计算智能

第13讲: 粒度计算

周水庚

计算机科学技术学院

2017-6-6

Introduction to Granular Computing

Outline

- Background
- Concepts
- Data Mining: in GrC view
- Conclusion

Introduction to Granular Computing

Background

Granular Computing

- Granular Computing (GrC)
  - An umbrella term to cover any theories, methodologies, techniques, and tools that make use of **granules** in problem solving.
  - A subset of the universe is called a **granule** in granular computing.
  - Basic ingredients of granular computing are subsets, classes, and clusters of a universe.

Historical Notes

Soft computing perspectives (fuzzy set perspectives):

- 1979, Zadeh first discussed the notion of fuzzy information granulation.
- 1997, Zadeh discussed information granulation again.
- 1997, the term "granular computing" (GrC) was suggested by T.Y. Lin, and a BISC special interest group (BSIC-GrC) is formed.
- 2004, IEEE NN (Computational Intelligence) Society, Task Force on Granular Computing is formed.

BISC: Berkeley Initiative in Soft Computing

Historical Notes

Rough sets perspectives

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

Historical Notes

- 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 have not been fully explored

The Concept of GrC is not New

- The basic ideas and principles of GrC have appeared in many fields
  - Artificial intelligence, Programming
  - Cluster analysis, Interval computing
  - Quotient space theory
  - Belief functions
  - Machine learning, Data mining
  - Databases, and many more.

Philosophy: Human Knowledge

- 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 and 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
- One can easily observe that granularity plays a key role in natural language. Some words are more general (in meaning) than some others

Technical Writings

- One can easily observe multiple levels of granularity in any technical writing:
- High level of abstraction
  - title, abstract
- Middle levels of abstraction
  - chapter/section titles
  - subsection titles
  - subsubsection titles
- Low level of abstraction
  - text

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 Problem Solving in Physics

Reif and Heller, 1982.

"Effective problem solving in a realistic domain depends crucially on the content and structure of the knowledge about the particular domain."

"Knowledge structures and problem-solving procedures of experts and novices differ in significant ways."

The knowledge about physics "specifies special descriptive concepts and relations described at various level of abstractness, is organized hierarchically, and is accompanied by explicit guidelines specifying when and how this knowledge is to be applied."

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

CS: Structured Programming

Top-down design and step-wise refinement:

- Design a program in multiple level of detail
- Formulation, verification and testing of each level

Top-down Theorem Proving

- Computer science: PROLOG, top-down theorem proving
- Mathematics: proving and writing proofs in multiple levels of detail

AI: Search

Quotient space theory (商空间理论，Zhang and Zhang 1992)

- Representation of state space at different levels of granularity
- Search a fine-grained space if the coarse-grained (quotient) space is promising.

AI: Hierarchical Planning

Planning in multiple levels of detail (Knoblock, 1993)

- An outline plan is structurally equivalent to a detailed plan
- It is related to hierarchical search

AI: A Theory of Granularity

Hobbs, 1985:

- “We look at the world under various grain sizes and abstract from it only those things that serve our present interest.”
- “Our ability to conceptualize the world at different granularities and to switch among these granularities is fundamental to our intelligence and flexibility.”
- “It enables us to map the complexities of the world around us into simpler theories that are computational tractable to reason in.”

AI: A Theory of Abstraction

Giunchigalia and Walsh, 1992.

- Abstraction may be thought as a process that "allows people to consider what is relevant and to forget a lot of irrelevant details which would get in the way of what they are trying to do."
- Levels of abstractions

AI: More

- Natural language understanding: granularity of meanings.
- Intelligent tutoring:
  granular structure of knowledge.
- Granulation of time and space:
  temporal and spatial reasoning.

Introduction to Granular Computing

Concepts

What is GrC?

- There does not exist a generally accepted definition of GrC.
- There does not exist a well formulated and unified model of GrC.
- Many studies focus on particular models/methods of GrC.
- Majority of studies of GrC is related to fuzzy sets and rough sets.

What is GrC?

- **GrC = Problem solving based on different levels of granularity (detail/abstraction)**.
- Level of granularity is essential to human problem solving.
- GrC attempts to capture the basic principles and methodologies used by human in problem solving. It models human problem solving qualitatively and quantitatively.

What is GrC?

- GrC provides a more general framework that covers many studies. It extracts the commonality from diversity of fields.
- GrC needs to move beyond fuzzy sets and rough sets.
- GrC is used as an umbrella term to label the study of a family of granule-oriented theories, methods and tools, for problem solving.

What is GrC?

GrC must be treated as a separate and interdisciplinary research field on its own right. It has its own principles, theories, and applications.

What is GrC?

- GrC leads to clarity and simplicity.
- GrC leads to multiple level understanding.
- GrC is more tolerant to uncertainty.
- GrC reduce costs by focusing on approximate solutions (solution at a higher level of granularity).

… … …

What is GrC?

GrC can be studied based on its own principles (understanding of GrC in levels).

- Philosophy level:

  GrC focuses on structured thinking.

- Implementation level:

  GrC deals with structured problem solving.

A Framework of GrC

- Basic components:
  - Granules
  - Granulated views
  - Hierarchies.
- Basic structures:
  - Internal structure of a granule
  - Collective structure of granulated view (a family of granules)
  - Overall structures of a family of granulated views

Granules

- Granules are regarded to as the primitive notion of granular computing.
- A granule may be interpreted as one of the numerous small particles forming a larger unit.
- A granule may be considered as a localized view or a specific aspect of a large unit.

Granules

- The physical meaning of granules become clearer in a concrete model.
- In a set-theoretic model, a granule may be a subset of a universal set (rough sets, fuzzy sets, cluster analysis, etc.).
- In planning, a granule may be a sub-plan.
- In theorem proving, a granule may be a sub-theorem.

