

Calcul formel et logiciels libres

20 octobre 2005

Résumé

Ce document essaye de faire une synthèse de la discussion qui s'est engagée à l'automne 2001, en particulier sur la liste de diffusion `calcul-formel-libre@math.cnrs.fr`, à propos du calcul formel et des logiciels libres.

Table des matières

| | | |
|----------|---|----------|
| 1 | Préface | 1 |
| 2 | Introduction | 2 |
| 2.1 | Enjeux scientifiques | 2 |
| 2.2 | Enjeux politiques et économiques | 3 |
| 3 | Projet | 4 |
| 4 | Journées « Calcul formel libre » | 5 |
| 5 | École d'été « Calcul et Algorithmique en liberté » | 5 |
| 6 | Besoins en calcul formel | 6 |
| 7 | Liens | 8 |

1 Préface

Je (Nicolas Thiéry) ne fait que maintenir cette page web ; le document ci-dessous a été écrit par plusieurs participants de la liste de diffusion `calcul-formel-libre@math.cnrs.fr`. Vous pouvez en consulter les archives sur : <http://www.math.cnrs.fr/archives/logiciels-formels/>.

2 Introduction

De tous temps, le développement de la science fondamentale a été étroitement lié au droit et à la possibilité d'échanger librement des connaissances.

Les méthodes commerciales de développement logiciel, créent un clivage : d'un côté les développeurs, en très petit nombre, de l'autre les utilisateurs, beaucoup plus nombreux, mais qui n'ont pas accès au contenu intime de la technologie. Le développement de logiciels libres de calcul permet de remettre tous les individus à égalité. Ceci peut ouvrir la voie à un travail collaboratif élargi, répondre à des impératifs de formation, permettre la réutilisation de parties de codes pour d'autres projets, etc.

Ce projet a pour objectif de créer les conditions permettant le développement de logiciels libres dans le domaine du calcul formel (création et/ou adaptation, évolution, maintenance, diffusion) et de contribuer plus généralement à une plateforme de calcul scientifique libre.

2.1 Enjeux scientifiques

Pour prouver des résultats scientifiques ou pour illustrer ces résultats, il peut être nécessaire de recourir à des calculs sur ordinateur. Lorsque l'on utilise à cet effet un logiciel dont le code source n'est pas disponible (ce qui est le cas des logiciels commerciaux en général), on est à la merci d'une erreur ou de l'oubli d'un cas particulier de l'auteur ou de l'éditeur du logiciel, sans pouvoir le vérifier. La rigueur scientifique exige dans ce cas de faire la preuve avec un logiciel dont les sources sont disponibles, voire modifiables (ce qui est le cas habituel des logiciels libres).

D'autre part, la recherche dans le domaine de l'algorithmique est directement liée aux progrès des logiciels de calcul scientifique. Prenons l'exemple du calcul formel. Le mode de diffusion habituel de nouveaux algorithmes, par le biais d'articles dans des revues scientifiques avec comité de lecture, est lent et ne permet souvent qu'une diffusion confidentielle. Il est aussi possible de diffuser ses résultats scientifiques par le biais d'archives électroniques de prépublications accessibles à tous (mais alors sans validation scientifique du contenu). Un nouveau mode de diffusion très utile pourrait être l'inclusion des algorithmes dans un logiciel de calcul formel libre ouvert et largement diffusé. Ceci permettrait aussi une évaluation plus objective des performances et de l'impact pratique des différents algorithmes proposés.

2.2 Enjeux politiques et économiques

Le contrôle de l'information numérique par des entités commerciales se pratique de différentes manières :

- par le format de donnée numérique des documents ; en sciences physiques ou en mathématiques cela concerne surtout les fichiers sources écrits dans des langages de programmation commerciaux
- en imposant des règles trop rigoureuses sur le droit d'auteur pour les documents (c'est la politique des éditeurs scientifiques) ou les bases de données
- par la diffusion massive de logiciels commerciaux
- par la standardisation systématique de formats d'échanges privés non-compatibles
- par les brevets (y compris les brevets logiciels ou brevets sur des méthodes intellectuelles dans les pays où ils sont légaux).

Lorsqu'une seule entité commerciale contrôle l'information dans son domaine par une des méthodes ci-dessus, elle est en situation de monopole et peut fixer les prix de ses services de manière arbitraire.

On constate à l'heure actuelle une baisse des coûts liés à l'achat des ordinateurs, malgré le fait qu'ils deviennent toujours plus performants, et une augmentation des coûts logiciels qui deviennent une part de dépense de plus en plus prédominante. Par exemple, une licence individuelle Maple coûte actuellement plus cher que l'ordinateur sur lequel le produit est installé. Cette question devient un problème majeur dans le financement des laboratoires de recherche et des institutions d'enseignement.

Examinons la situation en mathématique.

La situation des documents numériques en mathématique est satisfaisante puisque le format le plus utilisé, L^AT_EX, est libre, ce qui facilite les échanges et garantit la pérennité de ces documents. Il est probable que la généralisation du format de stockage XML permettra de plus aux moteurs de recherche de travailler sur des documents scientifiques, qui pourront devenir de véritables documents hypertextes.

