



RESEARCH CENTER
Nancy - Grand Est

FIELD

Activity Report 2015

Section Software

Edition: 2016-03-21

ALGORITHMICS, PROGRAMMING, SOFTWARE AND ARCHITECTURE

1. CAMUS Team	4
2. CAMEL Project-Team	8
3. CARTE Project-Team	11
4. CASSIS Project-Team	12
5. VEGAS Project-Team	15
6. VERDIS Project-Team	17

APPLIED MATHEMATICS, COMPUTATION AND SIMULATION

7. SPHINX Team	20
8. TOSCA Project-Team	21

DIGITAL HEALTH, BIOLOGY AND EARTH

9. BIGS Project-Team	22
10. CAPSID Project-Team	24
11. MIMESIS Team	27
12. NEUROSYS Project-Team	29
13. TONUS Team	31

NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING

14. COAST Project-Team	33
15. MADYNES Project-Team	34

PERCEPTION, COGNITION AND INTERACTION

16. ALICE Project-Team	37
17. LARSEN Team	40
18. MAGRIT Project-Team	41
19. MULTISPEECH Project-Team	42
20. ORPAILLEUR Project-Team	44
21. SEMAGRAMME Project-Team	50

CAMUS Team

6. New Software and Platforms

6.1. APOLLO

Automatic speculative POLYhedral Loop Optimizer
FUNCTIONAL DESCRIPTION

We are developing a framework called APOLLO (Automatic speculative POLYhedral Loop Optimizer), dedicated to automatic, dynamic and speculative parallelization of loop nests that cannot be handled efficiently at compile-time. It is composed of a static part consisting of specific passes in the LLVM compiler suite, plus a modified Clang frontend, and a dynamic part consisting of a runtime system. It has been extended in 2015 to apply on-the-fly any kind of polyhedral transformations, including tiling, and to handle nonlinear loops as while-loops referencing memory through pointers and indirections.

- Participants: Aravind Sukumaran-Rajam, Juan Manuel Martinez Caamaño, Luis Esteban Campostrini, Artiom Baloian, Willy Wolff and Philippe Clauss
- Contact: Juan Manuel Martinez Caamaño

6.2. CLooG

Code Generator in the Polyhedral Model
FUNCTIONAL DESCRIPTION

CLooG is a free software and library to generate code (or an abstract syntax tree of a code) for scanning Z-polyhedra. That is, it finds a code (e.g. in C, FORTRAN...) that reaches each integral point of one or more parameterized polyhedra. CLooG has been originally written to solve the code generation problem for optimizing compilers based on the polyhedral model. Nevertheless it is used now in various area e.g. to build control automata for high-level synthesis or to find the best polynomial approximation of a function. CLooG may help in any situation where scanning polyhedra matters. While the user has full control on generated code quality, CLooG is designed to avoid control overhead and to produce a very effective code. CLooG is widely used (including by GCC and LLVM compilers), disseminated (it is installed by default by the main Linux distributions) and considered as the state of the art in polyhedral code generation.

- Participant: Cédric Bastoul
- Contact: Cédric Bastoul
- URL: <http://www.cloog.org>

6.3. Clan

A Polyhedral Representation Extraction Tool for C-Based High Level Languages
FUNCTIONAL DESCRIPTION

Clan is a free software and library which translates some particular parts of high level programs written in C, C++, C# or Java into a polyhedral representation called OpenScop. This representation may be manipulated by other tools to, e.g., achieve complex analyses or program restructurations (for optimization, parallelization or any other kind of manipulation). It has been created to avoid tedious and error-prone input file writing for polyhedral tools (such as CLooG, LeTSeE, Candl etc.). Using Clan, the user has to deal with source codes based on C grammar only (as C, C++, C# or Java). Clan is notably the frontend of the two major high-level compilers Pluto and PoCC.

- Participants: Cédric Bastoul and Imèn Fassi
- Contact: Cédric Bastoul
- URL: http://icps.u-strasbg.fr/people/bastoul/public_html/development/clan/

6.4. Clay

Chunky Loop Alteration wizardrY

FUNCTIONAL DESCRIPTION

Clay is a free software and library devoted to semi-automatic optimization using the polyhedral model. It can input a high-level program or its polyhedral representation and transform it according to a transformation script. Classic loop transformations primitives are provided. Clay is able to check for the legality of the complete sequence of transformation and to suggest corrections to the user if the original semantics is not preserved.

- Participant: Cédric Bastoul
- Contact: Cédric Bastoul
- URL: http://icps.u-strasbg.fr/people/bastoul/public_html/development/clay/

6.5. IBB

Iterate-But-Better

FUNCTIONAL DESCRIPTION

IBB is a source-to-source xfor compiler which automatically translates any C source code containing xfor-loops into an equivalent source code where xfor-loops have been transformed into equivalent for-loops.

- Participants: Imen Fassi, Philippe Clauss and Cédric Bastoul
- Contact: Philippe Clauss

6.6. XFOR-Wizard

XFOR-Wizard

FUNCTIONAL DESCRIPTION

Xfor-Wizard is a programming environment for XFOR programs, assisting users in writing XFOR codes and applying optimizing transformations. Automatic dependence analysis and comparisons against a referential code (XFOR-loops or classic for-loops) are achieved to order to help the user in ensuring semantic correctness of the written code.

- Participants: Imen Fassi, Philippe Clauss and Cédric Bastoul
- Contact: Philippe Clauss

6.7. XFORGEN

XFOR code generator

FUNCTIONAL DESCRIPTION

XFORGEN is a tool to automatically generate an XFOR code that is equivalent to for-loops that have been automatically transformed using a static polyhedral compiler. The generated XFOR code exhibits the parameters of the transformations that have been applied and thus can be modified for further optimizations.

- Participants: Imen Fassi, Philippe Clauss and Cédric Bastoul
- Contact: Philippe Clauss

6.8. OpenScop

A Specification and a Library for Data Exchange in Polyhedral Compilation Tools

FUNCTIONAL DESCRIPTION

OpenScop is an open specification that defines a file format and a set of data structures to represent a static control part (SCoP for short), i.e., a program part that can be represented in the polyhedral model. The goal of OpenScop is to provide a common interface to the different polyhedral compilation tools in order to simplify their interaction. To help the tool developers to adopt this specification, OpenScop comes with an example library (under 3-clause BSD license) that provides an implementation of the most important functionalities necessary to work with OpenScop.

- Participant: Cédric Bastoul
- Contact: Cédric Bastoul
- URL: http://icps.u-strasbg.fr/people/bastoul/public_html/development/openscop/

6.9. ORWL and P99

ORWL is a reference implementation of the Ordered Read-Write Lock tools as described in [5]. The macro definitions and tools for programming in C99 that have been implemented for ORWL have been separated out into a toolbox called P99. ORWL is intended to become opensource, once it will be in a publishable state. P99 is available under a QPL at <http://p99.gforge.inria.fr/>.

Software classification: A-3-up, SO-4, SM-3, EM-3, SDL (P99: 4, ORWL: 2-up), DA-4, CD-4, MS-3, TPM-4

- Participants: Jens Gustedt, Mariem Saied, Daniel Salas
- Contact: Jens Gustedt
- <http://p99.gforge.inria.fr/>, <http://orwl.gforge.inria.fr/>

6.10. stdatomic and musl

We implement the library side of the C11 atomic interface. It needs compiler support for the individual atomic operations and provides library supports for the cases where no low-level atomic instruction is available and a lock must be taken.

- This implementation builds entirely on the ABIs of the gcc compiler for atomics.
- It provide all function interfaces that the gcc ABIs and the C standard need.
- For compilers that don't offer the direct language support for atomics it provides a syntactically reduced but fully functional approach to atomic operations.
- At the core of the library is a new and very efficient futex-based lock algorithm that is implemented for the Linux operating system.

A description of the new lock algorithm has been given in [24]. A short version of it has been accepted for SAC'16.

The primary target of this library is an integration into **musl** to which we also contribute. It is a re-implementation of the C library as it is described by the C and POSIX standards. It is *lightweight, fast, simple, free*, and strives to be correct in the sense of standards-conformance and safety. Musl is production quality code that is mainly used in the area of embedded device. It gains more market share also in other area, e.g. there are now Linux distributions that are based on musl instead of Gnu LibC.

- Participant: Jens Gustedt
- Contact: Jens Gustedt
- <http://stdatomic.gforge.inria.fr/>, <http://www.musl-libc.org/>

6.11. PolyLib

The Polyhedral Library
FUNCTIONAL DESCRIPTION

PolyLib is a C library of polyhedral functions, that can manipulate unions of rational polyhedra of any dimension. It was the first to provide an implementation of the computation of parametric vertices of a parametric polyhedron, and the computation of an Ehrhart polynomial (expressing the number of integer points contained in a parametric polytope) based on an interpolation method. Vincent Loechner is the maintainer of this software.

- Participant: Vincent Loechner
- Contact: Vincent Loechner
- URL: <http://icps.u-strasbg.fr/PolyLib/>

CAMEL Project-Team

6. New Software and Platforms

6.1. Belenios

Belenios - Verifiable online voting system

KEYWORD: E-voting

FUNCTIONAL DESCRIPTION

In collaboration with the CASSIS team, we develop an open-source private and verifiable electronic voting protocol, named BELENIOS. Our system is an evolution of an existing system, Helios, developed by Ben Adida, and used e.g., by UCL and the IACR association in real elections. The main differences with Helios are the following ones:

- In Helios, the ballot box publishes the encrypted ballots together with their corresponding voters. This raises a privacy issue in the sense that whether someone voted or not shall not necessarily be publicized on the web. Publishing this information is in particular forbidden by CNIL's recommendation. BELENIOS no longer publishes voters' identities, still guaranteeing correctness of the tally.
- Helios is verifiable except that one has to trust that the ballot box will not add ballots. The addition of ballots is particularly hard to detect as soon as the list of voters is not public. We have therefore introduced an additional authority that provides credentials that the ballot box can verify but not forge [27].

This new version has been implemented by Stéphane Glondou⁰. The first public release has been done in January 2014. Belenios has been used in Sep 2015 for the election of the new leader of the GT-C2 (Groupe de Travail Codes et Cryptographie) which is part of the GdR-IM (Groupement de Recherche Informatique Mathématique). The GT calcul formel of the GdR-IM plans to use Belenios in 2016 for the election of its new leader.

An online platform⁰ has been released in September 2015, so that setting up a new election can be done entirely from within a browser.

- Participants: Véronique Cortier, Pierrick Gaudry and Stéphane Glondou
- Contact: Stéphane Glondou
- URL: <http://belenios.gforge.inria.fr/>

6.2. CADO-NFS

Crible Algébrique: Distribution, Optimisation - Number Field Sieve

FUNCTIONAL DESCRIPTION

CADO-NFS is a complete implementation in C/C++ of the Number Field Sieve (NFS) algorithm for factoring integers and computing discrete logarithms in finite fields. It consists in various programs corresponding to all the phases of the algorithm, and a general script that runs them, possibly in parallel over a network of computers.

- Participants: Emmanuel Thomé, Pierrick Gaudry, Paul Zimmermann, Alexander Kruppa, François Morain, Cyril Bouvier.
- Contact: Emmanuel Thomé
- URL: <http://cado-nfs.gforge.inria.fr/>

⁰<http://belenios.gforge.inria.fr/>

⁰<https://belenios.loria.fr/>

In December 2015, a major new release of CADO-NFS, version 2.2.0, was published. It contains several bug fixes, efficiency improvements, and the computation of discrete logarithms is now almost “push-button”.

6.3. CMH

Computation of Igusa Class Polynomials

KEYWORDS: Mathematics - Cryptography - Number theory

FUNCTIONAL DESCRIPTION

Cmh computes Igusa class polynomials, parameterizing two-dimensional abelian varieties (or, equivalently, Jacobians of hyperelliptic curves of genus 2) with given complex multiplication.

- Participants: Emmanuel Thomé, Andreas Enge
- Contact: Emmanuel Thomé
- URL: <http://cmh.gforge.inria.fr/>

6.4. GF2X

FUNCTIONAL DESCRIPTION

GF2X is a software library for polynomial multiplication over the binary field, developed together with Richard Brent (Australian National University, Canberra, Australia). It holds state-of-the-art implementation of fast algorithms for this task, employing different algorithms in order to achieve efficiency from small to large operand sizes (Karatsuba and Toom-Cook variants, and eventually Schönhage’s or Cantor’s FFT-like algorithms). GF2X takes advantage of specific processor instructions (SSE, PCLMULQDQ).

- Participants: Pierrick Gaudry, Emmanuel Thomé and Paul Zimmermann
- Contact: Emmanuel Thomé
- URL: <https://gforge.inria.fr/projects/gf2x/>

6.5. GNU MPC

FUNCTIONAL DESCRIPTION

MPC is a C library for the arithmetic of complex numbers with arbitrarily high precision and correct rounding of the result. It is built upon and follows the same principles as MPFR. The library is written by Andreas Enge, Philippe Théveny and Paul Zimmermann.

- Participants: Andreas Enge, Paul Zimmermann, Philippe Théveny and Mickaël Gastineau
- Contact: Andreas Enge
- URL: <http://www.multiprecision.org/>

6.6. GNU-MPFR

KEYWORDS: Multiple-Precision - Floating-point - Correct Rounding

FUNCTIONAL DESCRIPTION

GNU MPFR is an efficient multiple-precision floating-point library with well-defined semantics (copying the good ideas from the IEEE-754 standard), in particular correct rounding in 5 rounding modes. GNU MPFR provides about 80 mathematical functions, in addition to utility functions (assignments, conversions...). Special data (Not a Number, infinities, signed zeros) are handled like in the IEEE-754 standard.

