Converting Spatiotemporal Data Among Multiple Granularity Systems

Muhao Chen
Department of Computer Science, University of California Los Angeles
muhaochen@ucla.edu

ABSTRACT

Spatiotemporal data are often expressed in terms of granularities in a granularity system to indicate the measurement units of the data. A granularity system usually consists of a set of granularities that share a "common refined granularity" (CRG) to ensure granular comparison and data conversion within the system. However, if data from multiple granularity systems need to be used in a unified application, it is necessary to extend the data conversion and comparison within a granularity system to those for multiple granularity systems. This paper proposes a formal framework to enable such an extension. The framework involves essentially some preconditions and properties for verifying existence of a CRG and unifying conversions of incongruous semantics, and supports the approach to integrate multiple systems into one processing granular interoperation across systems just like in a single system. Quantification of uncertainty in granularity conversion is also considered to improve the precision of granular comparison.

Categories and Subject Descriptors

H.2.1 [**Database Management**]: Logical Design—*Data models*; H.2.8 [**Database Management**]: Database Applications—*Spatial databases and GIS*;

General Terms

Theory; Design

Keywords

Spatiotemporal data, multiple granularity systems, granularity conversion, granular comparison, system combination

1. INTRODUCTION

In this decade where over 80% datasets have spatial features and are usually with temporal components [3], the notion of granularity has become significant for expressing and exchanging spatiotemporal data under specific units of measurement.

It is a common practice in literature to organize a group of granularities in a partial-order set or a lattice [2-5,7,9,11,15], where granularities are linked with a partial-order topological relation (hereafter *linking relation*) into a hierarchy set. Two major functions are normally associated with such a set, namely granularity conversion and granular comparison [4]. The former enables scaling for spatiotemporal data to express the same data in different measurement units, while the latter supports the topological or statistic analysis on spatially or temporally qualified information. These functionalities can be used in applications like multi-scaled retrieval, dynamic knowledge extraction [4,6], and in multi-dimensional databases to enhance various types of queries from big spatiotemporal datasets [14,18]. We may term a specific set of granularities as a *granularity system*.

The relevant literature has implicitly assumed that a single granularity system is sufficient for data in an application. However, when multiple applications need to be integrated or mashed into one, we may have a scenario where several granularity systems are simultaneously used for the data in the same spatial or temporal domain. Such coexistence usually results from different representation standards as well as separate data maintenance realms, causing *heterogeneity in granularities*. For example, time can be recorded in the solar calendar (a granularity system) and the lunar calendar (another granularity system). The government may collect civilian data bound with locations and time, managed separately by systems formed respectively with {climatic regions, provinces, direct municipalities, streets} and {economic regions, cities, districts, business areas}. Datasets can also be indexed by different multi-granular structures in different clusters of a distributed system. In addition, *heterogeneity in the linking relations* is reflected by the existing instances in the literature, such as FinerThan, GroupsInto, Partition, and CoarserThan [2,7,9,11,14,15]. A running example given as Example 3.1 in Section 3.2 will further illustrate such heterogeneity and subsequent problems.

Technically, realizing the interoperation of data across multiple granularity systems unifies current models from independent representation schemas to form a single global schema, essentially enabling reasoning and exchange of spatiotemporal data of multiple granularity systems. This will make it possible to reconstitute existing applications for new purposes, e.g., to support spatiotemporal queries and extraction of knowledge from various data sources or spatiotemporal-dependent resources regardless of how they are expressed in their respective original granularity systems.

However, this is a non-trivial problem with several new challenges to multi-granularity modeling. Besides accepting the system heterogeneity in a model, interoperation of data across systems inevitably requires extending the original in-system granularity conversion and granular comparison [4,9] to their inter-system equivalents. Heterogeneity of granularity systems often implies incongruous semantics of conversion, and indeterminacy of the existence of a common refined granularity (CRG) to support granular comparison. There also lacks a way to transform and organize all the granularities in an explicit unified structure. These challenges are currently without sufficient theoretical foundation to tackle.

Moreover, incongruity of linking relations further causes geometric uncertainty to conversion of data across granularity systems. Such uncertainty, which is not reflected by models in the current literature, causes imprecision to both representation and statistical analysis for spatiotemporal data. Thus, quantitative calculation of such uncertainty between two granularities aids in preserving the precision of granular data, which also makes it possible for us to find a certain composition of conversions with the least expected distortions when choosing an optimal CRG for granular comparison.

