Master Informatique - Parcours Informatique Théorique

R. Giroudeau et M. Montassier

31 mars 2015

1 Liste des UE

UE	Responsable
Complexité algorithmique	Bruno Durand
Théorie de l'information	Grégory Lafitte
Méthodes et algorithmique probabilistes	Anne-Elisabeth Baert
Calculabilité	Bruno Durand
Méthodes approchées	Rodolphe Giroudeau
Graphes et structures	Stéphane Bessy
Algorithmique distribuée	Rodolphe Giroudeau
Raisonnement par contraintes	Rémi Coletta
Calcul formel, codes et cryptographie	Pascal Giorgi
Optimisation combinatoire	Jean-Claude Konig
Graphes, algorithmique et complexité	Mickael Montassier
Théorie des langages et pavages	Gwenaël Richomme

2 Résumés

Complexité algorithmique Le cours présentera les fondements des notions de complexité des algorithmes. On s'intéressera à mesurer la complexité en temps et en espace sur un modèle de calcul, et on montrera que les modèles de calcul usuels sont polynomialement équivalents. Ensuite, on étudiera la complexité dans le pire des cas et en moyenne sur divers problèmes algorithmiques, en espace et en temps. La notion de réduction entre problèmes sera présentée afin d'établir les hiérarchies de la complexité classique. La NP-complétude sera abordée en détails afin que tout étudiant sache montrer la NP-complétude d'un problème élémentaire. Les exemples seront pris dans le domaine des graphes, dans celui des pavages, et de façon générale en algorithmique élémentaire. Enfin les jeux algorithmiques de type Arthur-Merlin et les preuves interactives seront abordées pour conclure sur le théorème IP=Pspace.

Théorie de l'information Ce module contient trois thèmes principaux qui s'entremêlent tout au long du module et qui présentent trois approches complémentaires et intimement liées de l'idée de quantifier le contenu en information d'un message ou d'un ensemble de messages. Le premier est la théorie de Shannon, avec les notions d'entropie, de codage de l'information et de modèles de communication. En particulier cette partie couvre tous les codes classiques et théorèmes associés, tels les différents codes de Huffman, Shannon-Fano, le théorème de l'inégalité de Kraft-McMillan et l'algorithme de Sardinas-Patterson. Le second thème est la théorie et la pratique des algorithmes et techniques de compression. Le troisième thème est une introduction à la théorie algorithmique de l'information de Kolmogorov. Ces deux derniers thèmes sont présentés en étant mis en relation avec la théorie de Shannon, ce qui nous permet d'étudier les limites de chacune des approches.

Méthodes et algorithmique probabilistes Ce module s'oriente autour de quatre axes :

Axe 1 Les fondements de l'analyse en moyenne avec une introduction à l'analyse en moyenne par séries génératrices et l'utilisation des propriétés analytiques de ses séries, les dénombrements asymptotiques.

Axe 2 Les processus Markovien : étude et utilisation des processus poissonniers et Markoviens, théorie des files d'attente.

Axe 3 Les graphes et autres structures aléatoires : méthode probabiliste, phénomènes de transition de phases, arbres et marches aléatoires et processus de branchement, applications aux réseaux (graphes petits mondes, graphes aléatoires géométriques).

Axe 4 Les algorithmes randomisés : algorithmes randomisés vs approximation, application à l'optimisation combinatoire et aux structures de données. Méthode de Monte Carlo.

Calculabilité La théorie de la calculabilité tourne autour de deux notions phares : les modèles de calcul et les réductions entre problèmes. Le cours commence par une présentation et des rappels sur différents modèles de calcul ayant diverses puissances de calcul, allant des automates finis aux machines de Turing en passant par la récursion primitive. Nous étudions alors différentes notions de réductions d'un problème à un autre pour ensuite présenter différents ensembles canoniques et les théorèmes principaux de la calculabilité naïve, tels l'ensemble diagonal, les ensembles d'indices de fonctions, les théorèmes de point-fixe de Kleene, le théorème de Rice, l'isomorphisme de Rogers, etc. Le cours se termine par un aperçu de la structure induite par les réductions sur les ensembles d'entiers quelconques et les ensembles récursivement énumérables. En particulier nous construisons différentes solutions au problème de Post.

Méthodes approchées Maitriser les techniques pour résoudre au mieux (méthodes exactes, méthodes approchées) des problèmes classés comme difficile au sens de la théorie de la complexité. Nous aborderons : branch and bound (algorithme A*), branch and cut, algorithmes approchés avec garantie de performance, programmation dynamique, PTAS, FPTAS, recherche locale,...