- Participants: Vincent Lefèvre, Guillaume Hanrot, Philippe Théveny and Paul Zimmermann
- Contact: Vincent Lefèvre
- URL: <http://www.mpfr.org/>

6.7. MPFQ

FUNCTIONAL DESCRIPTION

MPFQ is (yet another) library for computing in finite fields, with automatic generation of code for fields known at compile-time. It consists of roughly 18,000 lines of Perl code, which generate most of the C code. MPFQ is used in CADO-NFS, in particular for the linear algebra step during discrete logarithm computations.

- Participants: Emmanuel Thomé, Pierrick Gaudry and Luc Sanselme
- Contact: Pierrick Gaudry
- URL: <http://mpfq.gforge.inria.fr/>

6.8. Tinygb

Tinygb is a small software tool written in C++. Its aim is to provide an interface between several existing libraries (finite field arithmetic, linear algebra) for Gröbner bases computations occurring in problems investigated by the CAMEL group. The focus is not on the efficiency of the implementation, since this is already successfully achieved in other existing software such as *FGb* (developed by Jean-Charles Faugère) or in the CAS Magma (Gröbner bases algorithms are implemented by Alan Steel). The goal of Tinygb is to be a flexible research tool where variants of classical algorithms can be tested. Tinygb is still in development since it requires more testing and packaging before being released.

- Participants: Pierre-Jean Spaenlehauer

6.9. Platforms

6.9.1. CATREL cluster

Installed in 2013, the CATREL computer cluster now plays an essential role in providing the team with the necessary resources to achieve significant computations, which illustrate well the efficiency of the algorithms developed in our research, together with their implementations.

In 2015, the CATREL cluster was in particular used for the precomputations performed for the LOGJAM attack [15]. It was the main computing resource for a record discrete logarithm computation in finite fields of the form \mathbb{F}_{p^3} of 512 bits, and a larger computation for this kind of fields is currently running. It was also used intensively to optimize the sieving parameters of CADO-NFS for factoring numbers from 60 to 155 digits, in the preparation of the release 2.2.0. It was used to factor 47 large integers from nine aliquot sequences starting from 276 to 204828, the largest one being a 190-digit composite number from sequence 660. The current largest element known from an aliquot sequence has 197 digits (sequence 19560), and we expect the 200-digit frontier will be reached in 2016. Several experiments were also made with variations of the polynomial selection algorithm from [10] on RSA-896 and RSA-1024.

CARTE Project-Team

6. New Software and Platforms

6.1. CoDisasm

FUNCTIONAL DESCRIPTION

Codisasm is a new disassembly program which supports self-modifying code and code overlapping. Up to our knowledge, it is the first which copes both aspects of program obfuscation. The tool is based on the notion of “wave” developed in the group.

It is written in C and contains about 3k lines of code.

- Contact: Fabrice Sabatier
- URL: <http://www.lhs.loria.fr/wp/?p=289>

6.2. DynamicTracer

FUNCTIONAL DESCRIPTION

DynamicTracer is a new tool with a public web interface which provides run traces of executable files. The trace is obtained by recording a dynamic execution in a safe environment. It contains instruction addresses, instruction opcodes and other optional information.

It is written in C++ and contains about 2.5k lines of code.

- Contact: Fabrice Sabatier
- URL: <http://www.lhs.loria.fr>

6.3. Gorille

FUNCTIONAL DESCRIPTION

Gorille (formerly MMDEX) is a virus detector based on morphological analysis. It is composed of our own disassembler tool, of a graph transformer and a specific tree-automaton implementation. The tool is used in the EU-Fiware project and by some other partners (e.g., DAVFI project).

It is written in C and contains about 100k lines of code.

APP License, IDDN.FR.001.300033.000.R.P.2009.000.10000, 2009.

- Contact: Philippe Antoine
- URL: <http://www.lhs.loria.fr>

CASSIS Project-Team

6. New Software and Platforms

6.1. Protocol Verification Tools

Participants: Véronique Cortier, Stéphane Glondou, Pierre-Cyrille Héam, Olga Kouchnarenko, Steve Kremer, Michaël Rusinowitch, Mathieu Turuani, Laurent Vigneron.

6.1.1. *CL-AtSe*

We develop *CL-AtSe*, a Constraint Logic based Attack Searcher for cryptographic protocols, initiated and continued by the European projects *AVISPA*, *AVANTSSAR* (for web-services) and *Nessos* respectively. The *CL-AtSe* approach to verification consists in a symbolic state exploration of the protocol execution for a bounded number of sessions, thus is both correct and complete. *CL-AtSe* includes a proper handling of sets, lists, choice points, specification of any attack states through a language for expressing e.g., secrecy, authentication, fairness, or non-abuse freeness, advanced protocol simplifications and optimizations to reduce the problem complexity, and protocol analysis modulo the algebraic properties of cryptographic operators such as XOR (exclusive or) and Exp (modular exponentiation).

CL-AtSe has been successfully used to analyse protocols from e.g., France Telecom R&D, Siemens AG, IETF, Gemalto, Electrum in funded projects. It is also employed by external users, e.g., from the *AVISPA*'s community. Moreover, *CL-AtSe* achieves good analysis times, comparable and sometimes better than other state-of-the art tools.

CL-AtSe has been enhanced in various ways. It fully supports the Aslan semantics designed in the context of the *AVANTSSAR* project, including Horn clauses (for intruder-independent deductions, e.g., for credential management), and a large fragment of LTL-based security properties. A Bugzilla server collects bug reports, and online analysis and orchestration are available on our team server (<https://cassis.loria.fr>). Large models can be analysed on the TALC Cluster in Nancy with parallel processing. *CL-AtSe* also supports negative constraints on the intruder's knowledge, which reduces drastically the orchestrator's processing times and allows separation of duties and non-disclosure policies, as well as conditional security properties, like: i) an authentication to be verified iff some session key is safe; ii) relying on a leaking condition on some private data instead of an honesty predicate to trigger or block some agent's property. This was crucial for e.g., the Electrum's wallet where all clients can be dishonest but security guarantees must be preserved anyway.

6.1.2. *Akiss*

Akiss (Active Knowledge in Security Protocols) is a tool for verifying indistinguishability properties in cryptographic protocols, modelled as trace equivalence in a process calculus. Indistinguishability is used to model a variety of properties including anonymity properties, strong versions of confidentiality and resistance against offline guessing attacks, etc. *Akiss* implements a procedure to verify equivalence properties for a bounded number of sessions based on a fully abstract modelling of the traces of a bounded number of sessions of the protocols into first-order Horn clauses and a dedicated resolution procedure. The procedure can handle a large set of cryptographic primitives, namely those that can be modeled by an optimally reducing convergent rewrite system. The tool also include the possibility for checking everlasting indistinguishability properties [63].

The tool is still under active development, including optimisations to improve efficiency, but also the addition of new features, such as the possibility to model protocols using weak secrets, and the addition of support for exclusive or.

The *Akiss* tool is freely available at <https://github.com/akiss/akiss>.

6.1.3. *Belenios*

In collaboration with the Caramel project-team, we develop an open-source private and verifiable electronic voting protocol, named *Belenios*. Our system is an evolution and a new implementation of an existing system, Helios, developed by Ben Adida, and used e.g., by UCL and the IACR association in real elections. The main differences with Helios are a cryptographic protection against ballot stuffing and a practical threshold decryption system that allows to split the decryption key among several authorities, k out of n authorities being sufficient to decrypt. We will continue to add new cryptographic and protocol improvements to offer a secure, proved, and practical electronic voting system.

Belenios has been implemented (cf. <http://belenios.gforge.inria.fr>) by Stéphane Glondu and has been tested in December 2014 “in real conditions”, in a test election involving the members of Inria Nancy-Grand Est center and of the Loria lab (more than 500 potential voters) that had to elect the best pictures of the Loria. Since 2015, it is used by the CNRS for remote election among its councils. It has also been used to elect the leader of the C2 GdR-IM working group ⁰ (about 230 voters and 100 ballots cast). It has also been used in some smaller elections (eg to chose an invited speaker).

6.1.4. *SAPIC*

SAPIC is a tool that translates protocols from a high-level protocol description language akin to the applied pi-calculus into multiset rewrite rules, that can then be analysed using the Tamarin Prover.

Its aim is the analysis of protocols that include states, for example Hardware Security Tokens communicating with a possibly malicious user, or protocols that rely on databases. It has been succesfully applied on several case studies including the Yubikey authentication protocol.

A recent extension, *SAPIC** extends *SAPIC* by a Kleene star operator (*) which allows to iterate a process a finite but arbitrary number of times. This construction is useful to specify for instance stream authentication protocols. We used it to analyse a simple version of the TESLA protocol.

SAPIC is freely available at <http://sapic.gforge.inria.fr/>.

6.2. Testing Tools

Participants: Fabrice Bouquet, Frédéric Dadeau, Elizabeta Fourneret.

6.2.1. *Hydra*

Hydra is an Eclipse-like platform, based on Plug-ins architecture. Plug-ins can be of five kinds: *parser* is used to analyze source files and build an intermediate format representation of the source; *translator* is used to translate from a format to another or to a specific file; *service* denotes the application itself, i.e., the interface with the user; *library* denotes an internal service that can be used by a service, or by other libraries; *tool* encapsulates an external tool. The following services have been developed so far:

- BZPAnimator: performs the animation of a BZP model (a B-like intermediate format);
- Angluin: makes it possible to perform a machine learning algorithm (à la Angluin) in order to extract an abstraction of a system behavior;
- UML2SMT: aims at extracting first order logic formulas from the UML Diagrams and OCL code of a UML/OCL model to check them with a SMT solver.

These services involve various libraries (sometimes reusing each other), and rely on several *tool* plug-ins that are: SMTProver (encapsulating the Z3 solver), PrologTools (encapsulating the CLPS-B solver), Grappa (encapsulating a graph library). We are currently working on transferring the existing work on test generation from B abstract machines, JML, and statecharts using constraint solving techniques.

⁰<https://crypto.di.ens.fr/c2/election>

6.2.2. jMuHLPSL

jMuHLPSL [6] is a mutant generator tool that takes as input a verified HLPSL protocol, and computes mutants of this protocol by applying systematic mutation operators on its contents. The mutated protocol then has to be analyzed by a dedicated protocol analysis tool (here, the AVISPA tool-set). Three verdicts may then arise. The protocol can still be *safe*, after the mutation, this means that the protocol is not sensitive to the realistic “fault” represented by the considered mutation. This information can be used to inform the protocol designers of the robustness of the protocol w.r.t. potential implementation choices, etc. The protocol can also become *incoherent*, meaning that the mutation introduced a functional failure that prevents the protocol from being executed entirely (one of the participants remains blocked in a given non-final state). The protocol can finally become *unsafe* when the mutation introduces a security flaw that can be exploited by an attacker. In this case, the AVISPA tool-set is able to compute an attack-trace, that represents a test case for the implementation of the protocol. If the attack can be replayed entirely, then the protocol is not safe. If the attack can not be replayed then the implementation does not contain the error introduced in the original protocol.

The tool is written in Java, and it is freely available at: <http://members.femto-st.fr/sites/femto-st.fr/frederic-dadeau/files/content/pub/jMuHLPSL.jar>.

6.2.3. Praspel

Praspel is both a specification language, a test data generator and test execution driver for PHP programs. These latter are annotated to describe class (resp. method) contracts using invariants (resp. pre- and postconditions). Praspel contracts allow to describe data typing informations, by means of *realistic domains*. According to the contract-driven testing principles, the tool uses the contracts to both generate test data, using dedicated test generators (random for integer variables, grammar-based for strings, constraint-based for arrays), and establish the test verdict by checking the contract assertions at run-time.

The tool is open source and freely available at: <http://hoa-project.net>. It has been integrated into a PHP framework named Hoa, and coupled with the atoum tool (<https://github.com/atoum/atoum>) that can be used to execute the tests and report on their code coverage.

6.3. Other Tools

Several software tools described in previous sections are using tools that we have developed in the past. For instance BZ-TT uses the set constraints solver CLPS. Note that the development of the SMT prover haRVey has been stopped. The successor of haRVey is called veriT and is developed by David Déharbe (UFRN Natal, Brasil) and Pascal Fontaine (Veridis team). We have also developed, as a second back-end of AVISPA, TA4SP (Tree Automata based on Automatic Approximations for the Analysis of Security Protocols), an automata based tool dedicated to the validation of security protocols for an unbounded number of sessions.

We have also designed tools to manage collaborative works on shared documents using flexible access control models. These tools have been developed in order to validate and evaluate our approach on combining collaborative edition with optimistic access control.

VEGAS Project-Team

5. New Software and Platforms

5.1. ISOTOP

Topology and geometry of planar algebraic curves

KEYWORDS: Topology - Curve plotting - Geometric computing

Isotop is a Maple software for computing the topology of an algebraic plane curve, that is, for computing an arrangement of polylines isotopic to the input curve. This problem is a necessary key step for computing arrangements of algebraic curves and has also applications for curve plotting.

This software, registered at the APP in June 2011, has been developed since 2007 in collaboration with F. Rouillier from Inria Paris - Rocquencourt. The distributed version is based on the method described in [3], which presents several improvements over previous methods. In particular, our approach does not require generic position. This version is competitive with other implementations (such as ALCIX and INSULATE developed at MPII Saarbrücken, Germany and TOP developed at Santander Univ., Spain). It performs similarly for small-degree curves and performs significantly better for higher degrees, in particular when the curves are not in generic position.

We are currently working on an improved version integrating a new bivariate polynomial solver based on several of our recent results published in [11], [22], [27]. This version is not yet distributed.

- Contact: Sylvain Lazard & Marc Pouget
- URL: <http://vegas.loria.fr/isotop/>

5.2. SubdivisionSolver

KEYWORDS: Numerical solver - Polynomial or analytical systems

The software SubdivisionSolver solves square systems of analytic equations on a compact subset of a real space of any finite dimension. SubdivisionSolver is a numerical solver and as such it requires that the solutions in the subset are isolated and regular for the input system (i.e. the Jacobian must not vanish). SubdivisionSolver is a subdivision solver using interval arithmetic and multiprecision arithmetic to achieve certified results. If the arithmetic precision required to isolate solutions is known, it can be given as an input parameter of the process, otherwise the precision is increased on-the-fly. In particular, SubdivisionSolver can be interfaced with the Fast_Polynomial library (<https://bil.inria.fr/en/software/view/2423/tab>) to solve polynomial systems that are large in terms of degree, number of monomials and bit-size of coefficients.

