

Rapport de césure à l'IMDEA Madrid

Du 4 Avril au 12 Août 2011

Florence Clerc

Introduction : Présentation du contexte

L'informatique est un domaine qui permet des développements intéressants dans tous les autres secteurs. Aujourd'hui présente partout, elle est de moins en moins comprise malgré un potentiel considérable. C'est par les cours d'informatique que j'ai pu m'en apercevoir. Mon année de césure a donc été consacrée à préparer le M1 du Master Parisien de Recherche en Informatique. Centré sur l'informatique théorique, il permet de revenir aux bases de l'informatique, un sujet qui n'est au mieux que survolé à Centrale mais qui est nécessaire pour l'utiliser intelligemment dans nombre de cas. Je renvoie le lecteur vers mon précédent rapport de césure et vers le site du MPRI (<http://mpri.master.univ-paris7.fr/>) pour en savoir d'avantage. Mon second stage de césure s'est effectué en tant que stage de second semestre dans ce cadre. Je remercie donc M.Mauborgne pour m'avoir accueillie, M.Cousot pour m'avoir recommandée et Mme Delais pour m'avoir aidée tout au long du processus.

1 l'IMDEA : un laboratoire d'excellence

l'IMDEA est l'institut madrilène pour les hautes études. Il est divisé en plusieurs organisations dont l'IMDEA software fondé en 2006 dont le but est de poursuivre des recherches approfondies dans les méthodes, les langages et les outils permettant de réduire tous les coûts liés au développement de logiciel (que ce soit en efficacité, argent, sûreté etc). L'IMDEA software n'a donc pas vocation à générer de l'argent, mais travaille toutefois en lien avec des entreprises en restant de la recherche académique. Il correspond au concept de laboratoire d'excellence lancé récemment en France.

Je parlais précédemment d'applications de l'informatique dans tous les domaines, l'IMDEA se concentre sur quelques uns seulement qui à eux seuls sont cruciaux pour le développement : les systèmes embarqués et temps-réel (avec ARTEMIS notamment), les systèmes critiques (scanners IRM, voitures, centrales nucléaires etc), la sécurité et les architectures orientées-service. Ses principaux axes de recherche sont à la fois la modélisation, la sécurité de logiciels et de systèmes, la vérification et la validation de logiciels et les outils de programmation et d'optimisation.

Malgré sa création plutôt récente (moins de 5 ans), l'IMDEA attire de nombreux chercheurs du monde entier : tous les continents sont représentés. Cette richesse lui a permis sur l'année 2010 de publier 43 articles au total, 5 manuels et 3 doctorats pour 5 enseignant-chercheurs, 9 assistants, 7 post-doctorants,

10 doctorants et 5 stagiaires. L'institut se développe et pourra prochainement emménager dans de nouveaux locaux plus grands et plus adaptés à ses besoins.

2 la recherche académique : sujet du stage, environnement de travail et compétences requises

Depuis le début de mon master, j'ai eu la possibilité de m'intéresser à énormément de sujets. Cette expérience à l'IMDEA n'a fait que me convaincre un peu plus de l'immense richesse de la recherche académique par rapport à l'entreprise et me pousse encore plus qu'auparavant à clore mon parcours à centrale pour pouvoir compléter mon cursus.

J'ai eu l'opportunité de m'intéresser plus précisément à la représentation de la mémoire dans le cadre de l'exécution d'un programme, en m'appuyant sur l'analyse statique. La principale idée derrière l'analyse statique consiste à, sans exécuter le programme et uniquement en l'analysant, en déduire un certain nombre de propriétés comme par exemple le domaine dans lequel on est sûr de ne pas avoir d'erreur, les espaces mémoires modifiés et non modifiés etc. Cette approche est particulièrement intéressante dans certains domaines où l'erreur n'est pas permise comme l'aérospatiale afin de pouvoir tester le code sans avoir à le faire tourner sur tous les cas possibles.

Dans le cadre de ce stage, il s'agissait donc de pouvoir représenter la mémoire par un graphe afin de pouvoir en tirer des invariants permettant ainsi de déduire la façon dont le programme se comportait. De nombreuses études ont été faites sur le sujet, la plus proche étant l'approche TVLA basée sur de la logique à trois valeurs. Malheureusement, cette méthode a un gros inconvénient : à chaque étape de nouveaux graphes sont créés pouvant ainsi déboucher sur plus de mille graphes pour un simple problème, ceci aboutit bien entendu à des calculs pouvant être très longs.

