

Appel à candidatures pour un stage de recherche M2

Gestion de la variabilité de plusieurs familles de logiciels avec l'Analyse Polyadique de Concepts

Encadrement :

Giacomo Kahn, Jannik Laval (Université Lyon 2),
Jessie Galasso (Université de Montréal, Canada)

Etablissement : Lyon 2**Laboratoire:** laboratoire DISP**Profil recherché :**

Master 2 Informatique

Financement: DISP Lyon 2**Période de stage :** Février à Juillet 2023**Compétences souhaitées :**

Ingénierie Logicielle, Analyse Formelle de Concepts

Mots-clés: Software Product Lines, Formal Concept Analysis**Description du sujet :**

Des études menées au sein de plusieurs industries montrent que les éditeurs de logiciels ont tendance à produire des variantes de systèmes logiciels en réutilisant de manière ad-hoc des clones de projets précédents, aboutissant ainsi à des familles de logiciels similaires où chaque produit est géré et maintenu séparément. L'ingénierie des lignes de produits logiciels regroupe un ensemble de méthodes basées sur la réutilisation systématique d'artefacts logiciels afin de développer, maintenir et faire évoluer en parallèle un ensemble de variantes d'une même famille de logiciels. Cela permet de réduire le temps et le coût de développement et de maintenance de la famille de logiciels tout en augmentant sa qualité. La migration depuis une famille de produits logiciels vers une ligne de produits est donc une problématique essentielle partagée par de nombreux éditeurs de logiciels. La modélisation et la gestion de la variabilité, i.e., quels artefacts varient entre les variantes logicielles et comment, est un point central de ces méthodes.

L'analyse formelle de concepts est un cadre mathématique pour l'apprentissage automatique basé sur la théorie des treillis. Ce cadre permet de construire une structure de clustering hiérarchique à partir d'objets caractérisés par des descriptions. Il supporte des tâches d'analyse de données, d'extraction, de représentation et d'exploration des connaissances dans de nombreux domaines. Il a notamment été appliqué en recherche logicielle pour représenter et gérer la variabilité lors de la migration d'une famille de logiciels vers une ligne de produits.

L'ingénierie des lignes de produits multiples est un domaine émergent qui se consacre à la représentation et la gestion de la variabilité dans des familles de logiciels qui sont considérés comme complexes du fait de leur hétérogénéité, de leur taille ou encore de leur architecture. Cette complexité donne lieu à des données pouvant inclure de multiples dimensions (e.g., caractéristiques logicielles, contexte et environnement, équipe de développement, dépendances avec d'autres systèmes) ce qui présente un réel défi lorsque l'on cherche à les analyser avec les techniques actuelles. La migration depuis une famille de logiciels complexes vers une ligne de produits multiples nécessite donc de reconsidérer et d'adapter les techniques traditionnelles de rétro-ingénierie de la variabilité afin de manipuler des données plus complexes.

L'objectif de ce stage est d'utiliser une extension de l'analyse formelle de concepts, l'analyse polyadique (PCA), qui permet d'extraire des motifs et des informations de données multi-dimensionnelles, afin d'extraire la variabilité d'un ensemble de familles de logiciels complexes.

Il est demandé de :

- s'approprier les formalismes du PCA et de l'ingénierie des lignes de produits logiciels multiples,
- étudier les similitudes entre les structures extraites par PCA et les modèles de variabilités des lignes de produits logicielles complexes existantes,
- proposer des règles de rétro-ingénierie utilisant les structures construites par PCA.

Le stage est basé au laboratoire DISP de l'université Lyon 2. Vous serez entouré d'une équipe de chercheurs et travaillerez en lien fort avec l'équipe. Vous serez suivi pour planifier l'ensemble des résultats attendus sur la durée du stage.

Programme de travail proposé :

- Rédaction d'un état de l'art sur les interactions entre l'analyse formelle de concepts et l'ingénierie logicielle, appropriation des concepts et des formalismes. (1 mois)
- Mise en forme des lignes de produits logiciels complexes et analyse des structures extraites par PCA et des modèles de variabilités des lignes complexes. (2 mois)
- Proposition de règles de rétro-ingénierie pour l'extraction de variabilité. (1 mois)
- Rédaction d'un article scientifique pour valoriser les résultats obtenus. (2 mois)

Déposez votre candidature avant le 30/11/2022, par mail à giacomo.kahn@univ-lyon2.fr, jannik.laval@univ-lyon2.fr, jessie.galasso-carbonnel@umontreal.ca, en précisant la référence de l'offre DISP-2023A et en joignant un dossier composé de votre CV, une lettre de motivation et vos relevés de notes récents.

Call for applications for a M2 research internship

Managing the variability of complex software families using Polyadic Concept Analysis

Supervision:

Giacomo Kahn, Jannik Laval (Univ Lyon 2, France),
Jessie Galasso (Université de Montréal, Canada)

University: Lyon 2

Research laboratory: DISP laboratory

Required profile:

M2 in computer science

Funding: DISP Lyon 2

Internship period: Feb to Jul 2023

Required skills:

Software engineering, Formal Concept Analysis

Key words: Software Product Lines, Formal Concept Analysis

Topic description:

Industrial studies show that software editors tend to produce variant of software systems by reusing clone of other projects. This creates families of software where each product is managed and maintained separately. Software Product Line Engineering (SPLE) consists of a set of methods based on the systematic reuse of software artefacts to simultaneously manage, maintain, and develop variants from a same software family. This reduces the time, development cost and maintenance cost of the software family, while increasing its overall quality. To migrate from a family of software products to a software product line is an essential problem, shared by most software editors. To model and manage variability, i.e., which artefacts differ between software variants and how, is a central topic in SPLE.

Formal Concept Analysis is a mathematical framework for machine learning based on lattice theory. It allows for the computing of a hierarchical clustering structure from objects described by attributes. It is used in many fields for data analysis, pattern mining, knowledge representation and exploration. Formal Concept Analysis was applied in software research to represent and manage variability of a software family, to migrate to a software product line.

Multiple Software Product Line Engineering is an emerging field that aims at representing and managing variability in software families that are considered complex, by their size, heterogeneity, or architecture. This complexity is translated by high dimensional data (software characteristics, context and environment, development team, dependency with other systems), which is a real challenge for the existing methods. To migrate from a complex software family to a multiple software product line, one needs to reconsider and adapt the existing retro engineering methods to more complex data.

This internship's objective is to use an extension of Formal Concept Analysis, Polyadic Concept Analysis (PCA) to mine patterns and information from multidimensional data, to extract variability form a complex software family. The intern will have to:

- understand PCA's formalism and multiple software product line engineering,
- study similarities between conceptual structure extracted with PCA and variability models used in SPLE,
- infer retro-engineering rules that use the structures extracted with PCA.

This internship is located at the DISP laboratory, in its Lyon 2 site. You will be surrounded by a team of researchers in Lyon and Montréal, and work with tight ties to the team. You will be advised to plan all the expected results over the internship.

Suggested work program:

- Literature review on the interactions between software engineering and formal concept analysis, appropriation of formalisms and concepts. (1mo)
- Processing of a complex software family and analysis of the conceptual structures given by PCA and corresponding variability models (2mo)
- Proposition of retro engineering rules to extract variability from complex software families (1mo)
- Redaction of a scientific paper (2mo)

Submit your application by November 30, 2022, by sending an email to giacomo.kahn@univ-lyon2.fr, jannik.laval@univ-lyon2.fr, jessie.galasso-carbonnel@umontreal.ca. Precise the job offer reference (DISP-2023A) and attach a file composed of a CV, cover letter, and recent transcripts.