Ref:
Date:
Issue:
Running Page no:

Dissemination and Use Plan

Project number:	IST-1999-14186
Project acronym:	ALCOM-FT
Project title:	Algorithms and Complexity – *Future Technologies*

Authors: Aarhus.
List of Partners: Aarhus, Barcelona, Cologne, INRIA, MPI, Paderborn, CTI, Rome, Utrecht, Warwick.

1 Overview

1.1 Overview of Expected Results

All documents resulting from the ALCOM-FT project will be publicly available through our website at

`www.brics.dk/ALCOM-FT`,

and as such are part of our dissemination effort. In particular, this is the case for the scientific research reports produced continuously during the project. Two further websites with strong emphasis on dissemination will be set up during the project. In total, our deliverables concerned directly with dissemination are the following:

Number	*Name*	*Available*	*D/U*
D1	Research Reports	Cont.	D
D2	Project presentation web site	Sept. 1	D
D11	Algorithm Forum web site	June 2001	D
D14	PR Web Pages for algorithmics	Dec. 2001	D

Besides these, the dissemination effort of the ALCOM-FT project includes several other actions without explicit deliverables. These are described in Sections 1.2 and 2 below.

Regarding use, the following are the deliverables of the ALCOM-FT project for which we expect potential for exploitation:

Number	*Name*	*Available*	*D/U*
D16	Data mining demonstration package	June 2002	U
D20	Description language for high level constraints in optimization	June 2002	U
D23	Library of separation routines for ABACUS	Dec. 2002	U
D24	External memory experimental platform	June. 2003	U
D29	Complete software system for structured combinatorial optimization problems	June 2003	U

1.2 Approach to Dissemination and Use

Dissemination is an integral part of the ALCOM-FT project. The ten participating partners are all active members of the international research community in algorithmics, which by its very nature is based on the publication of results. Academic research reports and their publication by way of internationally renowned conferences and journals therefore form a natural part of the dissemination effort of the project.

However, the ALCOM-FT project is concerned not only with the *development* of new algorithmical knowledge and its reporting in scientific research reports, but also with the *transfer* of this knowledge to application domains. Our strategy for this consists of a number of action lines:

- To increase the awareness of advanced algorithmic methods outside the academic system, we will set up a PR web site for algorithmics, to visualize the methods used and the gains achievable in ways convincing to laypeople.
- To make advanced algorithmic methods more easily accessible to non-experts, we will producing software libraries implementing some of our developed results.
- Education is an important means of spreading knowledge of algorithmics. The partners of the consortium teach at leading institutions in Europe and our past deliverables LEDA, ABACUS, and DSP are used for teaching and academic projects within and outside the consortium. We will strengthen our focus on quality in education by
  - arranging a number of summer schools in advanced, current topics,
  - creating a web site with resources for teaching of algorithms,
  - cooperate in the education of PhD students in algorithmics.
- To increase the awareness within the algorithmic community of the interplay between theoretical results and real-life programs, we will arrange a public programming contest related to a set of concrete algorithmical problems.

A number of software packages will be produced as deliverables of the project. These will be made available to the academic community, to encourage further development and to benefit education.

Some of these software packages are expected to have exploitation potential. These are the ones given in the table in Section 1.1. Various approaches, possibly including spin-off companies and industrial contacts, will be used to test the commercial viability of these deliverables and to bring them to the market, if appropriate.

1.3 Market Projections

By experience from past deliverables, the expected commercial users of software deliverables are mostly large companies which have a research and development unit (e.g., Daimler-Chrysler, France Telecom, Zeiss, Intracom, Sony) or small companies whose founders came into contact with the results at academic institutions.

In this project, we will make an extra effort, through the dissemination actions listed above, to also reach medium and small size companies without prior exposure to advanced algorithmics.

2 Dissemination Plan

2.1 Conferences

During the project, ALCOM-FT partners will be responsible for the arrangement of several of the major conferences within the European algorithmic community. Among these are the following:

- ESA, WAE, and APPROX 2000 arranged by MPI.
- ESA and WAE 2001 arranged by Aarhus.
- ICALP 2001 arranged by CTI.
- ESA and WAE 2002 arranged by Rome.

2.2 Publications

The research reports of the ALCOM-FT project will appear at a wide range of scientific conferences, workshops, and refereed journals, including the most top-ranking within the theoretical computer science community worldwide. Among these are the conferences STOC, FOCS, ESA, SODA, ICALP, STACS, and the journals *Journal of the ACM*, *SIAM Journal on Computing*, *Random structures and Algorithms*, *Distributed Computing*, and *Algoritmica.*

All scientific reports from the ALCOM-FT project will be made publicly available at our project presentation website `www.brics.dk/ALCOM-FT`, and the evolving publication record will be tracked in the annual progress reports.

