## Fouille déclarative

## Exercice 1

Dans cet exercice, vous travaillerez avec les outils suivants :

- Choco-Mining : Une bibliothèque Java conçue pour résoudre des problèmes de fouille d'ensembles d'items, basée sur le framework Choco-solver.
- La bibliothèque SPMF : Un logiciel open-source basé sur Java, spécialisé dans la fouille de motifs (SPMF).

**Question 1** • Cloner le dépôt GitHub de Choco-Mining (lien).

Question 2 • Ouvrir le fichier ExampleClosedItemsetMining.java et réaliser les tâches suivantes :

- 1. Analyser le code en détail.
- 2. Exécuter la méthode main.
- 3. Tester avec d'autres ensembles de données comme mushroom ou chess.
- 4. Afficher le nombre de motifs trouvés.
- 5. Afficher le temps d'exécution.

**Question 3** • Ajouter la contrainte de fréquence :  $freq(P) \ge \alpha$ .

**Question 4** • Ajouter une contrainte sur la taille des motifs retournés :  $size(P) \ge lb$ .

**Question 5** • Réaliser maintenant les mêmes tâches avec SPMF. Exécuter le fichier . jar disponible dans votre dépôt local. L'objectif est d'exécuter LCM pour l'énumération des ensembles fermés, puis de relancer avec différents seuils de fréquence et de taille des motifs.

Question 6 • Ajouter une contrainte, appelée CategoryConstraint, dans le fichier 'ExampleClosedItemsetMining.java' afin de modéliser le problème suivant : Considérons un ensemble de données contenant n items, organisés en catégories de taille catSize (ex: produits ménagers, appareils électroménagers, etc.). L'ensemble de données est divisé en nbCat = n/catSize catégories, avec certains items n'appartenant à aucune catégorie (mais n'excédant pas la taille catSize). La figure 1 illustre un exemple avec 8 items, 2 catégories de taille 3, et 2 items hors catégorie. La tâche consiste à créer un modèle de contrainte qui énumère tous les ensembles fermés composés d'items appartenant à au moins m catégories:

$$\texttt{CategoryConstraint}(P) \equiv \sum_{i=1}^{nbCat} \prod_{j=1}^{catSize} P_i \geq m$$

Par exemple, dans l'ensemble de données illustré en figure 1, avec m=2, le motif produit est : BEF.

**Question 7** • Comment cette contrainte CategoryConstraint peut-elle être prise en compte dans SPMF?

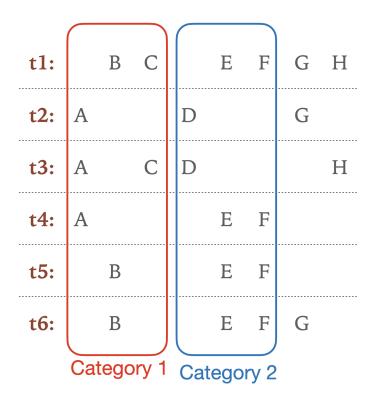


FIGURE 1 – Illustration des catégories d'items.