

Bases de Données II, Charleroi

Jef Wijssen

4 janvier 2010

Cahier ouvert. Durée : 1 heure

Nom et prénom
Année

La figure 2 montre une base de données XML concernant des peintures et des fresques. La base enregistre, entre autres, que la Mona Lisa est une peinture de Leonardo Da Vinci et que L'Ultima Cena (la Cène) est une fresque de Leonardo Da Vinci. Leonardo da Vinci, identifié par le code LV, est né en 1425 et mort en 1519. La figure 1 montre le DTD.

```
<!-- This file is called art.dtd -->
<!ELEMENT ART (ARTISTS, PAINTINGS, FRESCOS)>
<!ELEMENT ARTISTS (artist)*>
<!ELEMENT PAINTINGS (painting)*>
<!ELEMENT FRESCOS (fresco)*>
<!ELEMENT painting (title, artist)>
<!ELEMENT fresco (title, artist)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT artist (#PCDATA)>
<!ATTLIST artist id CDATA #REQUIRED>
<!ATTLIST artist sex CDATA #IMPLIED>
<!ATTLIST artist birth CDATA #IMPLIED>
<!ATTLIST artist death CDATA #IMPLIED>
<!ATTLIST painting year CDATA #REQUIRED>
<!ATTLIST fresco year CDATA #REQUIRED>
```

Figure 1: DTD.

```

<?xml version="1.0"?>
<!DOCTYPE ART SYSTEM "art.dtd">
<ART>
<ARTISTS>
  <artist id="PP" sex="M" birth="1881" death="1973">Pablo Picasso</artist>
  <artist id="LV" sex="M" birth="1452" death="1519">Leonardo Da Vinci</artist>
  <artist id="MA" sex="M" birth="1475" death="1564">Michelangelo</artist>
</ARTISTS>
<PAINTINGS>
  <painting year="1904"><title>The Madman</title>
    <artist id="PP"/>
  </painting>
  <painting year="1939"><title>Night Fishing at Antibes</title>
    <artist id="PP"/>
  </painting>
  <painting year="1507"><title>Mona Lisa</title>
    <artist id="LV"/>
  </painting>
  <painting year="1475"><title>The Annunciation</title>
    <artist id="LV"/>
  </painting>
  <painting year="1506"><title>Doni Tondo</title>
    <artist id="MA"/>
  </painting>
</PAINTINGS>
<FRESCOS>
  <fresco year="1498"><title>L'Ultima Cena</title>
    <artist id="LV"/>
  </fresco>
  <fresco year="1512"><title>Sistine Chapel ceiling</title>
    <artist id="MA"/>
  </fresco>
</FRESCOS>
</ART>

```

Figure 2: Fichier XML avec des informations sur des peintures et des fresques.

Question 1 Écrivez une expression XPath qui rend la date de décès de Leonardo Da Vinci.

.../6

Question 2 Écrivez une expression XPath qui rend le nom de l'artiste qui a peint la fresque *L'Ultima Cena*.

.../8

Question 3 Écrivez une expression XPath qui rend les titres des peintures de Pablo Picasso. Notez : l'expression doit utiliser le nom "Pablo Picasso", pas l'identifiant "PP". Voici le résultat :

```
<title>The Madman</title>  
<title>Night Fishing at Antibes</title>
```

.../8

Question 4 Écrivez un programme XSLT qui affiche pour chaque artiste les peintures et les fresques qu'il a peintes, dans le format qui suit :

```
<?xml version="1.0" ?>
<artistes>
  <artiste><nom>Pablo Picasso</nom>
    <peintures><titre>The Madman</titre>
      <titre>Night Fishing at Antibes</titre>
    </peintures>
    <fresques />
  </artiste>
  <artiste><nom>Leonardo Da Vinci</nom>
    <peintures><titre>Mona Lisa</titre>
      <titre>The Annunciation</titre>
    </peintures>
    <fresques><titre>L'Ultima Cena</titre>
    </fresques>
  </artiste>
  <artiste><nom>Michelangelo</nom>
    <peintures><titre>Doni Tondo</titre>
    </peintures>
    <fresques><titre>Sistine Chapel ceiling</titre>
    </fresques>
  </artiste>
</artistes>
```


Cahier fermé. Durée : 2 heures

Nom et prénom
Année

Situez chaque terme dans le cursus et expliquez de façon succincte mais précise.

Question 5 Cophenetic distance.

	.../5
--	-------

Question 6 Ordinal attribute.

	.../5
--	-------

Question 7 Oblique decision tree.

.../5

Question 8 Vertical data layout of transaction data set.

.../5

Expliquez chaque figure de façon détaillée.

Question 9 Figure 1.

.../10

Question 10 Figure 2.

.../10

Question 11 Figure 3.

.../10

Question 12 Figure 4.

