A Fuzzy-Graph-Based Approach to the Determination of Interestingness of Association Rules

Conference Paper *in* Lecture Notes in Computer Science · December 2002

DOI: 10.1007/3-540-36277-0_34 · Source: DBLP

2 authors:

B. Shekar
Indian Institute of Management Bangalore

Rajesh Natarajan
University Carlos III de Madrid

A Fuzzy-Graph-Based approach to the determination of Interestingness of Association Rules

B. Shekar and Rajesh Natarajan

Indian Institute of Management Bangalore,

Bangalore, India

Knowledge Management

- Knowledge- a significant organizational resource
- Knowledge Management Systems (KMS)
- Knowledge creation, storage/retrieval, transfer and application
- Tacit knowledge

Knowledge Discovery in Databases

Knowledge discovery in databases (KDD) is the non-trivial process of identifying **valid**, **novel**, **potentially useful** and **ultimately understandable patterns** in large databases (*Fayyad, et al., 1996*)

Association Rules

- Implication rules that bring out hidden relationships among attributes and/or items using **co-occurrence** as the basis.
- Bring out natural **affinity** between items purchased together in customer transactions
- Cannot indicate the exact nature of relationship
- Useful in exploratory studies

Association Rules (Cont'd.)

An Association Rule is defined to be an implication rule of the form

A ? B ,

where A, B ? I and A ? B = ?

i.e. A and B can be a single item or a set of items from I and there is no item that is common to the sets A and B.

Association Rules (Cont'd.)

Rule : A ? B

$$\text{Support} \;?\; \frac{\textbf{No. of transactions containing A and B}}{\textbf{Total no. of transactions in the database}}$$

$$\text{Confidence} \;?\; \frac{\textbf{No. of transactions containing A and B}}{\textbf{Total no. of transactions containing A}}$$

Association Rules (Cont'd.)

Bread* ? *Butter

{conf.= 80% support=10%} means

"that *80%* of the transactions that *contain bread also tend to contain butter*, and they *together occur in 10% of the overall transactions*"

Characteristics of Association Rule Mining Algorithms

- ✍Massive search through the database
- ✍Complete examination of the database
- ✍Parameters of the rules are exact, no estimation
- ✍Very High Demands on I/O, Memory and Computation
- ✍High Degree of Automation: Very little user interaction

Consequences

- Large number of Rules Discovered (**Rule Quantity problem**)
  - Example: Census database, No. of rules generated=23,712 (*Brin, Motwani, et al., 1997*)
  - Telephone company fault management (Simultaneous faults occurrence in a network)

    Rules generated=1,426 (*Mannila, et al., 1994*)
- Many of the rules are obvious to the user/domain expert (**Rule Quality problem**)

  e.g. Has Licence ? Age >18;

  "five year olds do not work"

  "Unemployed residents do not earn income from work" (*Brin, Motwani, et al.,1997*)

Consequences

- Really novel, relevant, useful and interesting rules are very few
  - **Example:** Rules from a Hurricane Database (*Major and Mangano, 1995*)
  - Started with 529 rules and ended up with only 19 rules that were really useful and relevant.
- Very difficult for the domain expert/user to get an overview of the domain
- Typical users have limited time at their disposal

Interestingness Measures

- ✍Quantify the amount of interest that a rule(pattern) is supposed to evoke in a user examining it
- ✍Elusive concept that is very difficult to capture (*Silberschatz, 1996*)
- ✍Objective and Subjective
- ✍Actionability, Unexpectedness, User Knowledge/Goals, Anticipation

Item Relatedness

- Relationships between items
  - Generic categories (primary functional purpose)
  - Secondary functionalities
  - Domains of application
- Item-pairs in Association Rules have high frequency of occurrence
- Item pairs that contain unrelated or weakly related items are interesting
- *Bread ? Butter*
- *Chocolate ? Paper Napkins, Chocolate ? Pencils*
- *Beer? Diapers*

Approach towards Finding Interestingness of Association rules

- Relationships between items are captured in the structure of a Fuzzy Taxonomy
- Define an Item-relatedness measure based on the structural aspects of the Fuzzy Taxonomy
- Find out the relatedness between item-pairs
- Combine relatedness to arrive at an interestingness value for the rule

Fuzzy Taxonomy

- Extension of the traditional concept hierarchy (taxonomy) tree
- Allows us to express the following kinds of relationships
  - An item belonging to two or more higher-level concepts
  - Items used to substitute the functions of other items to various extents (e.g. spoon, knife etc.)

