

Entretien avec Romain Reuillon

11 Avril 2017

JR Donc merci Romain Reuillon de nous accueillir pour cette interview - **RR** tu fais ça en mode radio du coup - **JR** qui sera à propos d'OpenMole, et peut être plus large. Donc juste pour introduire un peu le contexte, moi le point de vue dans lequel j'essaierai de positionner l'interview, c'est une réflexion un peu sur les liens entre outils, méthodes, et thématique, plutôt en géographie, mais je pense que les discussions seront plus larges, et d'autant plus intéressantes, et donc tout le lien entre la genèse de la théorie évolutive et la genèse d'OpenMole, comment les deux ont été reliés, ou pas, tout ça. C'est assez large. Donc j'ai préparé une grille, pareil c'est indicatif. Une première question assez globale pour commencer, est-ce que tu pourrais donner un peu un pitch, une vision synthétique d'OpenMole, positionnement, utilisation, personnes. . .

RR Pour moi OpenMole, c'est un logiciel, qui vise à, qui permet d'explorer des modèles de simulation. Donc en fait son but c'est d'instrumenter des exécutions de composants logiciels, qui sont en mode boîte noire, c'est à dire on regarde les entrées et les sorties, et essayer de trouver des régularités dans les comportements via l'analyse du résultat des sorties. Pour ça y'a, ça permet aux gens d'inclure leur composante logicielle dans ce truc là, le deuxième aspect c'est les méthodes, ça va fournir par exemple des algos génétiques pour trouver ces généralités, alors y'a des méthodes très simples, des plans directs, sampler directement l'espace des paramètres, pour essayer de regarder ce qu'il y a en sortie, et essayer d'extraire de la connaissance a posteriori par des traitements de données, et y'a des méthodes pour trouver des patterns, des régularités dans les comportements, sur lesquelles on a travaillé et qu'on a inclus dans la plateforme, et amener à développer et inclure, et rendre disponible à la communauté. Le troisième aspect de ça, c'est que pour être efficace, il faut exécuter de nombreuses fois le composant, souvent un modèle de simulation, dont l'exécution peut prendre plusieurs secondes, plusieurs heures, plusieurs jours même, donc l'exécuter sur des environnements de calcul distribué, pour permettre de mener à bien ces méthodes, en calcul distribué, type ce qu'on a vu là grille de calcul, des milliers de machines dispersées dans le monde entier, des clusters plutôt des centaines de machines au niveau d'un laboratoire unique ou d'un centre de calcul local, ou des environnements type cloud qui permettent d'acheter de la puissance de calcul chez des fournisseurs commerciaux, ou des calculs pair-à-pair basés sur des blockchain, on devrait bien pouvoir mettre en oeuvre dès que ça existera, les gens travaillent dessus actuellement, y'aura sûrement un marché de la puissance de calcul peer-to-peer, on espère bien qu'OpenMole permettra d'accéder à ce genre d'environnement.

JR D'accord donc c'est bien trois composantes, l'aspect embedded du modèle, les méthodes, et le calcul intensif.

RR On a beaucoup gambergé sur la façon de, OpenMole c'est pas un nouveau traitement de texte, c'est pas un logiciel qu'on a fait pour copier, on va fait faire ça mais en mieux, c'est vraiment quelque chose qui est sorti d'un besoin que moi, on a compris partiellement au départ, à force d'interagir avec des gens on a compris le besoin récurrent, et on étendu au fur et à mesure notre vision de ces besoins là, et du coup ça prend énormément de temps à formuler, à comprendre ce qui est essentiel là dedans, l'axe sur lequel on est d'accord depuis le début de l'année et qui marche le mieux, pour moi c'est ces trois aspects là, on spécialise OpenMole, c'est vraiment pour l'exploration de modèles, après il s'avère qu'il peut y avoir d'autres utilisations dérivée de cette chose là, notamment du traitement de données, la recherche de paramètre dans des algos de traitement d'image, on peut imaginer plein d'utilisations dérivées, mais ça, maintenant c'est des utilisations dérivées, qu'on va pas mettre au coeur de notre communication, le coeur de la communication c'est vraiment l'exploration de modèles, exploration de modèles de simulation.

JR D'accord. Donc au niveau du positionnement, du public auquel c'est destiné, a priori c'est assez large ?