.../10

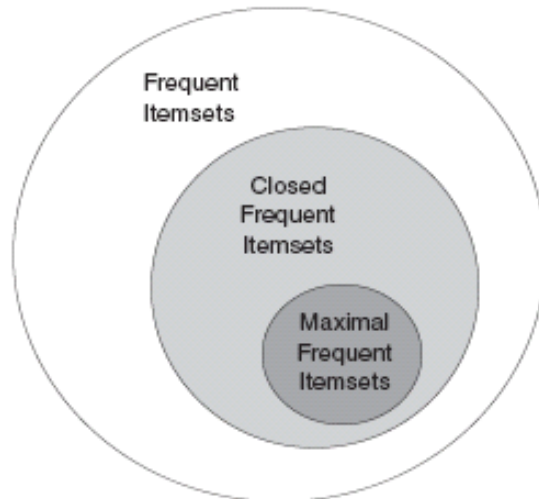


Figure 6.18. Relationships among frequent, maximal frequent, and closed frequent itemsets.

Figure 1

5. Decision trees provide an expressive representation for learning discrete-valued functions. However, they do not generalize well to certain types of Boolean problems. One notable example is the parity function, whose value is 0 (1) when there is an odd (even) number of Boolean attributes with the value *True*. Accurate modeling of such a function requires a full decision tree with 2^d nodes, where d is the number of Boolean attributes (see Exercise 1 on page 198).

Figure 2

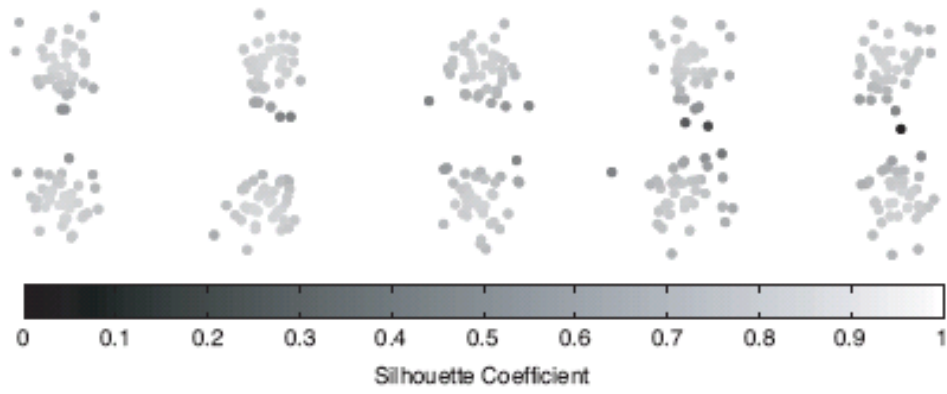


Figure 8.29. Silhouette coefficients for points in ten clusters.

Figure 3

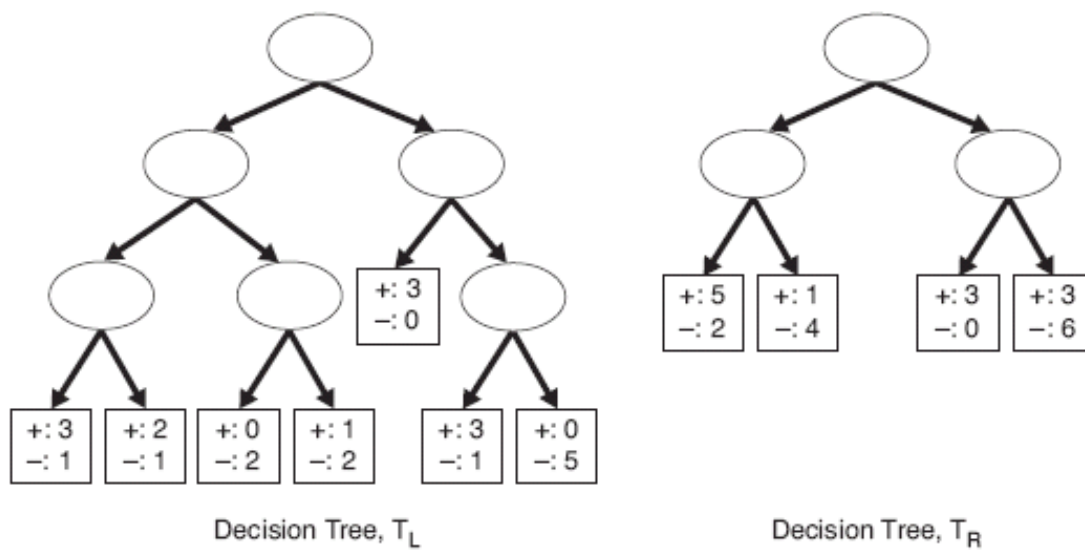


Figure 4.27. Example of two decision trees generated from the same training data.

Figure 4