

Stage Master 2 ou ingénieur :

Visualisation de fronts de Pareto et analyse de données multidimensionnelles

Contexte

Ce stage se déroulera dans le cadre d'une collaboration entre l'INRA et Télécom Paris Tech concernant la visualisation de données concernant la modélisation d'un processus œnologique. Ce projet, financé par une action de recherche inter équipes de l'INRA, l'AIC AromOpti, impliquant des modélisateurs (équipe MALICES de l'UMR GMPA de Versailles-Grignon) et des experts en œnologie (UE de Pech Rouge et LISBP), a pour objectif majeur de démontrer l'intérêt et la faisabilité expérimentale de nouvelles stratégies de conduite de la fermentation permettant d'optimiser le contenu final du vin en arômes fermentaires tout en minimisant la consommation de frigories.

Buts du stage

Sur la base des outils développés au cours de la première année de l'AIC, et de l'outil de visualisation multi-dimensionnel EvoGraphDice (initialement développé à l'INRIA), le stage a pour but de développer des techniques d'exploration interactives du comportement d'un modèle (front de Pareto, typiquement), afin de permettre des remises en question de différents choix de modélisation (nombre et structure des réactions, forme de la ou des fonctions à optimiser).

Nous exploiterons des simulateurs développés sous matlab à l'INRA. Ces simulateurs comportent de très nombreux paramètres, et l'apprentissage de paramètres sur des données expérimentales se fera par des méthodes d'apprentissage ou d'optimisation à base d'heuristiques stochastiques.

Algorithmes évolutionnaires, algorithmes génétiques

Les heuristiques évolutionnaires désignent un ensemble de techniques d'optimisation stochastique fondées sur une transposition informatique des principes d'évolution darwinienne (<http://evelyne.lutton.free.fr/EvolutionaryAlgorithms.html>). Ces méthodes consistent grossièrement à simuler l'évolution d'une population sous l'action de règles darwinistes (sélection, mutation, croisement). Des populations de points évoluent dans un espace de recherche sous l'action conjuguée de variations aléatoires (croisement par échange d'informations entre plusieurs points, mutation par perturbation locale sur un point) et d'un processus de sélection (les meilleurs gagnent, au regard d'une fonction définie sur l'espace de recherche considéré, dite fonction d'évaluation ou fitness). La souplesse de ces techniques permet de traiter des problèmes d'optimisation difficiles, inaccessibles aux techniques déterministes, au prix bien-sûr d'un coût algorithmique plus important.

Par ailleurs, la formulation de la résolution d'un problème d'optimisation complexe par le biais de l'évolution d'une population revient à effectuer un échantillonnage irrégulier de l'espace de recherche : ce mode d'exploration permet de gagner plus d'information sur la forme de l'espace de recherche et la nature du problème d'optimisation que la seule localisation d'un optimum global. L'équipe MALICES du GMPA a pu montrer qu'une analyse fondée sur des visualisations de l'échantillonnage de l'espace effectué par un algorithme évolutionnaire permettait de mettre en évidence la présence de corrélations

entre paramètres (Lutton et al, 2013).

Exploration interactive de front de Pareto

Les techniques évolutionnaires peuvent être aussi exploitées pour l'optimisation simultanée de plusieurs objectifs (évolution multi-objectif, Deb 2001), et permettent de produire un échantillonnage du front de Pareto, c'est-à-dire de l'ensemble des compromis « optimaux » entre objectifs (notion de dominance de Pareto). Dans le cadre de la modélisation, cette stratégie ouvre des possibilités très intéressantes pour l'exploration du comportement de modèles, mais cela nécessite le développement d'interfaces conviviales et adaptées (Lotov, 2008).

EvoGraphDice

On s'appuiera sur des travaux antérieurs concernant l'exploration visuelle évolutionnaire (Evolutionary Visual Exploration, <http://www.aviz.fr/EVE>), et notamment sur l'outil EvoGraphDice qui sera installé et expérimenté sur la plate-forme PiXLS de Télécom ParisTech (<http://www.telecom-paristech.fr/bibliotheque-crdn/espaces/digiscope.html>). EvoGraphDice combine en effet une technique interactive de visualisation (matrice de scatterplots) avec une optimisation évolutionnaire, pour étendre le concept d'analyse en composantes principales. Une "population" d'axes secondaires d'observation est évoluée par programmation génétique interactive, fournissant ainsi de nouveaux points de vues sur les données à base combinaisons non linéaires de paramètres.

