

# Équipe Modèles et Algorithmes (MoA)

1er janvier 2013 → 30 juin 2018

Visite du comité Hcéres

13 février 2019



# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Présentation de l'équipe Modèles et Algorithmes

- ▷ **Stéphane Vialette** (resp.) et **Arnaud Carayol** (resp. adj.).
- ▷ L'équipe MOA regroupe environ **25 permanents** et **15 doctorants**.
- ▷ Activité principale en **informatique fondamentale** (CNU 27, CoNRS 6 et 7).
- ▷ La recherche est organisée en trois axes principaux :
  - ▷ **Automates et logique.**
  - ▷ **Algorithmique.**
  - ▷ **Linguistique pour le traitement automatique des langues.**

# Effectifs

| Date      | Ens.-cherch.             | Cherch.                | Doc.     | Ing.     | Émérites |
|-----------|--------------------------|------------------------|----------|----------|----------|
| jan 2013  | <b>18</b><br>5 PU, 13 MC | <b>4</b><br>3 DR, 1 CR | <b>8</b> | <b>2</b> | <b>2</b> |
| juin 2018 | <b>19</b><br>5 PU, 14 MC | <b>7</b><br>4 DR, 3 CR | <b>7</b> | <b>1</b> | <b>2</b> |

## Arrivées

**L. Bulteau**, CR 06 CNRS, 2015  
**É. Colin de Verdière**, DR 07, 2016  
**X. Goac**, PU UPEM, 2013  
**A. Hubbard**, MC UPEM, 2016  
**N. Francis**, MC UPEM, 2017  
**V. Jugé**, MC UPEM, 2017  
**M. Weller**, CR 51 CNRS, 2017

## Départs

**G. Blin** PU U. Bordeaux  
**M. Constant** PU U. Lorraine  
**J.-P. Heymann** retraite

# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Production scientifique

## Publications

| MoA              | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | Total |
|------------------|------|------|------|------|------|------|-------|
| Journaux         | 27   | 23   | 21   | 31   | 29   | 18   | 149   |
| Art. de synthèse | 0    | 3    | 2    | 0    | 1    | 2    | 8     |
| Actes de conf.   | 34   | 22   | 24   | 27   | 18   | 16   | 141   |
| Actes de work.   | 6    | 7    | 1    | 3    | 3    | 3    | 23    |
| Edition actes    | 2    | 2    | 0    | 0    | 1    | 0    | 5     |
| Dir. ouvrages    | 1    | 2    | 0    | 0    | 0    | 0    | 3     |
| Chap. ouvrages   | 4    | 1    | 1    | 1    | 0    | 2    | 9     |
| Total            | 74   | 60   | 49   | 62   | 52   | 41   | 338   |

# Logiciels

Environ 15 logiciels.

- ▷ **Unitex/GramLab**. Analyse de corpus.  
<https://unitexgramlab.org/fr>
- ▷ **ProPhyle**. Classification métagénomique.  
<https://prophyle.github.io>
- ▷ **NewtonGF**. Bibliothèque de calcul symbolique et numérique.  
<http://perso.ens-lyon.fr/bruno.salvy/software/the-newtongf-package/>

# Formation et collaborations

## Étudiants et visiteurs

- ▷ 7 thèses en cours (juin 2018).
- ▷ 15 thèses soutenues.
- ▷ Pas d'HDR.
- ▷ 6 ATER.
- ▷ 6 Postdocs.
- ▷ 1 Ingénieur.
- ▷ 16 Visiteurs

## Formation

- ▷ Master Parisien de Recherche en Informatique (MPRI) : **É. Colin de Verdière, P. Gambette, G. Kucherov, A. Labarre, C. Nicaud et S. Vialette.** .
- ▷ Master Logique Mathématique et Fondements de l'Informatique (LMFI) : **D. Caucal.**
- ▷ Master Informatique (ENS Lyon) : **A. Carayol.**
- ▷ Double diplôme de master informatique avec l'Université de Palerme depuis 2015 : **G. Rindone.**



# Formation et collaborations (suite)

## Formation à l'étranger

- ▷ École CIMPA, Madagascar, 2016. **C. Pivoteau** et **C. Nicaud**.
- ▷ U. Fédérale du Rio Grande Sud, Brésil. **É. Laporte**.
- ▷ Shanghai Normal University, **A. Labarre**.

## Formation doctorale

- ▷ Membre du conseil de l'école doctorale MSTIC. **C. Nicaud**.
- ▷ Membre conseil académique et du conseil de formation doctorale d'UPE. **P. Gambette**.
- ▷ Membre CA de l'Association nationale des docteurs. **P. Gambette**.
- ▷ Responsable des comités de suivi doctoral. **É. Laporte**.

