Granular Computing for Machine Learning

JingTao Yao 姚静涛
Department of Computer Science,
University of Regina
jtyao@cs.uregina.ca
http://www.cs.uregina.ca/~jtyao

Introduction

- Granular computing (GrC) is an umbrella term to cover theories, methodologies, techniques, and tools that make use of granules in problem solving.
- Some basic concepts have been studied in other fields such as
  - belief functions, artificial intelligence, cluster analysis, chunking, data compression, databases, decision trees, divide and conquer, interval computing, machine learning from examples, structure programming, quantization, quotient space theory, and rough set theory.

GrC and Fuzzy Sets

- 1979, Zadeh first discussed the notion of fuzzy information granulation.
- 1997, Zadeh discussed information granulation again.
- 1997, T. Y Lin suggests the term “granular computing” (GrC), BISC special interest group (BSIC-GrC).
- 2004, IEEE Computational Intelligence Society, Task Force on Granular Computing.

GrC and Rough Sets

- 1982, Pawlak introduced the notion of rough sets.
- 1998, the GrC view of rough sets was discussed by many researchers (Lin, Pawlak, Skowron, Yao, and many more).
- Rough set theory can be viewed as a concrete example of granular computing.

Current

- Fuzzy set and rough set theories are the main driving force of GrC.
- Most researchers in GrC are from fuzzy set or rough set community.
- The connections to other fields and the generality, flexibility, and potential of GrC are under exploration.
- 1st IEEE GrC, July 25-27, 2005, Beijing

GrC Studies

- Two levels of study
  - Philosophical level
    - Conceptual, abstract, structured thinking
  - Application level
    - Problem solving techniques
- Basic issues
  - Components
    - granule, granulated views and levels, hierarchies, and granular structures
  - Operations on them
    - granulation, computing with granules.

Philosophy Level

- Human knowledge is normally organized in a multiple level of hierarchy.
- The lower (basic) level consists of directly perceivable concepts.
- The higher levels consists of more abstract concepts.

Concept Formation & Organization

- Concepts are the basic units of human thoughts that are essential for representing knowledge and its communication.
- Concepts are coded by natural language words.
- Granularity plays a key role in natural language. Some words are more general (in meaning) than some others.

Research Articles

- High level of abstraction
  - title, abstract
- Middle levels of abstraction
  - chapter/section titles
  - subsection titles
  - subsubsection titles
- Low level of abstraction
  - Words
- 八股文：起承转合，排比对偶

Human Problem Solving

- Human perceives and represents real world at different levels of granularity.
- Human understands real world problems, and their solutions, at different levels of abstraction.
- Human can focus on the right level of granularity and change granularity easily.

Knowledge Structure and Education

- Experts and novices differ in their knowledge organization.
- Experts are able to establish multiple representations of the same problem at different levels of granularity.
- Experts are able to see the connections between different grain-sized knowledge.
- 苏步青读书过程：把书读薄，再读厚

Structured Programming

- Top-down design and step-wise refinement:
  - Design a program in multiple level of detail.
  - Formulation, verification and testing of each level.

Granule

- Any subset, class, object, or cluster of a universe is called a granule
- *These granules are composed of finer granules that are drawn together by distinguishability, similarity, and functionality (Zadeh 1996)*
- Granules may have different formats and meaning when used in different particle models

Granulation

- *Granulation involves a decomposition of whole into parts. Organization involves an integration of parts into whole (Zadeh 96)*
- Extended concept:
  - granulation involves the process of two directions: construction and decomposition

Granulation

- Construction involves the process of forming a larger and higher level granule with smaller and lower level sub-granules.
  - This is a bottom-up process.
- The decomposition involves the process of dividing a larger granule into smaller and lower level granules.
  - This is a top-down process.

Granular Relationship

- Intrarelationship: relationships amongst sub-granules.
- A granule is a *clump of points drawn together by similarity, indistinguishability, and functionality.*
- Interrelationship: relationships amongst granules

Refinement and Coarsening

- Comparison of granules
- A granule X is a refinement of another granule Y if every sub-granule or element of X is contained in sub-granules of Y.
- X is finer than Y, Y is coarser than X

Partitions and Coverings

- Following way if a set-theoretical approach
- A partition of a set *U* is a collection of non-empty, and pairwise disjoint subset of *U* whose union is *U*.
- A covering of a set *U* is a collection of non-empty subset of *U* whose union is *U*.
- Partitions are a special case of coverings.
- A non-redundant covering
  - if any collection of subsets of *U* derived by deleting one or more granules from it is not covering.

