



Beyond cInQ - How to Go Forwards? Non-technical Remarks on the Application of IDB and Condensed Representations

Workshop on Inductive Databases
and Constraint Based Mining
Freiburg, 11.-13.3.2004

Mika Klemettinen, Nokia Research Center

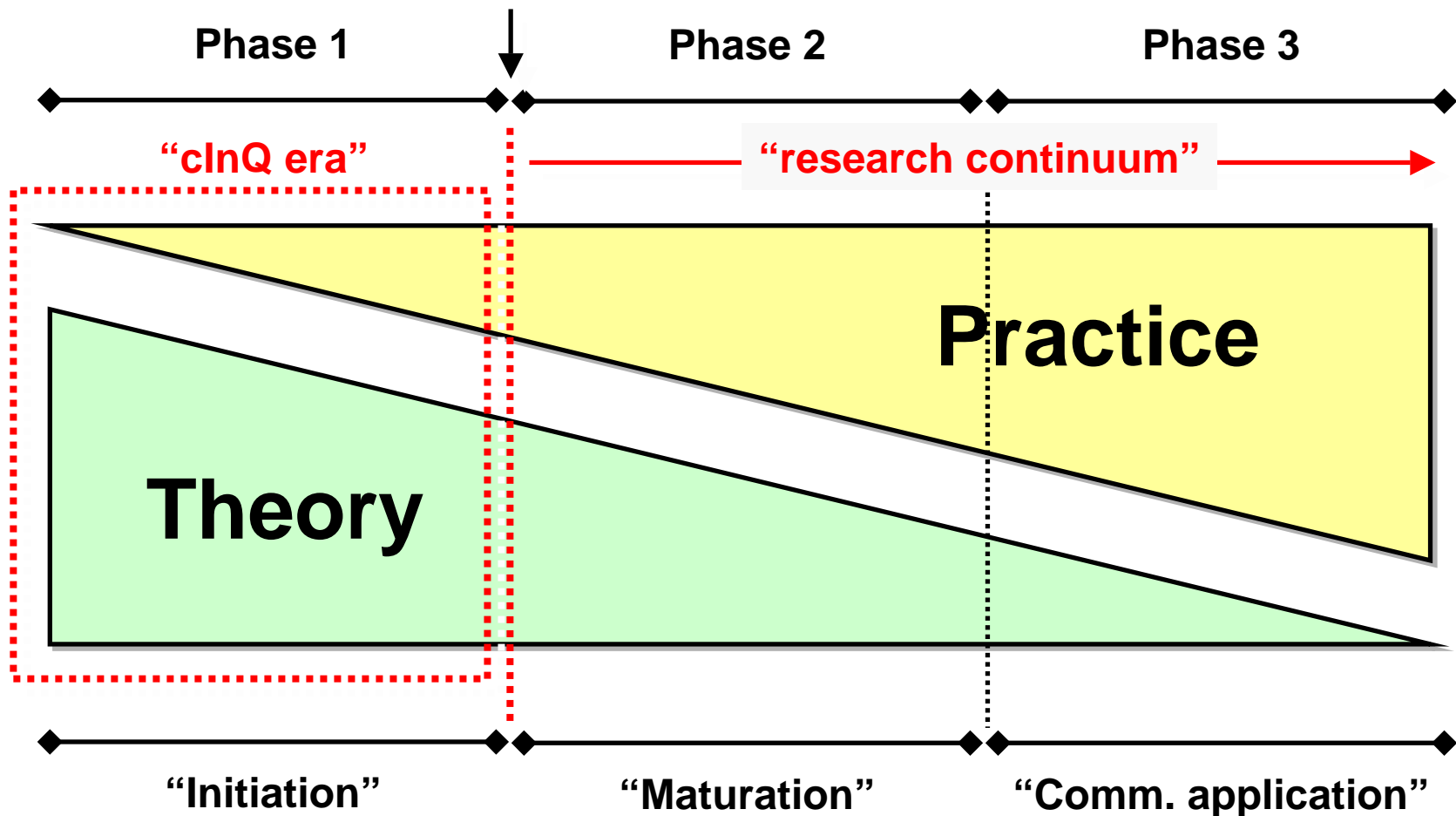


Project funded by the Future and Emerging Technologies arm of
the IST Programme FET-Open scheme



Beyond cInQ - How to Go Forwards?