Fuzzy Taxonomy

- A child need not be a full member of the category/concept of parent
- Connections restricted between leaf-level items and higher-level concept nodes
- Membership function: Extent to which the child node belongs to its parent ; µ ? [0,1]
- Membership Transfer: Child to parent to ancestors
- Membership Transfer: Membership grade of a child node ‘c’ in its ancestor ‘a’

  $? _{(c,a)} = \max\{ ? _{(c,a)} (k)\}$

  where k = 1, 2, ….N.

  $? _{(c,a)}$ (k) is the membership grade of the child node ‘c’ in its ancestor ‘a’ by the virtue of the path ‘k’.

Highest-level node of path [$H_{A, B}$ (p)]

- Node that occurs at the highest level (i.e. nearest to the root node) in the simple path 'p' connecting items A and B
- Plays a pivotal role in determining relatedness between items
- Closest common context that relates the two items

Highest-level Node Membership [$HM_{A,\,B}(p)$]

$HM_{A,\,B}\,(p) = \min\,[\;?_{\,A,\,H(A,\,B)}\,(p),\;\;?_{\,B,\,H(A,\,B)}\,(p)]$

example: $HM_{(\text{knife, shoes})}\,(p) = \min\,[0.5, 0.6] = 0.5$;
where the path ‘p’ is the fuzzy path connecting the two items

Why minimum?

Highest-level Relatedness [$HR_{A,B}(p)$]

- Two siblings at higher level of a taxonomy are less related to each other as compared to two siblings at a lower level of the taxonomy
- Longer path from root node to highest-level node implies higher relatedness (Why?)
- **$HR_{A,B}(p)$ = level [$H_{A,B}(p)$]**

Node Separation Relatedness [$NSR_{A,B}$ (p)]

- Length of path connecting two items indicates the conceptual distance between them.
- $NSR_{A,B}$ (p) = Length of the simple path ‘p’ connecting nodes A and B
- Inversely related to ‘relatedness’

Overall Relatedness

- Each path between two items contributes a component of relatedness between the two items

$$\mathbf{OR_{A,B}(p)} \;?\; \frac{\mathbf{(1 \;?\; HR_{A,B}(p))(HM_{A,B}(p))}}{\mathbf{NSR_{A,B}(p)}}$$

Total Relatedness

$$\mathbf{TR(A,B)}\ ?\ \underset{\mathbf{p}}{?}\ \frac{\mathbf{OR}_{\mathbf{A,B}}(p)}{\mathbf{K}}\ ?\ \underset{\mathbf{p}}{?}\ \frac{(\ \mathbf{1}\ ?\ \mathbf{HR}_{\mathbf{A,B}}\mathbf{(p)}\ )(\ \mathbf{HM}_{\mathbf{A,B}}\ \mathbf{(p)}\ )}{\mathbf{K}\ ?\ \mathbf{NSR}_{\mathbf{A,B}}\mathbf{(p)}}$$

Total Relatedness varies from 0 to p,
where ‘p’ is the maximum number of paths that exist between two items

Determination of total relatedness between *Knife and Shoes* (Table 1)

Path	A= Knife	B= Shoes	NSR	HR	HM	OR
I	Crisp (1.0)	Crisp (1.0)	3	1	1.0	0.166
II	Fuzzy (0.5)	Fuzzy (0.6)	5	1	0.5	0.05
III	Fuzzy (0.5)	Crisp (1.0)	6	0	0.5	0.021
IV	Crisp (1.0)	Fuzzy (0.6)	6	0	0.6	0.025

Table 2: A comparison of Item Relatedness for sample item-pairs.

