BRICS

Basic Research in Computer Science

An BDD-Based Implementation of the Allegro Software

Daniel Gudbjartsson
Anna Ingólfsdóttir
Augustin Kong

BRICS Notes Series **NS-02-4**

ISSN 0909-3206 **August 2002**

BRICS NS-02-4 Gudbjartsson et al.: An BDD-Based Implementation of the Allegro Software

See back inner page for a list of recent BRICS Notes Series publications.
Copies may be obtained by contacting:

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK–8000 Aarhus C
Denmark
Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

BRICS publications are in general accessible through the World Wide Web and anonymous FTP through these URLs:

`http://www.brics.dk`
`ftp://ftp.brics.dk`
This document in subdirectory `NS/02/4/`

An BDD-based implementation of the Allegro software

Daniel Gudbjartsson [*] Anna Ingólfsdóttir[*] [†‡] Augustin Kong[*]

With improved technology in collecting and processing biological data, the need for improved computational methods to interpret the corresponding results is rapidly growing. In particular this is the case for the field of linkage analysis. Existing algorithms or methods for linkage analysis are in most cases based either on the Elston-Stuart algorithm [ES71] or the algorithm due to Lander and Green, originally presented in [LG87]. As the complexity of the Elston-Stuart algorithm grows exponentially in the number of markers under investigation, it obviously does not suit for analysis based on modern techniques as they typically involve many markers. The Lander-Green algorithm, on the other hand, grows linearly in the number of markers but, which of course is a great improvement. But unfortunately the running time of the algorithm grows exponentially in the number of members of the pedigree being investigated and consequently it is only suitable for pedigrees of small or moderate sizes although many improvements have been suggested to ease the complexity of the method by recognizing symmetries when finding the single point probabilities for the single point case and applying fast Fourier transforms in the multi point calculations [KL95, KL96].

Recently researchers interest in analyzing large pedigrees has increased. This is for instance definitely the case at DeCode Genetics in Reykjavik, where huge amount of data is available and just waits for being investigated with improved methods.

In year 2000, Gudbjartsson et. al. [GJFK00] proposed a method that greatly improved the performance of existing methods. The main improvement from the computational point of view was that the authors, by traversing the pedigree in a top down fashion and correspondingly extending the inheritance pattern, relatively early discovered which patterns were inconsistent with the genotype data given as an input to the algorithm. Furthermore some more symmetries were recognized and applied to for further reductions of the state space that has to be investigated.

Recently Abecasis et. al. [ACCC02] improved further on this method by representing the inheritance patterns as sparse binary trees (but inheritance patterns are typically represented as binary vectors) where symmetries occur as identical subtrees which do not need to be recalculated every time they are referenced.

In theoretical computer science, or more precisely in software verification, software

---

[*] deCode Genetics, Reykjavík, Iceland

[†] on leave from Department of Computer Science, Aalborg University

[‡] Supported by **BRICS** (**B**asic **R**esearch **i**n **C**omputer **S**cience), Centre of the Danish National Research Foundation, DepartmXSent of Computer Science, Aalborg University, Fr. Bajersvej 7E, 9220 Aalborg Ø, Denmark. Corresponding email: `{annai}@cs.auc.dk` or `{annai}@decode.is` .

systems are often modelled as state machines. With more powerful computers, and at the same time more complicated problems that are to be solved by computers, the number of states that have to be explored is often huge. In particular this is the case when concurrent systems are described in this way. Therefore a large number of methods has been put forward in the literature to reduce this complexity, at least in many of the cases that occur in practice (but many of these problems are inherently complex to solve in worst case). One of the techniques used for this purpose was invented by Bryant and is reported in [Br86]. This method suggests that the problem is represented by means of a special data structure, Binary Decision Diagrams (BDD's), that was designed to express and manipulate Boolean functions with multiple inputs. BDD's keep track of symmetries that occur in the structure they describe in a systematic manner and in that way greatly reduce the average running time of many systems that implement queries and therefore can be encoded as Boolean functions. This technique has turned out to be very successful in many applications such as model checking and circuit verification.