Par contre, la diffusion des résultats scientifiques est faite par des éditeurs qui contrôlent le copyright des documents publiés aux dépens de la communauté mathématique qui fait pourtant la part la plus importante du travail (création, édition, référé) tout cela sans véritable contrepartie (ce qui est d'ailleurs contraire au droit d'auteur au moins en France). Les éditeurs pratiquent des hausses de prix constantes. Un détournement notable des ressources publiques pour la recherche s'opère ainsi au profit des éditeurs.

La situation n'est pas favorable non plus dans l'utilisation de logiciels de calcul en mathématiques : il n'y a pas de standard commun donc peu d'in-

teropérabilité, et pas de logiciel libre bien diffusé. Le produit de référence est parfois un logiciel commercial, comme Maple pour le calcul formel dans l'enseignement. La communauté scientifique fait la part initiale du travail (recherche des algorithmes, implantation dans un langage de logiciel commercial) mais ce sont les éditeurs de logiciels commerciaux qui le valorisent à leur seul bénéfice. La hausse récente des prix de licence de logiciels commerciaux est à mettre en parallèle avec la hausse des prix des revues. Il faudrait donc encourager l'utilisation de logiciels libres ce qui est possible comme le montre la standardisation autour du format L^AT_EX.

Pour toutes ces raisons, il est indispensable de disposer de logiciels libres dans tous les domaines du calcul scientifique, en particulier le calcul formel.

3 Projet

Le but du projet est de contribuer dans le domaine du calcul formel à une plate-forme scientifique entièrement libre, et dans laquelle les chercheurs pourraient mettre leurs résultats les plus récents à la libre disposition de tous.

Le projet comporte plusieurs aspects :

- création et/ou adaptation de logiciels de calcul formel libres ;
- intégration et diffusion, de telle sorte que le résultat soit utilisable par un large public, en particulier dans l'enseignement ;
- support, suivi, maintenance, documentation et formation.

Un comité s'est constitué en décembre 2001 afin d'organiser des journées de rencontre, en mai 2002, où seront présentés des projets passés et existants, y compris des projets réalisés dans d'autres domaines scientifiques, et où seront discutées les modalités pratiques du projet global vient de se constituer. Ce comité est composé de Daniel Duparc, Bernard Mourrain, Bernard Parrisé, Bernard Perrot, Fabrice Rouillier, Marie-Francoise Roy, Nicolas Thiéry, Paul Zimmermann. Il s'est réuni pour la première fois le 5 mars au LIP6, et sera dissout, ou renouvelé démocratiquement par les participants à l'issue des journées.

Sans préjuger des conclusions de ces journées, il sera très probablement indispensable :

- De créer une plateforme de travail collaboratif organisée autour d'un noyau dur d'animateurs du projet décidé.
- De prévoir les moyens matériels et surtout humains indispensables pour la réalisation du projet, notamment en identifiant les compétences nécessaires,
- De demander la mise à disposition du projet de moyens humains, selon des modalités adaptées et réalistes à définir (par exemple : participation

au projet de chercheurs ou d'ingénieurs, détachements partiel ou total d'enseignant-chercheurs, thèses, stages, contrats à durée déterminée, etc.).

4 Journées « Calcul formel libre »

Lieu : Université Lyon I, Domaine de Gerland

Date : mardi 21 - jeudi 23 mai 2002

Annonce : <http://www.lapcs.univ-lyon1.fr/~nthiery/CalculFormelLibre/workshop.html>

5 École d'été « Calcul et Algorithmique en liberté »

Lieu : Hyères (VVF de Giens)

Date : lundi 16 - vendredi 20 septembre 2002

Annonce : <http://www-sop.inria.fr/galaad/conf/ecole02.html>

6 Besoins en calcul formel

Ceci est un premier jet qui essaye de lister les besoins des différents types d'utilisateurs du calcul formel.

1. Utilisateurs dans l'enseignement :
 - Interfaces conviviales
Exemple : maple V est limitée (pas de menus) ; une réponse avec un `RootOf` n'est pas envisageable
 - Interpréteur
 - Documentation
 - Stabilité
 - Livres avec exercices
 - Internationalisation (documentation/système/interface/tutoriels/livres)
 - Formation des formateurs
 - Présentation attrayante du calcul formel
 - Formation des décideurs
 - Portabilité Windows / Mac
 - Compatibilité ascendante
2. Utilisateurs extra calcul formel (Physiciens, mathématiciens, chimistes, biologistes, ...)
 - (a) Milieu académique
 - Convivialité
 - Interfaces graphiques et texte pour le traitement automatique
 - On peut s'appuyer sur un plus grand bagage mathématique (pas de pb avec `RootOf`)
 - Formation au calcul formel
 - Tutoriaux sur l'utilisation d'outils spécialisés (ex : applications pratiques des bases de Gröbner, des fonctions symétriques, ...)
 - Gros calculs (on peut se restreindre à UNIX dans un premier temps) :
 - protocoles
 - parallélisme
 - échange de gros volumes de données
 - gestion des ressources
 - Échanges avec le monde extérieur
 - Appels à des bibliothèques spécialisées
 - Lapack, ... pour le calcul numérique
 - GAP
 - ex : mupad/scilab
 - Appel du calcul formel à partir d'autres systèmes