Cet outil sera étendu pour prendre en compte différentes sources, qualitatives et/ou quantitatives (informations statistiques, incertitudes, intervalles de confiance) pour guider l'exploration vers des vues pertinentes pour l'utilisateur. Une extension pour la visualisation de jeux de données en grandes dimensions sera aussi considérée (sélections de variables à visualiser), ainsi que la visualisation de modèles.

Prérequis

- Bonnes compétences en programmation (matlab, java), apprentissage, et développement d'interfaces.
- Un intérêt pour la recherche (ce stage pourra éventuellement déboucher sur une thèse).
- Une connaissance des méthodes évolutionnaires et de visualisation sera un plus.
- Il n'est pas nécessaire d'avoir des connaissances en chimie pour aborder ce sujet, mais l'étudiant aura à collaborer avec des chimistes et des biologistes.

Lieu de travail

UMR GMPA, Thiverval-Grignon.

Encadrement

Evelyne Lutton, UMR GMPA (<http://evelyne-lutton.fr>)
Nadia Boukhelifa, TélécomParisTech

Pour candidater

Envoyez CV et lettre de motivation à Evelyne.Lutton@grignon.inra.fr et Nadia.Boukhelifa@telecom-paritech.fr

Quelques références pour commencer

Avila, S.L. (2006), Optimisation multi-objectif et analyse de sensibilité appliquées à la conception de dispositifs. Application : Synthèse d'antennes à réflecteur embarquées dans un satellite, Thèse de L'Ecole Centrale de Lyon, 23 février 2006. (http://tel.archives-ouvertes.fr/docs/00/06/42/99/PDF/These_ECL2006-05_Avila.pdf)

Barichard, V. et Hao, J.K.(2004), Résolution d'un problème d'analyse de sensibilité par un algorithme d'optimisation multiobjectif, 5^{ème} Conférence Francophone de MOdélisation et SIMulation « Modélisation et simulation pour l'analyse et l'optimisation des systèmes industriels et logistiques » MOSIM'04, 1-3 septembre 2004, Nantes, France, (http://www.info.univ-angers.fr/~barichard/page_perso/publications/2004/barichard_hao-MOSIM2004.pdf)

Deb, K. (2001). Multi-objective optimization using evolutionary algorithms. Chichester, London: Wiley.

Lutton, E., Tonda, A., Gaucel, S., Fouquier, J., Riaublanc, A. and Perrot, N. (2013), Food model exploration through evolutionary optimization coupled with visualization: application to the prediction of a milk gel structure. In From Model Foods to Food Models. DREAM Project's International Conference, June 2013.

[Alexander V. Lotov](#), [Kaisa Miettinen](#), Visualizing the Pareto Frontier. [Multiobjective Optimization. Lecture Notes in Computer Science](#) Volume 5252, 2008, pp 213-243

Packham, I.S.J., Rafiq, M.Y., Borthwick, M.F. and Denham, S.L. (2005), Interactive visualisation for decision support and evaluation of robustness—in theory and in practice, *Advanced Engineering Informatics*, Volume 19, Issue 4, October 2005, Pages 263-280,

N. Boukhelifa, A. Bezerianos, W. Cancino and E. Lutton. Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search. *Evolutionary Computation Journal*, MIT press, 2015. [to appear] [\[bib\]](#) [\[pdf\]](#)

N. Boukhelifa, A. Bezerianos and E. Lutton. A Mixed Approach for the Evaluation of a Guided Exploratory Visualization System. *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3) 2015*, 25-26 May 2015. [\[bib\]](#) [\[pdf\]](#)

W. Cancino, N. Boukhelifa, A. Bezerianos and E. Lutton. Evolutionary Visual Exploration: Experimental Analysis of Algorithm Behaviour. *GECCO workshop on Genetic and Evolutionary Computation (VizGEC 2013)*, Jul 2013, Amsterdam, Netherlands. [\[bib\]](#) [\[pdf\]](#) [\[project page\]](#).

N. Boukhelifa, W. Cancino, A. Bezerianos and E. Lutton. Evolutionary Visual Exploration: Evaluation With Expert Users. *Computer Graphics Forum (EuroVis 2013, June 17--21, 2013, Leipzig, Germany)*, Eurographics Association, 2013, 32 (3). [\[bib\]](#) [\[pdf\]](#) [\[video\]](#) [\[project page\]](#).