

Connecting the dots

Giovanna Roda
Vienna, Austria
giovanna.roda@gmail.com

Abstract

Symbolic Computation is an area of computer science that after 20 years of initial research had its acme in the mid-1980s, when its many new algorithms were made available in "math systems for the masses" like Mathematica and Maple. Computational algebra and computational logic are the two main pillars on which this discipline is based.

Currently, the field experiences a new blossom by the integration and combination of new numeric, algebraic, geometric and logic algorithms made available in new, interactive and easy-to-use versions of these systems combined with web services. Spectacular new applications are reported in areas so different as elementary particle physics, cryptography, and automated software generation. One of the most outstanding outcomes of Symbolic Computation is the recently born Wolfram Alpha system, fully programmed in Mathematica.

The goal of this writing is to remind the Information Retrieval community of the chances and capabilities offered by Symbolic Computation.

1 Introduction

Symbolic Computation deals with the algorithmic solution of problems involving symbolic objects. These include algebraic objects that admit exact computations as well as any expression in a formal language, as opposed to numerical objects that are typically represented by an approximation.

This discipline exists since the early 1960s and it is being pushed forward by a small community referring to the yearly ISSAC conference and the ACM Special Interest Group on Symbolic and Algebraic Manipulation.

Mathematica, the system behind Wolfram Alpha, makes available all the power of symbolic computing in an integrated environment. We believe that there are still many open opportunities for synergies between the Symbolic Computation and the Information Retrieval communities. In order to motivate this, we list some of the features and functions of Mathematica that are relevant to IR.

After sketching a history of Symbolic Computation (section 2), we describe its scope and mention some of its areas of application (section 3). Finally, we present some of its capabilities with the help of the Mathematica system.

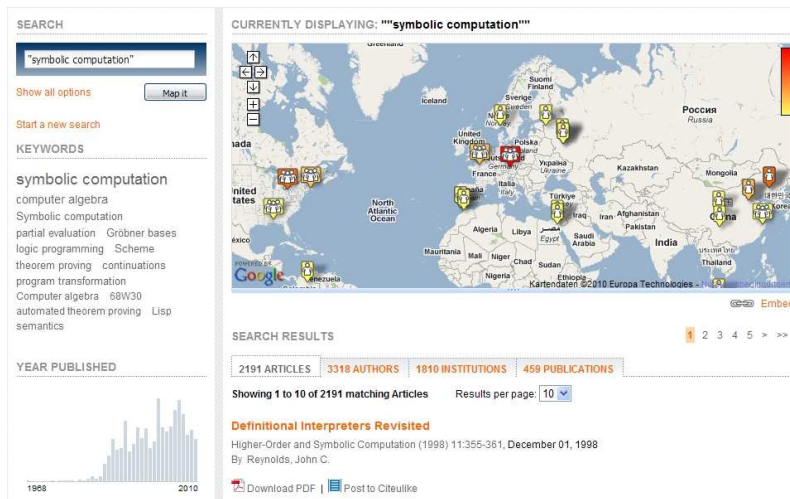


Figure 1: An AuthorMap for “symbolic computation”.

2 A brief history of Symbolic Computation

Symbolic computation (SC) is the name proposed by Bruno Buchberger in 1985 upon request from the Academic Press, London who solicited proposals for a journal in this emerging new field. SC encompasses all disciplines dealing with algorithmic methods that involve symbolic objects. In a very broad sense, any object that can be represented exactly on a computer can be considered a symbolic object. SC had been around, under various names (computer algebra, formula manipulation, symbolic and algebraic computation, computation in closed form, analytic computation, symbolic mathematics and many others) since approximately 1965 when it started with project like the Collins’ SAC system ([2]) or the MAXIMA system at MIT or Buchberger’s PhD thesis in which Gröbner bases were first introduced and a few others like first steps in computational group theory.

The name symbolic computation was preferred to more restrictive names (such as computer algebra) and broader and less mathematical names like artificial intelligence.

The International Symposium on Symbolic and Algebraic Computation (ISSAC), founded in 1966 by the then established ACM Special Interest Group on Symbolic and Algebraic Manipulation (SIGSAM) is the main annual event of the SC community.

3 Scope and applications

The editorial of the first issue of the Journal of Symbolic Computation (JSC) ([1]) provides a pragmatic definition of SC that is valid until today. SC is the combination of

- algorithmic mathematics in abstract mathematical domains (with a notion of “exact” computation) like computer algebra, computer analysis, computer geometry etc. (object level)
- plus computational logic (automated theorem proving, automated programming) etc. (meta-level).