- Interfaçage avec des formats d'entrées sorties existants :
 - Lecture / parsing de gros fichiers de données
 - Génération de documents LaTeX, HTML, MathML, Open-Math ?
 - Génération de code C, ...
 - Interfaces / modules spécialisés pour certains domaines (ex : mise en équation automatique d'un banc d'optique)
 - Bibliothèques graphiques (tracé de courbes, dessin vectoriel)
- (b) Milieu industriel
- Support technique
 - Service de conseil
 - Maintenance du code (il faut qu'il s'installe directement)
 - Validation et tests (Cf. NAG)
 - Portabilité
 - "Publicité" : faire connaître le calcul formel dans l'industrie
 - Transcription semi-automatique d'anciens codes

3. Chercheurs

- Interpréteur haut niveau (validation et prototypage d'algorithmes)
- Difficultés principales :
 - Devoir changer de langage
 - Échanger des données dans divers formats
- Calculs intensifs :
 - Échanges de gros volumes
 - Langage bas niveau, voire assembleur
 - Pêche absolue
 - Sauvegarde d'état (core) pour reprise ultérieure
 - Gestion de mémoire (ramasse-miettes)
 - Possibilité de séparer le noyau de l'interface
- Intégration d'autres contributions ("petites" bibliothèques); valorisation de telles contributions

7 Liens

- La Free Software Foundation et le projet GNU (dont GMP)
www.gnu.org
- Maxima, W. Schelter et al.,
<http://www.math.utexas.edu/users/wfs/>
- Axiom
<http://home.earthlink.net/~jgg964/axiom.html>
- Aldor
<http://aldor.org/>
- PARI, H. Cohen et al.,
<http://www.parigp-home.de>
- Giac, B. Parisse et al.,
<http://www-fourier.ujf-grenoble.fr/~parisse/giac.html>
- GiNaC, R. Kreckel et al.,
<http://www.ginac.de>
- Aubrey Jaffer
<http://www-swiss.ai.mit.edu/~jaffer/JACAL.html>
- NTL, V. Shoup,
<http://www.shoup.net>
- ALP, B. Mourrain et al.,
<http://www-sop.inria.fr/galaad/logiciels/ALP/>
- SYNAPS (Plateforme C++ intégrant ALP, Lapack, Gb, RS, MSolve, Lapack, GMP, MPFR, UDX, Interfaces avec MuPAD, Maple ...), B. Mourrain, F. Rouillier et al.
<http://www-sop.inria.fr/galaad/logiciels/synaps/>
- Zen, F. Chabaud et R. Lercier
<http://www.di.ens.fr/~zen/>
- Scilab, Inria,
<http://www-rocq.inria.fr/scilab>
- Octave, J. W. Eaton et al.,
<http://www.octave.org>
- GSL, M. Galassi, J. Theiler et al.,
<http://sources.redhat.com/gsl>
- Texmacs, J. van der Hoeven
<http://www.texmacs.org>
- Lyx
<http://www.lyx.org>
- ROOT, CERN (Rene Brun, Fons Rademakers)
<http://root.cern.ch/>
- DrGenius, H. Fernandes,

- <http://www.offset.org/drgenius/>
- Gnumeric (projet Gnome)
 - <http://www.gnome.org/projects/gnumeric/>
- Logiciels libres dans l'enseignement (CNDP, J.-P. Archambault)
 - <http://shalmaneser.sortilege.org/cndp/>
- Distribution GNU/Debian
 - <http://www.debian.org>
- Distribution Mandrake pour école d'ingénieurs,
 - <http://www.ensta.fr/~moncorge/mpi/yahoo.html>
- Association AsTeX, M. Lavaud,
 - <http://www.univ-orleans.fr/EXT/ASTEX/>
- Liste de liens et références sur le calcul formel, maintenue par M. Lavaud,
 - <http://www.univ-orleans.fr/EXT/ASTEX/astex/doc/fr/adrut/html/acfor.htm>
- L^AT_EX Navigator
 - <http://tex.loria.fr/index.html>
- Voir aussi
 - <http://www-sor.inria.fr/mirrors/sal/index.shtml>
- Rapport de J.P. Demailly (section IV) :
 - <http://www-fourier.ujf-grenoble.fr/~demailly/rapport.html>
- Linbox
 - <http://www.linalg.org/>
- GAP
 - <http://www-history.mcs.st-and.ac.uk/~gap/>
- OpenMath
 - www.openmath.org
- MuPAD-Combinat, F. Hivert, N. M. Thiéry et al.
 - <http://mupad-combinat.sourceforge.net>
- PerMuVAR, N. M. Thiéry
 - <http://permuvar.sourceforge.net>