Partial Ordering

- A granule X is a partial-refinement of another granule Y if some sub-granule or element of X is contained in sub-granules of Y.
- X is p-finer than Y, Y is p-coarser than X
- A fine relationship can be viewed as a special case of a p-fine relationship.

ISA Relationships

- ISA can be considered as a special case of refinement.
- If X is finer than Y and X inherit all properties of Y, we say X ISA Y.
  - children's hospital ISA hospital.
  - children's hospital is finer than hospital.
- Some refinement are not ISA
  - emergency department is finer than hospital
  - NOT emergency department ISA hospital

Similarity

- Inter and/or intra relationship
- A key measure to put elements/sub-granule in a granule
- Similarity measures: distance
- Similarity fuzzy sets

Fuzzy Relationships

- Zadeh's model
  - o = {X | X isr R}
  - where X is a value taken from a universe and R is a constraining relation.
  - Constraints: equality, possibility, probability, fuzzy and verity.
  - if X isr_1 A then Y isr_2 B

Type of Relationships

- Binary: finer
- u-nary: covering
- Hierarchical: hyperlink

Formal Concept Analysis

- A concept is a unit of thoughts consisting of two parts, the intension and extension of the concept.
- The intension of a concept
  - The sum of the attributes contained in a term
  - Consists of all properties or attributes that are valid for all those objects to which the concept applies.
  - Meaning, or its complete definition of a concept
- The extension of a concept
  - The class of objects designated by a specific term or concept
  - The set of objects or entities which are instances of the concept.
  - The collection, or set, of things to which the concept applies.

Formal Concept Analysis

- A concept is described jointly by its intension (a set of properties) and extension (a set of objects).
- The intension of a concept can be expressed by a formula, or an expression, of a certain language.
- The extension of a concept is presented as a set of objects satisfy the formula.

Data Mining

- A process extracting interesting information or patterns from large databases.
- Concept formation and concept relationship Identification are main tasks of knowledge discovery and data mining.

Machine Learning

- Machine learning refers to a system capable of the autonomous acquisition and integration of knowledge. This capacity to learn from experience, analytical observation, and other means, results in a system that can continuously self-improve and thereby offer increased efficiency and effectiveness. (AAAI)
- Machine learning usually refers to the changes in systems that perform tasks associated with artificial intelligence (AI). Such tasks involve recognition, diagnosis, planning, robot control, prediction, etc. ... (Nilsson 1996)

The Task of Classification

- A well-studied field of machine learning, data mining and pattern recognition.
- Two main purposes:
  - Describing the classification of labeled instances in the training dataset
  - Predicting the unseen new instances

Classification

- Task of classification occurs in a wide range of human activity.
- Forecasting or for decision making based on available information.
- Learning classification rules from examples.
- Forecast by patterns/rules
- Statistical, neural networks, machine learning

Information Tables

$$S = (U, A_t, L, \{V_a \mid a \in A_t\}, \{I_a \mid a \in A_t\})$$

- U: a finite nonempty set of objects.
- A_t: a finite nonempty set of attributes.
- L: a language defined using attributes in A_t.
- V_a: a nonempty set of values for $a \in A_t$
- $I_a : U \to V_a$ is an information function.

An Information Table

Object	height	hair	eyes	class
o_1	short	blond	blue	+
o_2	short	blond	brown	-
o_3	tall	red	blue	+
o_4	tall	dark	blue	-
o_5	tall	dark	blue	-
o_6	tall	blond	blue	+
o_7	tall	dark	brown	-
o_8	short	blond	brown	-

Concept Formation

- Atomic formula: $a=v$ ($a \in A_t$, $v \in V_a$)
  - If φ, ψ are formulas, so is $\varphi \wedge \psi$
  - If a formula is a conjunction of atomic formulas we call it a conjunctor.
- Meaning of a formula:
  - $m(\varphi)=\{x \in U \mid x \models \varphi\}$
  - $x \models a=v$ *iff* $I_a(x)=v$
- A definable concept is a pair (φ, $m(\varphi)$)
  - φ is the intension of the concept
  - $m(\varphi)$ is the extension of the concept

