

PREFACE

Hans Karlgren
coling@com.qz.se

The present publication contains a selection of new papers in computational linguistics. We have reason to believe that it gives a good view on recent research at its best in our domain with a good spread over subfields and geographical areas. Whether the reader is a participant of this year's Coling, one of those only too many who wanted to participate but for one reason or another could not do so or a member of the important category of those who so far did not consider computational linguistics a concern of theirs but who might in the future be originators or recipients of essential contributions, we hope he will find the reading inspiring.

This is a selection, not a full coverage compilation of everything worth reading. There is much more going on out there. What you see here is about a third of the set of papers which were offered to the 13th Coling conference. Having seen them all I can testify that the vast majority is good work on interesting issues, but practical arrangements set limits to what we could accommodate. A larger multi-channel performance might also jeopardize the intimate atmosphere which has proven so fruitful at Coling conferences.

The papers fall into two categories. Already in the the Call for Papers and in a number of subsequent communications we tried to establish a clear distinction between a Topical Paper on some crucial issue in computational linguistics and a Project Note. This is not a quality gradation. Their merits are almost complementary:

- A topical paper should focus on some new idea or disconfirm an old truth. It should have a minimum of narrative and self-biographical portions and should not attempt to bring the detailed account which rightly belongs to project documentation. It should promote the multilateral communication which is the scientific reason for meeting. It should contain some controversial element.

- Conversely, a project note should be very concrete; this requirement was emphasized by the insistence on a software demonstration accompanying its presentation.

The procedure for selecting the papers to be put on the schedule was the following.

For each paper three referees were chosen. The choice was made by me with the help of the advisors, in such a manner that the paper was read by referees who were known to be actively interested in the subject - i.e., those which would sympathize with the aim but would also be particularly competent to recognize flaws or duplication. This arrangement is intended to make the referees very actively engaged, and did have that effect.

For each topical paper each referee was asked to answer four questions:

1. Would it be a loss for Coling not to accept it as a topical paper?
2. Would it be a loss for Coling to accept it as a topical paper?
3. Would you be prepared to chair a 30 min session to discuss it?
4. Would it be appropriate to suggest that the author transforms his offer into one of a project note with demonstration?

"The third question is there to verify that there is some inflammable substance in the paper", the referees were told. "If no reviewer answers yes to this question, the paper will be rejected whatever its merits are otherwise claimed to be." Every paper therefore has a chairman, who wants to see the paper debated; that does not mean that he subscribes to its thesis! In some cases this is a virtual chairman because the referee found it impossible to go to Helsinki in August 1990.

For Project Notes the questions were essentially the same, except that Question 3 explored whether the note reported something worth looking at and that, obviously, Question 4 was modified.

The algorithm to select among papers with some Yes for question 3 was:

- with some Yes.No.*.* and no No.Yes.*.* the paper was accepted
- with some No.Yes.*.* and no Yes.No.*.* the paper was not accepted or moved over to the other category.

- with some Yes to Question 1 and some Yes to Question 2 additional reviewers were asked.
- with all answer sequences of the form No.No.*.* other considerations decided.

In the last two cases - with incompatible answers and with indecisive answers - the dialogue with the referees and with the advisors in many cases continued and other expertise was consulted. In quite a few cases there was a prolonged discussion over e-mail and telephone.

By Other considerations we mean that the concern for the balance of the conference influenced the final decision - the balance between types of approaches and kinds of questions. We had to say no to quite a few good papers to avoid what would have seemed to be duplication; that does not mean these were necessarily inferior in quality.

A hard task is to establish the scope of a conference of this kind. We have seen a number of very qualified papers which, after all, were judged to disqualify as essential contributions to computational linguistics. Thus, some good papers on formal logics and semantics were considered to be too uncomputational for Coling; it is a comforting thought that there is no lack of meetings dedicated to that kind of research. Some interesting reports on software design and skilfully programmed systems were also left out because their focus were considered to be too far off from linguistics. The demarkation is always disputable; cf. my own attempts at a formulation in my comment on Computational Linguistics in 1990 in this volume.

The large number of good papers not only made the work of the Programme Committee more time-consuming and demanding (as well as intellectually stimulating!) than expected. It also made all involved feel heavy at heart many times when we had to say No thanks to colleagues who have been engaged in serious research in our field and made great efforts to present it to us. I have tried to convey to everybody the message that when we say "No thanks" the 'thanks' is seriously meant. We appreciate your offer whether we can make use of it or not. I have also tried to give all writers some feedback. Some reviewers have written long and penetrating critical essays, others have made their point by mere yes's and no's and ranking, but I have tried to verbalise their reaction, or mine, whenever possible.

Large-scale use of e-mail in this connection helped to gather and relay information; unfortunately, I trusted this medium (as well as other

media and routines employed) too much so some colleagues did not get the intended messages appropriately in time. No referee or advisor can be blamed for whatever mistake was made in this extensive correspondence.

While we are confident that we have succeeded in the primary task to make a worth-while show in Helsinki and an enjoyable publication for the readers, we ask all those who feel disappointed not to overinterpret the individual decisions. We cannot hope to have made hundreds of decisions in a brief time all perfectly appropriate. It appears from what has been said about the procedure that it contains subjective and even arbitrary elements. It should be clear that the advisors have greatly improved the quality of the operation but that none of them can be held responsible for any mistake.

The papers in these volumes are classified - by their authors - into Topical Papers and Project notes. They are too rich in perspectives to make a disjunctive thematic classification meaningful; see, however, the remarks on trends and the distribution over topics in "Computational Linguistics in 1990" and the Subject Index. The papers are ordered so that the participants of the conference should typically need to carry with them each day only one volume.

Hans Karlgren

Programme Committee

ADVISORS

Eva Hajicová
Aravind K Joshi
Jürgen Kunze
Makoto Nagao
Don Walker

and all members of the International Committee for Computational Linguistics. Particular acknowledgment is due to Revekka Frumkina, who visiting Stockholm, helped in reviewing and processing many papers and reviews.