Non-technical Remarks on the Application of IDB and Condensed Representations





Beyond cInQ - How to Go Forwards?

Non-technical Remarks on the Application of IDB and Condensed Representations

- ◆ **Q1: Does the research answer the needs of the community, industry and/or the markets?**
 - Are the problems interesting from community/industrial research point of view? Are they interesting from other community/industrial needs point of view?
 - ◆ **Q2: Do the new algorithms and research results bring added value w.r.t. the available and existing solutions?**
 - Theoretically nice properties do not necessarily turn into exploitable solutions
 - ◆ How to convince research community and industry that IDB and condensed representations are useful (especially while the area is still evolving), and to ensure successful continuity?
-



Beyond cInQ - How to Go Forwards?

Non-technical Remarks on the Application of IDB and Condensed Representations

- 1. Disseminate the new results to research community and industry - create awareness and encourage new research**
 - ◆ Workshops/events (like this!), papers, "through the grapevine"
 - ◆ Make the technical/theoretical results available!
 - 2. Show how the new approaches can be applied and/or how they perform conceptually better than the "old ones"**
 - ◆ Well-defined real-life scenarios (Cyrille and Bart yesterday)
 - ◆ Examples of how/when to apply (Cyrille yesterday)
 - ◆ Show/demonstrate! (Industry: "Convince us in 2 minutes!")
 - 3. Show how the new approaches perform technically better than the "old ones"**
 - ◆ Benchmarks (Cyrille yesterday)
 - ◆ Show/demonstrate! (Industry: "Convince us in 2 minutes!")
-



Beyond cInQ - How to Go Forwards?

Non-technical Remarks on the Application of IDB and Condensed Representations

- ◆ **Explain the concept of IDB's and CR's in sound and understandable terms:**
 - Give non-theoretical descriptions and definitions
 - Give theory, theoretical results and algorithms
 - Books, papers, ...
 - ◆ **Show applicability of the IDB and CR approach:**
 - Show general and application area specific results (Bioinformatics, Telecommunications, Web-mining, Environmental/biomedical, ...)
 - Show, where the approach is applicable (areas/domains, types of problems, available systems/equipment or resources)
 - Show, where are the benefits, where are the drawbacks or problems
-