M.Mauborgne a donc développé une nouvelle approche, très similaire à TVLA, mais fondamentalement différente dans la façon d'aborder le problème. Détailons rapidement : un noeud représente plusieurs cases mémoires, le nombre exact est représenté par la cardinalité (un ensemble d'entiers), une arête entre deux noeuds représente juste une exécution possible (on s'appuie pour cela sur la notion de "memory state"). Une arête a bien entendu une cardinalité. Les noeuds peuvent porter des propriétés. On dispose de trois opérations de bases : focus, assign et union :

- Focus permet de s'assurer qu'en suivant une exécution dans le graphe on arrive sur un seul noeud de cardinalité 1 en créant éventuellement des noeuds pour les noeuds de cardinalité supérieure à 1 (par exemple un noeud de cardinalité {2,3} est séparé en plusieurs noeuds dont un de cardinalité {1}) ou des graphes quand plusieurs exécutions sont possibles qui amènent à des noeuds différents.
- Assign permet de représenter des assignments dans le programme sur le graphe.
- Union permet de faire l'union de graphes sans perdre d'information : beaucoup d'idées très différentes ont été abordées faisant intervenir du

partitionnement de graphe, des catégories, la notion de simulation et des probas.

C'est cette dernière étape qui permet d'espérer un meilleur coût que dans le cas de TVLA.

J'ai donc eu l'opportunité de formaliser ces idées. Bien que leur concept semble relativement simple, leur mise en oeuvre soulève de nombreux problèmes, ainsi l'union de graphe est plus complexe qu'initiallement prévue puisqu'il s'agit de ne pas perdre trop d'information. Il faut trouver un juste équilibre afin de ne pas perdre d'information importante dans le cadre souhaité mais sans toutefois garder une trop grosse structure couteuse en terme de performances. Ceci n'est qu'un exemple parmi d'autres, citons en vrac : le traitement des arêtes dans l'opération focus, celui des cardinalités, des propriétés, les noeuds de cardinalité possiblement nulle dans focus, le type de données dans assign etc. J'ai eu la chance que M.Mauborgne me laisse travailler en autonomie tout en restant ouvert à toutes mes suggestions (et il y en a eu plusieurs d'assez farfelues) et à toutes mes questions (des plus triviales aux plus complexes).

J'ai pu m'apercevoir que le métier de chercheur nécessitait à la fois rigueur, créativité, curiosité et candeur. Ce que j'ai appris lors de ce stage m'a permis de confirmer que je souhaite travailler dans la recherche académique.

3 Aspects culturels du stage

Il faut d'abord distinguer deux environnements totalement différents culturellement parlant.

Tout d'abord le laboratoire : j'ai eu l'occasion d'y rencontrer des gens d'origines et de parcours très différents et ayant des centres d'intérêt très variés. La langue la plus parlée était l'anglais suivie par l'espagnol puis le français et l'allemand. Les buffets surnommés "theory lunch" car suivi d'une conférence en général sur des travaux effectués à l'IMDEA sont une façon de se rencontrer et donc de discuter de sujets extrêmement variés allant des systèmes d'éducation à la nourriture.

Ensuite dans la vie de tous les jours : ce sont de nombreux légers détails qui font la différence par rapport à Paris. Ces petits détails sont peut-être liés au fait d'être en dehors de Madrid plutôt que dans la ville même. Peut-être sont-ils liés au climat, et j'aurais pu dire, dans ce cas, la même chose si j'avais effectué ce stage à toulouse ou marseille. Quoi qu'il en soit, l'environnement semble beaucoup moins stressé qu'à Paris. Les gens discutent longuement entre eux, préfèrent attendre l'escalator plutôt que prendre les escaliers etc.

Conclusion et projet professionnel

On l'aura compris : j'ai beaucoup apprécié ce stage qui a été enrichissant à la fois en terme de connaissances abordées et de projet personnel. Tout d'abord sur les connaissances, j'ai pu découvrir un sujet que je ne connaissais pas : l'abstraction et l'analyse statique. Cela m'a également permis d'utiliser des notions que j'avais vues au cours de l'année précédente.

Enfin, je suis impatiente de finir mes études à Centrale et mon master d'informatique théorique l'année prochaine afin de pouvoir enchaîner sur une

thèse.