

M2 Informatique UCBL

Examen TIW 2 - Interopérabilité

1 Mars 2021 - 9:45-11:45 (durée 2h) en distanciel

Attention: faire l'upload d'un fichier PDF sur Tomuss (et pas d'un fichier doc!)

Colonne Tomuss : "examen_dépôt_de_votre_pdf_avec_les_reponses"

Partie I "Intégration et Qualité de Données et BD Graphes" : pour cette partie, vous avez le droit et il est conseillé de consulter les transparents du cours. Faites des réponses courtes et synthétiques, pas de copie exhaustive de contenu de transparents, car la capacité de synthèse compte dans la note (moitié des points maximum si non respecté).

Cette partie consiste en 8 questions. Pour chaque question, le nombre de points est indiqué.
Barème à titre indicatif, susceptible de changer.

1. (3pt) Supposons d'avoir les s-t tgds suivantes:

$$S(X,Y), T(Y,Z), V(X,Z) \rightarrow W(X,Y,Z)$$

$$S(X,Y), V(X,Z) \rightarrow W(X,Y,Z)$$

et de partir de la solution source $I = \{S(1,2), T(2,3), V(1,3)\}$, quelle est la solution target J en appliquant l'algorithme de Chase Naïve?

Montrez les étapes du Chase sur les deux s-t tgds.

- En appliquant Chase sur la première s-t tgds on aura : $W(1,2,3)$.

- En appliquant chase sur la deuxième s-t tgds on aura : $W(1,2,3)$;

Donc $\text{Chase}(I) = \{W(1,2,3)\}$

et $j = \{W(1,2,3)\}$

2. (1pt) Donnez les principales différences entre les dépendances fonctionnelles (FDs) et les matching dependencies (MDs)
(faites une réponse synthétique en deux lignes maximum)

La Matching dependency c'est comme une dépendance fonctionnelle à l'exception que les égalités peuvent être relâchées à des similarités et que ça concerne deux relations.

A matching dependency is like an FD, except that equalities can be relaxed to similarities; and it relates to two relations.

3. (2pt) Étant donnée la FD suivante sur la relation
Étudiant(Identifiant, Nom, Prénom, Formation, Département, Diplôme)

Diplôme, Formation \rightarrow Département

Transformez cette FD dans une Conditional Functional Dependency de votre

choix.

[Diplôme = Master2, Formation = TIW] → [Département = Informatique]

4. (1pt) Quels sont les facteurs à considérer dans le traitement de la complétude dans une SGBD? Donnez une liste de facteurs et une brève explication pour chaque facteur (*1 ligne max pour chaque facteur*)

- Valeur manquante : Pour un enregistrement donné, la valeur d'un certain attribut n'est pas définie. (colonne nulle).
- Ligne inexistante : l'enregistrement répondant à la requête n'existe pas dans la Base, (la ligne n'existe pas du tout)

5. (1pt) Expliquer les nouveautés du modèle Property Graph par rapport aux modèles traditionnels de graphes étiquetés (*2 lignes max*).

Le Property Graph est organisé en : **nœuds** qui contiennent des attributs de forme key-value appelés propriétés et en **relations** qui sont des connexions dirigées nommées et pertinentes entre deux nœuds.

6. (2pt) Expliquez le concept de label-constrained reachability queries dans les bases de données graphes?

- Le label-constrained reachability queries consiste à trouver un chemin dans un graphe G entre un nœud s et t en utilisant uniquement des arêtes dont le libellé est L . Ce dernier est une généralisation du problème de reachability déjà très étudié. Et ça peut être formalisé de la manière suivante :

label-constrained reachability queries : Étant donné $x, y \in N$ et un ensemble de labels $I \subseteq L$, on détermine si (x, y) est dans l'ensemble :

$G * L = \{(s, t) \mid \text{Il existe un chemin dans } G \text{ allant de } s \text{ à } t \text{ en utilisant uniquement les arcs dont le label est dans } I\}.$

7. (3pt) Étant donné une requête ci-dessous et un graphe simplifié (avec des arêtes non dirigées et pas de propriétés sur les arcs/sommets) dans la Fig. suivante:

Example: Look for all pairs of friends and the city each friend lives in



- a. Trouvez les résultats de cette requête en adoptant:
- La sémantique de Subgraph Isomorphism
 - La sémantique de Subgraph Homomorphism
 - La sémantique de simulation

Isomorphism :

((Anne , Berlin) , (Chris, Leipzig))
((Mary, Berlin), (Chris, Leipzig))

Homomorphism :

((Anne , Berlin) , (Chris, Leipzig))
((Mary, Berlin), (Chris, Leipzig))
((Anne, Berlin), (Mary, Leipzig))

simulation :

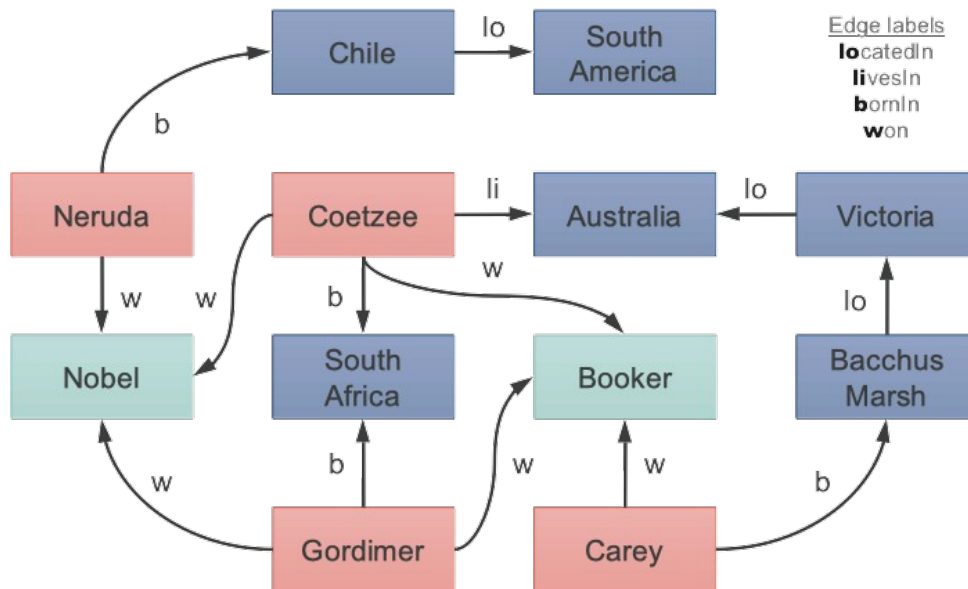
$\langle \rangle \leftarrow (?x , \text{friendOf} , ?y) , (?x , \text{LivesIn} , ?c1) , (?y , \text{LivesIn} , ?c2)$

Une simulation $S1 = \{(?x , \text{Anne}) , (?y , \text{Mary}) , (?c1 , \text{Berlin}) , (?c2 , \text{Berlin})\}$

Une simulation $S2 = \{(?x , \text{Anne}) , (?y , \text{Chris}) , (?c1 , \text{Berlin}) , (?c2 , \text{Leipzig})\}$

Une simulation $S3 = \{(?x , \text{Mary}) , (?y , \text{Chris}) , (?c1 , \text{Berlin}) , (?c2 , \text{Leipzig})\}$

8. Etant donné le graphe étiqueté suivant:



a. Ecrire comme une requête Datalog un exemple de votre choix pour chaque type de requête suivante (*conseil: vous pouvez enrichir de façon incrémentale la même requête*) :

- (RPQ) Regular Path Queries
- (CRPQ) Conjunctive Regular Path Queries (CRPQ)
- (UCRPQ) Unions of Conjunctive Regular Path Queries

- (RPQ) Regular Path Queries : Nom des écrivains qui ont gagné un Nobel

$\langle ?n \rangle \leftarrow (?n, w, \text{Nobel})$

- (CRPQ) Conjunctive Regular Path Queries : Nom des écrivains qui ont gagné un Nobel et les lieux où ils sont nés

$\langle ?n, ?l \rangle \leftarrow (?n, w, \text{Nobel}), (?n, b, ?l)$

- (UCRPQ) Unions of Conjunctive Regular Path Queries

$\langle ?\text{person}, ?c, ?\text{localisation} \rangle \leftarrow (? \text{person}, w, \text{Booker}), (? \text{person}, b, ?c), (? \text{person}, lo^*, ?\text{localisation})$

b. Pour chaque requête ci-dessus, donnez aussi les résultats sur le graphe.