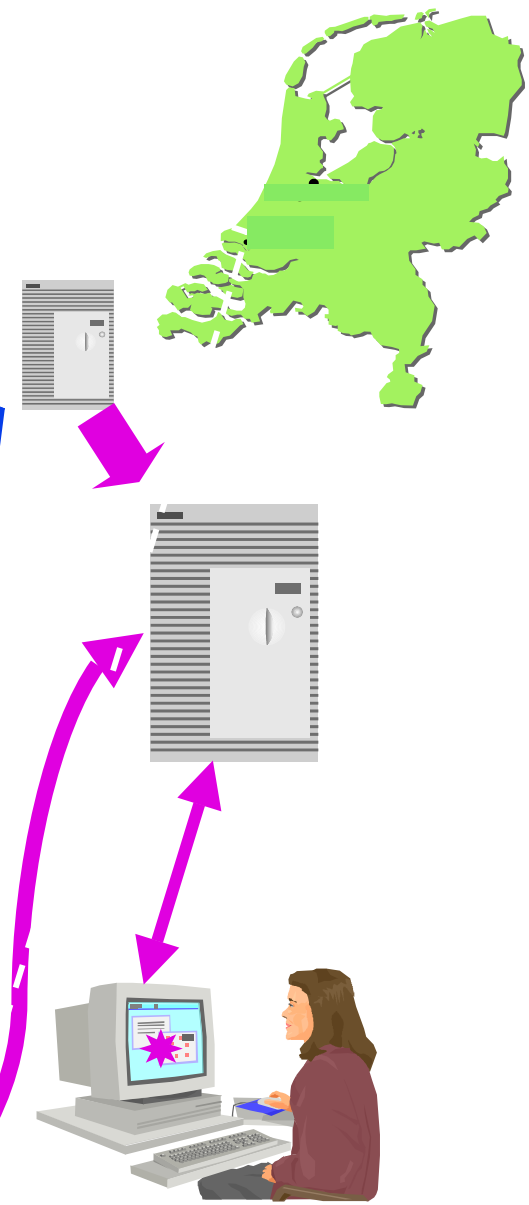
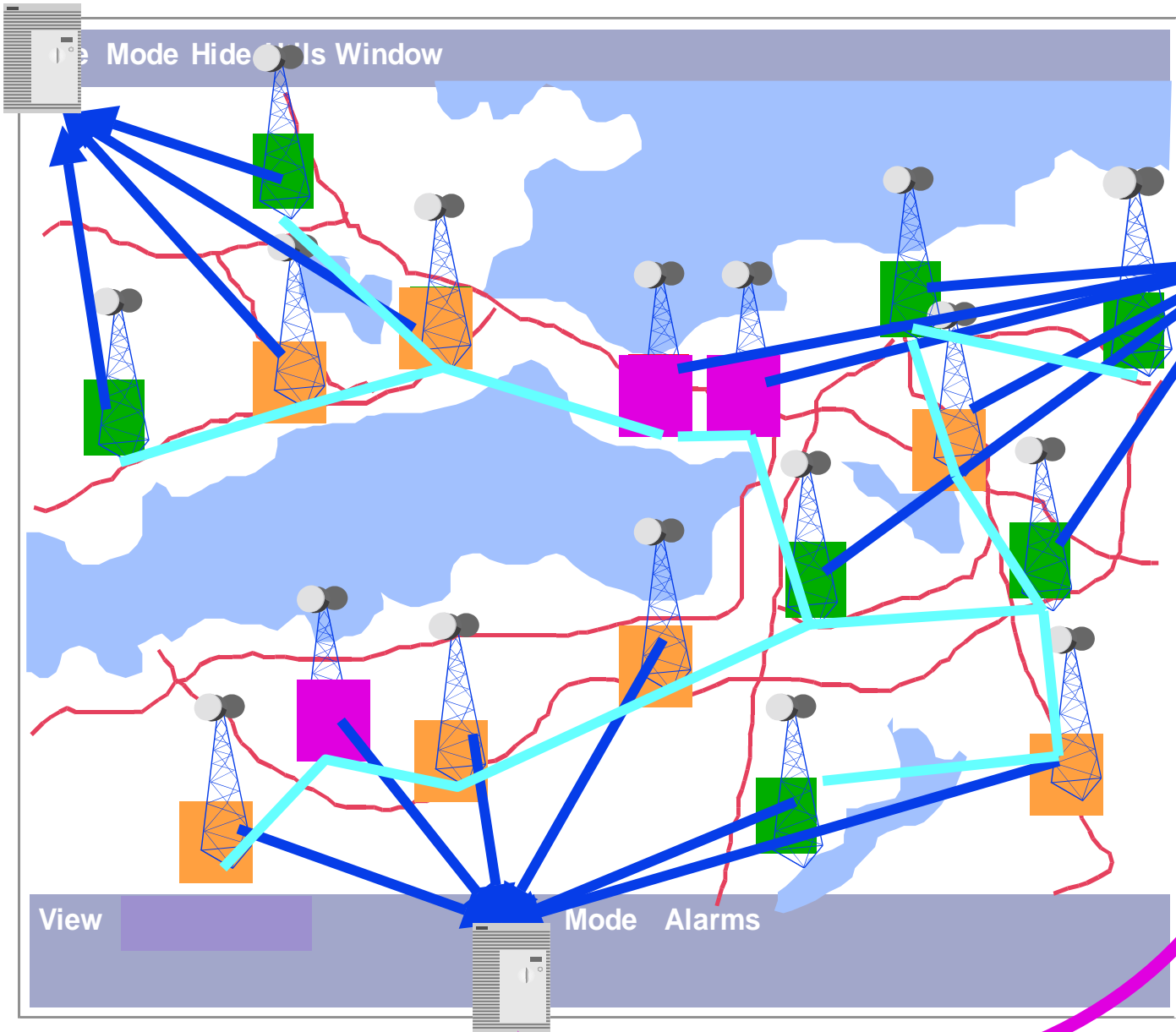
Beyond cInQ - How to Go Forwards?