## Écoles

- ▷ ALEA, 2014.
- ▷ EPIT, Algorithmique pour la bioinformatique, 2014.
- ▷ École d'hiver combinatoire et algorithmique de la convexité, 2015.
- ▷ RECOMB 2018,
- ▷ École d'été géométrie de basse dimension, 2018.

# Projets et contrats principaux

- ▷ ABS4NGS, Investissement d'avenir, **G. Kucherov** (PI), coordonné par l'Institut Curie, 2012-2017.
- ▷ Implication significative 10 projets ANR.
  - ▷ ANR PRC Aspag, **X. Goaoc** (coordinateur local), 2018-.
  - ▷ ANR JCJC AMIS, **A. Carayol** (porteur), **A. Meyer** et **C. Rispal**, 2011-2014.
  - ▷ ANR JCJC BIRDS, **G. Blin** (porteur), **I. Fagnot** et **S. Vialette**, 2011-2015.
- ▷ 1 FUI (2014-2017) Atos coordinateur, 3 financements de thèse.
- ▷ 6 PEPS.
- ▷ 4 PHC (Suisse, République Tchèque, Singapour et Allemagne).
- ▷ 2 super BQR.
- ▷ 1 PICS.

# Bilan

## Rayonnement

- ▷ Relations avec de nombreuses universités et instituts,
- ▷ Invitations et participation aux comités de programme de conférences internationales de référence de ses domaines de recherche.
- ▷ Publications dans des journaux de très bon niveau (★) : *Advances in Math.*, *Algorithmica*, *Discrete and C. Geometry*, *J. ACM*, *J. Discrete Algorithms*, *Siam Journal on Discrete Mathematics*, ...
- ▷ Participation comités éditoriaux : *Algorithms for Molecular Biology*, *European Journal of Combinatorics*, *Theoretical Computer Science* et *Lingvisticae Investigationes*.

## Distinctions

- ▷ Prix meilleur article (EGC et SoCG x 2),
- ▷ Prix meilleur article étudiant (CSR et ICALP),
- ▷ **C. David**, prix *Test of time award*, PODS 2016,
- ▷ **X. Goaoc**, membre junior IUF, 2014,
- ▷ **M. Crochemore**, docteur *Honoris causa*, U. Helsinki, 2013.

# Responsabilités

## Responsabilités locales

- ▷ **M.-P Béal.** Directrice du LIGM (-2015).
- ▷ **C. David.** Responsable L3 informatique (2012-).
- ▷ **X. Goaoc.** Responsable du Master (2014-2018).
- ▷ **T. Kyriacopoulou.** Directrice dpt Informatique IUT (2016-).
- ▷ **A. Meyer.** Responsable mention L Informatique (2014-).
- ▷ **C. Nicaud.** Directeur du LIGM (2015-), co-responsable du master *Mathématiques et informatique* (2018-).
- ▷ **D. Perrin.** Directeur de l'ESIEE (-2017).
- ▷ **C. Pivoteau.** Coordinatrice des matières IUT (2016-).
- ▷ **J. Thapper.** Responsable 1ère année ESIPE info. (2015-).

## Responsabilités nationales

- ▷ **M.-P Béal.** CNU 27 (-2015).
- ▷ **C. David.** CNU 27 (2015-).
- ▷ **S. Vialette.** Chargé de mission INS2I CNRS (2015-).

# Vie d'équipe

- ▷ Culture commune algorithmique.
- ▷ Séminaire hebdomadaire (mardi après-midi).
- ▷ Les chercheurs de l'équipe MoA sont tous regroupés dans le bâtiment Copernic (★).
- ▷ Préparation collégiale des candidats CNRS et EC.
- ▷ Co-encadrements de thèses.
- ▷ Réunions de travail.
- ▷ Projets étudiants.

# Outline

Présentation

Bilan

## Thèmes scientifiques

- Automates et logique

- Algorithmique

- Linguistique pour le traitement automatique des langues

Réponses à des questions du comité

Analyse SWOT

Projets et perspectives

# Thèmes scientifiques

## ▷ Automates et logique

- ▷ Bases de données (3 MC),
- ▷ Théorie algorithmique des modèles infinis (3 MC, 1 CR, 1 DR),
- ▷ Systèmes dynamiques (1 MC, 2 PU).