Première : {Gordimer}, {Neruda}

Deuxième : {Gordimer, south Africa}, {Neruda,Chile}

Troisième : (Coetzee, South Africa) – (Carey, Bacchus Marsh, Victoria, Australia) – (Gordimer, South Africa)

Partie II “Systèmes d’Informations Géographiques” : Pour cette partie, vous avez le droit et il est conseillé de consulter la documentation et accéder aux services disponibles sur Internet. Faites des réponses courtes et synthétiques, pas de copie exhaustive de doc, car la capacité de synthèse compte dans la note (moitié des points maximum si non respecté). Barème à titre indicatif, susceptible de changer.

1. (1pt) Quel est le système cartographique utilisé en France et quels sont ses points essentiels ? (faites une réponse synthétique en deux lignes maximum)

. La projection de Lambert est le SIG officiel de la France. Cette dernière est une projection conique qui conserve les angles.

- Elle est représentée avec des parallèles de référence sur l’hémisphère nord (gauche) et sur l’hémisphère sud (droite) et Les distances sont exactes uniquement le long des parallèles de référence - Tous les méridiens sont des lignes droites équidistantes convergeant vers un point commun, qui est le pôle le plus proche des parallèles de référence.
- Il ya eu la nouvelle version qui est la projection conique Lambert-93 en 9 zones (CC9) sur la France métropolitaine inspiré de (RGF93)

2.(1pt) Quelle est la différence entre le modèle d’Egenhofer et al. des 9 intersections 9IM et

celui de Clementini et al. normalisé par l'Open Geospatial Consortium et implémenté dans PostGis ? (faites une réponse synthétique en deux lignes maximum)

Dans le cas du modèle d'Egenhofer la caractérisation de l'intersection est binaire alors que le modèle de Clementini distingue si c'est un point, une ligne ou une surface.

3. a) (1pt) Regardez les données GEOJSON fournies en WFS par le Grand Lyon pour les voirie et équipements cyclables de Lyon et les Vélo'v. Écrivez les url et premier élément qui est fourni.

Vélo'v :

[https://download.data.grandlyon.com/wfs/grandlyon?
SERVICE=WFS&VERSION=2.0.0&request=GetFeature&typename=pvo_patrimoine_voirie.
pvostationvelov&outputFormat=GML3&SRSNAME=EPSG:4171&startIndex=0&count=100](https://download.data.grandlyon.com/wfs/grandlyon?SERVICE=WFS&VERSION=2.0.0&request=GetFeature&typename=pvo_patrimoine_voirie.pvostationvelov&outputFormat=GML3&SRSNAME=EPSG:4171&startIndex=0&count=100)

```
<wfs:member>
<ms:pvo_patrimoine_voirie.pvostationvelov
gml:id="pvo_patrimoine_voirie.pvostationvelov.1">
<gml:boundedBy>
<gml:Envelope srsName="urn:ogc:def:crs:EPSG::4171">
<gml:lowerCorner>45.748520 4.840525</gml:lowerCorner>
<gml:upperCorner>45.748520 4.840525</gml:upperCorner>
</gml:Envelope>
</gml:boundedBy>
<ms:msGeometry>
<gml:Point srsName="urn:ogc:def:crs:EPSG::4171">
<gml:pos>45.748520 4.840525</gml:pos>
</gml:Point>
</ms:msGeometry>
<ms:idstation>7024</ms:idstation>
<ms:nom>Bancel / Chevreul</ms:nom>
<ms:adresse1>Rue Bancel</ms:adresse1>
<ms:adresse2>Rue Chevreul</ms:adresse2>
<ms:commune>Lyon 7 ème</ms:commune>
<ms:numdansarrondissement>24</ms:numdansarrondissement>
<ms:nbbornettes>21</ms:nbbornettes>
<ms:stationbonus/>
<ms:pole>Itinéraire cyclable</ms:pole>
<ms:ouverte>Oui</ms:ouverte>
<ms:achevement/>
<ms:gid>1</ms:gid>
```

```
<ms:code_insee>69387</ms:code_insee>
</ms:pvo_patrimoine_voirie.pvostationvelov>
</wfs:member>
```