In this paper, we propose a formal framework to extend the granularity conversion and granular comparison of spatiotemporal data across multiple granularity systems. This framework first generalizes coexisting granularity systems to support their heterogeneity, and defines a graph model to represent their structure, and the semantics and uncertainty of conversions. It also defines two constraints for inter-system granularity conversions, namely *semantic preservation* and *semantic consistency*. We show that two granularity systems can be combined, or they have *combinabiliy*, only if they are semantically preserved or semantically consistent in addition to globally holding a CRG. An approach is given to combine multi-systems to a single lattice, where inter-system conversion and comparison can be processed transparently just like in a single system. Quantification of geometric and statistic distortion are also introduced for granularity conversions, so as to understand the (im)precision of granular comparison.

The rest of the paper is organized as follows. In the next section, we state the background with related work. In section 3, we model the granularity systems and related concepts. The combination of granularity systems and granularity conversion are discussed in section 4. Section 5 focuses on the uncertainty and common refined granularity search problem in granular comparison. In section 6, we provide a conceptual evaluation of our approach. And we conclude with section 7.

2. RELATED WORK

In computer science community, there are different definitions of granularities based on their modeling purposes. One common definition is the partition of a domain, e.g., the mapping from a domain to finite portions [7,9], and similar mapping with given semantic notion of "grain-size" in [15]. Some others add graph features to a spatial granularity to facilitate reasoning of granular data inside the granularity, e.g., the boundary pseudograph [10] and the labeled multi-digraph [2,3].

However the granularities are defined, most literature orders multiple granularities in a hierarchical structure either for multi-resolution representation, or for multi-scaling datasets. For example, the lattice of temporal granularities was proposed by Bettini et al in [5], which has later been transplanted to organize spatial granularities [4,7,9,11,15], as any of them benefits with a communal finest unit (i.e. the zero element) of representation and clarifies ordering of granularities w.r.t. their fine or coarse degree. Such hierarchical organization is also a support for scalable retrieval in stratified spatial or spatiotemporal datasets, like the pyramid structure in LARS [14]. Any instance of such a hierarchical set is a granularity system we discuss.

Usually a partial-order granularity relation associates granularities in a granularity system uniformly, such as FinerThan regulated in [7,9], GroupsInto and Partition additively considered in [2,4,5]. Based on that, the semantic and property of granularity conversion is defined. E.g. geometric congruity and topological consistency is guaranteed by Partition lattices, but not by FinerThan lattices [9]. Heterogeneity of such linking relations has been mentioned in [4]. Although these articles have respectively modeled granularity systems in the same or homeomorphous domains, none has considered the coexistence of multiple systems, nor the heterogeneity of systems which decides the property and semantics of significant in-system conversions, let along challenges when they are extended to inter-system. We accept the heterogeneity of multi-systems with a more general model to enable their coexistence.

Camossi et al has conducted *granularity conversion* and *granular comparison* as the fundamental challenges in current spatio-temporal multi-granularity research in [4]. In order to perform meaningful comparison, inter-granularity data must be converted to a CRG [4]. E.g. comparing the sales of two products bounded relatively with *months* and *weeks*, we are supposed to compare the aggregation values with them both refined to *days*. Thus, to enable inter-system granular comparisons, we must verify the existence of CRG for any pair of granularities from mult-systems. Due to the heterogeneity in granularities and linking relations, corresponding discussion, which is included in the proof of multi-system combinability, is essential to logically support inter-system granular comparison and enable O(1) implementation of such verification.

On the other hand, granularity conversion has been studied and designed in a few frameworks. Camossi et al [9], Moira et al [17] have respectively defined in-system granularity conversion operations w.r.t. the relation FinerThan, which preserves geometric correctness and topological consistency, and are compositional. Properties vary differently along with the conversion semantics in systems defined with other linking relations, e.g. the pyramid structure linked with CoveredBy [14]. To extend granularity conversion, the unpresented problem of semantic preservation and consistency are urged to be discussed so as to support the meaningful composition of granularity conversions across systems, which we focus on during the discussing of multiple systems combinability.