No	Item A	Item B	Path I (Crisp Crisp)	Path II (Fuzzy Fuzzy)	Path III (Fuzzy Crisp)	Path IV (Crisp Fuzzy)	TR (A, B)
1	Knife	Shoes	0.1668	0.05	0.020825	0.025	0.2626
2	Shoes	Safety Shoes	0.0417	0.0167	0.6	0.3	0.9583
3	Spoon	Shoes	0.1667	----	----	0.025	0.19175
4	Spoon	Safety Shoes	0.0417	----	----	0.0667	0.10835
5	Knife	Industri-al Knife	0.0417	0.0125	0.5	0.225	0.77918

Advantages

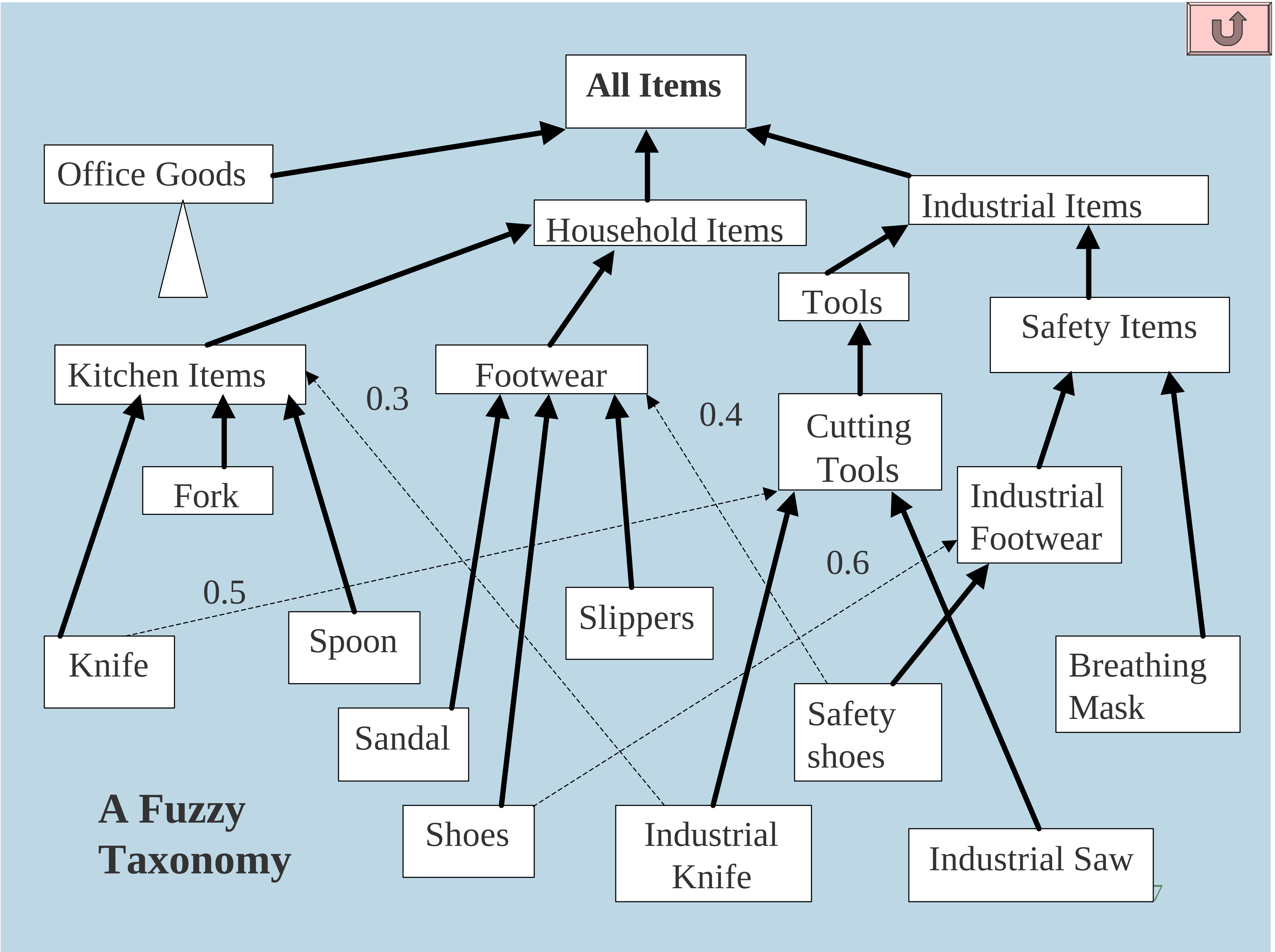
- Gradual development of the Fuzzy Taxonomy by an iterative process
- Easier for a manager to give his views
- Tacit knowledge of managers is brought out
- Only rules as identified as interesting need to be examined