The software is based on a classic branch and bound algorithm using interval arithmetic: an initial box is subdivided until its sub-boxes are certified to contain either no solution or a unique solution of the input system. Evaluation is performed with a centered evaluation at order two, and existence and uniqueness of solutions is verified thanks to the Krawczyk operator.

SubdivisionSolver uses two implementations of interval arithmetic: the C++ boost library that provides a fast arithmetic when double precision is enough, and otherwise the C mpfi library that allows to work in arbitrary precision. Considering the subdivision process as a breadth first search in a tree, the boost interval arithmetic is used as deeply as possible before a new subdivision process using higher precision arithmetic is performed on the remaining forest.

We used SubdivisionSolver for the experiments in [26], [14], see Section 6.3.2 .

- Contact: Rémi Imbach
- URL: <https://bil.inria.fr/fr/software/view/2605/tab>

5.3. CGAL Periodic Triangulations and Meshes

The **CGAL** library offers a package to compute the 3D periodic Delaunay triangulation of a point set in \mathbb{R}^3 , more precisely the Delaunay triangulation of a point set in the 3-dimensional flat torus with cubic domain [30]. The package has been used in various fields.⁰

We have been extending this package in three directions:

First, a few new small functions have been added to the Delaunay triangulation class and integrated in **CGAL** 4.7.

We have developed and documented some new classes allowing to compute *weighted* periodic Delaunay triangulations. They have been submitted to the **CGAL** editorial board and accepted for inclusion in **CGAL**. The code still needs some polishing, and the testsuite must be completed, before a public distribution in **CGAL**.

We have continued our work to use this package together with the 3D mesh generation package of **CGAL** [29], in order to propose a construction of meshes of periodic volumes. Although last year's preliminary results were already convincing [32], [33], the work is not ready yet for being submitted to **CGAL**: the code requires to be completed, documented, and extensively tested.

- Contact: Monique Teillaud
- In collaboration with Aymeric Pellé (Geometrica project-team)
- This work was done in the framework of the Inria ADT (*Action de Développement Technologique*) OrbiCGAL <http://www.loria.fr/~teillaud/ADT-OrbiCGAL/>

⁰see <http://www.cgal.org/projects.html>

VERIDIS Project-Team

6. New Software and Platforms

6.1. The Redlog Computer Logic System

FUNCTIONAL DESCRIPTION

Redlog is an integral part of the interactive computer algebra system Reduce. It supplements Reduce's comprehensive collection of powerful methods from symbolic computation by supplying more than 100 functions on first-order formulas. Redlog has been publicly available since 1995 and is constantly being improved. The name Redlog stands for Reduce Logic System. Andreas Dolzmann from Schloss Dagstuhl Leibniz-Zentrum für Informatik is a co-developer of Redlog.

Reduce and Redlog are open-source and freely available under a modified BSD license at <http://reduce-algebra.sourceforge.net/>. The Redlog homepage is located at <http://www.redlog.eu/>. Redlog generally works with interpreted first-order logic in contrast to free first-order logic. Each first-order formula in Redlog must exclusively contain atoms from one particular Redlog-supported theory, which corresponds to a choice of admissible functions and relations with fixed semantics. Redlog-supported theories include Nonlinear Real Arithmetic (Real Closed Fields), Presburger Arithmetic, Parametric QSAT, and many more.

Effective quantifier elimination procedures for the various supported theories establish an important class of methods available in Redlog. For the theories supported by Redlog, quantifier elimination procedures immediately yield decision procedures. Besides these quantifier elimination-based decision methods there are specialized, and partly incomplete, decision methods, which are tailored to input from particular fields of application.

In 2015 there was further significant progress with the identification of bifurcations in biochemical models based on real reasoning [17], [33]. With existential real quantifier elimination Redlog can now produce unsatisfiable cores in the infeasible case [27]. This is of considerable relevance in the course of using Redlog as a theory solver in SMT contexts, e.g., within the SMArT project (section 9.2).

Redlog is a widely accepted tool and highly visible in mathematics, informatics, engineering and the sciences. The seminal article on Redlog [4] has received more than 300 citations in the scientific literature so far.

- Participants: Thomas Sturm, Marek Kosta, and Maximilian Jaroschek
- Contact: Thomas Sturm
- URL: <http://www.redlog.eu/>

6.2. SPASS

FUNCTIONAL DESCRIPTION

SPASS is an automated theorem prover based on superposition that handles first-order logic with equality and several extensions for particular classes of theories. It has been developed since the mid-1990s at the Max-Planck Institut für Informatik in Saarbrücken. Version 3.8 is the final release of the SPASS first-order prover built on a traditional “select given loop” design; it is distributed under the FreeBSD license at <http://www.spass-prover.org>.

SPASS will be released in the future in the form of various reasoners for different logics, including combinations of first-order logic with background theories, in particular some forms of arithmetic. In 2015, we have continued our efforts to improve the superposition calculus as well as to develop dedicated arithmetic decision procedures for various arithmetic theories, in particular linear integer arithmetic. Our results are:

- new calculi and decidability results for finite domain fragments,
- specialized reasoning support for finite subsets,
- specialized decision procedures for linear real arithmetic with one quantifier alternation,
- new efficient and complete procedures for (mixed) linear integer arithmetic,
- decidability results and respective procedures for various combinations of linear arithmetic with first-order logic.
- Participants: Martin Bromberger, Thomas Sturm, Marco Voigt, Uwe Waldmann, Christoph Weidenbach
- Contact: Christoph Weidenbach
- URL: <http://www.spass-prover.org/>

6.3. The TLA+ Proof System

FUNCTIONAL DESCRIPTION

TLAPS, the TLA⁺ proof system developed at the Joint MSR-Inria Centre, is a platform for developing and mechanically verifying proofs about TLA⁺ specifications. The TLA⁺ proof language is hierarchical and explicit, allowing a user to decompose the overall proof into independent proof steps. TLAPS consists of a *proof manager* that interprets the proof language and generates a collection of proof obligations that are sent to *backend verifiers*. The current backends include the tableau-based prover Zenon for first-order logic, Isabelle/TLA⁺, an encoding of TLA⁺ as an object logic in the logical framework Isabelle, an SMT backend designed for use with any SMT-lib compatible solver, and an interface to a decision procedure for propositional temporal logic.

The current version 1.4.3 of TLAPS was released in June 2015, it is distributed under a BSD-like license. The prover fully handles the non-temporal part of TLA⁺. Basic temporal logic reasoning is supported through an interface with a decision procedure for propositional temporal logic that performs on-the-fly abstraction of first-order subformulas. Symmetrically, subformulas whose main operator is a connective of temporal logic are abstracted before being sent to backends for first-order logic.

A complete rewrite of the proof manager has started in 2015. Its objectives are to replace the ad-hoc parser used so far with an interface to SANY, the standard parser and semantic analyzer for TLA⁺, to extend the scope of the fragment of TLA⁺ that is handled by TLAPS, and general code refactoring and performance improvements.

TLAPS has been used in several case studies in 2015, including the proof of determinacy of PharOS (section 8.1) and the verification of the Pastry routing protocol (section 7.2). These case studies feed back into the standard library of the distribution.

- Participants: Stephan Merz, Martin Riener, Hernán Vanzetto
- Contact: Stephan Merz
- URL: <http://tla.msr-inria.inria.fr/tlaps/content/Home.html>

6.4. The veriT Solver

FUNCTIONAL DESCRIPTION

VeriT is an open, trustable and efficient SMT (Satisfiability Modulo Theories) solver developed in cooperation with David Déharbe from the Federal University of Rio Grande do Norte in Natal, Brazil, on leave for ClearSy. The solver can handle large quantifier-free formulas containing uninterpreted predicates and functions, and arithmetic over integers and reals. It features efficient decision procedures for uninterpreted symbols and linear arithmetic. It also has some support for user-defined theories, quantifiers, and lambda-expressions. This allows users to easily express properties about concepts involving sets, relations, etc. The prover can produce explicit proof traces when it is used as a decision procedure for quantifier-free formulas with uninterpreted symbols and arithmetic. To support the development of the tool, non-regression tests use Inria's grid infrastructure; it allows us to extensively test the solver on thousands of benchmarks in a few minutes. The veriT solver is available as open source under the BSD license at the [veriT Web site](#).

Efforts in 2015 have been focused on efficiency, stability, and expressiveness, with a new ability for handling non-linear arithmetic. The decision procedures for uninterpreted symbols and linear arithmetic have been further improved. The integration of the solver [Redlog](#) (section 6.1) for non-linear arithmetic in the context of the SMArT project (section 9.2) now works for quantifier-free formulas with non-linear real arithmetic, but is not yet complete for combinations.

The veriT solver participated in the SMT competition [SMT-COMP 2015](#) with decent results.

We target applications where validation of formulas is crucial, such as the validation of TLA⁺ and B specifications, and work together with the developers of the respective verification platforms to make veriT even more useful in practice. The solver is available as a plugin for the Rodin platform for discharging proof obligations generated in Event-B [53]; on a large repository of industrial and academic cases, this SMT-based plugin decreased by 75% the number of proof obligations requiring human interactions, compared to the original B prover.

- Participants: Pascal Fontaine, Pablo Dobal, David Déharbe, and Haniel Barbosa
- Partners: Université de Lorraine - Federal University of Rio Grande do Norte
- Contact: Pascal Fontaine
- URL: <http://www.veriT-solver.org>

SPHINX Team

6. New Software and Platforms

6.1. GPELab

Gross-Pitaevskii equations Matlab toolbox

KEYWORDS: 3D - Quantum chemistry - 2D

FUNCTIONAL DESCRIPTION

GPELab is a Matlab toolbox developed to help physicists for computing ground states or dynamics of quantum systems modeled by Gross-Pitaevskii equations. This toolbox allows the user to define a large range of physical problems (1d-2d-3d equations, general nonlinearities, rotation term, multi-components problems...) and proposes numerical methods that are robust and efficient.

- Contact: Xavier Antoine
- URL: <http://gpelab.math.cnrs.fr/>

6.2. GetDDM

KEYWORDS: Large scale - 3D - Domain decomposition - Numerical solver

FUNCTIONAL DESCRIPTION

GetDDM combines GetDP and Gmsh to solve large scale finite element problems using optimized Schwarz domain decomposition methods.

- Contact: Xavier Antoine
- URL: <http://onelab.info/wiki/GetDDM>

6.3. Platform: Vir'Volt

Vir'Volt is a prototype build in ESSTIN, an engineering school of Université de Lorraine, as part of a student project. The prototype enters low-consumption vehicle race, where the winner covers a given distance (depending upon the race, around 20 km) at a given average speed (around 25 km/h) with the lowest energy consumption. Thomas Chambrion has been in charge of the embedded automatic speed control of Vir'Volt for 6 years. In 2016, Vir'Volt will take part in the European Shell Eco Marathon organized in London. The sloping track (up to 5% uphill and 4% downhill) required a complete rebuild of the transmission parts. The proposed configuration has been obtained after intensive numerical simulations.

- Contact: Thomas Chambrion
- URL: http://www.ecomotionteam.org/blog/?page_id=3072

TOSCA Project-Team

6. New Software and Platforms

6.1. ExitBM

FUNCTIONAL DESCRIPTION

The ExitBM library provides methods to simulate random variables related to the first exit time and position of the Brownian motion from simple domains, namely intervals, squares and rectangles. This is a new software of 2015.

- Participants: Madalina Deaconu and Antoine Lejay
- Contact: Antoine Lejay
- URL: <http://exitbm.gforge.inria.fr/>

6.2. SDM

Stochastic Downscaling Method

FUNCTIONAL DESCRIPTION

The computation of the wind at small scale and the estimation of its uncertainties is of particular importance for applications such as wind energy resource estimation. To this aim, we have developed a computer code belonging to the family of codes of atmospheric flow calculation, in the atmospheric boundary layer. SDM especially concerns the simulation of wind at small space scales (meaning that the horizontal resolution is one kilometer or less), based on the combination of an existing Numerical Weather Prediction model providing a coarse prediction, and a Lagrangian Stochastic Model for turbulent flows.

This year we added to SDM a software tool for Configuration Interface and Visualization (CIV) of the SDM simulations. This dedicated GUI restitutes the 3D simulation view of all SDM outputs (including the rendering of interactions with mills). It is also a key environment tool to visualize a coarse resolution input, to extract time boundary condition of any chosen subdomain simulation for a NetCDF (Network Common DataForm) input file, to prepare the compilation procedure of any simulation case of SDM, to execute codes.

- Participants: Mireille Bossy, Sélim Kraria
- Contact: Mireille Bossy
- URL: <http://windpos.inria.fr>

6.3. Triton

KEYWORDS: Image analysis - Oceanography

FUNCTIONAL DESCRIPTION

The Triton software aims at providing a toolbox to analyze nearshore waves images recorded by a camera on the beach. More precisely, it aims at estimating the height, length and speed of waves, to find speed and direction of currents, and to reconstruct the bathymetry from these images. This is a new software of 2015.

- Participants: Stanislas Larnier, Rafael Almar and Antoine Lejay
- Contact: Antoine Lejay

BIGS Project-Team

6. New Software and Platforms

6.1. SesIndexCreator: An R Package for Socioeconomic Indices Computation and Visualization

KEYWORDS: socioeconomic status, multidimensional index, principal component analysis, hierarchical classification, R

SCIENTIFIC DESCRIPTION

In order to study social inequalities, indices can be used to summarize the multiple dimensions of the socioeconomic status. As a part of the Equit'Area Project, a public health program focused on social and environmental health inequalities, a statistical procedure to create (neighborhood) socioeconomic indices was developed. This procedure uses successive principal components analyses to select variables and create the index. In order to simplify the application of the procedure for non-specialists, the R package SesIndexCreator was created. It allows the creation of the index with all the possible options of the procedure, the classification of the resulting index in categories using several classical methods, the visualization of the results, and the generation of automatic reports. [8]

FUNCTIONAL DESCRIPTION

This package allows computing and visualizing socioeconomic indices and categories distributions from datasets of socioeconomic variables (These tools were developed as part of the EquitArea Project, a public health program).

