



Activity Report 2017

Team TAPDANCE

Theory and Practice of Nanoscale Computing Engines

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Paris

THEME
Computational Biology

Table of contents

1. Personnel	1
2. Overall Objectives	1
3. Research Program	2
4. Highlights of the Year	3
5. New Software and Platforms	3
5.1. Sanakirja	3
5.2. Thrussh	3
5.3. Pijul	3
5.4. SeqDesign	3
6. Partnerships and Cooperations	3
6.1. European Initiatives	3
6.2. International Research Visitors	4
7. Dissemination	4
7.1. Promoting Scientific Activities	4
7.1.1. Scientific Events Selection	4
7.1.2. Journal	4
7.1.2.1. Guest Editor for Journal Special Issue	4
7.1.2.2. Reviewer - Reviewing Activities	4
7.1.3. Invited Talks	4
7.1.4. Contributed Talks (with no associated publication)	4
7.1.5. Conference Programme Committees	4
7.2. Teaching - Supervision - Juries	4
8. Bibliography	5

Team TAPDANCE

Creation of the Team: 2016 June 03

Keywords:

Computer Science and Digital Science:

A1.1.12. - Non-conventional architectures

A1.3. - Distributed Systems

A2.2. - Compilation

A8.1. - Discrete mathematics, combinatorics

Other Research Topics and Application Domains:

B5.3. - Nanotechnology

B5.6. - Robotic systems

1. Personnel

Research Scientists

Pierre Etienne Meunier [Inria, Starting Research Position]

Damien Woods [Inria, Advanced Research Position, Team Leader]

Intern

Tristan Stérin [Ecole Normale Supérieure Lyon, from Dec 2017; Also visiting student from Aug 2017 until Nov 2017]

Administrative Assistant

Helene Milome [Inria]

External Collaborator

David Doty [UC Davis, from Sep 2017; Also visitor from Jun 2017 until Jul 2017]

2. Overall Objectives

2.1. Overall Objectives

In biological systems we see extraordinarily sophisticated growth processes, where molecular self-assembly is combined with active molecular components. Indeed, biological systems consume energy (e.g. ATP) and exhibit phenomena such as rapid growth in cell size and numbers, reconfiguration of internal components, molecular motors that push and pull large structures around, as well as molecular complexes, cells and whole organs that actively respond to the environment. Computer science gives us tools and methodologies to think about and design systems with large number of interacting components. Our goal is to bring these ideas together to design computational molecular systems.

The work of the TAPDANCE team is concerned with the theory and practice of active DNA nanostructures that build structures and compute, all at the nanoscale.

We focus on:

1. Proposing and analysing models of computation for nanoscale biomolecular systems. This includes finding new models for the systems we wish to build, proving theorems (e.g. about their computational power), as well as developing the theory of existing models.
2. Implementing these models in the wet-lab, primarily using DNA.
3. Software to design these kinds of systems (e.g. DNA sequence design) as well as coarse-grained molecular models for system analysis. Software tools are one of the main ways we bridge the gap between theory and experiments.

3. Research Program

3.1. Recent results and ongoing work

Theoretical work by Meunier and Woods “The non-cooperative tile assembly model is not intrinsically universal or capable of bounded Turing machine simulation” was published in 2017 at the conference STOC, Montreal, Canada, and later presented as a poster and chosen for a short oral talk at DNA23, UT Austin, TX, USA, (presentations by Meunier). In this model, called the noncooperative (or temperature 1) abstract Tile Assembly Model, square tiles assemble into structures in the discrete plane (\mathbb{Z}^2) where each tile binds to a growing structure if one of its 4 coloured edges matches the colour of some available edge on the growing structure. It has been conjectured since 2000 that this model is not capable of computation or other sophisticated forms of growth. We show two results. One of our results states that time-bounded Turing machine computation is impossible in this model if we require the simulation to occur in a bounded rectangle in the plane. This result has a short proof that essentially follows from our other main result (with a more involved proof) which states that this model is not “intrinsically universal”. This latter result means that there is no single tileset in this model that can simulate any instance of the model, answering a question from [Meunier, Patitz, Summers, Theyssier, Winslow, Woods; SODA 2014] and contrasting a result [Doty, Lutz, Patitz, Schweller, Summers, Woods; FOCS 2012] for the more general cooperative (temperature 2) model.