Concept Examples

- Formulas:
  - **hair** = dark, **eyes** =blue ∧ **hair**=blond
- Meanings:
  - m(**hair** = dark) = $\{o_4,o_5,o_7\}$
  - m(**eyes** =blue ∧ **hair**=blond)=$\{o_1,o_6\}$
- A concept:
  - (**height** = tall ∧ **hair**=dark, $\{o_4,o_5,o_7\}$)

Partition

- A partition of a set *U* is a collection of non-empty, and pairwise disjoint subset of *U* whose union is *U*.
- The subsets in a partition are called blocks or equivalence granules.

Covering

- A covering of a set *U* is a collection of non-empty subset of *U* whose union is *U*.
- A non-redundant covering
  - if any collection of subsets of *U* derived by deleting one or more granules from it is not covering.
- The subsets in a partition are called blocks.

(Conjunctively) Definable Granule

- A subset $X \subseteq U$ is called a definable granule in an information table *S* if there exists at least one formula φ such that $m(\varphi) = X$.
- A subset $X \subseteq U$ is a conjunctively definable granule in an information table *S* if there exists a conjunctor φ such that $m(\varphi) = X$.
- (Conjunctively) definable partition.
- (Conjunctively) definable covering.

Refinement

- A partition π_1 is refinement of another partition π_2, or equivalently, π_2 is a coarsening of π_1, denoted by $\pi_1 \preccurlyeq \pi_2$, if every block of π_1 is contained in some block of π_2.
- Covering refinement (substitute with τ)
- $\tau \preccurlyeq \pi$ holds

Different Level of Measures

- For a single granule.
  - Generality.
- For a pair of granules.
  - Confidence, covering.
- For a granule and a family of granules.
  - Conditional entropy

Classification Problems

- Assume that each object is associated with a unique class label.
- Objects are divided into disjoint classes which form a partition of the universe.
- The set of attributes is expressed as $A_t = F \cup \{class\}$, where F is the set of attributes used to describe the objects.
- To find classification rules of the form, $\varphi \Rightarrow class = c_i$, where φ is a formula over F and c_i is a class label.

Solution to Classification Problems

- The partition solution to a consistent classification problem is a conjunctively definable partition π such that $\pi \preccurlyeq \pi_{class}$.
- The covering solution to a consistent classification problem is a conjunctively definable covering τ such that $\tau \preccurlyeq \pi_{class}$.

Classification Algorithms

- Attribute oriented approaches
  - ID3, C4.5: Entropy based selection
  - Rough set: approximation equality
  - Average
  - Multi-class
- Attribute value oriented approaches
  - PRISM
  - Best granule
  - Local
  - Single class
- Granular computing approach: Nature combination

An Example

- $\pi_{class} = \{\{o_1,o_3,o_6\}\ \{o_2,o_4,o_5,o_7,o_8\}\}$
- $\pi = \{\{o_1,o_6\}, \{o_2,o_8\}, \{o_3\}, \{o_4,o_5,o_7\}\}$
  - $\pi \preccurlyeq \pi_{class}$
  - **eyes** =blue ∧ **hair**=blond ⇒ **class** = +
  - **height** =short ∧ **eyes** =brown ⇒ **class** = -
  - **hair** =red ⇒ **class** = +
  - **height** =tall ∧ **hair**=dark ⇒ **class** = -
- $\tau = \{\{o_1,o_6\}, \{o_2,o_7,o_8\}, \{o_3\}, \{o_4,o_5,o_7\}\}$
  - $\tau \preccurlyeq \pi_{class}$
  - **eyes** =brown ⇒ **class** = -

Granule Networks

- Modification of decision tree
- Each node is labelled by a subset of objects
- The arc leading from a larger granule to a smaller granule is labelled by an atomic formula
- The smaller granule is obtained by selecting those objects of the larger granule that satisfy the atomic formula

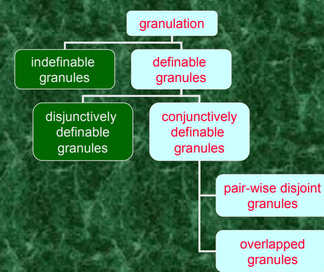
Granule Networks

- The pair (a=v, m(a=v)) is called a basic concept
- Each node is a conjunction of some basic granules, and thus a conjunctively definable granule.
- The granule network for a classification problem can constructed by a top-down search of granules.

