Concept Hierarchies

By Susan L. Miertschin
What is a Concept Hierarchy?

• A way to organize concepts defined in a knowledge domain
• Multilevel organization of concepts
• Links represent relationships between two concepts
  – can be two-way (bidirectional) or one-way
Concept Hierarchy Example - Book Genre Taxonomy
Data Mining Concept Hierarchies (ESX)

- Three levels
  - Root
  - Concept
  - Instance
Concept Hierarchy Data Structure

- Concept 1
  - Summary Statistics (Instances Belonging to Concept 1)
    - Instance 1-1
      - Raw Data
    - Instance 1-2
      - Raw Data
    - Instance 1-k1
      - Raw Data

- Concept 2
  - Summary Statistics (Instances Belonging to Concept 2)
    - Instance 2-1
      - Raw Data
    - Instance 2-2
      - Raw Data
    - Instance 2-1
      - Raw Data

- Concept n
  - Summary Statistics (Instances Belonging to Concept n)
    - Instance n-1
      - Raw Data
    - Instance n-2
      - Raw Data
    - Instance n-kn
      - Raw Data
Evaluation of Concept Hierarchy

• Class resemblance scores
  – A measure of overall similarity based on the exemplars (instances) making up each concept class

• Compute the average class resemblance (sum of all resemblances, divided by number of classes)

• Classify a new instance into a class to achieve highest (or lowest) change in average resemblance

• Supervised » classes are pre-determined

• Unsupervised » classes can be added to raise average resemblance score
Concept Hierarchies

By Susan L. Miertschin