## ▷ Algorithmique

- ▷ Analyse d'algorithmes (1 MC, 1 PU),
- ▷ Géométrie algorithmique et combinatoire (1 MC, 1 PU, 1 DR),
- ▷ Algorithmique pour la bioinformatique (2 MC, 2 CR, 2 DR),
- ▷ Algorithmique des structures combinatoires (2 MC, 1 CR, 1 PU, 2 DR).

## ▷ Linguistique pour le traitement automatique des langues (2 IE, 1 MC, 2 PU)

**De nombreux posters illustrent quelques aspects de nos recherches.**

# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives



# Bases de données

## Participants

- ▷ **O. Curé**, MC UPEM
- ▷ **C. David**, MC UPEM
- ▷ **N. Francis**, MC UPEM

## Points clés

- ▷ Bases de données incomplètes.
- ▷ Sémantique formelle de langages de requêtes pour les graphes de donnée (*cypher*, Neo4j).
- ▷ Graphes de données RDF.
- ▷ Échange et cohérence de contraintes pour les données semi-structurées (graphes et arbres).

# Théorie algorithmique des modèles infinis

## Participants

- ▷ **A. Carayol**, CR 06 CNRS
- ▷ **D. Cauca**, DR 06 CNRS
- ▷ **V. Jugé**, MC UPEM
- ▷ **A. Meyer**, MC UPEM
- ▷ **C. Rispal**, MC UPEM

## Points clés

- ▷ Automates infinis et théorie des langages.
- ▷ Vérification de programmes d'ordre supérieur.
- ▷ Vérification de systèmes dynamiques temporisés.

# Théorie algorithmique des modèles infinis

## Longs mots évitant l'inévitable

*« Le problème de l'équivalence des automates à pile déterministes consiste à décider si deux tels automates acceptent le même langage et joue un rôle crucial en vérification (équivalence syntaxique des programmes récurifs).*

*La terminaison du meilleur algorithme connu pour ce problème repose sur l'apparition de motifs, appelés les motifs de Zimin, dans tout mot suffisamment long. »*

[A. Carayol and S. Göller. STACS. 2017.]

$$Z_1 = x_1$$

$$Z_2 = x_2 x_1 x_2$$

$$Z_3 = x_2 x_1 x_2 x_3 x_2 x_1 x_2$$

$$\vdots$$

$$Z_{n+1} = Z_n x_{n+1} Z_n$$

# Systèmes dynamiques

## Participants

- ▷ **M.-P. Béal**, PU UPEM
- ▷ **D. Perrin**, PU UPEM
- ▷ **D. Revuz**, MC UPEM
- ▷ **G. Rindone**, MC UPEM

## Points clés

- ▷ Systèmes dynamiques symboliques.
- ▷ Théorie des codes.
- ▷ Synchronisation d'automates.

# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Analyse d'algorithmes

## Participants

- ▷ **C. Nicaud**, PU UPEM
- ▷ **V. Jugé**, MC UPEM
- ▷ **C. Pivoteau**, MC UPEM

## Points clés

- ▷ Comportements des algorithmes dans les cas typiques.
- ▷ Mécanismes de prédiction de branchements.
- ▷ Textes aléatoires.



# Géométrie algorithmique et combinatoire

## Participants

- ▷ **É. Colin de Verdière**, DR 07 CNRS
- ▷ **X. Goaoc**, PU UPEM → PU U. Lorraine
- ▷ **A. Hubard**, MC UPEM

## Points clés

- ▷ Algorithmique des graphes plongés sur les surfaces.
- ▷ Plongeabilité des complexes simpliciaux dans les variétés.
- ▷ Convexité combinatoire : structures aléatoires et extrémales.

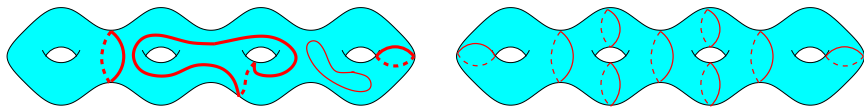


# Géométrie algorithmique et combinatoire

## Inégalités systoliques, cadres discrets et continus

« Nous relient les cadres discrets et continus et résolvons au passage une conjecture de 1993 : la largeur en arêtes d'une surface de genre  $g$  à  $n$  triangles est  $O\left(\log(g)\sqrt{n/g}\right)$ . »

[É. Colin de Verdière, A. Hubard and A. de Mesmay.  
Discrete and Computational Geometry, 2015.]