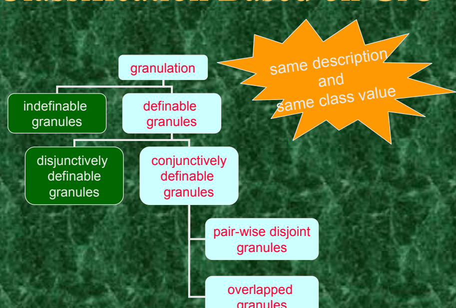
A Construction Algorithm

- **Construct** the family of basic concept with respect to atomic formulas:
  - $BC(U) = (a=v, m(a=v)) \mid a \in F, v \in V_a\}$
- **Set** the granule network to $GN = (\{U\}, \theta)$, which is a graph consists of only one node and no arc.
- **While** the set of inactive nodes is not a non-redundant covering solution of the consistent classification problem:
  - **Select** the active node with the maximum value of activity**.**
  - **Compute** the fitness of each unused basic concept.
  - **Select** the basic concept $bc=(a=v, m(a=v))$ with maximum value of fitness with respect to the selected active node.
  - **Modify** the granule network GN by adding bc to the selected active node; **connect** the new nodes by arcs labelled by $a = v$.

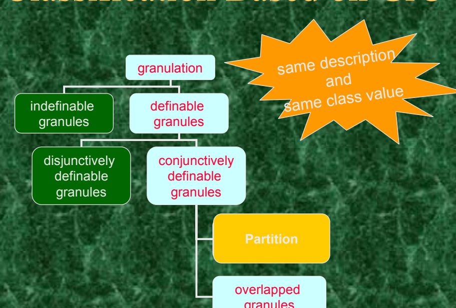
Granular Computing (GrC)