RR Le public le plus évident, le public premier c'est la communauté des Systèmes Complexes, enfin les modélisateurs dans les systèmes complexes. - **JR** D'accord, donc des gens qui ont déjà une connaissance technique - **RR** Ca c'est la communauté historique en fait, c'est des gens comme toi, qui ont un problème de modélisation, on dit non tu vas pas tester ça à la main, regarde y'a des choses qui existent, et c'est intégré dans le logiciel. Ou alors a tu fais ça, c'est presque faire ça, attends on va essayer comment intégrer dans le logiciel et composer ça avec l'environnement de calcul distribué. Et après on essaye d'élargir au maximum, j'aimerais bien élargir ça, au niveau académique, ça resterait dans les modélisateurs, mais pas forcément systèmes complexes, là ma réponse va être moins synthétique, mais quand on regarde OpenMole, les modèles de simulation, comment plus dire, les plus bizarres, c'est en général des systèmes avec des entités en interaction. Comment dire les problèmes des physiciens, en modélisation, ils sont assez différents par rapport à ceux là. Donc ça c'est la communauté académique, des gens qui font de la modélisation, et après je pense qu'il y a une communauté qui va naître je pense, et qu'on a déjà rencontré, qui va se renforcer, des industriels et des PME, qui vont utiliser la modélisation, soit comme produit pour faire de l'innovation, donc recherche et innovation au niveau industriel, mais aussi comme manière de créer des services, et ça c'est le travail qu'on avait commencé avec Forcity, eux ils vendent à la fin directement une plateforme qui est basée sur des modèles, donc ils nécessitent à la fois du travail de R& D pour développer leurs modèles, qui peuvent nécessiter OpenMole, et aussi dans la plate-forme y'a l'étude de scénarios. - **JR** Ah vous bossez avec ForCity ! Parce que j'avais vu leur truc, mais y'a longtemps, y'a trois ans, et à l'époque ils le vendaient, c'était un peu une arnaque, ils disaient on a modélisé toute la ville et tout, et je crois que je leur avais dit mais il est pas du tout calibré votre modèle, enfin ils superposaient des couches de modèle, et ils disaient c'est un modèle intégré... **RR** Oui eux on a mis du temps à leur faire comprendre ça, on a pas mal discuté avec eux, car un des fondateurs c'est un des potes de Matthieu, du coup on a beaucoup discuté avec eux, et ils disaient moi je prends les modèles tels qu'ils viennent de l'académique, ils ont fait le travail d'évaluation, moi je couple et je mets en avant, mais rien que faire le travail de couplage c'est faire un nouveau modèle, t'as pas le choix, à force de discuter et tout il a rencontré les problèmes dont on lui avait parlé, on a finit par lancer un projet avec eux, qui est bien financé. Paul et Guillaume là, ils sont financés sur ce projet.

JR Je pense qu'on reviendra sur l'histoire du couplage et tout. C'est au centre un peu de la problématique d'exploration et tout. Mais du coup j'avais pensé plutôt à une grille un peu temporelle, un peu l'histoire, essayer de revenir sur l'histoire d'OpenMole. Est-ce que tu as en tête le début, je crois qu'on en avait parlé une fois rapidement au relais-fac, un moment précis qui marque la naissance d'OpenMole, un acte fondateur, des précurseurs avant.

RR Ouais, alors c'est un mélange des deux forcément, si je te fais l'histoire un petit peu longue, je dirais que ça a vraiment commencé, les premiers prémisses de la chose c'est en 2003, l'été, quand j'ai fait mon premier stage de recherche avec des physiciens qui débutaient leur travaux sur la grille de calcul, la grille de calcul était un sujet tout nouveau à l'époque.

...

Ils m'expliquaient leurs algos, PSE ça vient de leur vision de la diversité en fait.

Après en 2012, on a été recrutés par Denise, et on a commencé à avoir en seulement quelques mois on a eu des premières versions sur lesquelles on pouvait calibrer le modèle, les informaticiens disaient voilà votre modèle marche à ce moment là, ça donnait des pistes, déjà ça été un résultat, on a été un peu libre quand on a réussi à montrer que le modèle fonctionnait, voilà on a trouvé des endroits sur lesquels ça fonctionne vraiment, alors qu'on le savait pas. - **JR** Alors est-ce qu'il y a eu une période sur laquelle il y a eu du prototypage, essayage à la main et tout, ou vous êtes direct partis en exploration. **RR** Je pense que Seb avait testé les trajectoires à la main, mais pas trouvé de trajectoire satisfaisante, il avait même pas d'indicateur en fait. Il savait visuellement. Donc ça ça été le premier travail, avec Clara, l'idée de comment calculer des indicateurs.

