamelessly explore this area. I invite people interested in these problems to get in touch, as I am considering setting up a Thematic Series on *Evaluation and Visualization* issues that will focus on the issues of comparing results from computational models and human/animal data.

In the age of Big Data, with teams of people around the whole planet tackling very similar problems, or even exactly the same problem using exactly the same data set, the traditional ideas of significance are becoming a bit dated and misleading. Not that the area has ever been anything but controversial. We would encourage researchers to think in terms of the ‘New Statistics’ that encourages people to present data in a way that permits meta-analysis - that is to combine data from the published work of multiple groups to get a bigger more accurate and more diverse sample. In particular, we want to see effect sizes with standard deviations, and standard errors, and when significance is reported numerically a precise p -

value should be given rather than an α -band. Furthermore, the sample size should always be specified.

We recommended showing standard deviation and standard error in tables and plots.

The reason for this is that if many groups do similar experiments, and each use an $\alpha = 0.05$ threshold of significance, as soon as we have a few groups or a dozen tests, we have a good chance of seeing a ‘significant’ result by chance. If people have test beds where they explore many algorithms against many datasets, then we expect one in twenty to be ‘significantly better’ just by chance, unless they correct for the multiple testing or graphically align *all* the relevant results (dropping the ‘bad’ ones is also a form of biasing the selection just as much as showing only the ‘good’ results).

Unfortunately errorbars and \pm notations are not only used to show standard errors, but are also employed to display of larger confidence intervals that represent the confidence that a mean lies in the indicated range, and is set as some multiple of the standard error according to some specific, but often implicit, model of significance. Even worse is using the whisker notation to display standard deviations.

Some basic principles...

Start with the standards that are appropriate to the disciplines you have been trained in – we expect that every set of authors will include some computational or mathematical training and some training in at least one cognitive discipline. Show effect sizes not just p -values, and show actual p -values not broad α -bands - p -values do not tell us how strong the effect is, just how likely it is to be a chance result, and α should be determined *a priori*. For a properly performed experiment using a reasonable model with plausible assumptions, failure to achieve an effect that is of significance is still of interest and worth reporting, whether this is due to the effect being of low magnitude or formally failing a statistical test.

Where data is paired (within subject, repeated measures), show differences and treat that as your effect, and calculate and present means, standard deviations and standard errors on them.

In terms of the underlying measure, pay particular attention as to whether it is meaningful, and whether you should be evaluating using chance-correct techniques. Generally single class measures like Recall, Precision and F-measure are inappropriate for *CCS* (Powers 2008a, b), uncorrected Accuracy is inappropriate for imbalanced or variable prevalence data, and various forms of Kappa and DeltaP or Youden J statistic have been proposed for the dichotomous case (Powers, 2012). It is important to consider both Sensitivity and Specificity, and it turns out that balanced optimization of the pair optimizes Informedness, the probability of an informed decision. Informedness is useful in the multiclass case, being a generalization

of Youden's J and a form of Kappa. These statistics are appropriate for a single direction of prediction, with Accuracy assuming all *instances* are weighted or costed equally, while Informedness assumes all *classes* are weighted or costed equally.

If there is no preferred direction of prediction, then Correlation is an appropriate statistic - this is the geometric mean of the two directions of Informedness, and is also closely related to common tests for Significance (Powers 2008b). If groups or clusterings are being compared, and there is not even a known or fixed number of classes, then there are a large number of clustering comparison techniques to choose from (Pfitzner et al. 2009).

While tables of results are useful, it is usually easier to understand results when the best results and the significant results are highlighted, and the results are supplemented with a visualization.

Graphs and plots

For discrete conditions or events histograms are appropriate, possibly grouped where that makes sense. I like to show (Powers 2013a, b) the standard deviation around the mean with different coloured parts of the bar, as well as showing the standard error in the mean with fine whiskers or error bars. In these graphs I also include an underlay showing the baseline case (e.g. a random or naïve approach) and the treeline case (the best achievable with some standard approach), also showing colour-coded standard deviations.