Applications

- Shelf-space design
- Display design
- Discount coupons design
- Selection of products for product bundling to induce sales

Thank You

Questions, Clarifications, Comments, Suggestions,..... are **Most Welcome**


All Items
Office Goods
Household Items
Industrial Items
Tools
Safety Items
Kitchen Items
Footwear
0.3
0.4
Cutting Tools
Industrial Footwear
Fork
0.6
0.5
Slippers
Knife
Spoon
Safety shoes
Breathing Mask
Sandal
A Fuzzy Taxonomy
Shoes
Industrial Knife
Industrial Saw

Algorithm

Compute all paths between an item and root node

Path[A,B,k]=

{[Ipath(A,i)? Ipath(B,j)]-[Ipath(A,i)? Ipath(B,j)]}

? H[Ipath(A,i)? Ipath(B,j)]

where

Ipath(A,i)= ith path between A and root node

Ipath(B,j)=jth path between B and root node

H [Ipath(A,i)? Ipath(B,j)]=highest-level node

Use of Relatedness to Compute Interestingness of ARs

- Interestingness and Relatedness are opposing notions
- Relatedness:Item Pairs
- AR: A ? B
- Antecedent Item Pairs
  - ({a_i, a_j}, where i?j and a_i ? A)
- Consequent Item Pairs
  - a pair {b_i, b_j} where i?j and b_i ? B

Use of Relatedness to Compute Interestingness of ARs (Cont'd)

- Antecedent-Consequent Item Pairs
  - a pair $\{a_i , b_j \}$ where a_i ? A and b_j ? B, A ? B = ? and A, B? I.
- Different Item-pairs might have different Interestingness for Users
- Relatedness Values for all such pairs are computed

Interestingness Measure

$$\mathbf{I} ? \left(\frac{\mathbf{m_1}}{\mathbf{m_1} ? \mathbf{m_2} ? \mathbf{m_3}} \right) \left(\frac{\mathbf{1}}{\mathbf{minTR(AP)}} \right) ? \left(\frac{\mathbf{m_2}}{\mathbf{m_1} ? \mathbf{m_2} ? \mathbf{m_3}} \right) \left(\frac{\mathbf{1}}{\mathbf{minTR(CP)}} \right)$$

$$? \left(\frac{\mathbf{m_3}}{\mathbf{m_1} ? \mathbf{m_2} ? \mathbf{m_3}} \right) \left(\frac{\mathbf{1}}{\mathbf{minTR(ACP)}} \right)$$

min TR(AP) = minimum (TR(a_i, a_j)) ? {a_i, a_j} /a_i ? A and i?j ;

min TR(CP) = minimum (TR (b_i, b_j)) ? {b_i, b_j}/ b_i ? B and i?j ;

min TR(ACP) = minimum (TR(a_i , b_j)) ? {a_i , b_j}/ a_i ? A and b_j ? B, A ? B = ? and A, B? I;