Algorithmique distribuée Le but de ce module est d'introduire un nouveau paradigme : l'algorithmique distribuée. Un algorithme réparti (ou distribué) est généralement un algorithme réparti sur plusieurs sites. Chaque site calcule et produit des résultats qui sont transmis à d'autres sites via le réseau. Dans ce module nous allons concevoir et analyser des algorithmes distribués, en présentant quelques grands problèmes liés à ce nouveau cadre : temps logique versus temps physique, élection (c'est à dire la possibilité de revenir temporairement à un système maître-esclave), le problème de la terminaison (comment garantir qu'une application distribuée est bien terminée avec aucun processus actif et aucun message en transit), les généraux byzantins (comment garantir une fiabilité des systèmes), le problème de la tolérance aux pannes (comment le système peut continuer à fonctionner malgré quelques pannes), le problème d'exclusion mutuelle (comment accéder à une ressource partagée critique). La complexité d'un algorithme sera mesurée en fonction du nombre de messages échangés.

Graphes et structures Nous présenterons à travers ce cours des résultats classiques de la théorie des graphes. Notre objectif sera d'appréhender des preuves parfois complexes manipulant des objets non triviaux. Nous reviendrons sur les raisonnements par l'absurde, par récurrence, par réduction, ... Entre autres, nous nous intéresserons aux :

- problème du couplage dans les graphes bipartis (théorème de Hall), dans les graphes en général (théorème de Tutte), flots...
- problèmes de colorations de graphes avec notamment le problème des quatre couleurs et plus précisément le théorème des cinq couleurs (argument de Kempe (preuve par l'absurde), Lovász (preuve par réduction), Thomassen (preuve par induction))...
- problèmes sur les graphes orientés avec les théorèmes de Gallai-Roy, Gallai-Milgram, d'Edmonds.

Également seront abordées les problématiques de connectivité, de décompositions, d'hypergraphes,...

Graphes, algorithmique et complexité Ce module se situe à l'interface de la combinatoire des structures discrètes, de l'algorithmique et de la théorie de la complexité. Il comportera donc plusieurs volets complémentaires tels que :

- *Théorie des graphes.* L'objectif est ici de présenter les développements récents et les techniques émergentes pour les problèmes classiques (e.g. coloration, domination. . .) mais aussi les théories des décompositions de graphes et des mineurs ou encore des graphes géométriques / topologiques.
- *Algorithmes combinatoires.* Les propriétés structurelles des graphes (décomposition, plongement dans les surfaces) permettent souvent, sur des classes de graphes particulières, soit de prouver l'existence d'algorithmes efficaces ou de concevoir de tels algorithmes. Nous étudierons essentiellement les méthodes algorithmiques (polynomiales, paramètrées, exponentielles) permettant de calculer des solutions exactes à des problèmes d'optimisation combinatoire.

— *Algorithmique paramétrée.* Au delà de la théorie de la complexité classique, nous nous intéresserons en particulier à la complexité paramétrée et aux méthodes de pré-processing polynomial (kernelization) en lien avec la logique (model checking) et les automates.

Théorie des langages et pavages Les thèmes abordés dans ce cours sont amenés à varier chaque année.

— *Pavages.* L'étude des pavages, notamment par des tuiles colorées, intéressent diverses communautés, les logiciens ont les premiers développé des théories générales, notamment sur les problèmes de pavabilité. Curieusement, ces résultats et techniques sont liés à la notion même de calcul : un pavage a l'air d'être un objet géométrique dévolu aux mathématiques ou à la physique des cristaux mais en fait leur structure éventuellement compliquée vient de ce qu'ils expriment naturellement de l'algorithmique. Les pavages peuvent aussi être vus comme des mots en dimension 2 vérifiant des contraintes locales, mais la théorie des langages habituelle en dimension 1 ne se généralise pas aisément à la dimension 2.

— *Langages.* Dans de nombreux domaines de l'informatique (algorithmique du texte, compression, géométrie discrète, langages formels et automates, codage, gestion de files d'attente, ...) ou d'autres disciplines (placement de robots, algorithmique du génome, théorie des nombres, systèmes dynamiques, ...), une compréhension de la combinatoire des mots ou de langages (ensembles de mots) définis par des contraintes particulières ou par des mécanismes engendrant est nécessaire. Le cours illustrera quelques exemples en liaison avec des problèmatiques d'actualités.

Raisonnement par contraintes Le raisonnement par contraintes permet de résoudre de nombreux problèmes combinatoires (emplois du temps, affectation de personnel, etc.). Une solution à un problème de satisfaction de contraintes (ou CSP) est une affectation de valeurs à des variables soumises à des restrictions (= contraintes) sur les combinaisons de valeurs qu'elles peuvent prendre. Nous aborderons rapidement la modélisation de problèmes en CSP et décrirons les principales techniques pour résoudre un CSP. Nous insisterons sur la propagation de contraintes et nous étudierons le concept de contrainte globale, utilisé dans tous les logiciels de résolution de contraintes. Ensuite nous présenterons différentes extensions apportées au modèle CSP pour répondre à des questions plus complexes comme la recherche de la "meilleure" solution selon certains critères. Une partie du cours sur la programmation par contraintes sur intervalles montrera comment ces techniques peuvent s'appliquer dans le monde continu pour résoudre des systèmes de contraintes d'égalités et d'inégalités sur les nombres réels et effectuer de l'optimisation globale sous contraintes.

