

Report of the ALGODAN Scientific Advisory Board

Johannes Gehrke and Volker Tresp

February 3, 2009

Executive Summary

We report on the progress of the ALGODAN Centre of Excellence (CoE). In summary, we were very impressed by the CoE and its activities. We believe that ALGODAN is among the leading CoEs worldwide, that the research carried out in the CoE is of high quality, and that it already had and very likely will continue to have high impact in the future.

Over the next year, we suggest formulating a more focused goal and vision, e.g., in terms of an application or application area, such that the program, to an even larger degree, becomes more than the sum of its excellent parts. This would improve visibility and facilitate advertising the impact of the center.

Scientific Activity, Production and Progress of the CoE

We believe that the CoE is poised for success. First, the CoE has excellent leadership through its director, Professor Esko Ukkonen, and its vice-director, Academy Professor Heikki Mannila, and its three other team leaders. Second, the research in the CoE encompasses many scientific fields ranging from algorithms, data mining machine learning to probabilistic models. We find it very laudable that these are not isolated teams, but that the teams work together with ad hoc collaborations that span several teams, and we commend the participants on the interdisciplinary nature of the research. Third, the CoE involves many PhD students and postdocs, training the next generation of scientists for Finland.

The level of scientific activity that we encountered during our site visit was exemplary. Our site visit included a poster session where we had the possibility to personally talk with many of the PhD students involved in the research. From what we saw during our review meeting, we believe that the center has a very strong projection into the future.

We very much liked the focus of the Center: Internationally competitive academic excellence.

Development of the CoE profile

The research groups seem to be well-integrated in a way that comes naturally out of overlap in the projects that they work on. We do not think that further overlap needs to be forced; we rather think that this should happen naturally.

Synergies of being a CoE

The CoE plans to address problems from biology, medicine, the telecom industry, environmental studies, linguistics, and neuroscience. Problems in these areas are best solved by interdisciplinary teams, and the researchers in the CoE with complementary skills in algorithms, data mining, machine learning, and statistics are poised to make significant advances during the life of the CoE. The CoE will house experts from different application areas (such as computational ecology and sustainability) and will also provide a home for novel types of interdisciplinary students. Joint seminars between research groups will foster an exchange of language and culture. We also believe that the center is well on its way to provide a “home” for novel types of interdisciplinary students that do not squarely fit into a single area.

We also find the general technical direction of this confluence of people and research areas – the combination of combinatorial and statistical techniques in data analysis – highly innovative and we believe that it has the potential to lead to real breakthroughs.

Institutional Support for the CoE/Availability of Facilities

The institutional support for the center is very strong; it is housed at a joint strategic research institute (HIIT), and both universities provide some money and funding for PhD students. We liked that there is little administrative overhead in running the center, and thus the funds from the Academy of Finland will be mainly used as intended, for world-class research.

International visibility of the CoE

The CoE has no visible international profile yet because it just started, but the SAB was not concerned. We believe that citing center support in publications, and creating a nice website with pointers to the ongoing research, downloadable publications, and other information about the center and its research will increase international visibility.

Researcher career development

We believe that career development of the next generation of scientists is a strength of the center. First, the faculty involved in the center have already established a pipeline of high-quality postdocs who take highly competitive jobs, and thus they have the precedent of graduating very high-quality people. Second, the existing and planned research of both the senior members and junior members of the faculty was impressive.

We like the approach to get postdocs from abroad and then send participants abroad. This strategy is good; PhD students from abroad might be hard to get.

We know that the participants are aware of the gender imbalance which might be a hard problem to address.

International and national co-operations

The PIs of the ALGODAN CoE have very good contacts to industry within Europe and the US. We think the current level of co-operation of the CoE at both the international and national level is sufficient and adequate.

CoE's Impact on Society

The center is just at the beginning, but the CoE's general direction of computational sustainability and computational ecology are very important topics with possibly huge payoffs for society. Interactions on problems related to computational biology and medicine might also have big impact.

Attention for the Next Three Years

The CoE just started, and we believe it is very well on track.

We suggest that in the course of the program ALGODAN should attempt to formulate a focused goal or vision, e.g., in terms of an application or application area, such that the program, to an even larger degree, becomes more than the sum of its excellent parts. One way to phrase our suggestion is: How will the world be different after ALGODAN is completed?

Other Comments

The SAB was impressed by the big challenge problems that define parts of this effort. However, the big challenge problems were presented mainly at the technical level instead of also pointing out their overall benefit to the public. We suggest adding some discussion to put everything into context.

We suggest to the CoE to also consider software usage as a possible indicator.

In summary, we believe that the funds of the Academy are very effectively used, and we commend the center on its progress.