- Participants: Benoît Lalloué, Severine Deguen, Jean-Marie Monnez and Nolwenn Le Meur
- Contact: Benoît Lalloué
- URL: http://www.equitarea.org/documents/packages_1.0-0/

6.2. Angio-Analytics

KEYWORDS: Health - Cancer - Biomedical imaging

SCIENTIFIC DESCRIPTION

Angio Analytics at l'APP under identification N° FR001.280027.000.R.P.2015.000.10000 - Software for image analysis and statistical analysis of parameters issued from these images

Angiogenesis is the phenomenon by which new blood vessels are created from preexisting ones. But this natural process is also involved, in a chaotic way, in tumor development. Many molecules have shown particular efficiency in inhibiting this phenomenon, hopefully leading to either: (i) a reorganization of the neovessels allowing a better tumor uptake of cytotoxic molecules (as chemotherapy) or (ii) a deprivation of the tumor vascular network with the view to starve it. However, characterizing the anti-angiogenic effects of a molecule remains difficult, mainly because the proposed physical modeling approaches have barely been confronted to in vivo data, which are not directly available. We have developed an original approach to characterize and analyze the anti-angiogenic responses in cancerology that allows biologists to account for spatial and dynamical dimensions of the problem. The proposed solution relies on the association of a specific biological in vivo protocol using skinfold chambers, image processing and dynamic system identification. An empirical model structure of the anti-angiogenic effect of a tested molecule is selected according to experimental data. Finally the model is identified and its parameters are used to characterize and compare responses of the tested molecule. The solution has been implemented in a software developed in The Matlab environment.

FUNCTIONAL DESCRIPTION

Angio-Analytics allows the pharmacodynamic characterization of anti-vascular effects in anti-cancer treatments.

- Participant: Thierry Bastogne
- Contact: Thierry Bastogne

6.3. In silico design of nanoparticles for the treatment of cancers by enhanced radiotherapy

KEYWORDS: Bioinformatics - Cancer - Drug development

FUNCTIONAL DESCRIPTION

To speed up the preclinical development of medical engineered nanomaterials, we have designed an integrated computing platform dedicated to the virtual screening of nanostructured materials activated by X-ray making it possible to select nano-objects presenting interesting medical properties faster. The main advantage of this in silico design approach is to virtually screen a lot of possible formulations and to rapidly select the most promising ones. The platform can currently handle the accelerated design of radiation therapy enhancing nanoparticles and medical imaging nano-sized contrast agents as well as the comparison between nano-objects and the optimization of existing materials.

- Participant: Thierry Bastogne
- Contact: Thierry Bastogne

CAPSID Project-Team

6. New Software and Platforms

6.1. Kpax

KEYWORDS: Protein Structure Alignment

SCIENTIFIC DESCRIPTION

Kpax is a program for flexibly aligning two or more protein structures and for searching databases of protein structures.

FUNCTIONAL DESCRIPTION

The Kpax program exploits the fact that each amino acid residue has a carbon atom with a highly predictable tetrahedral geometry. This allows the local environment of each residue to be transformed into a canonical orientation, thus allowing easy comparison between the canonical orientations of residues within pairs of proteins using a novel scoring function based on Gaussian overlaps. Kpax is now used by the KBDOCK web server [3] to find structural templates for docking which might be beyond the reach of sequence-based homology modeling approaches. In 2015, the Kpax program was extended to allow the flexible alignment and superposition of multiple protein structures, and a new multiple alignment quality measure has been developed. According to this quality measure, Kpax gives higher quality multiple structural alignments than all other published approaches. A journal article describing these new developments is under review.

- Contact: David Ritchie
- URL: <http://kpax.loria.fr>

6.2. KBDOCK

KEYWORDS: Protein Binding Sites

SCIENTIFIC DESCRIPTION

KBDOCK is a database of all known protein-protein interactions that have experimentally determined 3D structures. In 2015, we used the latest version of KBDOCK in several rounds of the community-wide “CAPRI” blind docking experiment [36]. A journal article has been accepted for publication in *Proteins*.

FUNCTIONAL DESCRIPTION

KBDOCK combines coordinate data from the PDB with the Pfam protein domain family classification [28] in order to describe and analyze all known protein-protein interactions for which the 3D structures are available.

- Contact: David Ritchie
- URL: <http://kbdock.loria.fr>

6.3. Hex

KEYWORDS: Protein Docking - 3D rendering - 3D interaction

SCIENTIFIC DESCRIPTION

Hex is an interactive protein docking and molecular superposition program. The underlying approach uses our polar Fourier correlation technique to accelerate the search for close-fitting orientations of the two protein molecules.

FUNCTIONAL DESCRIPTION

Hex understands protein and DNA structures in PDB format, and it can also read small-molecule SDF files. Hex will run on most Windows, Linux and Mac OS X computers. The recent versions include CUDA support for Nvidia GPUs. On a modern workstation, docking times range from a few minutes or less when the search is constrained to known binding sites, to about half an hour for a blind global search (or just a few seconds with CUDA). On multi-processor Linux systems, docking calculation times can be reduced in almost direct proportion to the number of CPUs and GPUs used. In 2015, the Hex code base was re-organised to separate the GUI and computational components into separate libraries. The computational library is now used in our Sam and Kpax software.

- Contact: David Ritchie
- URL: <http://hex.loria.fr>

6.4. Sam

KEYWORDS: Protein Symmetry Assembly - Protein Docking

SCIENTIFIC DESCRIPTION

Sam is a program for building models of protein complexes having arbitrary point group symmetry. The Sam program was developed in the frame of the ANR “PEPSI” project with The Nano-D team at Inria Grenoble – Rhône Alpes. A journal article describing Sam has been accepted for publication in the Journal of Applied Crystallography [16].

FUNCTIONAL DESCRIPTION

The underlying approach makes use of multiple one-dimensional polar Fourier correlations (implemented in the Hex code-base) to search rapidly a symmetry-constrained rigid body protein docking search space. The approach may be used to build symmetrical multi-component protein complexes having a given cyclic (C_n), dihedral (D_n), tetrahedral (T), octahedral (O) or icosahedral (I) point group symmetry.

- Contact: David Ritchie
- URL: <http://sam.loria.fr>

6.5. EC-DomainMiner

KEYWORDS: Protein Domain Annotation

SCIENTIFIC DESCRIPTION

EC-DomainMiner is a recommender-based approach for associating EC (Enzyme Commission) numbers with Pfam domains.

FUNCTIONAL DESCRIPTION

EC-DomainMiner uses a statistical recommender-based approach to infer EC-Pfam relationships from EC-sequence relationships that have been annotated previously in the SIFTS and Uniprot databases.

- Contact: David Ritchie
- URL: <http://ecdm.loria.fr>

6.6. MD-Kmean

KEYWORDS: Molecular Dynamics Analysis

SCIENTIFIC DESCRIPTION

MD-Kmean is a fast program for the analysis of large numbers of Molecular Dynamics frames. The accurate comparison of different protein structures plays important roles in structural biology, structure prediction and functional annotation. The root-mean-square-deviation (RMSD) after optimal superposition is the predominant measure of similarity due to the ease and speed of computation. MD-Kmean was designed to perform both the RMSD and the clustering step necessary to compare large numbers of protein 3D structures stored in large datasets and was applied to a set of 2 microsecond MD simulations producing 2 million frames to be compared and clustered.

FUNCTIONAL DESCRIPTION

We have implemented a very fast version of RMSD for graphics processing units (GPUs) using a quaternion method for calculating the optimal superposition and RMSD that is designed for parallel applications. This acceleration in speed allows RMSD calculations to be used efficiently in computationally intensive applications such as the clustering of large number of molecular dynamics frames. MD-Kmean is 50 times faster on a Nvidia GPU, on average, than the original single-threaded CPU implementation on an Intel quad-core processor.

- Contact: Bernard Maigret

6.7. Protein-Marshmallow

KEYWORDS: Coarse-Grained Representation

SCIENTIFIC DESCRIPTION

A Protein-protein interaction may be considered in terms of physical interaction between two deformable objects. The description at the atomic level of such complex objects is beginning to be feasible by MD simulations, but this requires the use of petaflop machines which are out of reach of most laboratories.

FUNCTIONAL DESCRIPTION

The Protein-Marshmallow program represents the surface of a protein as “coarse grained” 3D triangle mesh. In this mesh, each triangle is colored according to some biological property. In this way, a large complex object may be represented by a much smaller number of samples in a 3D mesh. The Marshmallow program describes deformations of such meshes under the influence of an external force field to simulate the strains that one object may undergo over time due to the interaction with another one.

- Contact: Bernard Maigret

6.8. Platforms

6.8.1. *The MBI Platform*

The MBI (Modeling Biomolecular Interactions) platform (<http://bioinfo.loria.fr>) was established to support collaborations between Inria Nancy – Grand Est and other research teams associated with the University of Lorraine. The platform is a research node of the Institut Français de Bioinformatique (IFB), which is the French national network of bioinformatics platforms (<http://www.france-bioinformatique.fr>).

- Contact: Marie-Dominique Devignes

MIMESIS Team

5. New Software and Platforms

5.1. The SOFA Framework

5.1.1. Description

SOFA⁰ is an open-source software framework targeted at real-time multi-physics simulation, with an emphasis on medical simulation. The idea of SOFA was initiated by members of the MIMESIS team, strongly supported by Inria and still actively developed within the MIMESIS team. Based on C++, the SOFA engine provides many algorithms, physiological models and anatomical data, made available within a plugin architecture. With its high level of modularity, SOFA appears to be an efficient tools to benchmark and develop new medical technologies using existing algorithms.



Figure 7. Logo of the SOFA framework

The SOFA framework relies on a multi-model representation which allows for having several representations (e.g. mechanical, thermal and visual) of the same object. Those different representations are connected together through a mechanism called mapping. With these features, it is also possible to have models of very different natures interacting with each other, for instance rigid bodies, deformable objects, and fluids. CPU and GPU implementations can be transparently combined to exploit the computational power of modern hardware architectures.

SOFA is at the heart of a number of research projects, including cardiac electro-physiology modeling, interventional radiology planning and guidance, planning for cryosurgery and deep brain stimulation, robotics, percutaneous procedures, laparoscopic surgery, non-rigid registration, etc. As a proof of its success, SOFA has been downloaded nearly 150,000 times, and is used today by many research groups around the world, as well as a number of companies. The mailing list used to exchange with the community includes several hundreds of researchers, from about 50 different institutions. SOFA is currently used by several industrial partners (Siemens Corporate Research, Epona Medical, Moog, SenseGraphics, etc.) and also provides the key technology on which our newly created start-up (InSimo) is relying. We strongly believe that today SOFA has become a reference for academic research, and is increasingly gaining recognition for product prototyping and development. The best illustration of this worldwide positioning is the role of SOFA in the challenge set by the HelpMeSee foundation to win the contract for the development of a very ambitious and high-risk project on cataract surgery simulation.

⁰More information about SOFA at <http://www.sofa-framework.org>

5.1.2. The SOFA Consortium

SOFA started ten years ago as an Inria collaborative research project. Now, SOFA includes many different functionalities, several companies rely on the framework as a physics engine and a large community rose over the years. To better meet the expectations of the community, Inria and the SOFA architects decided to create the SOFA Consortium in which the MIMESIS is strongly involved. The official kick-off of the Consortium took place in Strasbourg on the 25. November 2015.



Figure 8. The SOFA Consortium was created around the SOFA platform in 2015: here are some fields of application of SOFA

The **objectives of the SOFA Consortium** can be defined as:

- Represent the identity of SOFA,
- Structure and develop the community,
- Coordinate the development of SOFA to make it always more efficient and stable.

The Consortium has to represent the identity of SOFA. As a consequence, the first mission of the Consortium is to promote SOFA in conferences, forums or any other event. The Consortium must present SOFA to researchers and industrials and inform about all activities around the simulation platform and the available applications. By advertising all this work, the Consortium will bring more visibility to the entire SOFA community, encourage partnership and stimulate technology transfer.

Second, the Consortium now becomes a privileged contact point for any question or request. Members, users, beginners or any interested partner can contact us. We will find the answer to their needs and thus increase the interactions outside and within the community.

Third, the Consortium is in charge of coordinating the developments made in SOFA. Through regular meetings, and bi-annual technical committee, the Consortium makes sure the development follows the road map. Moreover, the Consortium sticks to the vision of SOFA as an open-source software, that has to become more and more stable and easy to use.

Finally, a free support is provided by the Consortium on the public version of SOFA, with the help of the entire SOFA community.

NEUROSYS Project-Team

6. New Software and Platforms

6.1. AnaesthesiaSimulator

KEYWORDS: General anaesthesia - Spiking neural networks - Health

FUNCTIONAL DESCRIPTION

AnaesthesiaSimulator simulates the activity of networks of spiking neurons subject to specific receptor dynamics. The tool is a platform to test effects of anaesthetics on neural activity and is still in its first stage of development. The neural activity can be visualized in a 2D and 3D-plot evolving in time. It is written in Python, open-source and involves heavily the simulation package BRIAN⁰.