- Elle est représentée avec des parallèles de référence sur l'hémisphère nord (gauche) et sur l'hémisphère sud (droite) et Les distances sont exactes uniquement le long des parallèles de référence
- Tous les méridiens sont des lignes droites équidistantes convergeant vers un point commun, qui est le pôle le plus proche des parallèles de référence.
- Il ya eu la nouvelle version qui est la projection conique Lambert-93 en 9 zones (CC9) sur la France métropolitaine inspiré de (RGF93)

Voix et équipement cyclable :

```
https://download.data.grandlyon.com/wfs/grandlyon?
SERVICE=WFS&VERSION=2.0.0&request=GetFeature&typename=pvo_patrimoine_voirie.
pvoamenagementcyclable&outputFormat=GML3&SRSNAME=EPSG:4171&startIndex=0&count=100
```

```
<wfs:member>
<ms:pvo_patrimoine_voirie.pvoamenagementcyclable
gml:id="pvo_patrimoine_voirie.pvoamenagementcyclable.1">
<gml:boundedBy>
<gml:Envelope srsName="urn:ogc:def:crs:EPSG::4171">
<gml:lowerCorner>45.764920 4.770609</gml:lowerCorner>
<gml:upperCorner>45.765795 4.772414</gml:upperCorner>
</gml:Envelope>
</gml:boundedBy>
<ms:msGeometry>
<gml:LineString srsName="urn:ogc:def:crs:EPSG::4171">
<gml:posList srsDimension="2">
45.764920 4.772414 45.764959 4.772365 45.765077 4.772056 45.765261 4.771660
45.765436 4.771309 45.765640 4.770908 45.765692 4.770808 45.765795 4.770609
</gml:posList>
</gml:LineString>
</ms:msGeometry>
<ms:nom>Boulevard du Valvert</ms:nom>
<ms:commune1>ECULLY</ms:commune1>
<ms:commune2/>
<ms:reseau>Réseau secondaire</ms:reseau>
<ms:typeamenagement>Piste Cyclable</ms:typeamenagement>
<ms:typeamenagement2/>
<ms:positionnement>Unilatérale unidirectionnelle</ms:positionnement>
<ms:senscirculation>Sens Unique</ms:senscirculation>
<ms:environnement>Voie de circulation</ms:environnement>
<ms:localisation>Sur chaussée</ms:localisation>
<ms:typologiepiste>Piste sur chaussée</ms:typologiepiste>
<ms:revetementpiste>
Matériaux liés (asphaltes, enrobés, bétons et nouveaux liants)
</ms:revetementpiste>
```

```
<ms:domanialite>Métropole</ms:domanialite>
<ms:reglementation>Vélo obligatoire</ms:reglementation>
<ms:zonecirculationapaisee/>
<ms:anneelivraison/>
<ms:identifiant>9</ms:identifiant>
<ms:observation/>
<ms:gid>1</ms:gid>
</ms:pvo_patrimoine_voirie.pvoamenagementcyclable>
</wfs:member>
```

3.b) (1pt) On souhaiterait stocker localement ces données pour les afficher d'abord sur Qgis et à terme sur une application avec un fond GoogleMaps. On exécute d'abord :

```
ogr2ogr -lco geometry_name=g -f "PostgreSQL" PG:"host=localhost dbname=d user=u
password=p" ./amcycl.json -nln ac
ogr2ogr -lco geometry_name=g -f "PostgreSQL" PG:"host=localhost dbname=d user=u
password=p" ./velov.json -nln vv
```

Faut-il vérifier et se méfier de certaines choses, voire intervenir pour modifier ?

Oui , s'assurer que les données sont compatible avec les données du systeme destinations.

3.c) (1pt) Quel premier tuple affichera `SELECT *, ST_ASTEXT(g) FROM ac` ? Votre réponse doit être lisible, formatez ainsi :

nom_attribut1	type	valeur
nom_attribut2	type	valeur
nom_attribut3	type	valeur
Etc...		