# Algorithmique pour la bioinformatique

## Participants

- ▷ **L. Bulteau**, CR 06 CNRS
- ▷ **P. Gambette**, MC UPEM
- ▷ **G. Kucherov**, DR 06 CNRS
- ▷ **A. Labarre**, MC UPEM
- ▷ **S. Vialette**, DR 06 CNRS
- ▷ **M. Weller**, CR 51 CNRS

## Points clés

- ▷ Séquençage haut débit.
- ▷ Phylogénie.
- ▷ Protonthérapie.
- ▷ Génomique comparative.

# Analyse de données et apprentissage (★)

## Positionnement

- ▷ Beaucoup de nos travaux se placent dans le contexte du *big data* : filtres de Bloom, sketching, ...
- ▷ L'apprentissage prend une place importante aujourd'hui en bioinformatique mais ce n'est pas notre cœur de métier.
- ▷ Les méthodes combinatoires restent encore largement dominantes en bioinformatique.
- ▷ Importance des structures de données efficaces.
- ▷ Beaucoup de nos travaux se placent en amont des applications.

# Algorithmique des structures combinatoires

## Participants

- ▷ **L. Bulteau**, CR 06 CNRS
- ▷ **M. Crochemore**, PU UPEM
- ▷ **I. Fagnot**, MC P7
- ▷ **G. Kucherov**, DR 06 CNRS
- ▷ **J. Thapper**, MC UPEM
- ▷ **S. Vialette**, DR 06 CNRS

## Points clés

- ▷ Algorithmique du texte.
- ▷ Algorithmique de graphes et complexité paramétrée.
- ▷ Satisfaction de contraintes.

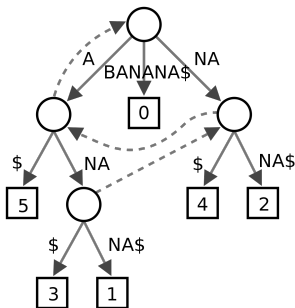
# Algorithmique des structures combinatoires

## Mise à jour en temps réel d'un index pour la recherche de motifs

« Dans la version online de la recherche d'un motif  $P$ , nous recevons un flot de caractères et nous devons pouvoir calculer à tout moment les occurrences de  $P$ .

Nous avons proposé une structure de données (basée sur l'arbre des suffixes) permettant : (i) une mise à jour très rapide, et (ii) de calculer toutes les occurrences en temps  $O(|P| + \text{occ})$ . »

[G. Kucherov and Y. Nekrich. Algorithmica. 2017.]



# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Linguistique pour le traitement automatique des langues

## Participants

- ▷ **M. Constant**, MC UPEM → PU U. Lorraine
- ▷ **J.-P. Heymann**, IE CNRS → Retraite
- ▷ **P.T. Kyriacopoulou**, PU UPEM
- ▷ **É. Laporte**, PU UPEM
- ▷ **T. Nakamura**, IE CNRS

## Points clés

- ▷ Description syntaxique et sémantique.
- ▷ Production de bases de données lexicales et grammaticales.
- ▷ Analyse syntaxique probabiliste.
- ▷ Extraction d'informations.

# Linguistique pour le traitement automatique des langues

## VerbNet

« Nous avons contribué au développement du VerbNet du français à partir du modèle du VerbNet de l'anglais. Le résultat de ce travail a une couverture quasiment équivalente à celle du VerbNet de l'anglais et encode les mêmes informations syntaxiques et sémantiques. »

[L. Danlos, Q. Pradet, L. Barque, T. Nakamura and M. Constant. Traitement Automatique des Langues, 2016.]

### Classe settle-89

- Membres de VerbNet: **accorder s'** **arranger s'** **entendre s'** **allier s'**
- Roles :Agent(<-organization|animate), Co-Agent(<-organization|animate), Goal

#### Canonique

|            |                                           |
|------------|-------------------------------------------|
| Exemple    | Luc s'est entendu avec Marie sur le prix. |
| Surfacique | NP V PP.co-agent PP.goal                  |
| Syntaxe    | Agent V (avec) Co-Agent (sur) Goal        |
| Sémantique | agree(result[E], Agent, Co-Agent, Goal)   |

#### NP V PP.co-agent pour V-Inf

|            |                                                      |
|------------|------------------------------------------------------|
| Exemple    | Luc s'est entendu avec Marie pour proposer un projet |
| Syntaxe    | Agent V (avec) Co-Agent (pour) Goal<-VAgent-Inf>     |
| Sémantique | agree(result[E], Agent, Co-Agent, Goal)              |



# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Collaborations au sein de MoA (★)

## Avec publications communes

- ▷ **A. Carayol** et **C. Nicaud**. CIAA (2017).
- ▷ **X. Goaoc** et **C. Nicaud**. European Journal of Combinatorics, 2013.
- ▷ **P. Gambette**, **É.Laporte** et **P.T. Kyriacopoulou**. RNTI-SHS, 2014 et Colloque AnimalHumanité, 2016.