Non-technical Remarks on the Application of IDB and Condensed Representations

- ◆ **Provide guidelines/help for...**
 - using IDB's
 - evaluating IDB's
 - viewpoints:
 - End-user exploiting/using (or intending to use) an IDB for some specific task
 - Specialist to evaluate the complexity/capabilities of an IDB (performance etc. benchmarks)
 - General instructions and a checklist?
 - Examples/scenarios to help applying the guidelines?

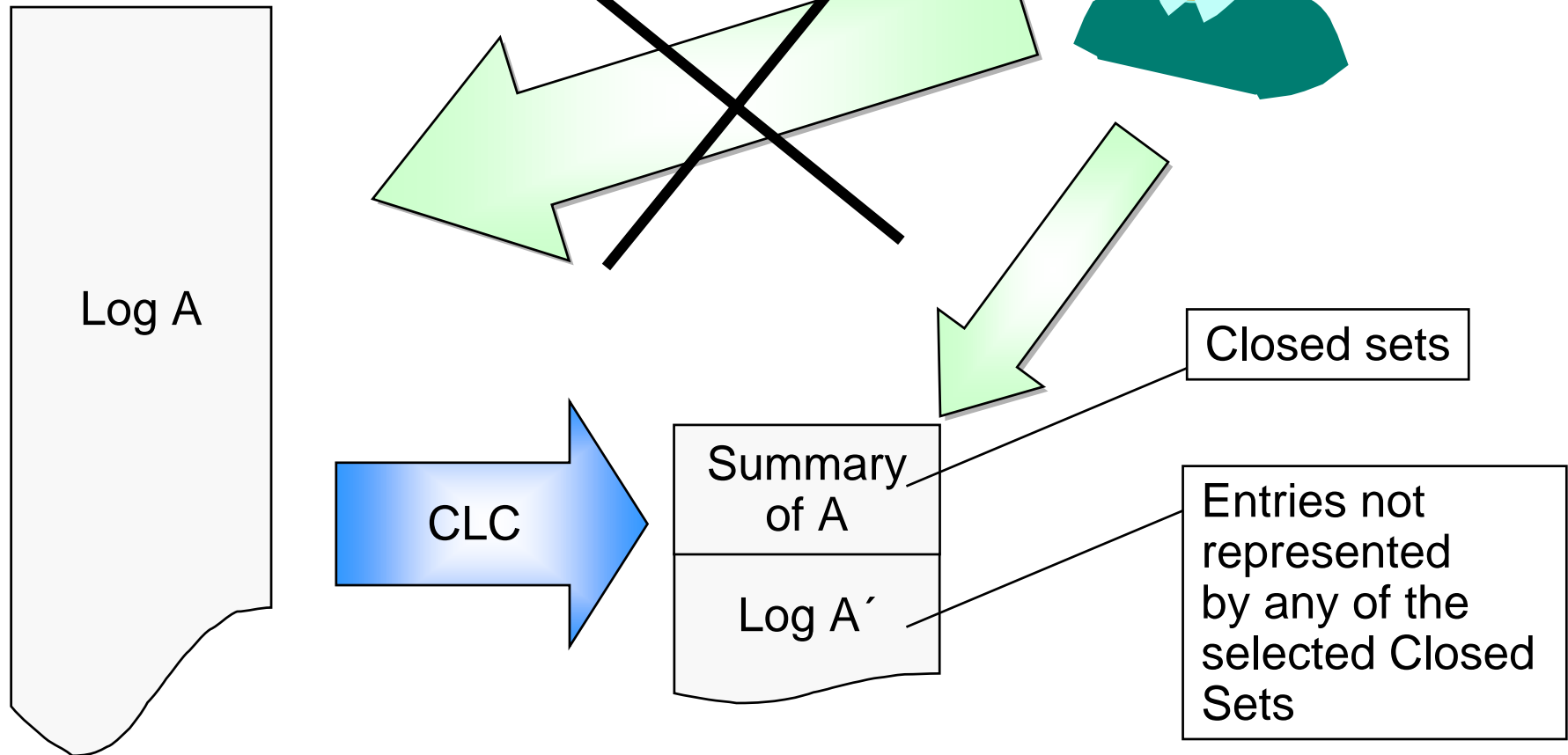
 - ◆ **The above actions might help all IDB and CR related research, and help our way forwards...!**
-

