

DOI – EXAMEN (EXTRAIT)

1 Itemsets fréquents et règles d'association (5,5 points)

La table 1 indique la présence (1) ou l'absence (0) de personnages dans huit romans de Terry Pratchett. Pour simplifier l'écriture, chaque personnage est représenté par une initiale (A, C, G, H, R, S ou T).

| | Angua von Überwald | Carrot Ironfounderson | Granny Weatherwax | Havelock Vetinari | Rincewind | Samuel Vimes | Twoflower |
|---------------------|--------------------|-----------------------|-------------------|-------------------|-----------|--------------|-----------|
| Livre | A | C | G | H | R | S | T |
| The Colour of Magic | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Wyrd Sisters | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Guards! Guards! | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| Witches Abroad | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Men at Arms | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| Interesting Times | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| Night Watch | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| Thud! | 1 | 1 | 0 | 1 | 0 | 1 | 0 |

TABLE 1 – Apparitions de personnages dans huit romans

Pour les questions de cet exercice, lorsqu'une valeur de support est demandée, vous pouvez choisir de calculer, au choix, le support relatif ou le support absolu.

Q1.1 Calculer le support des itemsets suivants :

- {H}
- {A, C}
- {C, G}
- {A, C, H}

Q1.2 Donner une interprétation textuelle de la valeur de support de l'itemset {C, G}.

Q1.3 Lister les itemsets fréquents pour un support minimum de 3, avec leur valeur de support.

Sur cette même table 1, les itemsets fréquents maximaux calculés pour un support minimum de 1 sont les suivants :

- {G} (support = 2)
- {H, R, T} (support = 1)
- {A, C, H, S} (support = 2)

Q1.4 À partir de ces fréquents maximaux *uniquement*, donner, en justifiant votre réponse, une borne inférieure ou supérieure de la valeur de support des itemsets suivants :

- {H}
- {G, T}
- {A, C}

Q1.5 En vous basant sur votre réponse à la question Q1.3, lister les itemsets fréquents clos pour un support minimum de 3.

Q1.6 Calculer la confiance et la mesure de Kulczynski des règles d'association suivantes :

- $A, C \rightarrow H$
- $H \rightarrow A, C$
- $C \rightarrow S$
- $S \rightarrow C$

La table 2 donne les thèmes principaux associés à chaque roman.

| Livre | Thème |
|---------------------|------------|
| The Colour of Magic | Rincewind |
| Wyrd Sisters | Witches |
| Guards! Guards! | City Watch |
| Witches Abroad | Witches |
| Men at Arms | City Watch |
| Interesting Times | Rincewind |
| Night Watch | City Watch |
| Thud! | City Watch |

TABLE 2 – Thèmes principaux des romans

Q1.7 Comment feriez-vous pour étudier les règles d'association qui existent entre les personnages et les thèmes des romans?

2 Denial constraints (1,5 points)

La table 3 (notée E dans la suite de l'exercice) liste les examens de la période B.

| Matière | Date | Option | Durée | Salle | Coefficient |
|---------|------------------|--------|-------|-------|-------------|
| LOS | 2021-03-22 9:30 | IA | 120 | A106 | 1 |
| SAV | 2021-03-22 13:45 | IA | 90 | B140 | 0.5 |
| APM | 2021-03-22 16:30 | IA | 90 | B140 | 0.5 |
| DOI | 2021-03-23 9:30 | IA-D | 120 | B374 | 1 |
| VAT | 2021-03-23 9:30 | IA-S | 120 | A102 | 0.5 |
| SSE | 2021-03-23 16:45 | IA-S | 120 | B140 | 0.5 |

TABLE 3 – Examens de la période B

Q2.1 Donner, pour chacune des *denial constraints* suivantes : une interprétation textuelle de la règle, sa validité par rapport aux données (règle vraie ou fausse) et un contre-exemple si la règle est fausse.

- $\forall t_1, t_2 \in E, \neg(t_1.\text{date} = t_2.\text{date} \wedge t_1.\text{option} = t_2.\text{option})$
- $\forall t_1, t_2 \in E, \neg(t_1.\text{durée} < t_2.\text{durée} \wedge t_1.\text{coefficient} > t_2.\text{coefficient})$
- $\forall t \in E, \neg(t.\text{option} = \text{IA} \wedge t.\text{salle} \neq \text{B140})$
- $\forall t_1, t_2 \in E, \neg(t_1.\text{option} = t_2.\text{option} \wedge t_1.\text{coefficient} < t_2.\text{coefficient} \wedge t_1.\text{date} > t_2.\text{date})$