## Collaborations

- ▷ **C. Athénosy**, **A. Carayol**, **C. David**, **P. Heller**, **R. Kyriakoglou**, **C. Nicaud** et **C. Pivoteau**. Organisation CIAA (2017).
- ▷ **A. Carayol** et **C. Nicaud**. Encadrement M2.
- ▷ **C. Nicaud** et **S. Vialette**. ISAAC (2013) et Encadrement M2.
- ▷ **A. Carayol** et **S. Vialette**.
- ▷ **A. Carayol**, **C. David**, **N. Francis** et **V. Marsault**.

# Évolutions structurelles à moyen/long/très long terme (★)

*« La diversité des thèmes de recherche permet en effet des échanges continuels et fructueux qui participent à la grande polyvalence de l'équipe et favorisent grandement l'intégration des nouveaux chercheurs. »*

- ▷ Spectre large.
- ▷ Culture informatique commune (algorithmique et automates).
- ▷ Le découpage en 3 *thèmes* principaux n'est pas rigide.
- ▷ Éviter un cloisonnage qui n'existe pas aujourd'hui.
- ▷ **Linguistique pour le traitement automatique des langues.**
- ▷ C'est une question que nous évoquons régulièrement et librement.

# Priorités thématiques de recrutement par rapport aux thématiques isolées (★)

## **Linguistique pour le traitement automatique des langues**

- ▷ Concours en général ouvert à tous les thèmes de l'équipe.
- ▷ Ce thème est représenté par deux PU et un IE CNRS.
- ▷ Soutien financier du laboratoire.
- ▷ Les candidats MC TAL n'étaient pas au niveau ou ne sont finalement pas venus.

## **Automates et logique**

- ▷ **V. Jugé**, 2017.
- ▷ **N. Francis**, 2017.
- ▷ **V. Marsault**, 2018.

## Moyens de calculs (★)

- ▷ Les moyens de calcul sont essentiellement locaux.
  - ▷ Intel(R) Xeon(R) CPU E5640 @ 2.67GHz, 8 cœurs, 256 Go.
  - ▷ Intel(R) Xeon(R) CPU E5-2630 v2 @ 2.60GHz, 24 cœurs, 64 Go.
  - ▷ Serveurs de calculs (2 x Intel Xeon E5-2683 V4 16 coeurs 256G RAM chacune).
- ▷ Pas de cluster de calcul.

# Outline

Présentation

Bilan

Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

Réponses à des questions du comité

**Analyse SWOT**

Projets et perspectives

# Analyse SWOT

## Points forts

- ▷ Richesse des thématiques.
- ▷ Attractivité forte.
- ▷ Publications.
- ▷ Reconnaissance nationale et internationale.
- ▷ Recrutements externes.

## Points à améliorer

- ▷ Absence d'ERC.
- ▷ HDR.
- ▷ Collaboration avec le thème **Linguistique pour le traitement automatique des langues.**

# Analyse SWOT (suite)

## Risques liés au contexte

- ▷ Recrutement de doctorants.
- ▷ Géométrie algorithmique.
- ▷ Implication dans l'administration de l'enseignement.
- ▷ Déménagement.

## Possibilités liées au contexte

- ▷ Labex Bézout.
- ▷ Laboratoire Analyse et Mathématiques Appliquées (LAMA).
- ▷ UMI Relax, Chennai, Inde.



# Outline

## Présentation

## Bilan

## Thèmes scientifiques

Automates et logique

Algorithmique

Linguistique pour le traitement automatique des langues

## Réponses à des questions du comité

## Analyse SWOT

## Projets et perspectives

# Projets et perspectives

## Collaborer

- ▷ Labex Bézout.
- ▷ Intensifier et diversifier les collaborations entre les différents thèmes de l'équipe MoA et les autres équipes du LIGM.
- ▷ Intensifier et diversifier les collaborations avec le LAMA.

## Consolider

- ▷ Bases de données. **V. Marsault**, → CR CNRS.
- ▷ Géométrie algorithmique et combinatoire. **X. Goaoc**, PU UPEM → PU U. Lorraine.
- ▷ Linguistique pour le traitement automatique des langues.

## Développer et accompagner

- ▷ Humanités numériques.
- ▷ Algorithmique et complexité.
- ▷ Requêtes dans les bases de données massives, LRI, U. Paris-Sud 11, IFB, LIGM.