For parametric visualization of effects where distributions approximate normal, it is important to always show mean (or another a measure of central tendency), standard deviation (how much variation is natural -68 % probability of a value being in the Standard Range) and standard error (how accurately you know your mean -68 % probability that the true mean lies within the Error Range). Cumming and Finch (2005) report that most researchers use and interpret confidence intervals incorrectly and give a series of Rules of Eye for interpreting error bars and confidence intervals (based on a two-sided test with $\alpha = 0.05$ and group size $N \geq 10$).

For non-parametric visualization where the data don't seem to be normally distributed, we suggest the box plot (Tukey, 1977) which shows median, quartiles (50 % Inter Quartile Range around the median), and generally whiskers marking the 2 and 98 % and/or the 9 and 91 % percentiles: the 7-mark model having edges and whiskers that appear equally spaced if normal, with points outside the whisker range being shown as outliers. Means are sometimes shown as diamonds, and notches in box plots function similarly to error bars in normal plots and non-overlap is intended to indicate significant difference. However, the appropriate (notch or errorbar) widths in both cases depend on distributional and experimental assumptions as above. The breadth of the box is sometimes set to the root group size,

\sqrt{N} , and a variety of simple and complex box plot variants are discussed by Benjamini (1998).

Other forms of plot may be used in special cases, including scatter plots (perhaps overlaid with classes or clusters), and ROC or LIFT or Cost charts (Hernández-Orallo et al. 2013; Powers 2015), but other connected line graphs can also be used if the data can meaningfully be interpolated. However, connected lines or curves are not appropriate for nominal data or discrete data in general.

For biomedical and multimodal data some kind of image will be appropriate, but the same principles apply. For example for brain images, colour coding can usefully be used to show the effect size (power or voltage or current density with appropriate units), but *p*-values should *not* be plotted, although contours corresponding to an α -threshold (e.g. $p = 0.05$) may usefully be plotted.

The future

Welcome one again to *Computational Cognitive Science*. We look forward to an exciting future and trust the *Computational Cognitive Science*.

Received: 10 May 2015 Accepted: 22 May 2015

Published online: 07 July 2015

References

- Benjamini, Y. (1998). Opening the box of a boxplot. *The American Statistician*, 42(4), 257–262.
- Cummin, G., & Finch, S. (2005). Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist*, 60(2), 170–180. doi:10.1037/0003-066X.60.2.170.
- Hernández-Orallo, J., Flach, P., & Ferri, C. (2013). ROC curves in cost space. *Machine Learning*, 93, 71–91. doi:10.1007/s10994-013-5328-9.
- Lindh-Knuutila, T., & Honkela, T. (2015). Exploratory analysis of semantic categories: comparing data-driven and human similarity judgments. *Computational Cognitive Science*, 1(1). doi:10.1186/s40469-015-0001-1.
- Pfitzner, D., Leibbrandt, R.E., & Powers, D.M.W. (2009). Characterization and evaluation of similarity measures for pairs of clusterings. *Knowledge and Information Systems*, 19(3), 361–394.
- Powers, D.M.W. (2008a). Evaluation evaluation. The 18th European Conference on artificial intelligence, pp843–844. Publisher's 2 page version as published - doi:10.3233/978-1-58603-891-5-843.
- Powers, D.M.W. (2008b). Evaluation Evaluation. The 18th European Conference on Artificial Intelligence, pp843–844. Author's 5 page version as submitted - doi:10.13140/RG.2.1.1165.5207.
- Powers D.M.W., (2012). *The problem with kappa*, *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 345–355).
- Powers, D.M.W. (2013). *AdaBook & MultiBook: Adaptive Boosting with Chance Correction*, *10th International Conference on Informatics in Control, Automation and Robotics (ICINCO)* (pp. 349–359).
- Powers, D.M.W. (2013b). *A computationally and cognitively plausible model of supervised & unsupervised learning* (pp. 145–156). Beijing: Brain Inspired Cognitive Systems.
- Powers, D.M.W. (2015). *Visualization of tradeoff in evaluation from Precision-Recall & PN to LIFT, ROC & BIRD*, *Technical Report KIT-14-002*. Adelaide: Computer Science, Engineering & Mathematics, Flinders University
- Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.