

What Makes Patterns Interesting  
in  
Knowledge Discovery Systems

Avi Silberschatz and Alexander Tuzhilin

Jerry Jarrard

# Introduction

- Knowledge Discovery systems
  - Glut of patterns
  - most uninteresting to user
- Measures of interestingness
  - Objective
    - Structure and underlying data
    - Rule  $A \longrightarrow B$ 
      - Support
      - confidence
  - Subjective
    - Structure and underlying data
    - User

# Classification of Interestingness

- Unexpectedness
  - Is it surprising to user
- Actionability
  - Can user act on it to his advantage

Database of student evaluations  
Evaluation(Term, Year, Course,  
Section, Instructor,  
Instructor\_rating, Course\_rating)

- Example 1
- For most course evaluations  
 $\text{instructor\_rating} > \text{course\_rating}$
- For course A  
 $\text{instructor\_rating} < \text{course\_rating}$ 
  - unexpected
  - actionable

- Example 2
- For most courses: 60%-90% response rate.
- For course B: 8% responded
  - unexpected
  - not actionable
  
- Example 3
- Considering Prof. Y for “teacher-of-the-year award.”
- Consistently high ratings.
- Latest rating also high.
  - expected
  - actionable

# Unexpectedness vs. Actionability

- Key concept is Actionability
  - elusive
  - difficult to capture formally
- Most actionable items are unexpected
- Most unexpected items are actionable
- Address actionability through unexpectedness

# Unexpectedness and beliefs

- Hard Belief
  - Can't change with new evidence
  - Example: number of responses
- Soft belief
  - Changes with new evidence
  - assigned a degree that specifies the strength of belief

# Bayesian approach

- Degree of Belief  $\alpha$  :  $P(\alpha|\xi)$
- Given new evidence

$$P(\alpha|E, \xi) = \frac{P(E|\alpha, \xi)P(\alpha|\xi)}{P(E|\alpha, \xi)P(\alpha|\xi) + P(E|\neg\alpha, \xi)P(\neg\alpha|\xi)}$$

# Interestingness of Patterns

- Hard Belief
  - Contradictions always interesting
- Soft Belief

$$I(p, B, \xi) = \sum_{\alpha_i \in B} w_i |d(\alpha_i | p, \xi) - d(\alpha_i | \xi)|$$

$$\sum_{\alpha_i \in B} w_i = 1$$

# Proposed Discovery Scheme

- New data  $\longrightarrow$  Revise degrees of all beliefs
- if some change  $>$ threshold  $\longrightarrow$  interesting patterns in the data
- Launch discovery process

# Summary

- Interestingness: objective, subjective
- Subjective: Unexpected, actionable
- Capture actionability via unexpectedness
- Defined interestingness in terms of affect upon existing belief system.
- Use interestingness to discover interesting patterns in the data.