2.3 Web Presence

Three web sites will be produced as part of the ALCOM-FT project:

- A *project presentation* web site. This has already been started, and can be found at the address `www.brics.dk/ALCOM-FT`. It contains descriptions of the project plan, participants, and deliverables, and an electronic archive of ALCOM-FT scientific reports is currently being added.
- *PR-web-pages for algorithmics.* The pages will describe algorithmic problems and their solutions in a way that it is easily grasped by laypeople, will illustrate the functionality of our algorithms through

animations (JAVA applets), and will give indications of their performance. The main goal is to demonstrate the gains achievable by using modern algorithmic techniques.

- An *Algorithm Forum* collecting texts and tools supporting the dissemination of information related to algorithms. The forum will be web-based and will be administrated by the consortium. It is open to contributions from inside and outside the project; we will use contributions from within the project to get the forum started. The forum will collect information related to teaching algorithms (lecture notes, teaching software, animations of algorithms, classroom projects) and using algorithms (links to systems). We expect that the forum will grow into *the* entry point for information related to algorithms.

2.4 Other

As mentioned in section 1.2, we will organize schools on important topics in algorithmics. We expect to organize two per year. A school on *Foundations of Computer Science* has already been held at MPI, and the next will be in Rome on the topic *Algorithmic Engineering*.

We will also organize an algorithmic competition, to attract the interest of a broader group of algorithm researchers, and to strengthen the focus on actual implementations of theoretical techniques within the community.

3 Use Plan

In the following, we describe the nature and the potential applications of the deliverables (D16, D20, D23, D24, D29) believed to have potential for exploitation.

There are no current plans for Intellectual Property Right actions (patents, trademarks) for these deliverables. The commercial viability will be tested through industrial contacts and through our existing spin-off companies from previous projects.

3.1 Data Mining Demonstration Package (D16)

We focus part of our work on the variants of the Apriori algorithm. This algorithm is routinely used in most, if not all, data mining activities by companies worldwide, and it is already part of all commercial data mining tools.

There are several problems considered as unsolved by these tools. In most of them, the Apriori algorithm is applicable, but it confronts cases in which the running time becomes unaffordable.

We are designing a number of algorithmic ideas for these problems; eventually, they might fit together into a single prototype.

First, we have started from the DIC algorithm, which is a faster generalization of Apriori that has the disadvantage of requiring the user to tune a parameter that is crucial for its efficiency. Then, we have found a new algorithm, Ready&Go, that builds on the same ideas but pursues them further, obtaining two advantages:

- There is no longer the need to tune the extra parameter by the user.
- It runs consistently faster, beating DIC by 20%.

Third, we are performing a novel reengineering of the up to now standard data structures employed in the algorithm, by considering external memory adaptations as promised in the project description. This work is still in progress.

Additionally, we are analyzing carefully the user-dependent parameters that are well-known to be a difficulty for the application of data mining tools, including criteria such as deviation from independence over probabilities and automated adjustment of window width in the problem of finding frequent episodes in sequences. We have a new proposal that is being studied and implemented.

Although we do not think that these advances warrant a spin-off company or a patent application at their current state, we are ready to test them in real cases, and are in contact with an industrial advisor (an European database company which also has data mining activities) to see the direct applicability of our advances.

3.2 SILC (D20, D23, and D29)

The deliverables D20, D23 and D29 comprise a system for structured combinatorial optimization problems (SILC). It consists of a high-level description language and a library of separation routines for the branch-and-cut system ABACUS. We will disseminate our work through three channels:

- publications in relevant conferences and journals,
- public releases of SILC on our web sites,
- information targeted specifically at the users of ABACUS and LEDA, the systems underlying SILC.

Partners Cologne and MPI use integer linear programming extensively in projects outside ALCOM-FT. In the case of MPI, these projects concern problems in bioinformatics and reverse engineering. In the case of Cologne, these projects concern problems in statistical physics, graph drawing and transportation planning. We will use SILC in these projects.

ABACUS and LEDA are in use at many sites world-wide. Our belief is that some of the users of ABACUS and/or LEDA will pick up SILC and use it in projects. If SILC is successful, it can be made available to commercial users through the channels already established for ABACUS and LEDA.

3.3 External Memory Experimental Platform (D24)

A major part of this deliverable will consist of a library for basic services for parallel external memory. This is mainly intended as a prototyping tool. However, we also plan to use it as a vehicle for dissemination. It can be used to demonstrate the usefulness of our algorithmic results like random redundant storage and optimal scheduling algorithms. Our plan is to approach research groups and companies working on storage area networks, network attached storage, RAID-Servers, database runtime systems, parallel file systems, video servers, or scientific computing to find possible cooperations. The advantage of this approach is that integrating our techniques into large existing software systems would obviate the need to rewrite large quantities of software for example to implement a database system. Currently, we have preliminary discussions in this direction with the database group at the University of Saarbrücken, Philips Research Eindhoven, NEC Research in St. Augustin, and a group at the University of Karlsruhe that develops a parallel file system for Compaq.