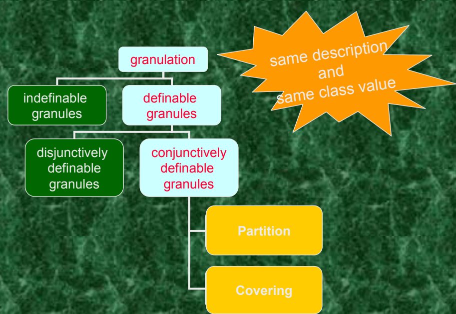
Classification Based on GrC



Classification Based on GrC

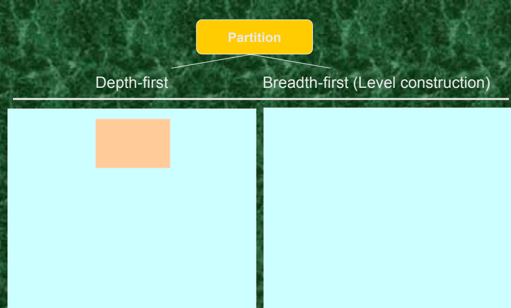


Classification Based on GrC

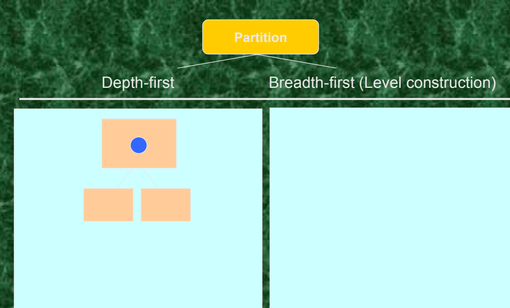


Partition-based Classification

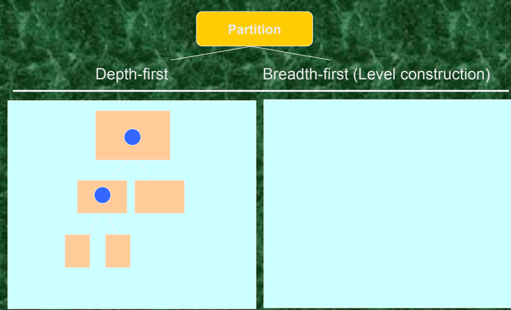
Partition-based Classification



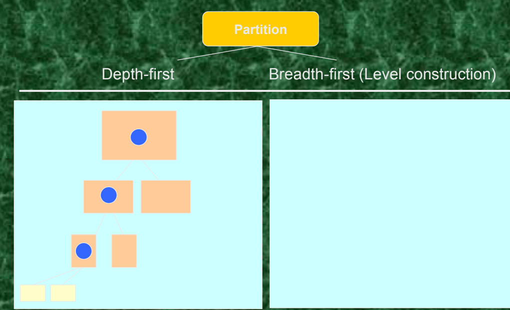
Partition-based Classification



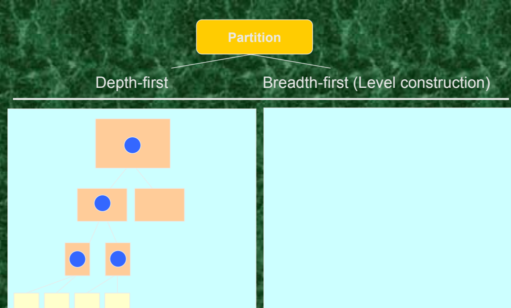
Partition-based Classification



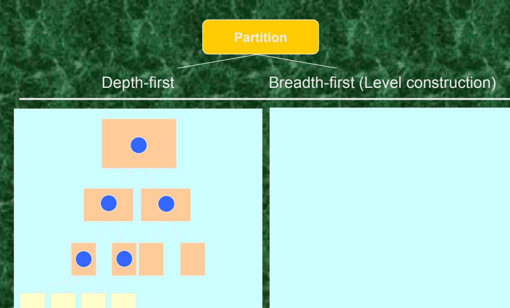
Partition-based Classification

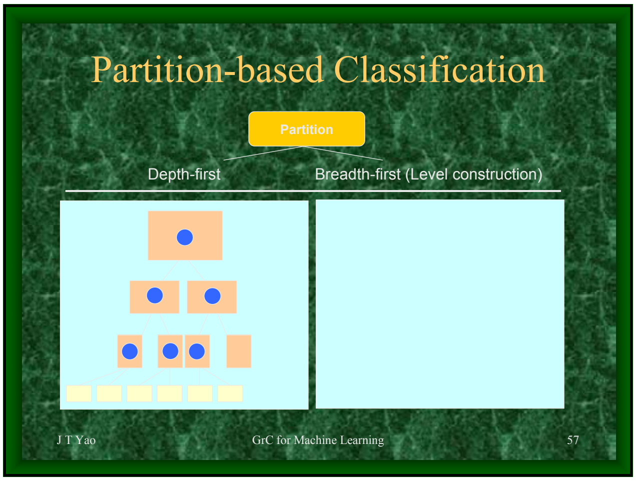


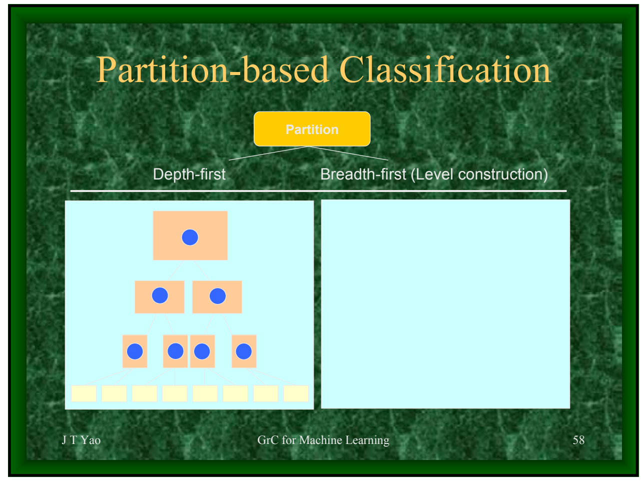
Partition-based Classification

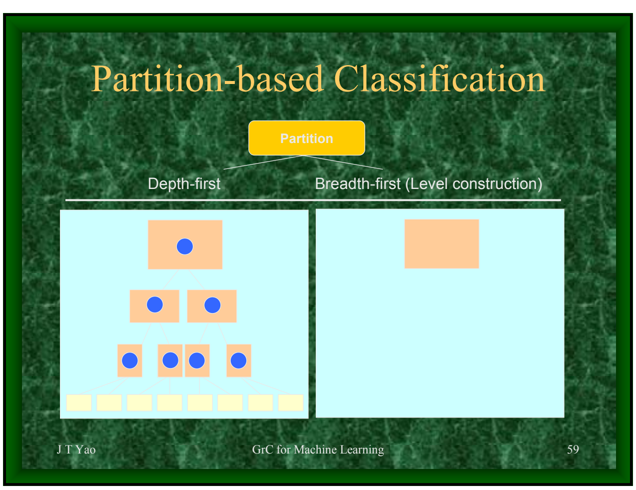


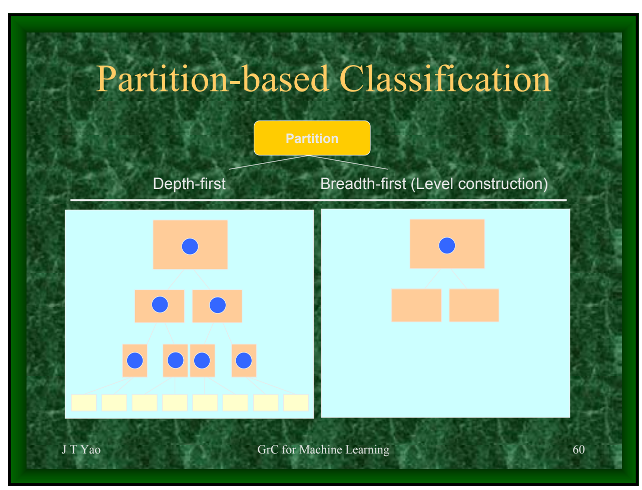
Partition-based Classification



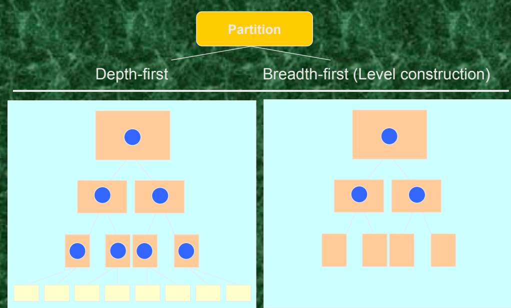

Partition-based Classification
Partition
Depth-first
Breadth-first (Level construction)
J T Yao
GrC for Machine Learning
57


Partition-based Classification
Partition
Depth-first
Breadth-first (Level construction)
J T Yao
