



**KTH Computer Science
and Communication**

Kerstin Frenckner, tel 08-790 9754, e-mail: kfrenck@csc.kth.se2

February 12, 2009

Copyright CSC, KTH

OPPOSITION FOR MASTER'S PROJECT

The duties of an opponent are to:

- Critically review the report in question
- Pay particular attention to the problem approach, the methodology chosen and to the interpretation/evaluation of results
- Make annotations on the report of clerical errors, other minor errors, incomprehensible or ambiguous text
- Complete this Opponent Record (use a computer or black ink)
- In advance – at the time stipulated – give this record to the persons stipulated in the instructions for your exjobb subject.
- Orally present your general opinion of and comments on the work during about 5 minutes after the author's presentation of the work
- Put questions to the author of the report following his/her presentation: you may put forward the questions set down in the Opponent Record, or some of these questions, but it is also reasonable to expect the presentation to generate new questions.
- Give the Opponent Record and the annotated report to the author at the conclusion of the seminar

You may contact the person responsible for the degree project, e.g. to test programs.

The Opponent Record can be completed either using a computer or manually. If writing by hand, use red or black ink and write distinctly. The Record copies must be legible but not necessarily aesthetically pleasing.

Master's projects vary considerably. Consequently, at times not all of the questions will be relevant to the project you are opposing. It can be appropriate to rephrase the questions to fit the project. You may also introduce one or two additional questions.

Attempt to answer the questions in the Opponent Record in relative detail. Answers such as **Yes** and **Good** are insufficient.

OPPONENT RECORD

Thesis compiled by:

Kristoffer Hallqvist and Dmitrij Lioubartsev

Title of thesis:

Data-driven learning of the meaning of route descriptions

Opponent:

Max Malmgren

Was it easy to understand the underlying purpose of the project? Comments.

Yes, it was clearly explained in section 3 (Procedure). It is not concisely stated how it relates to the broader area of research, but understanding emerges when reading section 3 with the context of section 2 (Background / introduction).

Do you consider that the report title justly reflects the contents of the report?

I feel that it does indeed do so.

How did the author describe the project background? Was there an introduction and general survey of this area?

There was evidence of thorough background research, given by the noting of several reports in the same area of research. They described the project as having notable previous research, even one report that seemed to touch upon much the same issues, but more thoroughly. The introduction was understandable.

To what degree did the author justify his/her choice of method of tackling the problem?

They did not directly justify the method.

Did the author discuss the extent to which the prerequisites for the application of such a method are fulfilled?

The authors briefly notes that for the chosen algorithm, it is important to prevent ambiguous words for the same concepts, since that would confuse the program. They also state, and I quote, "In a real world environment, the tf-idf technique is probably not optimal".

Is the method adequately described?

The method is described in length in sections 3.3 – 3.6. I found the description easy to understand and read.

Has the author set out his/her results clearly and concisely?

Yes, the data is formatted clearly into tabular representation and are presented in sections with information describing their relevance.

Do you consider the author's conclusions to be credible?

Yes. They do not seem to overreach, and considering the positive nature of their data, it is no surprise that they came to the conclusions they did. They note that the results might be less accurate in more complex situations (using multi-word concepts for example).

What is your opinion of the bibliography? What types of literature are included? Do you feel they are relevant?

One of the references is used as data for the report, and as such is very relevant to the report. The three scientific reports and Introduction to Information Retrieval might have less relevance. Introduction to Information Retrieval is not referenced even once, but perhaps their knowledge of the tf-idf algorithm was acquired from it.

Which sections of the report were difficult to understand?

The section about tf-idf was somewhat difficult to comprehend, specifically how the formula is denoted using subscripts that are not present in the explanation of the applicable term.

Other comments on the report and its structure.

I found the report enjoyable to read, the structure was easy to follow and presented the report in a logical manner.

What are the stronger features of the work/report?

I think the presentation and structure of the report is very good, as well as the thorough explanation of the algorithm designed.

What are the weaker features of the work/report?

Perhaps the news value as noted below, and the use of the cited references.

What is your estimation of the news value of the work?

Very difficult to answer, since I have no in depth knowledge of the research area or similar work. It seems like the report *Learning to Follow Navigational Directions* might precede and even extend further than this report, given that the authors themselves note that '[The authors] work with very similar test data, but their methodology is too complex to take much from in this small project'.

Summarize the work in a few lines.

The authors has experimented with an algorithm to deduce connections between navigational concepts (geographic directions and landmarks) and their representation in a natural language, by analysing the movement pattern of the mouse pointer and transcribed directions synchronised in time. This data is analysed using a type of word / document statistical formula used to recognize words typical for a certain section. The results show that there is potential in the method used.

Questions to author:

1. Was there anything, during the work or research, or about the results that surprised or astounded you?
2. What could using the method you designed in the report mean for real-world applications, or for AI research in general? Future work?
3. How did you arrive at using Term Frequency – Inverse Document Frequency as the statistical formula? Were there any other methods reviewed / any other considerations?