Moreover, due to finite precision of spatial granularities and possible incongruity of geometric properties satisfied by different linking relations, granularity conversions can lead to uncertainty (more precisely, vagueness [12]), which has been discussed by Wang and Liu in [11] with a classification based on granular coverage. But no quantization of such uncertainty for the granularity conversion is considered in literatures, even though it's essential for granular comparison and related data analysis. Thus not only the geometric distortion [4] of granules is uncontrollable, but also the imprecision of granular comparison and quantitive analysis of granular data that exists among systems and even among granularities, cannot be evaluated and controlled. We quantify such geometric or statistic distortion, and take the expectation of such distortion as edge-weight in weighted granularity graphs. In that way we can evaluate the geometric/statistic distortion in conversions and search for the optimal common refined granularity (OCRG) which has greatest expectation of geometric and statistical precision during granular comparison by finding the LCA [16].

3. MODELING GRANULARITY SYSTEMS

To raise against the problem, we begin with the modeling of spatial and temporal granularity systems.

3.1 Granularities and Granularity Relations

Granularities and granularity relations are the two major constituents of a granularity system. The former provides the units to measure or scale dimensional data, and the latter verifies topological associations between any pair of granularities.

A granularity forms with a partition on a Euclidean domain. The spatial and temporal granularities are defined as follows.

Definition 3.1 (*Spatial Granularity*): *A spatial granularity is defined with a mapping* $G_S: N \rightarrow P(S)$.

- $S \subseteq R^2$ *is a spatial extent of the granularity in the spatial domain* R^2*, while R is the real number field, and N is the natural number field. P(S) is the power set of S.*
- $\forall i \in N$, $G_S(i)$ *is a granule iff* $G_S(i)$ *is not empty.* $\forall i,j \in N$ *that* $i \neq j$*, if* $G_S(i)$ *and* $G_S(j)$ *are none-empty, then* $G_S(i) \cap G_S(j) = \emptyset$.

Definition 3.2 (*Temporal Granularity*): *A temporal granularity is defined with a mapping* $G_T: N \rightarrow P(T)$.

- $T \subseteq R$ *is the temporal extent of the granularity, while R is the real number field, and N is the natural number field. P(T) is the power set of T.*
- $\forall i \in N$, $G_T(i)$ *is a granule iff* $G_T(i)$ *is not empty.* $\forall i,j \in N$ *s.t.* $i<j$*, if* $G_T(i)$ *and* $G_T(j)$ *are none-empty, then each element of* $G_T(i)$ *is*

less than all elements of $G_T(j)$. $\forall i,j,k\epsilon N$ s.t. $i<j<k$, if $G_T(i)$ and $G_T(k)$ are non-empty, then $G_T(j)$ is non-empty.

A spatial granularity divides a spatial extent to finite or denumerable disjoint regions called *spatial granules*, which set the irresoluble base units for spatially qualified information. E.g., continents, nations, and climatic regions form several granularities on the world map. Similarly a temporal granularity divides a time extent to ordered and continuous intervals, known as *temporal granules*. E.g., years, months and weeks form temporal granularities on the timeline.

Given a granularity G, we refer G(i) to its i^{th} granule, $G(i)^o$, ∂G(i), $\overline{G}$(i) respectively to the interior, boundary and exterior of G(i), and |G| to the number of non-empty granules at G.

In [3,8] have respectively listed several topological granularity relations for spatial granularities. We could classify them into *partial-order relations* and *symmetrical relations*.

Partial-order relations
GroupsInto(G,H): *each granule of H is equal to the union of a set of granules of G. The converse is* ***GroupedBy (H,G).***
FinerThan(G,H): *each granule of G is contained in one granule of H. The converse is* ***CoarserThan(H,G)***.
SubGranularity(G,H): *for each granule of G, there exists a granule in H with the same spatial extent*.
Partition(G,H): *G groups into and is finer than H. The converse is* ***PartitionedBy(H,G)***.
CoveredBy(G,H): *each granule of G is covered by some granules of H. The converse is* ***Covers(H,G)***.
Symmetric relations
Disjoint(G,H): *any granule of G is disjoint with any granule of H.*
Overlap(G,H): *some granules of G and H overlap.*

These relations can be adopted for the temporal granularities followed by two more partial-order relations defined for time [4,5].

GroupsPeriodicallyInto(G,H): *G groups into H. $\exists n, m \in N$ where. $n<m$ and $n<|H|$, s.t. $\forall i \epsilon N$, if $H(i)=\bigcup_{r=0}^{k} G(j+r)$ and $H(i+n) \neq \emptyset$ then $H(i+n)=\bigcup_{r=0}^{k} G(j+r+m)$.*

GroupsUniformlyInto: *G groups periodically into H, as well as that m=1 in the above definition of* **GroupsPeriodicallyInto**.