- Participants: Axel Hutt and Laure Buhry
- Partner: University of Auckland
- Contact: Axel Hutt
- URL: <https://gforge.inria.fr/projects/anasim/>

6.2. NeuralFieldSimulator

NeuralFieldSimulator

KEYWORDS: Neurosciences - Simulation - Health

FUNCTIONAL DESCRIPTION

The NeuralFieldSimulator computes numerically activity in two-dimensional neural fields by solving integral-differential equations involving transmission delays and visualizes the spatio-temporal activity. The tool includes a GUI that allows the user to choose field parameters. It is written in Python, open-source and is aimed to be promoted to become a major graphical visualization tool in the domain of neural field theory. We aim to establish this simulation software as the first open-source standard simulator for the neural field research community.

- Participants: Axel Hutt and Eric Nichols
- Partner: Kevin Green, University of Ontario, Canada
- Contact: Axel Hutt
- URL: <https://gforge.inria.fr/projects/nfsimulator/>

6.3. OpenViBE

KEYWORDS: Neurosciences - Interaction - Virtual reality - Health - Real time - Neurofeedback - Brain-Computer Interface - EEG - 3D interaction

FUNCTIONAL DESCRIPTION

OpenViBE is a software platform for real-time neurosciences (that is, for real-time processing of brain signals). It can be used to acquire, filter, process, classify and visualize brain signals in real time from various signal sources. OpenViBE is free and open source software. It works on Windows and Linux operating systems.

⁰<http://briansimulator.org/>

This year, Neurosys was in charge of the linux release. the main development efforts concern machine learning: Multi-layer perceptrons have been added as classification methods and evaluation plugins allow to get standalone visualization for Kappa coefficient, Receiving Operative Curve (ROC) curve and general statistics.

- Participants: Yann Renard, Anatole Lécuyer, Fabien Lotte, Bruno Renier, Vincent Delannoy, Laurent Bonnet, Baptiste Payan, Jozef Legény, Jussi Tapio Lindgren, Alison Cellard, Loïc Mahé, Guillaume Serrière, Marsel Mano, Maureen Clerc Gallagher, Théodore Papadopoulo, Laurent Bougrain, Jérémy Frey and Nathanaël Foy
- Partners: INSERM - CEA-List - GIPSA-Lab
- Contact: Anatole Lécuyer
- URL: <http://openvibe.inria.fr>

6.4. BRIAN Contributions

KEYWORDS: Spiking neurons models - Neurosciences - BRIAN

FUNCTIONAL DESCRIPTION

6.4.1. Spiking Neuron Templates

BRIAN-compatible libraries has been developed by the team for various ionic currents to be assembled together to create template neurons. The purpose of these is to speed up simulation set-up time and reduce code duplication across simulation scripts. Template neurons are defined by the ionic currents that flow through their membrane. Implemented templates include:

- Hodgkin-Huxley pyramidal neuron (leak, sodium and potassium)
- Hodgkin-Huxley pyramidal neuron with calcium-activated non-specific (CAN) receptors (leak, sodium, potassium, m-current, calcium, CAN)
- Hodgkin-Huxley fast-spiking inhibitory hippocampal (leak, sodium, potassium, m-current)

Implemented ionic current libraries include:

- Traub and Miles Hodgkin-Huxley (I_{Leak} , I_K , I_{Na}) implementation;
- M-Current (I_M) implementation;
- Calcium current (I_L) implementation;
- Calcium pump mechanisms ($\frac{dCa}{dt}$) implementation;
- Calcium-activated non-selective current (I_{CAN}) implementation;
- Wang and Buszáki inhibitory Hodgkin-Huxley (I_{Leak} , I_K , I_{Na}) implementation;

The current library is easily extensible by third-party users due to its hierarchical design. The template neurons and their currents are defined as YAML⁰ files, which are conveniently parsed by a Python library which acts as an interface to the BRIAN simulator API's.

6.4.2. BRIAN Simulation Parameters

A python library which acts as a wrapper for BRIAN simulations, allowing the user to define simulation parameters in external YAML files, which are then parsed and forwarded to the BRIAN simulator. In addition, simulation parameters can be overridden via the command-line when invoking the python script containing the simulation.

- Participants: Francesco Giovannini
- Contact: Francesco Giovannini
- URL: <http://www.briansimulator.org>

⁰<http://www.yaml.org/>

TONUS Team

6. New Software and Platforms

6.1. SCHNAPS

Participants: Emmanuel Franck, Pierre Gerhard, Philippe Helluy [correspondent], Michel Massaro, Malcolm Roberts, Bruno Weber.

Solveur pour les lois de Conservation Hyperboliques Non-linéaires Appliqué aux PlasmaS
SCIENTIFIC DESCRIPTION

It is clear now that future computers will be made of a collection of thousands of interconnected multicore processors. Globally, it appears as a classical distributed memory MIMD machine. But at a lower level, each of the multicore processors is itself made of a shared memory MIMD unit (a few classical CPU cores) and a SIMD unit (a GPU). When designing new algorithms, it is important to adapt them to this kind of architecture. Our philosophy will be to program our algorithms in such a way that they can be run efficiently on this kind of computers. Practically, we will use the MPI library for managing the coarse grain parallelism, while the OpenCL library will efficiently operate the fine grain parallelism.

We have invested for several years until now into scientific computing on GPUs, using the open standard OpenCL (Open Computing Language). We were recently awarded a prize in the international AMD OpenCL innovation challenge thanks to an OpenCL two-dimensional Vlasov-Maxwell solver that fully runs on a GPU. OpenCL is a very interesting tool because it is an open standard now available on almost all brands of multicore processors and GPUs. The same parallel program can run on a GPU or a multicore processor without modification. OpenCL programs are quite complicated to construct. For instance it is difficult to distribute efficiently the computation or memory operations on the different available accelerators. StarPU <http://starpu.gforge.inria.fr/> is a runtime system developed at Inria Bordeaux that simplifies the distribution of tasks on heterogeneous compute units. We have started to use this software tool in SCHNAPS.

Because of the envisaged applications, which may be either academic or commercial, it is necessary to conceive a modular framework. The heart of the library is made of generic parallel algorithms for solving conservation laws. The parallelism can be both fine-grained (oriented towards GPUs and multicore processors) and coarse-grained (oriented towards GPU clusters). The separate modules allow managing the meshes and some specific applications. With our partner AxesSim, we also develop a C++ specific version of SCHNAPS for electromagnetic applications.

FUNCTIONAL DESCRIPTION

SCHNAPS is a generic Discontinuous Galerkin solver, written in C, based on the OpenCL, MPI and StarPU frameworks.

- Partner: AxesSim
- Contact: Philippe Helluy
- URL: <http://schnaps.gforge.inria.fr/>

6.2. Selalib

Participants: Sever Adrian Hirstoaga, Michel Mehrenberger [correspondent], Pierre Navaro, Laurent Navoret, Thi Trang Nhung Pham, Christophe Steiner.

SEmi-LAgrangian LIBrary

KEYWORDS: Plasma physics - Semi-Lagrangian method - PIC - Parallel computing - Plasma turbulence

SCIENTIFIC DESCRIPTION

The objective of the Selalib project (SEmi-LAgrangian LIBrary) is to develop a well-designed, organized and documented library implementing several numerical methods for kinetic models of plasma physics. Its ultimate goal is to produce gyrokinetic simulations.

Another objective of the library is to provide to physicists easy-to-use gyrokinetic solvers, based on the semi-Lagrangian techniques developed by Eric Sonnendrücker and his collaborators in the past CALVI project. The new models and schemes from TONUS are also intended to be incorporated into Selalib.

FUNCTIONAL DESCRIPTION

Selalib is a collection of modules conceived to aid in the development of plasma physics simulations, particularly in the study of turbulence in fusion plasmas. Selalib offers basic capabilities from general and mathematical utilities and modules to aid in parallelization, up to pre-packaged simulations.

- Partners: Max Planck Institute - Garching - IRMA, Université de Strasbourg - IRMAR, Université Rennes 1 - LJLL, Université Paris 6
- Contact: Michel Mehrenberger
- URL: <http://selalib.gforge.inria.fr/>

6.3. Django

Participants: Emmanuel Franck [correspondent], Boniface Nkonga, Ahmed Ratnani.

- Scientific description:
The JOREK code is one of the most important MHD codes in Europe. This code written 15 years ago allows to simulate the MHD instabilities which appear in the TOKAMAK. Using this code the physicist has obtained some important results. However to run larger and more complex test cases it is necessary to evolve the numerical methods used.
In 2014, the DJANGO code has been created, the aim of this code is double: have a numerical library to implement, test and validate new numerical methods for MHD, fluid mechanics and Electromagnetic equations in the finite element context and prepare the future new JOREK code. This code is a 2D-3D code based on implicit time schemes and IsoGeometric (B-Splines, Bezier curves) for the spatial discretization.
- Functional description:
DJANGO is a finite element implicit solver written in Fortran 2003 with a Basic MPI framework. The code is coupled with the PETSC library for the linear solvers and the code CAID (A. Ratnani) for the mesh.
- Authors:
Ahmed Ratnani (Max Planck Institut of Plasma Physic, Garching, Germany), Boniface NKonga (University of Nice and Inria Sophia-Antipolis, France), Emmanuel Franck (Inria Nancy Grand Est, TONUS Team)
- Contributors:
Laura Mendoza, Mustafa Gaja (PhD), Jalal Lakhilili, Celine Caldini-Queiros, Matthias, Hoelzl, Eric Sonnendrücker (Max Planck Institut of Plasma Physic, Garching, Germany), Ayoub Iaagoubi (ADT), Hervé Guillard (University of Nice and Inria Sophia-Antipolis, France), Virginie Grandgirard, Guillaume Latu (CEA Cadarache, France)
- Year 2015:
The year 2015 is an important year for the JOREK code. Indeed, after the year 2014 where the IsoParametric (Bezier curves) finite element approach in 2D have been implemented for basic elliptic equations, in 2015 we have extended the code for more complex problems. Now the code can treat some hyperbolic, parabolic and elliptic models with different approaches (IsoParametric/IsoGeometric approach, Splines for triangle) in 2D and 3D by tensor product. The compilation of the code is more stable and some regression test cases have been added. To finish, two realistic MHD models (which come from to the JOREK code) have been implemented and must be validated. The Year 2016 will be the year of the first physical and realistic results.

COAST Project-Team

4. New Software and Platforms

4.1. Replication Benchmark

Participants: Pascal Urso [contact], Mehdi Ahmed Nacer, Gérald Oster.

The Replication Benchmark is a performance evaluation framework for optimistic replication mechanisms used in collaborative applications. It contains a library of implementation of several CRDT (Commutative Replicated Data Type) and OT (Operational Transformation) algorithms for different data types: text, set, trees. The framework is able to evaluate the performance of comparable algorithms on different corpus of events traces. These events traces can be produced randomly according to different parameters, can be extracted from real real-time editing session that have been recorded, or can be automatically extracted from distributed version control repositories such as the one produced with Git. Performances of the algorithms are measured in term of execution time, memory footprint and merge result quality (compared to manual merge history stored in git repositories). The source code of this evaluation framework is available at <https://github.com/coast-team/replication-benchmark/>.

4.2. MUTE

Participants: Claudia-Lavinia Ignat, Luc André, François Charoy, Gérald Oster [contact].

MUTE (Multi-User Text Editor) is a web-based text editing tool that allows to edit documents collaboratively in real-time. It implements our recent work on collaborative editing algorithms and more specifically the LOGOOT-SPLIT+ approach [17]. Compared to existing web-based collaborative text editing tool this editor does not require a powerful central server since the server is not performing any computation and acts as a simple broadcast server. Our editor offers support for working offline while still being able to reconnect at a later time. This prototype is distributed under the term of GNU GPLv3 licence and is freely available at <https://github.com/coast-team/mute-demo/>. A demo server is hosted at <http://www.coedit.re/>.

4.3. OpenPaaS POC

Participants: Olivier Perrin [contact], Ahmed Bouchami.

The OpenPaaS Proof of Concept was presented during the final review of the FSN OpenPaaS project. Our contribution was dedicated to the authentication mechanism, and to the authorization framework. These two functionalities was delivered thanks to two RESTful services. The authentication service wrapped the LemonLDAP::NG product within a fully RESTful service, while the authorization service uses the reputation of people within the Enterprise Social Network, and computes the decision/rejection of access based on the trust level of the subject requesting the access. The source code can be obtained via request addressed to Olivier Perrin.

MADYNES Project-Team

6. New Software and Platforms

6.1. Distem

KEYWORDS: Large scale - Experimentation - Virtualization - Emulation

It can be used to transform an homogenous cluster (composed of identical nodes) into an experimental platform where nodes have different performance, and are linked together through a complex network topology, making it the ideal tool to benchmark applications targeting such environments.

- Participants: Luc Sarzyniec, Lucas Nussbaum and Tomasz Buchert
- Partners: CNRS - Université de Lorraine - Loria - Grid'5000 - Inria
- Contact: Lucas Nussbaum
- URL: <http://distem.gforge.inria.fr>

6.2. Escape

KEYWORDS: Security - Web - Privacy - TLS (Transport Layer Security)

Escape is a Firefox web browser add-on that offers the ability to bypass HTTPS firewalls that filter websites based on the SNI value of the TLS connection. In addition, it can be used to bypass legacy filtering of DNS requests. The extension is implemented in JavaScript and is based on another security add-on named Convergence.

- Participants: Shbair Wazen, Thibault Cholez, Antoine Goichot and Isabelle Chrisment
- Contact: Thibault Cholez
- URL: <http://madynes.loria.fr/Research/Software#toc1>

6.3. Flowoid

KEYWORDS: Android - NetFlow - Monitoring

Flowoid is a flow based monitoring probe dedicated to Android environments. It uses the NetFlow protocol to send to a collector, information related to the network activities of running Android applications. The information about each connection are grouped into records containing traditional properties including source IP address, destination IP address, bytes, packets, etc. In addition, Flowoid associates and sends for each NetFlow record a set of information related to the geographic location of the device, the name of the application that established the connection, the state of the device screen (ON, OFF, locked, unlocked) and the type of the traffic (foreground, background).