Calcul formel, codes et cryptographie L'objectif de ce module est de présenter les concepts et outils mathématiques de la théorie de l'information, notamment la cryptographie et la théorie des codes correcteurs. Dans un monde de plus en plus interconnecté, la cryptographie est devenu un ingrédient indispensable à la sécurisation des données et des communications. Elle s'attache à protéger des messages pour assurer confidentialité, authenticité et intégrité. La théorie des codes permet de reconstruire des messages en cas de modifications (erreurs) pendant la transmission sur un canal bruité (par ex. communications par satellites). Ces deux disciplines de la théorie de l'information s'appuient fortement sur des notions et outils de la théorie des nombres.

La première partie du cours est consacrée à une introduction à ces outils fondamentaux d'arithmétique et de calcul formel. Nous présentons les notions de divisibilité, les algorithmes de multiplication rapide et de pgcd, les anneaux et corps finis, ainsi que quelques notions d'algèbre linéaire. La seconde partie est dédiée aux applications de ces concepts à la cryptographie asymétrique moderne et aux codes correcteurs d'erreurs. Nous présentons aussi les méthodes permettant d'attaquer les problèmes difficiles sous-jacents comme la factorisation entière et le calcul de logarithme discret. Pour l'ensemble du module, des applications sont proposées permettant de mettre en œuvre les concepts vus en cours.

Optimisation combinatoire Nous approfondirons tout d'abord les techniques classiques de conception d'algorithmes/schémas d'approximation (analyse d'algorithmes gloutons, programmation linéaire et arrondis, recherche locale, approximation duale, techniques de "guess",..), ainsi que des résultats de base sur l'inapproximabilité (typiquement introduction / transfert de gap). Nous nous intéresserons ensuite au cas de l'analyse on-line (les données ne sont pas connues à l'avance), de l'analyse multicritère (plusieurs critères antagonistes sont considérés). Enfin, nous présenterons quelques résultats d'approximation

polynomiale liés à l'utilisation du rapport différentielle, alternative à la mesure classique. Ces techniques seront dans un premier temps introduites sur les problèmes de base du domaine (vertex cover and set cover, independent set, sac à dos, problèmes de packing,...). Puis, nous verrons comment les utiliser sur des problèmes plus spécifiques dans les réseaux et dans l'ordonnancement. Entre autres, nous étudierons certains problèmes de recouvrement sous contraintes (recouvrement des graphes, problème de Steiner sous contraintes de degré des sommets, multiples contraintes dans l'optimisation, ...).

3 Parcours Informatique Théorique

S1.	Complexité algorithmique	MI	5 ECTS
	Théorie de l'information		5 ECTS
	Algorithmique du texte (parcours BCD)	MI	5 ECTS
	Méthodes et algorithmique probabilistes	MI	5 ECTS
	Anglais		5 ECTS
	1 option parmi :		5 ECTS
	Réseaux et télécommunications (parcours AIGLE)		
	Ingéniérie logicielle (parcours AIGLE)		
	Compilation et interprétation (parcours AIGLE)		
	Base de données avancées (parcours DECOL)		
	Intelligence artificielle (parcours DECOL)		

S2.	Calculabilité	MI	5 ECTS
	Méthodes approchées	MI	5 ECTS
	Graphes et structures	MI	5 ECTS
	Algorithmique distribuée	MI	5 ECTS
	TER		5 ECTS
	1 option parmi :		5 ECTS
	Spécifications formelles, vérification, validation (parcours AIGLE)		
	Technologies de la langue (parcours DECOL)		
	Algorithmes d'exploration de mouvement (parcours IMAGINA)		
	Algorithmes géométriques et géométrie discrète (parcours IMAGINA)		

S3.	6 modules à choisir parmi les 7 suivants : 6×5 ECTS	
	Graphes, algorithmique et complexité	MI
	Théorie des langages et pavages	MI
	Raisonnement par contraintes	MI
	Calcul formel, codes et cryptographie	MI
	Théorie des bases de connaissances (parcours DECOL)	MI
	Algorithmique pour la bioinformatique (parcours BCD)	MI
	Optimisation combinatoire	MI

S4.	Étude bibliographique	5 ECTS
	Stage académique	20 ECTS
	Enjeux juridiques et déontologiques de l'informatique	5 ECTS

Remarque Les modules estampillés MI sont communs au parcours **Mathématiques Informatique**.