A number of popular models of computation for molecular computing are kinetic in nature: the rules of the systems describe how system evolves, step-by-step over time. However, such models allow one to program structures and behaviours that contradict how molecular systems would behave on long enough (e.g. infinite) time-scales. Recently, Doty, Rogers, Soloveichik, Thachuk and Woods proposed a thermodynamic based model called Thermodynamic Binding Networks. One programs the model by specifying a multiset of molecules, and then the ‘output’ is defined as the (or a, if many) multiset of polymers deemed most ‘thermodynamically stable’ out of all possible multisets of polymers definable in the model. In order to clarify the roles of fundamental thermodynamic concepts in molecular computing, the model makes a number of simplifying assumptions including a lack of geometry and (essentially) an infinite biasing of enthalpy over entropy. This model was published at the DNA23 conference at the University of Texas at Austin in 2017, along with some results: The authors show how to program the model to evaluate Boolean AND/OR formulas, and how to efficiently self-assemble simple structures (a binary counter). Also, limitations (upper bounds) on the size of objects self-assemblable are shown. It is hoped this work will lead to new ways of thinking about computation with molecules.

In experimental work, Thubagere, Li, Johnson, Chen, Doroudi, Lee, Izatt, Wittman, Srinivas, Woods, Winfree and Qian, implemented a molecular walker made out of DNA. The walker randomly around on a 2D nanoscale testing ground, picking up one of two types of cargos (also DNA molecules) and dropping them off at specific goal locations. The work shows the power of simple randomised algorithms in molecular systems. This work was led by Thubagary and Qian at Caltech; Woods’ contribution occurred while at Caltech (before joining Inria). The work was published in Science.

Experimental work by Woods and collaborators Doty, Myhrvold, Hui, Zhou, Yin and Winfree has focused on experimentally implementing a wide class of Boolean circuits of a certain form. Experiments were mostly carried out at Caltech, with data analysis and paper writeup being carried out at Inria, UC Davis and Caltech. A publication is in preparation.

ENS student Tristan Stérin is leading theoretical work on analysing the computational power of the previously mentioned iterated Boolean circuit model of Woods, Doty, Myhrvold, Hui, Zhou, Yin, Winfree. Preliminary results were presented as a poster at DNA23 (UT Austin, TX, USA) where Tristan won “DNA23 best poster award”. The work is in preparation for a conference publication.

There are a number of ongoing projects along the lines of topics above in Overall Objectives.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Tristan Stérin won a best poster award at the conference DNA 23.

5. New Software and Platforms

5.1. Sanakirja

KEYWORD: Databases

FUNCTIONAL DESCRIPTION: Sanakirja is a fully transactional (all operations are atomic) key-value dictionary stored in a file (usable at the low-level layer of a more full-featured database engine), with a zero-copy fork operation (fork is in time and space $O(\log n)$, where n is the number of keys in the file). This project is written in Rust.

- Contact: Pierre-Etienne Meunier
- URL: https://nest.pijul.com/pijul_org/sanakirja

5.2. Thrussh

KEYWORD: Security

FUNCTIONAL DESCRIPTION: Purely asynchronous SSH library in Rust.

- Contact: Pierre-Etienne Meunier

5.3. Pijul

KEYWORDS: Data structures - Distributed systems

FUNCTIONAL DESCRIPTION: Patch-based distributed version control system using category theory. This solves a number of problems in other systems (such as Git), like:

- Scaling up to giant repositories (as used by Facebook, Google and Mozilla). - Being easy to understand and use, because based on a solid theory.

- Contact: Pierre-Etienne Meunier
- URL: <https://pijul.org>

5.4. SeqDesign

KEYWORDS: Chemistry - Molecular simulation

FUNCTIONAL DESCRIPTION: SeqDesign is a free energy calculation tool for DNA secondary structures. We use it to design sequences capable of self-assembling in a designed way.

- Contact: Pierre-Etienne Meunier

6. Partnerships and Cooperations

6.1. European Initiatives

6.1.1. FP7 & H2020 Projects

Woods applied for an ERC Consolidator award. The application was successful and begins in 2018.

6.2. International Research Visitors

6.2.1. Visits of International Scientists

David Doty (UC Davis) visited the team several times in 2017.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Selection

7.1.1.1. Reviewer

Both Damien Woods and Pierre-Étienne Meunier blind-reviewed papers for a number of computer science theory conferences (and do not wish to disclose further information related to blind peer review).

7.1.2. Journal

7.1.2.1. Guest Editor for Journal Special Issue

Damien Woods was guest editor (along with Yannick Rondelez, CNRS & ESPCI) of a special issue of the journal *Natural Computing* dedicated to the conference DNA22 (2016). The special issue will appear in 2018.