- Participants: Abdelkader Lahmadi, Frédéric Beck, Julien Vaubourg and Olivier Festor
- Contact: Abdelkader Lahmadi

6.4. Grid'5000 testbed

Grid'5000 is a scientific instrument designed to support experiment-driven research in all areas of computer science related to parallel, large-scale or distributed computing and networking. It gathers 10 sites, 25 clusters, 1200 nodes, for a total of 8000 cores. It provides its users with a fully reconfigurable environment (bare metal OS deployment with Kadeploy, network isolation with KaVLAN) and a strong focus on enabling high-quality, reproducible experiments.

- Participants: Luc Sarzyniec, Jérémie Gaidamour, Arthur Garnier, Clement Parisot, Emmanuel Jeanvoine, Lucas Nussbaum and Émile Morel
- Contact: Lucas Nussbaum
- URL: <https://www.grid5000.fr/>

6.5. Kadeploy

Kadeploy is a scalable, efficient and reliable deployment (provisioning) system for clusters and grids. It provides a set of tools for cloning, configuring (post installation) and managing cluster nodes. It can deploy a 300-nodes cluster in a few minutes, without intervention from the system administrator. It plays a key role on the Grid'5000 testbed (see below), where it allows users to reconfigure the software environment on the nodes, and is also used on a dozen of production clusters both inside and outside Inria.

- Participants: Emmanuel Jeanvoine, Lucas Nussbaum and Luc Sarzyniec
- Partners: CNRS - Université de Lorraine - Loria - Grid'5000 - Inria
- Contact: Emmanuel Jeanvoine
- URL: <http://kadeploy3.gforge.inria.fr>

6.6. MECSYCO suite

KEYWORDS: Modeling - Simulation - Simulator - Multi-model - Co-simulation - Multi-agent - Agent - Artefact

The MECSYCO suite (formerly AA4MM) is aiming at the modeling and simulation of complex systems. It provides concepts and tools to describe and then simulate a system as a set of heterogeneous models (namely a multi-model). It features MECSYCO-RE-java the Java implementation of the central part (core) and VisuAA4MM a visualization tool.

- Participants: Vincent Chevrier (Former LORIA Maia team), Laurent Ciarletta, Julien Siebert, Yannick Presse, Benjamin Segault, Benjamin Camus, Victorien Elvinger, Julien Vaubourg, Christine Bourjot, Benjamin Vouillaume and David Michel
- Partners: Université de Lorraine - Inria
- Contact: Vincent Chevrier

6.7. MPIGate

KEYWORDS: Internet of things - Home Automation - Health - Smart home - Home care - Connected object
MPIGate stands for Multi Protocol Interface GATEway for Tele-care, Environment Monitoring and Control. It is a set of softwares aiming at facilitating the development of both home automation and ambient assisted living applications thanks to the abstraction of heterogeneous sensor data and the facility of access to read and write functions over the devices plugged to the networks. This year, its evolution has mainly been carried out within SATELOR project and LAR project. Bluetooth Low Emission (BLE) has been integrated this year. It can be used by people working on home automation and ambient assisted living applications.

- Participants: Mandar Harshe and Ye-Qiong Song
- Contact: Yeqiong Song
- URL: <http://mpigate.loria.fr/>

6.8. Ruby-cute

KEYWORDS: Experimentation - HPC - Cloud

Ruby-Cute is a set of Commonly Used Tools for Experiments, or Critically Useful Tools for Experiments, depending on who you ask. It is a library aggregating various Ruby snippets useful in the context of (but not limited to) development of experiment software on distributed systems testbeds such as Grid'5000.

- Contact: Lucas Nussbaum
- URL: <http://ruby-cute.github.io/>

6.9. WISCAN

KEYWORDS: Network monitoring - Scanning

This tool allows to scan the entire IPv4 space in an efficient way. It relies on the Zmap (smap.io) while optimizing the randomness of the scanned IP addresses to avoid overloading destination networks and reduce the probability to scan two successive addresses of the same IP address block. Our approach to generate random permutation of IP addresses outperforms the approach used by Zmap in terms of the distribution of distance between successive IP addresses. Besides, our scan methodology can be distributed from multiple sources with few efforts.

- Contact: Jérôme François
- URL: <http://gforge.inria.fr/projects/wiscan>

6.10. XPFlow

XPFlow is an implementation of a new, workflow-inspired approach to control experiments involving large-scale computer installations. Such systems pose many difficult problems to researchers due to their complexity, their numerous constituents and scalability problems. The main idea of the approach consists in describing the experiment as a workflow and execute it using achievements of Business Process Management (BPM), workflow management techniques and scientific workflows.

- Participants: Tomasz Buchert and Lucas Nussbaum
- Contact: Lucas Nussbaum
- URL: <http://xpflow.gforge.inria.fr/>

6.11. Platforms

6.11.1. SCADA and IoT security assessment platform

Participants: Abdelkader Lahmadi [contact], Jérôme François, Olivier Festor.

SCADA *Supervisory Control and Data Acquisitions* refers to a centralized control and monitoring system for a variety of machinery and equipment involved with many industrial activities including: power generation and distribution, transportation, nuclear plants, manufacturing processes, etc. The most threaten accidents in SCADA networks are caused by targeted attacks, where adversaries exploit vulnerabilities available in software or network protocols components to disturb and make damage to the physical process. Therefore, it is important to provide new methods and tools for protecting SCADA networks from malicious cyber attacks targeting physical processes and infrastructures.

We are developing and maintaining a platform to assess and analyse the security of SCADA systems. The current version of the testbed combines real hardware Programmable Logic Controllers (PLCs) and simulation tools of physical processes. It also provides a set of tools that we have developed to capture and analyse control messages exchanged between a PLC and the physical processes. During the year 2015, we have received a regional funding to extend our platform with more devices and off-the-shelf solutions for home automation.

We have also extended the platform with IoT devices dedicated to home automation solutions (smart plugs, home boxes, lighting systems , door locks and detectors, etc). Our main goal is to rely on Software Defined Radio solution to evaluate the security of these devices and finding their communication protocol vulnerabilities.

ALICE Project-Team

6. New Software and Platforms

6.1. Vorpaline

Participants: Dobrina Boltcheva, Bruno Lévy.

Vorpaline is a commercial software / programming library. The Vorpaline software takes a new approach to 3D mesh generation, based on the theory of numerical optimization. The optimal mesh generation algorithm, developed as part of the European Research Council GOODSHAPE project, globally and automatically optimizes the mesh elements with respect to geometric constraints. It is the subject of two patents. The mathematical foundations of this algorithm, *i.e.*, the minimization of a smooth energy function, result in practice in a faster algorithm, and—more importantly—in a higher flexibility. For instance, it will allow automatic generation of the aforementioned “hex-dominant” meshes. Vorpaline is based on Geogram (see below). It adds some specialized components targeted to specific industrial usage, such as 3D gridding for the oil and gas industry. It includes our latest research results in automatic meshing. Vorpaline is licensed under a proprietary license.

6.2. IceSL

Participants: Jérémie Dumas, Jean Hergel, Sylvain Lefebvre, Frédéric Claux, Jonas Martínez Bayona, Samuel Hornus.

IceSL exploits parallel algorithms running on the GPU to afford for interactive modeling of objects described by a Constructive Solid Geometry⁰ scripting language. This also enables direct slicing for additive manufacturing by considering the printer bed as a screen onto which each object slice has to be drawn as quickly as possible. During display and slicing the CSG model is converted on the fly into an intermediate representation enabling fast processing on the GPU. Slices can be quickly extracted, and the tool path is prepared through image erosion. The interactive preview of the final geometry uses the exact same code path as the slicer, providing an immediate, accurate visual feedback.

IceSL allows practitioners to design and combine complex objects with unprecedented ease. Our latest version can combine meshes as well as analytical primitives (*i.e.*, shapes described by an equation), and outputs printer instructions for filament printers as well as stereolithography printers and laser cutters.

We also augmented IceSL with a modern UI that allows users to immediately visualize changes made to the script, as well as expose a set of parameters to non-expert users who are interested in customizing a model created with IceSL. https://youtu.be/I2y_yZ4VEgk.

IceSL is the recipient software for our ERC research project “ShapeForge”, led by Sylvain Lefebvre and includes several research results from the project.

6.3. Graphite

Participants: Dobrina Boltcheva, Samuel Hornus, Bruno Lévy, Nicolas Ray.

Graphite is an experimental 3D modeler, built in top of the Geogram programming library. It has data structures and efficient OpenGL visualization for pointsets, surfacic meshes (triangles and polygons), volumetric meshes (tetrahedra and hybrid meshes). It has state-of-the-art mesh repair, remeshing, reconstruction algorithms. It also has an interface to the Tetgen tetrahedral mesh generator (by Hang Si). This year, Graphite3 was released. It is a major rewrite, based on Geogram, with increased software quality standards (zero warnings on all platforms, systematic documentation of all classes / all functions / all parameters, dramatically improved performances). It embeds Geogram (and optionally Vorpaline) with an easy-to-use Graphic User Interface. Graphite is licensed under the GPLv3.

⁰Boolean operations between solids: difference, union, intersection.

6.4. GraphiteLifeExplorer

Participant: Samuel Hornus.

GLE is a 3D modeler, developed as a plugin of Graphite, dedicated to molecular biology. It is developed in cooperation with the Fourmentin Guilbert foundation and has recently been renamed "GraphiteLifeExplorer". Biologists need simple spatial modeling tools to help in understanding the role of the relative position of objects in the functioning of the cell. In this context, we develop a tool for easy DNA modeling. The tool generates DNA along any user-given curve, open or closed, allows fine-tuning of atomic positions and, most importantly, exports to PDB (Protein Data Bank) file format.

The development of GLE is currently on hold, but it is still downloaded (freely) regularly. We plan to add some functionalities in 2016.

6.5. OpenNL - Open Numerical Library

Participants: Bruno Lévy, Nicolas Ray, Rhaleb Zayer.

OpenNL is a standalone library for numerical optimization, especially well-suited to mesh processing. The API is inspired by the graphics API OpenGL, this makes the learning curve easy for computer graphics practitioners. The included demo program implements our LSCM [36] mesh unwrapping method. It was integrated in **Blender** by Brecht Van Lommel and others to create automatic texture mapping methods. OpenNL is extended with two specialized modules :

- **CGAL parameterization package:** this software library, developed in cooperation with Pierre Alliez and Laurent Saboret, is a **CGAL** package for mesh parameterization.
- **Concurrent Number Cruncher:** this software library extends OpenNL with parallel computing on the GPU, implemented using the CUDA API.

6.6. Geogram

Participant: Bruno Lévy.

Stemming from project GOODSHAPE (ERC Starting Grant) and project VORPALINE (ERC Proof of Concept) **Geogram** is a programming library of geometric algorithms. It includes a simple yet efficient mesh data structure (for surface and volumetric meshes), exact computer arithmetics (a-la Shewchuck, implemented in `GEO::expansion`), a predicate code generator (PCK: Predicate Construction Kit), standard geometric predicates (orient/insphere), Delaunay triangulation, Voronoi diagram, spatial search data structures, spatial sorting, and less standard ones (more general geometric predicates, intersection between a Voronoi diagram and a triangular or tetrahedral mesh embedded in n dimensions). The latter is used by FWD/WarpDrive, the first algorithm that computes semi-discrete optimal transport in 3D that scales up to 1 million Dirac masses (see `compute_OTM` in the example programs). Geogram is licensed under the three-clauses BSD license.

Geogram Pluggable Software Modules: Some users are interested in small subsets of Geogram. Following the principle that Geogram should be as easy to use/compile as possible, some subsets of functionalities are alternatively available as a standalone pair of (header,implementation) files, automatically extracted/assembled from Geogram sourcetree. This makes the functionality usable with 0 dependency: client code that uses a PSM just need to insert the header and the implementation file into the project (rather than linking with the entire Geogram library). The Pluggable Software Modules are licensed under the three-clause BSD license. These Pluggable Software Modules include:

- **OpenNL:** a library of easy-to-use numerical solvers for sparse matrices,
- **MultiPrecision:** a number-type that can be used for computations in arbitrary precision, based on Shewchuk's arithmetic expansion [43]. It is shipped with wrapper classes to use it as a number type for CGAL,
- **Predicates:** implementation of exact and symbolically perturbed predicates with arithmetic filters (based on a combination of Meyer and Pion's arithmetic filter generator [38] and our MultiPrecision library).

6.7. LibSL

Participant: Sylvain Lefebvre.

LibSL is a Simple library for graphics. Sylvain Lefebvre continued development of the LibSL graphics library (under CeCill-C licence, filed at the APP). LibSL is a toolbox for rapid prototyping of computer graphics algorithms, under both OpenGL, DirectX 9/10, Windows and Linux. The library is actively used in both the REVES / Inria Sophia-Antipolis Méditerranée and the ALICE / Inria Nancy - Grand Est teams.

LARSEN Team

6. New Software and Platforms

6.1. New Platforms

6.1.1. Experimental Room for Robotics

We collaborate on this experimental platform with Olivier Rochel (SED Inria Nancy - Grand Est).



Figure 2. Overview of the new experimental room.

A new room has been installed for the experiments of the ResiBots ERC project and of the Codyco FP7 project (Figure 2). This 100 m² room contains a 5.5 × 6 m experimental “arena” made with aluminium trusses.

It is equipped with:

- a 6D motion capture system (Optitrack), with 8 gigabit cameras (Prime 13);
- 4 high-power, studio lights;
- 3 mobile 19” racks (on wheels), which host the power supplies and the computers to control the robots;
- a 6-legged robot, used by the ResiBots project;
- an omnidirectional wheeled robot (Kuka Youbot), used by the ResiBots project;
- a hybrid, wheel-legged robot, used by the ResiBots project (loan by the Pierre and Marie Curie University);
- a Kinova robotic arm, used by the Codyco project.

The trusses support the motion capture system and the lights, and hold all the cables (network, power, etc.).

This room will also host the iCub humanoid robot that should be received in March 2016.