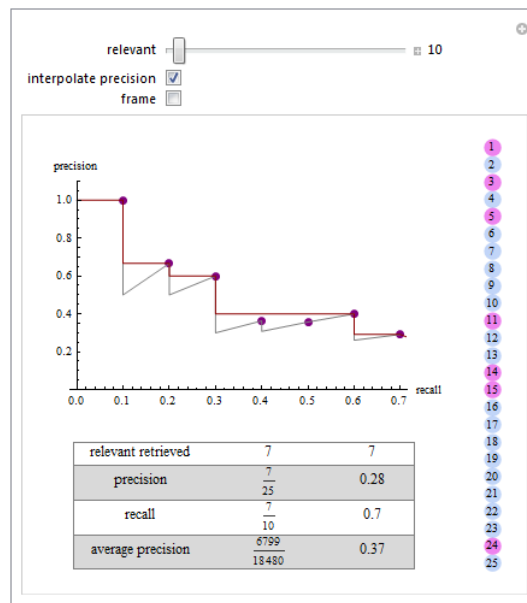


Figure 2: A sample educational application of Mathematica.

The object level often gives methods on the meta-level by making logic “algebraic”, i.e. pressing the logic level down to the “mathematics” level. Here are some of the topics tackled by SC researchers that are within the scope of the JSC (for a complete list see [1]):

- algebraic combinatorics
- computational geometry and differential geometry
- interfacing symbolic and numerical algorithms
- algebraic algorithms in coding and cryptography
- automated theorem proving
- automatic program verification
- impact of symbolic computation on mathematical education

Figure 1 shows an overview of the field of Symbolic Computation generated with the online AuthorMapper¹ service by Springer.

Symbolic Computation has practical applications in many areas of science and technology, such as elementary particles physics, chemistry, computational biology, cryptography, robotics, image processing, statistical experimental design and many others.

Symbolic computation systems are best suited for education thanks to a representation of objects that reflects the way a human thinks about them. A typical example is the capability of symbolic computation systems to carry on computations with algebraic numbers (such as $\sqrt{2}$) without having to use an approximation but keeping it as a symbolic parameter. Furthermore, they allow interactive explorations supported by graphics and animations and can be used for rapid prototyping. Figure 2

¹<http://www.authormapper.com>

shows an educational application generated in Mathematica for plotting interactively a precision-recall curve (available as a Wolfram demonstration²).

In addition to the mainstream commercial products such as Maple or Mathematica, algorithms for symbolic computation are also implemented in a few highly specialized, public domain systems such as CoCoA, Macaulay2, or Singular.

4 The next killer application of Symbolic Computation?

Mathematica, the system used to build Wolfram Alpha, provides all known state-of-the-art algorithms of symbolic computation. We concentrate on this system because of its many features that are relevant to IR research.

Here is the list of capabilities that are provided out-of-the-box in the Mathematica system that are relevant to IR:

string processing this includes regular expression matching, sequence alignment, dictionary lookup;

XML processing the symbolic XML format allows to manipulate an XML document as a Mathematica expression;

distance and similarity measures all standard measures from Manhattan to Damerau-Levenshtein distance - any other arbitrary measure can be defined symbolically;

clustering unsupervised learning methods for classifying data hierarchically or by local optimization;

curated data genes, proteins, chemical structures, geographic and geospatial data, linguistic data (albeit limited to English) that includes properties of words and networks of relationships between them;

image processing filtering and neighborhood processing, mathematical morphology.

5 Epilogue

We presented a short overview of Symbolic Computation and of some of its capabilities. With Wolfram Alpha a new awareness of the power of symbolic computation has been raised in the IR community. Still, we believe that both communities would benefit from a closer interaction.

6 Acknowledgements

Thanks to Bruno Buchberger for pointing me to the Editorial of the first issue of the Journal of Symbolic Computation.

²<http://demonstrations.wolfram.com/ThePrecisionRecallCurveInInformationRetrievalEvaluation>

References

- [1] Bruno Buchberger. Symbolic Computation (An Editorial). *Journal of Symbolic Computation*, 1:1–6, 1985.
- [2] George E. Collins. The SAC-1 system: An introduction and survey. In *SYM-SAC '71: Proceedings of the second ACM symposium on Symbolic and algebraic manipulation*, pages 144–152, New York, NY, USA, 1971. ACM.