Granules

- The size of a granule may be considered as a basic property.
- It may be interpreted as the degree of abstraction, concreteness, or details.
- In a set-theoretic setting, the cardinality may be used to define the size of a granule.

Granules

- Connections and relationships between granules can be modeled by binary relations.
- They may be interpreted as dependency, closeness, overlapping, etc.
- Based on the notion of size, one can define order relations, such as “greater than or equal to”, “more abstract than”, coarser than”, etc.

Granules

- Operations can also be defined on granules.
- One can combine many granules into one or decompose a granule into many.
- The operations must be consistent with the relationships between granules.

Granulated Views and Levels

Marr, 1982

A full understanding of an information processing system involves explanations at various levels.

Many studies used the notion of levels.

Granulated Views and Levels

Foster, 1992

Three basics issues:

the definition of levels,

the number of levels,

relationships between levels.

Granulated Views and Levels

Foster, 1992

- A level is interpreted as a description or a point of view.
- The number of levels is not fixed.
- A multi-layered theory of levels captures two senses of abstractions:

    concreteness,

    amount of details.

Granulated Views and Levels

- A level consists of a family of granules that provide a complete description of a problem.
- Each entity in a level is a granule.
- Level = Granulated view

  = a family of granules

Granulated Views and Levels

- Granules in a level are formed with respect to a particular degree of granularity or detail.
- There are two types of information or knowledge encoded by a level:
  - a granule captures a particular aspect;
  - all granules provide a collective description.

Hierarchies

- Granules in different levels are linked by the order relations and operations on granules.
- The order relation can be used to define order relations on levels.
- The ordering of levels can be described by hierarchies.

Hierarchies

- A higher level may provide constraint to and/or context of a lower level.
- A higher level may contain or be made of lower levels.
- A hierarchy may be interpreted as levels of abstraction, levels of concreteness, levels of organization, and levels of detail.

Hierarchies

- A granule in a higher level can be decomposed into many granules in a lower level.
- A granule in a lower level may be a more detailed description of a granule in a higher level.

Granular Structures

Internal structure of a granule:

- At a particular level, a granule is normally viewed as a whole.
- The internal structure of a granule need to be examined. It provides a proper description, interpretation, and the characterization of a granule.
- Such a structure is useful in granularity conversion.

Granular Structures

The structure of a granulated view:

- Granules in a granulated view are normally independent.
- They are also related to a certain degree.
- The collective structure of granules in a granulated view is only meaningful is all granules are considered together.

Granular Structures

Overall structure of a hierarchy:

- It reflects both the internal structures of granules, and collective structures of granules in a granulated view.
- Two arbitrary granulated views may not be comparable.

Basic Issues of GrC

Two major tasks:

- Granulation;
- Computing and reasoning with granules.

Basic Issues of GrC

Algorithmic vs. semantic studies:

- Algorithmic studies focus on procedures for granulation and related computational methods.
- Semantics studies focus on the interpretation and physical meaningfulness of various algorithms.

Granulation

Granulation criteria:

- Why two objects are put into the same granule.
- Meaningfulness of the internal structure of a granule.
- Meaningfulness of the collective structures of a family of granules.
- Meaningfulness of a hierarchy.

Granulation

Granulation methods:

- How to put objects together to form a granule?
- Construction methods of granules, granulated views, and hierarchies.

Granulation

Representation/description:

- Interpretation of the results from a granulation method.
- Find a suitable description of granules and granulated views.

Granulation

Qualitative and quantitative characterization:

Associate measures to the three components, i.e., granules, granulated views, and hierarchy.

Computing with Granules

Mappings:

The connections between different granulated views can be defined by mappings. They links granules together.

Computing with Granules

Granularity conversion:

- A basic task of computing with granules is to change granularity when moving between different granulated views.
- A move to a detailed view reveals additional relevant information.
- A move to a coarse-grained view omits some irrelevant details.

Computing with Granules

Operators:

- Operators formally define the conversion of granularity.
- One type of operators deals with refinement (zooming-in).
- The other type of operators deals with coarsening (zooming-out).

Computing with Granules

Property preservation:

- Computing with granules is based on principles of property preservation.
- A higher level must preserve the relevant properties of a lower level, but with less precision or accuracy.

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
  - 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 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.

Introduction to Granular Computing

Data Mining in GrC View

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.

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 $a \in A_t$
- $I_a : U \rightarrow V_a$ 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:
  - $m(\varphi)=\{x \in U \mid x \vDash \varphi\}$
  - $x \vDash a{=}v$ *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 $\wedge$ **hair**=blond
- Meanings:
  - m(**hair** = dark) = $\{o_4,o_5,o_7\}$
  - m(**eyes** =blue $\wedge$ **hair**=blond)=$\{o_1,o_6\}$
- A concept:
  - (**height** = tall $\wedge$ **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 $\tau \preccurlyeq \pi_{class}$.

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 $\wedge$ **hair**=blond $\Rightarrow$ **class** = +
  - **height** =short $\wedge$ **eyes** =brown $\Rightarrow$ **class** = -
  - **hair** =red $\Rightarrow$ **class** = +
  - **height** =tall $\wedge$ **hair**=dark $\Rightarrow$ **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 $\Rightarrow$ **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)) | a ∈ F, v ∈ V_a}
- **Set** the granule network to GN = ({U},θ), 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.

Introduction to Granular Computing

Conclusion

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.

Concluding Remarks

- By using GrC as an example, we want to demonstrate the one needs to move beyond the typical algorithm oriented study.
- One need to study a topic at various levels.
- The conceptual level study, although extremely important, has not received enough attention.