MAGRIT Project-Team

5. New Software and Platforms

5.1. Ltrack

The Inria development action LTrack aims at developing an Android platform in order to facilitate the transfer of some of our algorithms onto mobile devices. This year we finished developing an application that performs tracking by synthesis using the camera and the sensors of a mobile phone. User tests will start in January 2016 and we expect to submit the application to the Android Market in the middle of 2016.

- Contact: Marie-Odile Berger, Gilles Simon.

5.2. PoLAR

PoLAR (Portable Library for Augmented Reality) is a software library that offers powerful and state of the art visualization solutions under an API that is adapted and easy to use for a computer vision scientist. An ADT, also named PoLAR, started in October, 1st 2014 to sustain its development: a software engineer, Pierre-Jean Petitprez, was hired for two years.

After the code was made independent from our other research codes (RALib), the library was ported to up-to-date versions of the supporting libraries: OpenSceneGraph 3.2 and Qt5.4. Heavy code refactoring was also carried out to set the core functionalities in conformity with the standards of the supporting libraries.

PoLAR was made available to the public in October 2015, and can be used under Linux or Windows at the moment.

Also this year, a research branch was developed to add the management of physics engines in PoLAR: so far, Bullet and Vega deformation engines were considered, the former being well integrated and the latter still being a work in progress.

- Contact: Erwan Kerrien, Pierre-Frédéric Villard.
- URL: <http://polar.inria.fr>

5.3. RAlib

RAlib is a library which contains the team's research development on image processing, registration (2D and 3D) and visualization. The library was extended over the period to integrate the Java code developed by Maxime Malgras during his Master's internship. Several applications either used internally or to demonstrate the team's work have been designed with this library.

- Contact: Erwan Kerrien, Gilles Simon

5.4. Reproducible research

Matlab software implementing the algorithms described in published articles is publicly available: NESIF (noise estimation by stacking images affected by illumination flickering) [15], ARPENOS (automated removal of quasi-periodic noise using frequency domain statistics) [14], and AC-ARPENOS (a-contrario automated removal of quasi periodic noise using frequency domain statistics) [22].

MULTISPEECH Project-Team

6. New Software and Platforms

6.1. ANTS - Automatic News Transcription System

FUNCTIONAL DESCRIPTION: ANTS is a multipass system for transcribing audio data, and in particular radio or TV shows. The audio stream is first split into homogeneous segments that are decoded using the most adequate acoustic model with a large vocabulary continuous speech recognition engine (Julius, Sphinx or Kaldi). Further processing passes are run in order to apply unsupervised adaptation processes on the features and/or on the model parameters, or to use Speaker Adaptive Training based models. Latest version include DNN (Deep Neural Network) acoustic modeling.

- Participants: Dominique Fohr, Odile Mella, Irina Illina and Denis Jouvét
- Contact: Dominique Fohr

6.2. ASTALI - Automatic Speech-Text Alignment

FUNCTIONAL DESCRIPTION: ASTALI is a software for aligning a speech signal with its corresponding orthographic transcription (given in simple text file for short audio signals or in .trs files as generated by transcriber for longer speech signals). Using a phonetic lexicon and automatic grapheme-to-phoneme converters, all the potential sequences of phones corresponding to the text are generated. Then, using acoustic models, the tool finds the best phone sequence and provides the boundaries at the phone and at the word levels. The web application makes the service easy to use, without requiring any software downloading. Also, the software is currently under integration in the EQUIPEX ORTOLANG platform.

- Participants: Dominique Fohr, Odile Mella, Antoine Chemardin and Denis Jouvét
- Contact: Dominique Fohr
- URL: <http://astali.loria.fr/>

6.3. JCorpusRecorder

FUNCTIONAL DESCRIPTION: JCorpusRecorder is a software for the recording of audio corpora. It provides an easy tool to record with a microphone. The audio input gain is controlled during the recording. From a list of sentences, the output is a set of wav files automatically renamed according to textual information given in input (nationality, speaker language, gender, ...). An easy to use tagging allows for displaying a textual/visual/audio context for guiding the speaker, along with the text of the sentence to pronounce. Several text encodings are enabled (allowing for instance Chinese texts). The sentences can be presented in a random order. The last version can record up to 8 synchronous channels (8 channels under Linux and 2 channels under Windows). The software is developed in Java, and is currently used for the recording of sentences in several projects.

- Contact: Vincent Colotte

6.4. JSnoori

FUNCTIONAL DESCRIPTION: JSnoori is written in Java and uses signal processing algorithms developed within the WinSnoori software with the double objective of being a platform independent signal visualization and manipulation tool, and also for designing exercises for learning the prosody of a foreign language. Thus JSnoori currently focuses the calculation of F0, the forced alignment of non native English uttered by French speakers and the correction of prosody parameters (F0, rhythm and energy). Several tools have been incorporated to segment and annotate speech. A complete phonetic keyboard is available, several levels of annotation can be used (phonemes, syllables and words) and forced alignment can exploit pronunciation variants. In addition, JSnoori offers real time F0 calculation which can be useful from a pedagogical point of view. Besides the traditional graphic interface, JSnoori can now be used via scripts written in Jython.

- Participants: Yves Laprie, Slim Ouni, Julie Busset, Aghilas Sini and Ilef Ben Farhat
- Contact: Yves Laprie
- URL: <http://www.loria.fr/~laprie/WinSnoori/>

6.5. VisArtico - Visualization of EMA Articulatory Data

FUNCTIONAL DESCRIPTION: VisArtico is a user-friendly software which allows visualizing EMA data acquired by an articulograph (AG500, AG501 or NDI Wave). This visualization software has been designed so that it can directly use the data provided by the articulograph to display the articulatory coil trajectories, synchronized with the corresponding acoustic recordings. Moreover, VisArtico not only allows viewing the coils but also enriches the visual information by indicating clearly and graphically the data for the tongue, lips and jaw. In addition, it is possible to insert images (MRI or X-Ray, for instance) to compare the EMA data with data obtained through other acquisition techniques. The last version of VisArtico can handle multimodal data, not articulatory data only. In fact, it is possible to visualize motion capture data from Vicon or kinect-like systems (PrimeSense and RealSense). It is possible to generate video from the visualized trajectories. A derived version from VisArtico is also used in the ADT Plavis as a tool to visualize and process the audiovisual data. The software is used by more than 170 researchers around the world.

- Participants: Slim Ouni, Loïc Mangeonjean and Illef Ben Farhat
- Contact: Slim Ouni
- URL: <http://visartico.loria.fr>

6.6. Xarticulators

KEYWORD: Medical imaging

FUNCTIONAL DESCRIPTION: The Xarticulators software is intended to delineate contours of speech articulators in X-ray images, construct articulatory models and synthesize speech from X-ray films. This software provides tools to track contours automatically, semi-automatically or by hand, to make the visibility of contours easier, to add anatomical landmarks to speech articulators and to synchronize images with the sound. In addition we also added the possibility of processing digitized manual delineation results made on sheets of papers. Xarticulators also enables the construction of adaptable linear articulatory models from the X-ray images and incorporates acoustic simulation tools to synthesize speech signals from the vocal tract shape. Recent work was on the possibility of constructing a velum model and incorporating it into the area functions.

- Contact: Yves Laprie

6.7. Platform EMA - Electromagnetic Articulography Acquisition

FUNCTIONAL DESCRIPTION: Since the purchase of the articulograph AG500 in 2007, we have built a strong experience with respect to the acquisition technique and we have developed an acquisition protocol. The platform has been improved by acquiring the latest articulograph AG501 funded by the EQUIPEX ORTOLANG project. The AG501 allows tracking the movement of 24 sensors at reasonable high frequency (250Hz) up to a very high frequency (1250Hz). In addition, we have continued improving VisArtico (cf. 6.5), a powerful tool to visualize articulatory data acquired using an articulograph. This year we have used the system to acquire articulatory data for the tongue, jaw and lips to study stuttering speech disorder (informal collaboration with F. Hirsch, Praxiling (UMR 5267)). We have also used the EMA platform to acquire motion capture data for the lips, to be used in the context of audiovisual speech synthesis [82].

- Contact: Slim Ouni

6.8. Platform MRI - Magnetic Resonance Imaging

KEYWORDS: Health - Medical imaging

FUNCTIONAL DESCRIPTION: Magnetic Resonance Imaging (MRI) takes an increasing place in the investigation of speech production because it provides a complete geometrical information of the vocal tract. We thus initiated a cooperation with the IADI laboratory (Imagerie Adaptive Diagnostique et Interventionnelle) at Nancy Hospital, which studies in particular magnetic resonance imaging. This year the work focused on the development of compressed sensing algorithms and the reconstruction of good quality images to acquire cineMRI at a sampling rate between 25 and 60 Hz. The algorithms were implemented on the 3T GE research MRI machine of the Nancy Hospital.

- Contact: Yves Laprie

ORPAILLEUR Project-Team

6. New Software and Platforms

6.1. Symbolic KDD Systems

6.1.1. The Coron Platform

- Contact: Amedeo Napoli
- URL: <http://coron.loria.fr/site/index.php>
- KEYWORDS: Data mining, Closed itemset, Frequent itemset, Generator, Association rule, Rare itemset

FUNCTIONAL DESCRIPTION.

The Coron platform [102], [96] is a KDD toolkit organized around three main components: (1) Coron-base, (2) AssRuleX, and (3) pre- and post-processing modules. The software was registered at the “Agence pour la Protection des Programmes” (APP) and is freely available (see <http://coron.loria.fr>).

The Coron-base component includes a complete collection of data mining algorithms for extracting itemsets such as frequent itemsets, closed itemsets, generators and rare itemsets. In this collection we can find APriori, Close, Pascal, Eclat, Charm, and, as well, original algorithms such as ZART, Snow, Touch, and Talky-G [103]. AssRuleX generates different sets of association rules (from itemsets), such as minimal non-redundant association rules, generic basis, and informative basis. In addition, the Coron system supports the whole life-cycle of a data mining task and proposes modules for cleaning the input dataset, and for reducing its size if necessary.

The Coron toolkit is developed in Java, is operational, and was already used in several research projects.

6.1.2. Orion: Skycube Computation Software

- Contact: Chedy Raissi
- URL: <https://github.com/leander256/Orion>
- KEYWORDS: Skyline, skycube.

FUNCTIONAL DESCRIPTION.

This program implements the algorithms described in a research paper published at VLDB 2010 [100]. The software provides a list of four algorithms discussed in the paper in order to compute skycubes. This is the most efficient –in term of space usage and runtime– implementation for skycube computation.

6.1.3. OrphaMine – Data mining platform for orphan diseases

- Partners: INSERM - MoDYCo CNRS - Délégation régionale Ile-de-France, secteur ouest et nord - Greyc Université de Caen - Basse-Normandie
- Contact: Chedy Raissi
- URL: <http://webloria.loria.fr/~mosmuk/orphamine/>
- KEYWORDS: Bioinformatics, data mining, biology, health, data visualization, drug development.

FUNCTIONAL DESCRIPTION.

The OrphaMine platform, developed as part of the ANR Hybrid project, enables visualization, data integration and in-depth analytics. The data at the heart of the platform is about orphan diseases and is extracted from the OrphaData ontology (<http://www.orpha.net>).

We aim to build a true collaborative portal that will serve the different actors of the Hybrid project: (i) A general visualization of OrphaData data for physicians working, maintaining and developing this knowledge database about orphan diseases. (ii) The integration of analytics (data mining) algorithms developed by the different academic actors. (iii) The use of these algorithms to improve our general knowledge of rare diseases.

6.1.4. PoQeMON Analytics: Platform for Quality Evaluation of Mobile Networks

- Partners: Altran, DataPublica, GenyMobile, HEC, Inria Nancy-Grand Est, IP-Label, Next Interactive Media, Orange, Université Paris-Est Créteil
- Contact: Chedy Raissi
- URL: <https://members.loria.fr/poqemon/>
- KEYWORDS: Data mining, data visualization.

FUNCTIONAL DESCRIPTION.

PoQEMoN is a quality evaluation platform for mobile phone networks. The quality measures include the coverage, availability and network performances. Multiple methods are implemented in this platform, either in visualization or in data anonymization to make on-line analytics as simple as possible.

6.2. Stochastic systems for knowledge discovery and simulation

6.2.1. The CarottAge System

- Contact: Jean-François Mari
- URL: http://www.loria.fr/~jfmari/App/index_in_english.html
- KEYWORDS: Stochastic process, Hidden Markov Models.

FUNCTIONAL DESCRIPTION.

The system CarottAge is based on Hidden Markov Models of second order and provides a non supervised temporal clustering algorithm for data mining and a synthetic representation of temporal and spatial data [97]. CarottAge is currently used by INRA researchers interested in mining the changes in territories related to the loss of biodiversity (projects ANR BiodivAgrim and ACI Ecoger) and/or water contamination. CarottAge is also used for mining hydromorphological data proved to give very interesting results for that purpose.

CarottAge is freely available under GPL license (see <http://www.loria.fr/~jfmari/App/>). A special effort is currently aimed at designing interactive visualization tools to provide the expert a user-friendly interface.

6.2.2. The ARPEntAge System

- Contact: Jean-François Mari
- URL: http://www.loria.fr/~jfmari/App/index_in_english.html
- KEYWORDS: Stochastic process, Hidden Markov Models.

FUNCTIONAL DESCRIPTION.

ARPEntAge, for “Analyse de Régularités dans les Paysages : Environnement, Territoires, Agronomie” (<http://www.loria.fr/~jfmari/App/>) is a software based on stochastic models (HMM2 and Markov Field) for analyzing spatio-temporal data-bases [98]. ARPEntAge is built on top of the CarottAge system to fully take into account the spatial dimension of input sequences. It takes as input an array of discrete data in which the columns contain the annual land-uses and the rows are regularly spaced locations of the studied landscape. It performs a Time-Space clustering of a landscape based on its time dynamic Land Uses (LUS). Displaying tools and the generation of Time-dominant shape files have also been defined.