Interestingness Measure

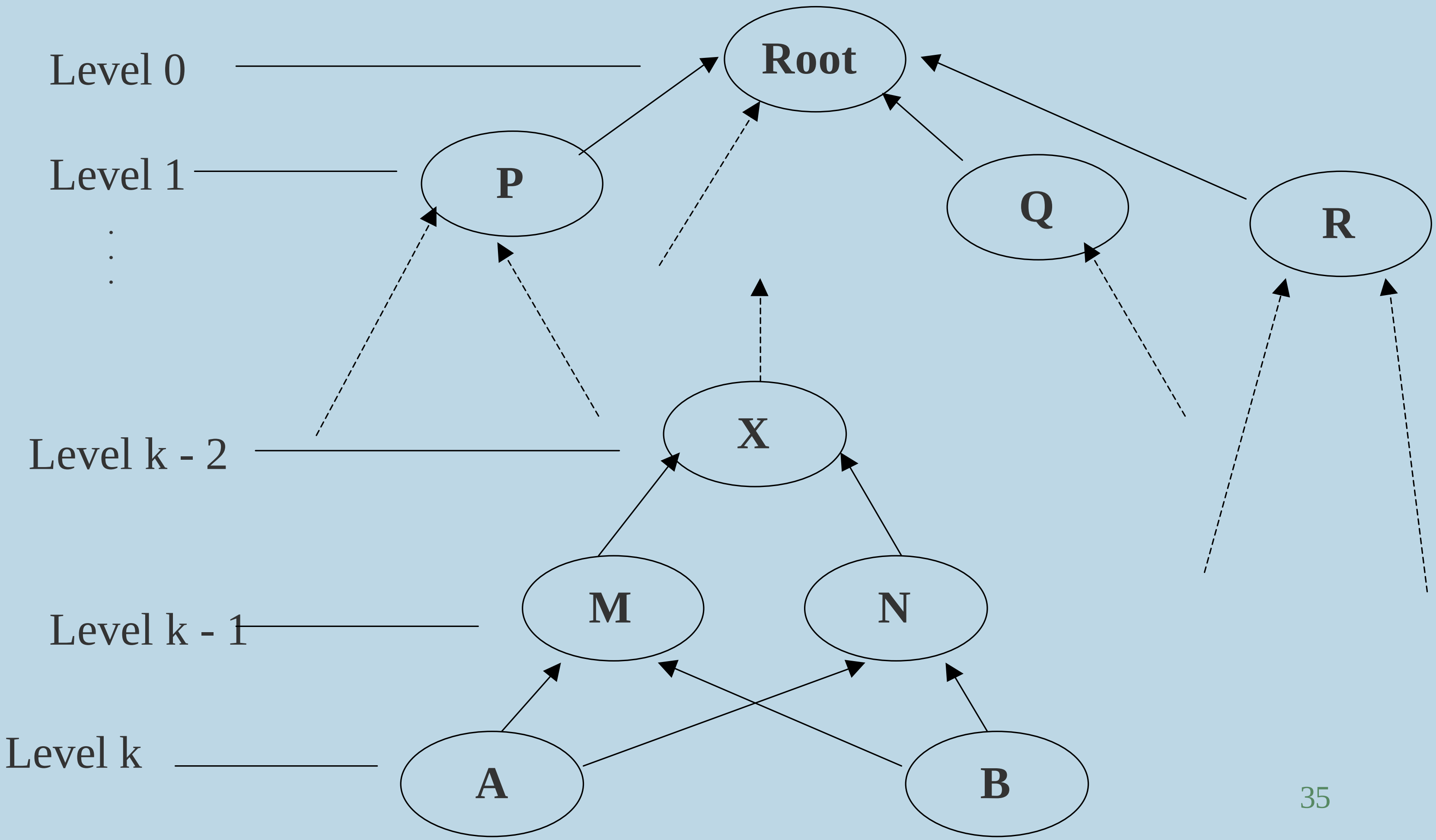
- ***m_1, m_2 and m_3*** are three numbers used to appropriately weigh the interestingness of antecedent, consequent and antecedent-consequent pairs respectively
- Here, ***$m_1 = m_2 = m_3 = 1$***
- Implicit assumption: Greater number of conditions gives more information; leads to greater interestingness
- Direction of rule not consequential