7.1.2.2. Reviewer - Reviewing Activities

Both Damien Woods and Pierre-Étienne Meunier reviewed papers for a variety of scientific journals (and do not wish to disclose further information related to blind peer review).

7.1.3. Invited Talks

- Woods. Centre for Research and Interdisciplinarity (CRI), Masterclass, Paris, France, 2017.
- Woods. Paris Sciences et Lettres (PSL institute), group on Origins and Conditions of Appearance of Life, Observatoire de Paris, France. 2017.
- Meunier, RustFest 2017 in Kiev, Ukraine, April 2017.

7.1.4. Contributed Talks (with no associated publication)

- Woods (speaker). FNANO 2017: 14th Annual Conference on Foundations on Nanoscience: Self-Assembled Architectures and Devices. Snowbird, UT. [Woods, Doty, Myhrvold, Hui, Zhou, Yin and Winfree. *Iterated circuit computation by DNA self-assembly in a field-programmable gate array architecture*]
- Woods. DNA23: The 23rd International Conference on DNA Computing and Molecular Programming, 2017. UT Ausin, Texas [Woods, Doty, Myhrvold, Hui, Zhou, Yin and Winfree. *Iterated circuit computation by DNA self-assembly in a field-programmable gate array architecture*]
- Meunier (speaker). DNA 23: The 23rd International Conference on DNA Computing and Molecular Programming. UT Austin Texas, USA. [Meunier, Woods. *The non-cooperative tile assembly model is not intrinsically universal or capable of bounded Turing machine simulation*]

7.1.5. Conference Programme Committees

- Woods. CiE 2017. Computability in Europe. Turku, Finland.
- Woods. DNA23: The 23rd International Conference on DNA Computing and Molecular Programming, 2017. UT Ausin, Texas
- Woods. UCNC 2017. Unconventional Computation and Natural Computation. Fayetteville, AR.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

Woods taught a 1-week course on molecular computing to Masters students at ENS Lyon (with N. Schabanel, Y. Rondelez, C. Moskalenko, L. Bellon). Course title: ER02: Molecular programming: Theory & Wet Lab Nano-Scale Computation. <https://www.irif.fr/~nschaban/2017-ER02/>

8. Bibliography

Major publications by the team in recent years

- [1] A. THUBAGERE, W. LI, R. JOHNSON, Z. CHEN, S. DOROUDI, Y. LEE, S. W. G. IZATT, N. SRINIVAS, D. WOODS, E. WINFREE, L. QIAN. *A cargo-sorting DNA robot*, in "Science", 2017, vol. 357, n^o 6356, eaan6558 p.

Publications of the year

Articles in International Peer-Reviewed Journals

- [2] L. KARI, S. KOPECKI, P.-É. MEUNIER, M. J. PATITZ, S. SEKI. *Binary Pattern Tile Set Synthesis Is NP-Hard*, in "Algorithmica", May 2017, vol. 78, n^o 1, pp. 1 - 46 [DOI : 10.1007/s00453-016-0154-7], <https://hal.inria.fr/hal-01659779>

International Conferences with Proceedings

- [3] D. WOODS, D. DOTY, T. A. ROGERS, D. SOLOVEICHIK, C. THACHUK. *Thermodynamic Binding Networks*, in "DNA 2017 - The 23rd International Conference on DNA Computing and Molecular Programming", Austin, United States, DNA 2017: DNA Computing and Molecular Programming, September 2017, vol. 10467, pp. 249-266, <https://arxiv.org/abs/1709.07922> [DOI : 10.1007/978-3-319-66799-7_16], <https://hal.inria.fr/hal-01662285>
- [4] D. WOODS, P.-É. MEUNIER. *The non-cooperative tile assembly model is not intrinsically universal or capable of bounded Turing machine simulation*, in "STOC 2017 - Theory Fest 49th Annual ACM Symposium on the Theory of Computing", Montréal, Canada, STOC 2017 Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, ACM, June 2017, pp. 328-341, <https://arxiv.org/abs/1702.00353> - This submission is an extended (arxiv) version of a STOC 2017 conference paper, with arxiv identifier arxiv:1702.00353 (or version arxiv:1702.00353v2 at the time of submission on HAL) [DOI : 10.1145/3055399.3055446], <https://hal.inria.fr/hal-01659730>

Other Publications

- [5] C. GEARY, P.-É. MEUNIER, N. SCHABANEL, S. SEKI. *Folding Turing is hard but feasible*, December 2017, <https://arxiv.org/abs/1508.00510> - working paper or preprint, <https://hal.inria.fr/hal-01659777>