ARPEntAge is freely available (GPL license) and is currently used by INRA researchers interested in mining the changes in territories related to the loss of biodiversity (projects ANR BiodivAgrim and ACI Ecoger) and/or water contamination. In these practical applications, CarottAge and ARPEntAge aim at building a partition –called the hidden partition– in which the inherent noise of the data is withdrawn as much as possible. The estimation of the model parameters is performed by training algorithms based on the Expectation Maximization and Mean Field theories. The ARPEntAge system takes into account: (i) the various shapes of the territories that are not represented by square matrices of pixels, (ii) the use of pixels of different size with composite attributes representing the agricultural pieces and their attributes, (iii) the irregular neighborhood relation between those pixels, (iv) the use of shape files to facilitate the interaction with GIS (geographical information system).

ARPEntAge and CarottAge were used for mining decision rules in a territory showing environmental issues. They provide a way of visualizing the impact of farmers decision rules in the landscape and revealing new extra hidden decision rules.

6.2.3. *The GenExp System*

- Contact: Florence Le Ber
- URL: http://orpailleur.loria.fr/index.php/GenExp-LandSiTes:_KDD_and_simulation
- KEYWORDS: Simulation, Hidden Markov Models.

FUNCTIONAL DESCRIPTION.

In the framework of the project “Impact des OGM” initiated by the French Ministry of Research, we have developed a software called GenExp-LandSiTes for simulating bidimensional random landscapes, and then studying the dissemination of vegetable transgenes. The GenExp-LandSiTes system is linked to the CarottAge system, and is based on computational geometry and spatial statistics. The simulated landscapes are given as input for programs such as “Mapod-Maïs” or “GeneSys-Colza” for studying the transgene diffusion. Other landscape models based on tessellation methods are under studies. The last version of GenExp allows an interaction with R and deals with several geographical data formats.

6.3. KDD systems in Biology

6.3.1. *IntelliGO Online*

- Contact: Malika Smaïl-Tabbone
- URL: <http://plateforme-mbi.loria.fr/intelligo/>
- KEYWORDS: Bioinformatics, genomics.

FUNCTIONAL DESCRIPTION.

The IntelliGO measure computes semantic similarity between terms from a structured vocabulary (Gene Ontology: GO) and uses these values for computing functional similarity between genes annotated by sets of GO terms [82]. The IntelliGO measure is available on line (<http://plateforme-mbi.loria.fr/intelligo/>) to be used for evaluation purposes. It is possible to compute the functional similarity between two genes, the intra-set similarity value in a given set of genes, and the inter-set similarity value for two given sets of genes.

6.3.2. *WAFObI: KNIME Nodes for Relational Mining of Biological Data*

- Contact: Malika Smaïl-Tabbone
- KEYWORDS: Bioinformatics, genomics.

FUNCTIONAL DESCRIPTION.

KNIME (for “Konstanz Information Miner”) is an open-source visual programming environment for data integration, processing, and analysis. The KNIME platform aims at facilitating the data mining experiment settings as many tests are required for tuning the mining algorithms. Various KNIME nodes were developed for supporting relational data mining using the ALEPH program (<http://www.comlab.ox.ac.uk/oucl/research/areas/machlearn/Aleph/aleph.pl>). These nodes include a data preparation node for defining a set of first-order predicates from a set of relation schemes and then a set of facts from the corresponding data tables (learning set). A specific node allows to configure and run the ALEPH program to build a set of rules. Subsequent nodes allow to test the first-order rules on a test set and to perform configurable cross validations.

6.3.3. MODIM: M^Odel-driven Data Integration for Mining

- Contact: Malika Smaïl-Tabbone
- URL: <https://gforge.inria.fr/projects/modim/>
- KEYWORDS: Data integration, workflow, data modeling.

FUNCTIONAL DESCRIPTION.

The MODIM software (M^Odel-driven Data Integration for Mining) is a user-friendly data integration tool which can be summarized along three functions: (i) building a data model taking into account mining requirements and existing resources; (ii) specifying a workflow for collecting data, leading to the specification of wrappers for populating a target database; (iii) defining views on the data model for identified mining scenarios.

Although MODIM is domain independent, it was used so far for biological data integration in various internal research studies and for organizing data about non ribosomal peptide syntheses.

6.4. Knowledge Systems in Health and Cooking

6.4.1. The Kasimir System for Decision Knowledge Management

- Contact: Jean Lieber
- KEYWORDS: Classification-based reasoning, case-based reasoning, decision knowledge management, knowledge edition, knowledge base maintenance, semantic portal

FUNCTIONAL DESCRIPTION.

The objective of the Kasimir system is decision support and knowledge management for the treatment of cancer. A number of modules have been developed within the Kasimir system for editing treatment protocols, visualization, and maintenance. Kasimir is developed within a semantic portal, based on OWL. KatexOWL (Kasimir Toolkit for Exploiting OWL Ontologies, <http://katexowl.loria.fr>) was developed in a generic way and is applied to Kasimir. In particular, the user interface EdHibou of KatexOWL is used for querying the protocols represented within the Kasimir system. In [86], this research is presented, together with an extension of Kasimir for multi-viewpoint case-based reasoning.

Cabamaka (case base mining for adaptation knowledge acquisition) is a module of the Kasimir system. This system performs case base mining for adaptation knowledge acquisition and provides information units to be used for building adaptation rules. Actually, the mining process in Cabamaka is based on a frequent close itemset extraction module from the Coron platform (see §6.1.1).

The Oncologik system is a collaborative editing tool aiming at facilitating the management of medical guidelines. Based on a semantic wiki, it allows the acquisition of formalized decision knowledge also includes a graphical decision tree editor called KcatoS. A version of Oncologik was released in 2012 (<http://www.oncologik.fr/>).

6.4.2. Taaable: a System for Retrieving and Creating New Cooking Recipes by Adaptation

- Contact: Emmanuel Nauer
- URL: <http://intoweb.loria.fr/taable3ccc/>
- KEYWORDS: Knowledge acquisition, ontology engineering, semantic annotation, case-based reasoning, hierarchical classification, text mining.

FUNCTIONAL DESCRIPTION.

The objectives of the Taaable system are to retrieve textual cooking recipes and to adapt these retrieved recipes whenever needed [84]. Suppose that someone is looking for a “leek pie” but has only an “onion pie” recipe: how can the onion pie recipe be adapted?

The Taaable system combines principles, methods, and technologies such as case-based reasoning (CBR), ontology engineering, text mining, text annotation, knowledge representation, and hierarchical classification. Ontologies for representing knowledge about the cooking domain, and a terminological base for binding texts and ontology concepts, were built from textual web resources. These resources are used by an annotation process for building a formal representation of textual recipes. A CBR engine considers each recipe as a case, and uses domain knowledge for reasoning, especially for adapting an existing recipe w.r.t. constraints provided by the user, holding on ingredients and dish types.

The Taaable system is available on line since 2008 at <http://intoweb.loria.fr/taaaable3ccc/>, and is constantly evolving. Since 2014, Taaable is based on Tuurbine, a generic ontology guided CBR engine over RDFS, and Revisor, an adaptation engine implementing various revision operators. This year, new features have been added to the Taaable system in order to address the new challenges of the 8th Computer Cooking Contest at ICCBR 2015. Firstly, FCA was used to improve the ingredient substitution, by taking into account ingredient combinations in a large set of recipes. Secondly, an approach based on mixed linear optimization has been used to adapt ingredient quantities, in order to be more realistic with a real cooking setting.

6.4.3. Tuurbine: a Generic Ontology Guided Case-Based Inference Engine

- Contact: Emmanuel Nauer
- URL: <http://tuurbine.loria.fr/>
- KEYWORDS: case-based reasoning, inference engine, knowledge representation, ontology engineering, semantic web

FUNCTIONAL DESCRIPTION.

The experience acquired since 5 years with the Taaable system conducted to the creation of a generic case-based reasoning system, whose reasoning procedure is based on a domain ontology [91]. This new system, called Tuurbine (<http://tuurbine.loria.fr/>), takes into account the retrieval step, the case base organization, and also an adaptation procedure which is not addressed by other generic case-based reasoning tools. Moreover, Tuurbine is built over semantic web standards that will ensure facilities for being plugged over data available on the web. The domain knowledge is represented in an RDF store, which can be interfaced with a semantic wiki, for collaborative edition and management of the knowledge involved in the reasoning system (cases, ontology, adaptation rules). The development of Tuurbine was supported by an Inria ADT funding until October 2013.

6.4.4. BeGood: a Generic System for Managing Non-Regression Tests on Knowledge Bases

- Contact: Emmanuel Nauer
- URL: <https://github.com/kolflow/begood>
- KEYWORDS: Tests, non-regression, knowledge evolution.

FUNCTIONAL DESCRIPTION.

BeGood is a system allowing to define test plans, independent of any application domain, and usable for testing any system answering queries by providing results in the form of sets of strings. BeGood provides all the features usually found in test systems, such as tests, associated queries, assertions, and expected result sets, test plans (sets of tests) and test reports. The system is able to evaluate the impact of a system modification by running again test plans and by evaluating the assertions which define whether a test fails or succeeds. BeGood is used by the Taaable system for managing the evolution of the knowledge base used by the CBR system.

6.4.5. Revisor: a Library of Revision Operators and Revision-Based Adaptation Operators

- Contact: Jean Lieber
- URL: <http://revisor.loria.fr/>
- KEYWORDS: Belief revision, adaptation, revision-based adaptation, case-based reasoning, inference engines, knowledge representation.

FUNCTIONAL DESCRIPTION.

Revisor is a library of inference engines dedicated to belief revision and to revision-based adaptation for case-based reasoning. It is open source, under a GPL license and available on the web (<http://revisor.loria.fr/>). It gathers several engines developed during the previous years for various knowledge representation formalisms (propositional logic, with or without the use of adaptation knowledge, conjunction of linear constraints, and qualitative algebras [89]). Some of these engines are already used in the Taaable system. Current developments on Revisor aim at defining new engines in other formalisms. In particular, a study on a revision operator in the propositional closure of linear constraints (with integer and real number variables) has been carried out [70]: definition, properties and algorithm.

SEMAGRAMME Project-Team

5. New Software and Platforms

5.1. Abstract Categorial Grammar Development Toolkit (ACGtk)

The current version of the ACG development toolkit prototype focuses on providing facilities to develop grammars. To this end, the type system currently implemented is the linear core system plus the (non-linear) intuitionistic implication, and a special attention has been paid to type error management. Since 1.0b released in Feb. 2014, ACGtk allows for transformations both from abstract terms to object terms, and from object terms to abstract terms (ACG parsing). The parsing algorithm follows a method which is being implemented for second-order ACGs. It is based on a translation of ACG grammars into Datalog programs and is well-suited to fine-grained optimization.

However, since we are interested not only by recognizability (hence whether some fact is provable) but also by the parsing structure (hence the proof), the Datalog solver has been adapted to produce not only yes/no answer to queries, but also all the proofs of the answers to the queries. The next steps concern optimization and efficiency. Note however that in the general case, the decidability of translating an object term to an abstract one is still an open problem.

We also have enriched the ACG development toolkit with graphical output. The new module includes a small functional OCaml library for manipulating images which enables users to customize the rendering of formulas as pictures.

The software is implemented in OCaml and is available as OPAM⁰ package. Version 1.3.0 was released on November 30th.

- Contact: Sylvain Pogodalla
- URL: <http://www.loria.fr/equipes/calligramme/acg/>

5.2. Grew

Grew is a Graph Rewriting tool dedicated to applications in NLP. Grew takes into account confluent and non-confluent graph rewriting and it includes several mechanisms that help to use graph rewriting in the context of NLP applications (built-in notion of feature structures, parametrization of rules with lexical information).

A online version of Grew for graph matching was presented as a demo in the TALN conference [19].

- Contact: Bruno Guillaume
- URL: <http://grew.loria.fr>

5.3. ZombiLingo

Crowdsourcing is nowadays a way of constructing linguistic resources which is more and more used. In the crowdsourcing area, one of the way to motivate a large amount of people to contribute to a project is to present it as a game. Games used in this particular way are called GWAPs (Game With A Purpose). ZombiLingo is a GWAP where gamers have to give linguistic information about the syntax of French natural language sentence.

At the end of 2015, 460 players are registered on the game website and they have produce 63,000 annotations.

⁰<https://opam.ocaml.org/>

In 2015, an Inria ADT started based on the prototype built in 2014. The engineer (Nicolas Lefebvre) worked on this project since October 2015. The main improvements were: migration towards a new framework (Laravel) and code refactoring, integration of new designs into the game and internationalization of the interface to prepare the game for application to other natural languages.

- Participants: Bruno Guillaume, Karën Fort (Université Paris Sorbonne) and Nicolas Lefebvre
- Contact: Bruno Guillaume
- URL: <http://zombilingo.org/>

5.4. SLAMtk

A management chain of the transcriptions of interviews for the SLAM project which produces a full anonymized randomized version of the resources. Some extensions have been implemented based on Distagger (disfluencies) and MELt (POS and lemma) and propose different analyses of repartition, mainly during the Théophile De Logivière internship.

- Contact: Maxime Amblard
- URL: <http://slam.loria.fr>

5.5. Dep2pict

Dep2pict is a program for drawing graphical representation of dependency structures of natural language sentences.

- Contact: Bruno Guillaume
- URL: <http://dep2pict.loria.fr>

5.6. LEOPAR

Leopar is a parser for natural languages which is based on the formalism of Interaction Grammars. The main features of the current version of the software are: automatic parsing of a sentence or a set of sentences, dependency and parse-tree representation of sentences, interactive parsing (the user chooses the couple of nodes to merge) and visualization of grammars produced by XMG-2 or of sets of description trees associated to some word in the linguistic resources.

- Participants: Guillaume Bonfante, Bruno Guillaume and Guy Perrier
- Contact: Bruno Guillaume
- URL: <http://leopar.loria.fr>