Network



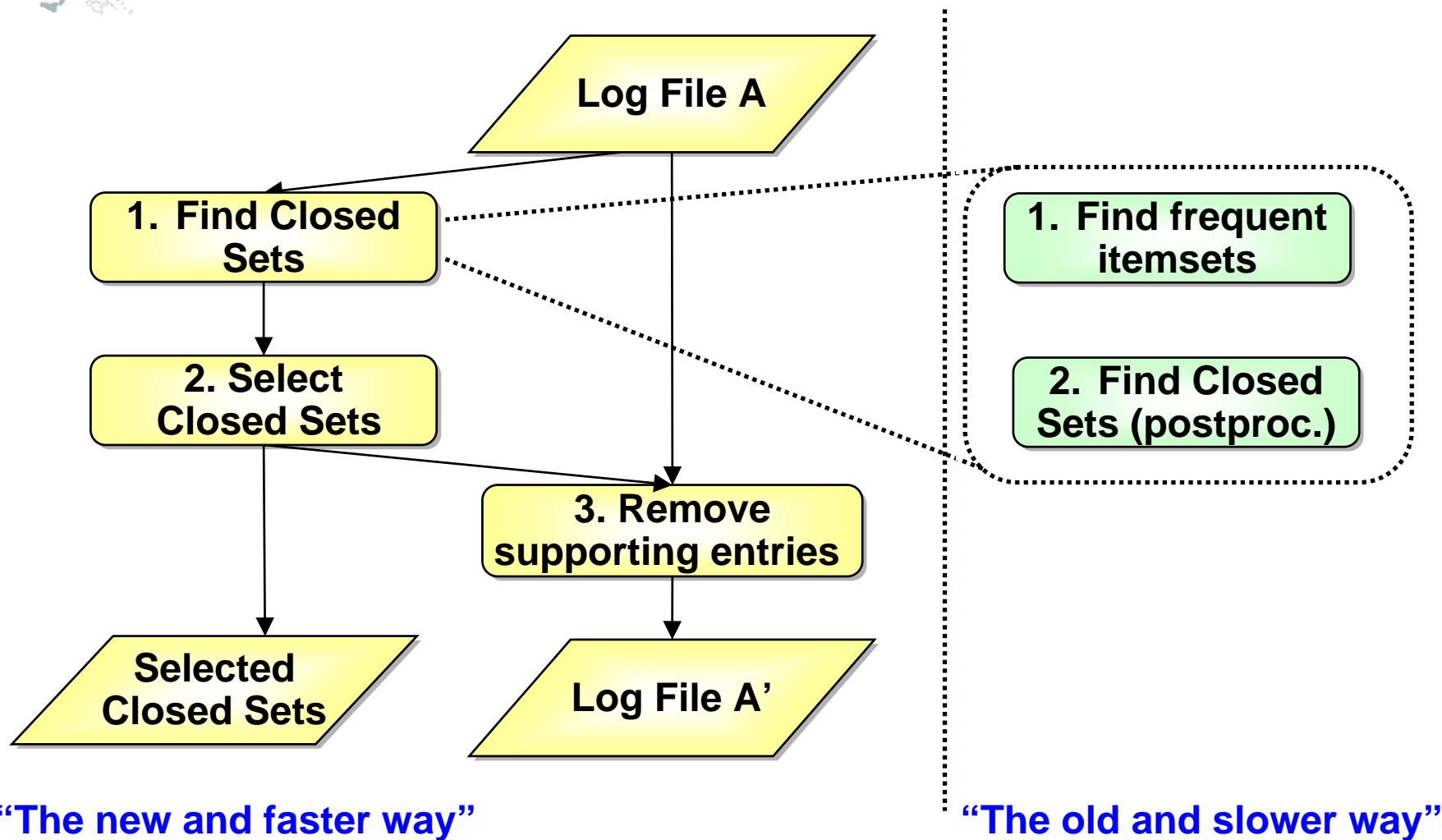


Comprehensive Log Compression



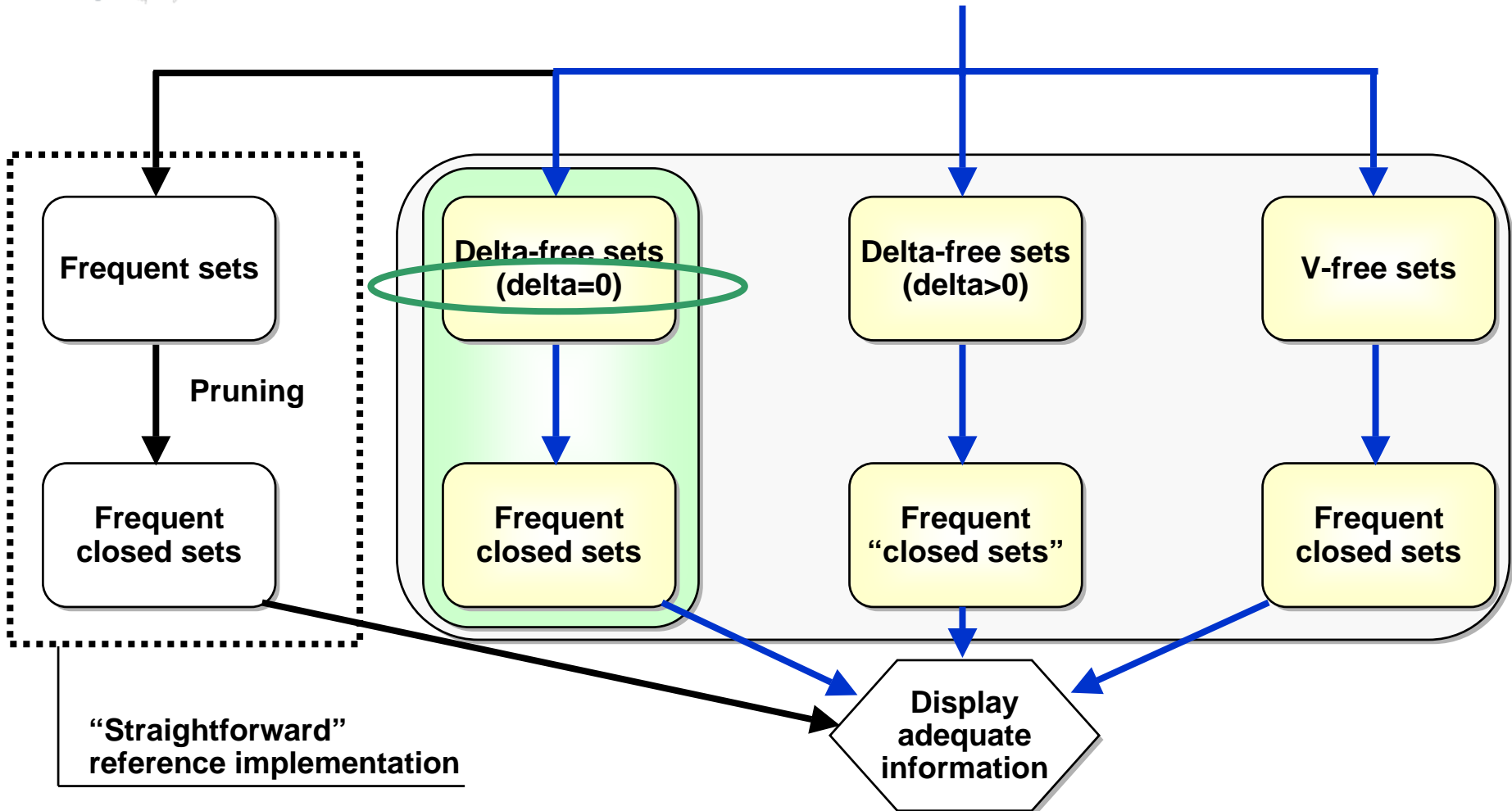


Comprehensive Log Compression





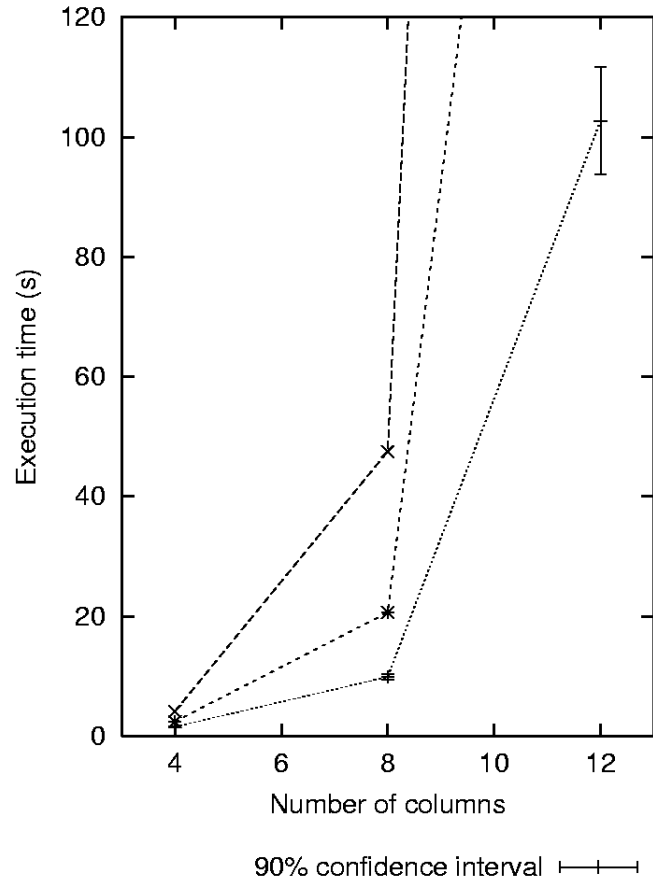