JR Alors est-ce qu'il y a un rapport avec la thèse de Thomas, il avait fait je sais plus quel Simpop, et après ils cherchaient un stagiaire pour, du coup la thèse de Seb, plateforme.

RR On a pris un stagiaire, mais ça n'a pas marché pour diverses raisons. C'est moi qui encadrait la partie info - **JR** C'était quoi, une plateforme qui couplait OpenMole avec la visualisation de données ? **RR** Alors ça y'a eu un autre stagiaire qui, ouais, mais c'était trop tôt, OpenMole était pas mûr. Je pense que maintenant ça commencerait à voir du sens dans quelques années, je pense que la visu, interactive ça serait intéressant, mais pour le moment c'est vrai que. Pour le point de vue d'OpenMole, c'est quand même pouvoir poser une question, et directement qu'OpenMole fournisse les données pour y répondre ; et après quand tu sais pas quoi poser comme question, en mode interactif. C'était une expérience, mais c'était compliqué.

Quand on est arrivés sur le front de Pareto, on était rassurés car le modèle marchait, globalement, on pouvait voir qu'un critère était satisfait par rapport à un autre, mais on avait aucun retour sur l'espace des entrées. Et les profils ça a marché du feu de dieu quand même. Moi j'ai eu une intuition un peu informaticien, qu'est ce qu'on demande à l'utilisateur. On lui demande une distance dans l'espace des entrées, il agrège des choux et des carottes, c'est un peu compliqué. Même si les informaticiens font souvent ça, je me serais retrouvé un peu honteux d'arriver devant un public et dire voilà, vous définissez une fonction avec des coefficients, je savais pas trop comment. Du coup peut être qu'une solution c'était d'enlever cette fonction là, et chercher que ce qui est différent. Ce qui est différent, on va essayer de l'apprendre, un truc interactif, ou alors donner une fonction de distance, mais c'est encore agréger les choux et les carottes. Du coup j'ai su convaincre Denise que c'était une bonne piste, comme on avait eu des bons résultats avant je crois qu'elle m'a fait confiance, j'ai recruté Guillaume j'ai l'impression d'avoir un peu forcé, mais

à la fin les résultats étaient là. Bon je sais pas si Denise avait vu l'ampleur de ce résultat là qui était vraiment du lourd, et tous les gens à l'ISC l'ont vraiment vu.

JR Parce que du coup l'arrivée de Guillaume ça été PSE, le travail avec Clem tout ça ?

RR Ouais voilà, puis en parallèle, alors je sais plus quand c'est arrivé cette histoire, je crois que c'était après les profils, je voulais faire de l'exploration dans l'espace des mécanismes, de manière plus systématique pour les géographes, le fait d'avoir une grille avec les dimensions. La décomposition, j'aimerais bien revenir un peu là dessus, là je vais avoir du temps je pense, exploration hybride entre paramètres et mécanismes. Les profils, la manière dont on a fait le calibrage, ça devrait pas être plus compliqué que ça. Après j'aimerais aussi voir les méthodes MCMC, ils génèrent des configurations de bâti, pourquoi pas faire ça avec les mécanismes, et du coup explorer l'espace des possibles, avec que des modifications locales comme ça. Une des manières de faire des espaces de mécanismes qui seraient gigantesques. Parce que pour l'instant, les mécanismes qu'on peut envisager c'est quelques milliers de modèles, parce qu'en fait on les calibre tous, en mémoire ça prend pas mal, mais là si on peut avoir une combinatoire de mécanismes quasiment infinie, bah du coup l'algo il pourrait lui aller chercher là dedans, pas avoir à aller chercher, pas tout générer pour savoir ce que tu veux, mais vraiment aller parcourir l'arbre des mécanismes possibles, j'ajoute ce mécanisme, je l'enlève. **JR** Donc un peu le chemin - **RR** Dans les possibilités, oui, savoir à un moment quand je rajoute ça ça donne plus rien, alors je sais pas trop comment.

JR OK, historique, interaction avec thématique. Sur un peu le présent. On en a un peu parlé, mais, en gros quels sont les projets et collaboration en cours en ce moment, les communautés qui l'utilisent. Peu être les géographes un peu moins à présent, que Geodiversity est fini. Est ce qu'il y a d'autres champs.