Sr. No.	(Table 3) Rule	I $m_1, m_2, m_3 = 1$
1	{Knife}? {Shoes}	1.2693
2	{Spoon} ? {Shoes}	1.7384
3	{Industrial Knife} ? {Safety Shoes}	1.8601
4	{Knife}? {Safety Shoes}	1.9043
5	{Knife} ? {Shoes, Safety Shoes}	1.9391
6	{Industrial Knife} ? {Shoes}	2.0305
7	{Knife, Industrial Knife} ? {Safety Shoes}	2.3321
8	{Industrial Knife} ? {Shoes, Safety Shoes}	2.3783
9	{Knife, Industrial Knife} ? {Shoes}	2.4583
10	{Knife, Industrial Knife} ? {Shoes, Safety Shoes}	2.8061
11	{Spoon} ? {Safety Shoes}	3.0764

Related Work

- Taxonomies: Generalized ARs (Srikant et al.,1995), Negative ARs(Navathe, et al.,1998)
- Fuzzy Taxonomy (Chen et al. 2000)
- Interestingness in Fuzzy Taxonomies (Graff et al.,2000, 2001)
- Interestingness of Summaries (Hamilton et al.,1995)
- Text Mining (Basu, et al., 2001)
- Noun Sense Disambiguation (Lin et al., 2002)

Maximum Relatedness Contributed by a path



Maximum Relatedness Contributed by a path

- Maximum relatedness contributed by a path in a fuzzy taxonomy of depth k is ‘k’
- Use ‘k’ to normalize the measure
- Each contribution varies from 0 to 1

Approaches in Literature

- ✍Rule Templates (*Klemettinen et al., 1994*)
- ✍Rule Covers (*Toivonen, 1995*)
- ✍Mining specific kinds of association patterns (Negative Association rules, Profile Association rules, Cyclic rules etc.)
- ✍Summarization (GSE patterns)
- ✍Grouping and Clustering
- ✍Interestingness Measures

Rule Templates

- ✍General structure of the rule given by the user in the form of inclusive/restrictive templates
- ✍Neural Networks ? Design and Analysis of Algorithms (0.48, 0.02)
- ✍Template: Graduate Course, Any Course ? Design and Analysis of Algorithms
- ✍User explicitly gives what is interesting and what is not.

General Rules, Summaries and Exceptions (Liu, Hu et al. 2000)

✍**Example of a GSE pattern**

(General Rule, summary and Exceptions)

A1=a => X (sup=41%, conf=77%)

Except R7: A1=a, A2=b => Y

(sup=10%, conf=83%)

Objective Measures of Interestingness

- ✍Depend on the structure of the pattern
- ✍Underlying data used in the discovery process
- ✍Domain Independent
- ✍Each measure looks at a specific aspect of the data
- ✍Cannot capture all the complexities of the pattern discovery process

Objective Measures of Interestingness: Examples

- Discovery of Surprising patterns like Simpson's Paradox
- Statistical Measures like: support, confidence, conviction, implication, rule interest etc.
- Information theoretic measures like Entropy, Gini, J-Measure, Hellinger Measures etc.

