

Letter to the Editor

Re Sparck Jones Re Ballard on the Need for Careful Description

I read with interest the recent response by Karen Sparck Jones (1983) to my suggestions (1983) on improving the extent to which published accounts of NL prototypes convey an accurate impression of the true capabilities of the systems they describe. She begins by confirming that I am "right to call for better standards", and her subsequent remarks indicate that we are in principle in agreement. I completely agree with her that "raising the standard of reporting raises the standard not only of the reader's work but that of the writer's", so any differences in perception lie in methods rather than goals. Accordingly, Sparck Jones proceeds to voice concerns that

- ▶ what I ask for might prove excessively voluminous;
- ▶ I might have considered papers that come close to what I'm asking for; and
- ▶ I ought to "show the way" by providing the type of report I call for.

Before responding to each of these concerns, let me remind the reader that my proposal called for providing

- (1) overview,
- (2) sample session,
- (3) system overview,
- (4) example trace,
- (5) features,
- (6) theory,
- (7) detailed description,
- (8) readiness, and
- (9) prospects.

Concerning the potential length of reports adhering to my proposal, I'm in partial agreement with Sparck Jones: an accurate and responsible account of a complex piece of AI software may take some space, more than is available in a conference paper, and much more than one can provide in an "extended abstract". But I do believe a standard journal-length paper can provide most of what I have asked for. In fact, much of what I call for (for example, topics 1, 2, 8, and 9) can be accommodated by a simple reorganization of material or by adding just a few paragraphs. Beyond this, I have clearly recognized that "it may be useful for some papers to gloss over a particular topic or make it the subject of a separate report". As to Sparck Jones' inference that I would expect a complete grammar and dictionary, I suggest that a well-designed theory, followed by a well-thought-out implementa-

tion, will lead to a system whose capabilities can be quite well conveyed by carefully chosen but informal statements, especially when complemented with representative examples. To my mind, it's the sloppy, uncertain, ad hoc theories and systems that can only be truly understood through exhaustive detail. We must endeavor, where possible, to provide capabilities whose interactions are uniformly provided for, rather than dealing independently with specific and perhaps arbitrary combinations of features. This will lead to systems that are easier to understand, to experiment with, and to learn from.

Concerning Sparck Jones's suggestion that I consider some papers that come close to what I'm asking for, I admit that I have my own favorite authors and papers, but I suspect my case would in fact be weakened by my singling out authors for what might be construed as political reasons. Instead I've settled for an idealized benchmark which to varying degrees we all fall short of. Incidentally, I also have my favorite examples of bad papers that, willfully or not, lure the reader into grossly exaggerated impressions of what a theory or system has achieved. This makes it all the more difficult for subsequent researchers to build honestly on previous work and convince the research community of their having gone further than earlier workers, as observed in McDermott (1976).

Concerning the request that I "take some system and provide the kind of account of it [I] am looking for", I reiterate my basic agreement with Sparck Jones that the problem of accurately communicating on paper the capability of a complex piece of AI software, assuming an accurate account is indeed our intent, is a difficult matter. Though I'm out of reprints, I am willing to direct the interested reader to our latest paper (Ballard, Lusth, and Tinkham 1984) for an idea of what I think we can do in the short term. Since this paper appears in a "professional" as opposed to a fully "technical" journal (such as this one), some of what I've called for is intentionally omitted, for example, detailed descriptions of processing methods (topics 5 and 7). Other information (topics 4 and 8) is given in simplified form due to the limitations of the version of the system being written up. The humbling experience of having written a fairly extensive paper that compromises on what with the same pen I proposed to readers of this journal further convinces me of the need for intensified attention to some of the problems I raised.

In conclusion, I suspect that, were Karen and I to sit down to discuss these matters, as I hope we soon shall, there would be more concensus than difference of opinion. To the extent that the dissemination of her thoughts and mine might help to improve the clarity and credibility of published reports of NL systems, and thus contribute indirectly to more productive research, I am grateful.

Bruce W. Ballard
Department of Computer Science
Duke University
Durham, NC 27706

References

- Ballard, Bruce 1983 On the Need for Careful Description of NL Prototypes. *American Journal of Computational Linguistics* 9(1): 23-24.
- Ballard, Bruce; Lusth, John; and Tinkham, Nancy 1984 A Transportable, Knowledge-Based Natural Language Processor for Office Environments. *ACM Trans. on Office Information Systems* 2(1): 1-25.
- Sparck Jones, Karen 1983 Re Ballard on the Need for Careful Description. *American Journal of Computational Linguistics* 9(2): 92.
- McDermott, Drew 1976 Artificial Intelligence Meets Natural Stupidity. *SIGART Newsletter*, No. 57.