GrC for Machine Learning
58

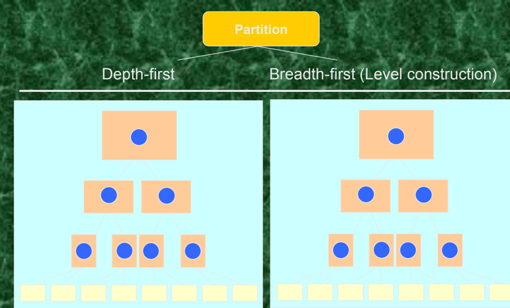

Partition-based Classification
Partition
Depth-first
Breadth-first (Level construction)
J T Yao
GrC for Machine Learning
59


Partition-based Classification
Partition
Depth-first
Breadth-first (Level construction)
J T Yao
GrC for Machine Learning
60

Partition-based Classification



Partition-based Classification



Concluding Remarks

- GrC is an interesting research area with great potential.
- One needs to focus on different levels of study of GrC.
  - The conceptual development.
  - The formulation of various concrete models (at different levels).

Concluding Remarks

- The philosophy and general principles of GrC is of fundamental value to effective and efficient problem solving.
- GrC may play an important role in the design and implementation of next generation information processing systems.

Granular Computing for Machine Learning

JingTao Yao 姚静涛
Department of Computer Science,
University of Regina
jtyao@cs.uregina.ca
http://www.cs.uregina.ca/~jtyao