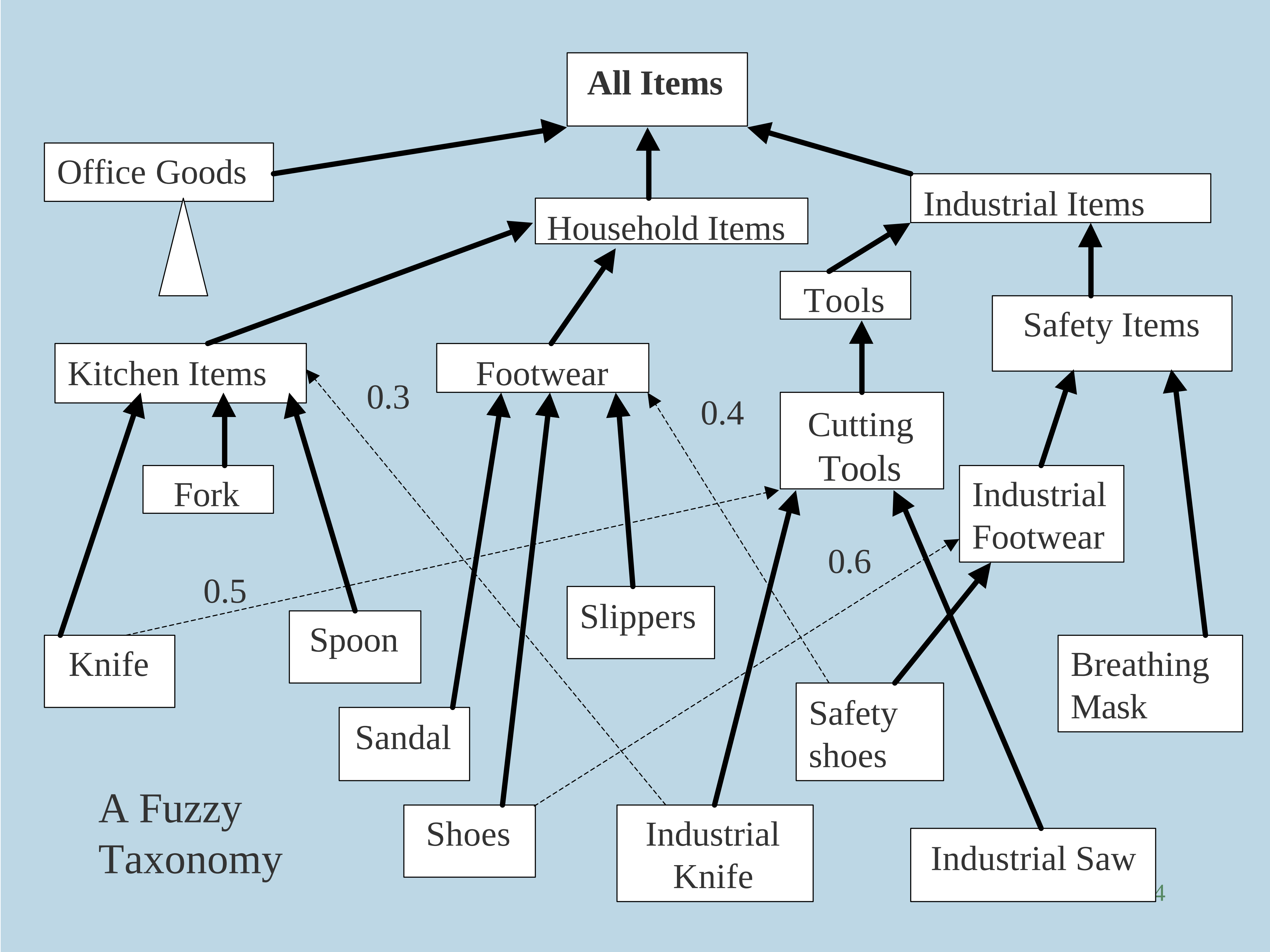
Subjective Measures of Interestingness

- ✍Also depend on the **User** who examines the pattern in addition to structure of the pattern and the characteristics of the data/data generating process
- ✍Mostly Domain Dependent
- ✍Unexpectedness (*Silberschatz, 1996*)
- ✍Actionability (Usefulness) (*Adomavicius, et al. 1997*)
- ✍Anticipation (*Roddick and Rice, 2001*)
- ✍User Goals/User Knowledge (*Ram,A. ;1990*)

Subjective Measures of Interestingness

General Approach

1. Elicit user opinions/beliefs
2. Express them in a form suitable for computation/Measure for comparison
3. Mine patterns from the database
4. Compare User Opinions with mined patterns
5. Patterns that deviate the most from user beliefs are deemed interesting


All Items
Office Goods
Household Items
Industrial Items
Tools
Safety Items
Kitchen Items
Footwear
Cutting Tools
Industrial Footwear
Fork
Spoon
Knife
Sandal
Shoes
Slippers
Industrial Knife
Safety shoes
Industrial Saw
Breathing Mask
0.3
0.4
0.5
0.6
A Fuzzy Taxonomy