

Analyse Formelle de concepts

Approfondissement : règles d'implications versus structures

1 Présentation

Dans ces travaux, nous prolongerons l'analyse menée sur le jeu de données précédent. L'objectif sera de mener des réflexions d'ordre général et également de dégager ce qu'apporte plus spécifiquement chacune des représentations (structure conceptuelle ou ensemble de règles d'implications).

Vous rédigerez un compte-rendu d'expérience incluant toutes vos données et tous vos résultats (il peut être rendu par groupes de 1 à 3 étudiants, si possible les mêmes que pour le TP précédent).

Les étapes du travail consistent à :

- Construire avec COGUI si vous ne l'aviez pas déjà fait : l'AOC-poset et la base d'implications de Duquenne-Guigues.
- Elaborer autour des questions proposées plus loin et ajouter vos propres réflexions.
- Rendre les éléments suivants sur Moodle (pas de lien extérieur vers des dépôts cette fois-ci) :
 - Fichier d'entrée de COGUI en format csv préférentiellement.
 - AOC-poset issu de COGUI en format .dot et image (.pdf ou .png).
 - Dossier issu de COGUI des fichiers .txt de règles de la base Duquenne-Guigues.
 - Document rédigé contenant vos conclusions.

2 Questions

- Sur l'AOC-poset
 - Comment trouver rapidement un ou plusieurs objets ayant un groupe d'attributs ?
 - Comment trouver les objets les plus proches (en termes d'attributs) d'un objet donné ?
 - Observez-vous des schémas dans les intensions des concepts, un exemple de schéma peut être `Year = XX, #Citations = YY, autres attributs` ? Qu'est-ce que cela peut vous apprendre ?
 - Comment observer facilement des règles sur l'AOC-poset (et quel type de règles) ?
- Sur la base de Duquenne-Guigues
 - Les règles ne sont pas redondantes entre elles, mais observez-vous des redondances internes dans les règles (dans leurs prémisses et/ou dans leurs conclusions), par exemple, lorsque vous observez les deux règles : $A, B, C \Rightarrow D$ et $A \Rightarrow B$, elles sont non redondantes entre elles, mais B n'est pas utile (d'un point de vue logique) dans la prémisse de la première. Est-ce gênant pour un expert si cela arrive en prémisse ? en conclusion ? suivant le type d'attribut ?
 - Observez-vous des schémas de règles, par exemple, un schéma peut s'appuyer sur l'apparition en prémisse ou en conclusion de valeurs de certains attributs. Un exemple de schéma peut être `Year = XX, autres attributs \Rightarrow #Citations = YY, autres attributs`. Est-il facile d'énumérer certains ou tous ces schémas et est-ce que cela peut servir dans l'analyse ?

- Observez-vous des règles qui ne font que traduire votre codage des données (par exemple si vous avez quelque chose comme **Year in [2014, 2015] \Rightarrow Year in [2013,2018]**, et qu'en pensez-vous ?
- Comment faire la correspondance entre une règle et les objets satisfaisant cette règle ?
- Sur la comparaison des deux approches ; quels sont les procédés, avantages et inconvénients d'utiliser l'AOC-poset ou la base d'implications pour :
 - Connaître la fréquence d'apparition d'un groupe d'attributs.
 - Trouver des catégories.
 - Trouver des attributs qui n'apparaissent jamais ensemble.
 - Trouver des attributs qui apparaissent systématiquement ensemble.
- Énoncez des questions possibles pour les méthodes de fouille (ex. Quelles sont les propriétés de la méthode la plus citée ? Y a-t-il une méthode semblable à la méthode XX mais plus récente ? Quelles nouvelles combinaisons de méthodes seraient intéressantes à tester - et pourquoi ? / complétez) et indiquez si vous utiliseriez plutôt les implications, plutôt l'AOC-poset ou indifféremment les deux pour y répondre, ou si aucune de ces méthodes ne vous semble satisfaisante.