A generalization of the BDD data structure , the so called Multi Teminal BDD's (MTBDD's) [CFMMY93] allow for encoding of and reasoning about function that take Boolean as input and but where the output value is not necessarily Boolean. MTBDD's enjoy many of the nice properties of the original BDD's and have also turned out to be successful in many applications.

We have applied MTBDD's to parts of the algorithms sitting inside the Allegro software. The results have been really promising and we have been able to handle considerably larger pedigrees than with other methods. This implementation was performed by students at Aalborg University under the supervision of the second author and is currently being incorporated in the software at DeCode Genetics where it will be more thoroughly tested.

References

[ACCC02] G R Abecasis, S S Cherny, W O Cookson and L R Cardon *Nature Genet.* 30, 97 - 101 (2002)

[Br86] Bryant, R. E., G R Abecasis, S S Cherny, W O Cookson and L R Cardon *Nature Genet.* 30, 97 - 101 (2002)

[CFMMY93] Clarke, E.M., Fujita, M., McGeers, P.C., McMillan, K., Yang, J.C.Y.. IWLS'93. May 1993.

[ES71] Elston, R. C, Stewart, J. A. *Hum. Hered.* 21, 523-542 (1971)

[GJFK00] D F Gudbjartsson, K Jonasson, M L Frigge and A Kong, *Nature genet.* 25, 12 –13

[KL95] Kruglyak L. & Lander, E. S. *Am. J. Hum. Genet.*, 57, 439–454, 1995.

[KL96] Kruglyak L. & Lander, E. S. *Am. J. Hum. Genet.*, 58,1347–1363, 1996.

[LG87] Lander, E: S. & Green, P. *Proc. Natl acad. Sci. USA* 84,2363-2367 (1987).

Recent BRICS Notes Series Publications

NS-02-4 Daniel Gudbjartsson, Anna Ingólfsdóttir, and Augustin Kong. *An BDD-Based Implementation of the Allegro Software*. August 2002. 2 pp.

NS-02-3 Walter Vogler and Kim G. Larsen, editors. *Preliminary Proceedings of the 3rd International Workshop on Models for Time-Critical Systems, MTCS '02,* (Brno, Czech Republic, August 24, 2002), August 2002. vi+141 pp.

NS-02-2 Zoltán Ésik and Anna Ingólfsdóttir, editors. *Preliminary Proceedings of the Workshop on Fixed Points in Computer Science, FICS '02,* (Copenhagen, Denmark, July 20 and 21, 2002), June 2002. iv+81 pp.

NS-02-1 Anders Møller and Michael I. Schwartzbach. *Interactive Web Services with Java: JSP, Servlets, and JWIG.* April 2002. 99 pp.

NS-01-8 Anders Møller and Michael I. Schwartzbach. *The XML Revolution (Revised)*. December 2001. 186 pp. This revised and extended report superseeds the earlier BRICS Report NS-00-8.

NS-01-7 Patrick Cousot, Lisbeth Fajstrup, Eric Goubault, Jeremy Gunawardena, Maurice Herlihy, Martin Raußen, and Vladimiro Sassone, editors. *Preliminary Proceedings of the Workshop on Geometry and Topology in Concurrency Theory, GETCO '01,* (Aalborg, Denmark, August 25, 2001), August 2001. vi+97 pp.

NS-01-6 Luca Aceto and Prakash Panangaden, editors. *Preliminary Proceedings of the 8th International Workshop on Expressiveness in Concurrency, EXPRESS '01,* (Aalborg, Denmark, August 20, 2001), August 2001. vi+139 pp.

NS-01-5 Flavio Corradini and Walter Vogler, editors. *Preliminary Proceedings of the 2nd International Workshop on Models for Time-Critical Systems, MTCS '01,* (Aalborg, Denmark, August 25, 2001), August 2001. vi+ 127pp.

NS-01-4 Ed Brinksma and Jan Tretmans, editors. *Proceedings of the Workshop on Formal Approaches to Testing of Software, FATES '01,* (Aalborg, Denmark, August 25, 2001), August 2001. viii+156 pp.