Experiments and Evaluation



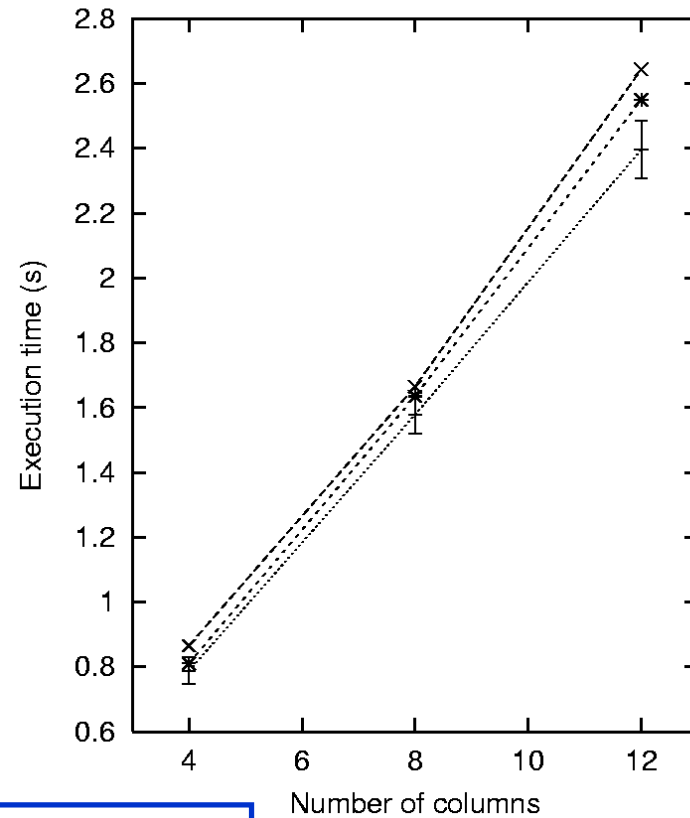


Results: Apriori as a Baseline

Mining on different minimum supports with Apriori



Mining on different minimum supports with acminer



**Different scaling
on the Y-axis**

Threshold 1% ---x---
Threshold 5% ---*---
Threshold 10%#.....



Thank You for Your Attention!



Project funded by the Future and Emerging Technologies arm of
the IST Programme FET-Open scheme