RR Ouais, alors j'ai jamais une vision globale de la communauté, c'est super dur, d'un point de vue technique, ce que j'aimerais faire c'est une espèce d'instance multi-utilisateur sur laquelle on peut se connecter, comme ça moi en plus j'aurai un aperçu de l'utilisation. **JR** Le truc avec Docker du coup ? - **RR** Ouais, ça c'est vraiment pas loin, déjà lancé, on va bientôt faire le front end sur lequel on pourra se connecter. Du coup l'idée d'avoir une instance à l'ISC qu'on peut héberger, visibilité de qui utilise OpenMole. Et après y'a aussi l'idée de vendre un service d'ici pas longtemps, du côté entreprise, qu'on puisse avoir une instance comme ça mais payante. Donc ça c'est du cloud, mais plutôt payant. **JR** Donc ça c'est un service payant ? **RR** C'est un service payant, la valorisation, trouver comment proposer un mode de valorisation basé sur le logiciel libre et le développement du logiciel libre, scientifique, à l'ISC et ailleurs, pour pourvoir quand même recruter des ingénieurs, des commerciaux. Y'a une valeur ajoutée de fou dans ces plateformes. Mais nous on a un problème de temps, un problème de temporalités de post-docs, de gens qui voilà, et des gens à l'extérieur qui seraient motivés pour payer des ingés. C'est encore un peu flou mais c'est l'idée.

JR Et du coup, par rapport à la réticence des gens de l'analytique, soit des économistes, soit des physiciens, qui veulent que le modèle soit résolu.

RR Ca c'est autre chose, une autre manière d'approcher les choses, dans des cas limites. C'est sûrement intéressant, après je pense que ça rentre pas en compte dès que tu veux faire de l'étude

de scenario un peu complexe, coupler ça avec machin, et boum ça devient un cas qui n'est plus un cas limite, et après même sur des approches comme ça dès que tu as un peu de stochasticité, par exemple Giulia là elle fait un modèle de physicien, elle veut regarder plein de configurations différentes, elle va paralléliser ça, elle se sert pas encore des lagos génétiques mais y'a carrément moyen. C'est un truc qui génère des formes, des cercles, une ville qui génère une ville. Elle a quelque paramètres, elle veut regarder les transitions. C'est de l'exploration directe. Meme ces modèles là rentrent dans ce cadre.

JR Donc là c'est une des questions, comment réussir à continuer à faire vivre ça, pour la nouvelle génération, est-ce qu'il faudrait pas de la formation, pour les jeunes.

RR Je pense que y'aurait besoin d'un bon... Thomas il va bosser, ouais c'est vrai que moi je suis beaucoup engagé à l'ISC, du coup je suis pas du tout, je mets plus du tout les pieds à Géocités, du coup le fond dont les gens ont bénéficié, dont toi t'as bénéficié, il est un peu bancal là. Donc je sais pas, si toi t'as plus de vision que moi là dessus, je sais pas comment on pourrait faire là dessus ? Je peux pas m'engager à aller très souvent à Géocités, parce que globalement j'ai tellement de taf ici, tellement de trucs à organiser ici. On peut essayer de faire les formations Jedi.

JR Peut être au niveau des masters, du coup au niveau de Geoprisme, leur faire déjà des initiations. Je pense que c'est en amont, c'est un avis subjectif, mais j'ai l'impression qu'il y a un manque de culture sur ces points là - **RR** Ils font déjà de la modélisation ? - **JR** Bah ils pourraient en fait. Pas beaucoup, plus dans Carthageo, dans la géomatique. Ils pourraient en fait, c'est exactement la formation. Je pense qu'il faudrait faire plus à ce niveau là, comme ça quand t'arrives, renouveler le vivier de. Enfin j'ai l'impression que des fois y'a des réticences. Du coup garder un peu cette culture qui est la culture Géocités, qui est un peu l'ADN.

RR Oui y'a des trucs intéressants là dedans. Nous on va continuer à montrer nos résultats avec Julie tout ça, mais c'est vrai que d'un point de vue pragmatique sur les nouveau doctorants, faudra bien. Peut être ça serait une bonne idée en effet, modélisation, un peu OpenMole. - **JR** Ouais donner envie. Petit atelier, pour les masters. - **RR** Atelier, le jeudi aprèm.

JR Ok je pense qu'on a fait le tour, y'a rien qui te vient en tête ? Super !