These granularity relations are essential to manage the granularities in a granularity system. Algorithms to verify the satisfaction of each relation can be implemented with topological relation reasoning in many spatial database extensions [8].

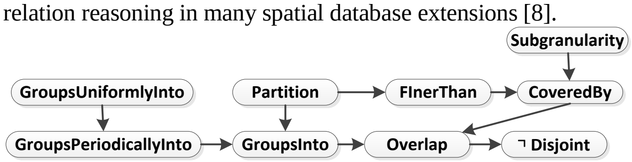


Fig. 1. **The Hasse diagram of the logical inference of above granularity relations (do not include the converse relations)**

Property 3.1 gives the logical inference of these relations, whose transitive closure can be proved complete. Fig.1 shows the Hasse diagram for the inference. This is a premise to discuss semantic constraints of granularity conversion later.

Property 3.1 (*Logical Inference of Granularity Relations*): *Let G, H be two granularities, the topological relations from G to H follow such logical inferences:*

- *GroupsInto(G,H)⊢Overlap (G,H)*
- *FinerThan(G,H)⊢CoveredBy(G,H)*
- *Partition(G,H)⊢FinerThan(G,H)∧GroupsInto(G,H)*
- *FinerThan(G,H)∧GroupsInto(G,H)⊢Partition(G,H)*
- *SubGranularity(G,H) ⊢CoveredBy(G,H)*
- *CoveredBy(G,H)⊢Overlap(G,H)*
- *GroupedBy(G,H)⊢Overlap(G,H)*
- *Covers(G,H)⊢Overlap(G,H)*
- *CoarserThan(G,H)⊢Covers(G,H)*
- *PartitionedBy(G,H)⊢CoarserThan(G,H)∧GroupedBy(G,H)*
- *Disjoint(G,H)⊢ ¬Overlap(G,H)*
- *Overlap(G,H)⊢ ¬Disjoint(G,H)*
- *GroupsPeriodicallyInto(G,H)⊢GroupsInto(G,H)*
- *GroupsUniformlyInto(G,H)⊢GroupsPeriodicallyInto(G,H)*

3.2 Granularity Systems

Various granularities are adopted in a granularity system. Many literatures organize them as a lattice with a specific granularity relation (linking relation) [4,5,7,9,11,13] as such an algebraic structure to guarantee the significant granular conversion and comparison in a granularity system [4]. We have addressed that, among multiple cases of these lattices, the *heterogeneity* generally lies in linking relations as well as granularities incl. identity and zero elements. We hereby generalize above features in our definition.

Definition 3.3 (*Granularity System*): *A (multi-)granularity system is a set of granularities over a domain hierarchically linked with a granularity relation. It is defined by the quintuple GS(D, {G}, ≤, G_0, G_1).*

- *D: The definition domain of the granularities in GS.*
- *{G}: The set of granularities in GS.*
- *≤: The partial-order linking relation that manages the granularities in {G}. It is the granularity relation which forms ({G}, ≤) as a partial-order lattice.*
- *G_0: N→P(D) is the granularity G_0 ϵ{G} known as the zero element, i.e. ∀Gϵ{G}, G_0≤G always holds.*
- *G_1: N→P(D) is the granularity G_1 ϵ{G} known as the identity element, i.e. ∀Gϵ{G}, G≤G_1 always holds.*

Over a domain, a group of granularity systems can be constructed simultaneously. The notion *D-system group* is to denote the universal set of granularity systems on domain D.

Definition 3.4 (*D-system Group*): *A D-system group $\mathcal{E}_D$ is a group of granularity systems over the domain D.*

Several heterogeneous granularity systems are allowed to coexist in $\mathcal{E}_D$. When we discuss multi-system combination and inter-system conversions, each system involved is from one $\mathcal{E}_D$. Example 3.1 forms a D-system Group with three heterogeneous granularity systems from real-world systems.

Example 3.1. *$\mathcal{E}_D$={GS_1, GS_2, GS_3} includes three GSs formed with granularities fetched respectively from Wikipedia, GeoNames and TGN[20] (We modified to better illustrate our problem).*

- *GS_1 is fetched from Wikipedia's Places category: Granularities {g_{11}=Continents, g_{12}=Nations}; the linking relation* ***GroupsInto*** *applies as {(g_{12}, g_{11})};*
- *GS_2 is fetched from TGN: Granularities {g_{21}=Subcontinent, g_{22}=Climatic regions, g_{23}=Provinces, g_{24}=Districts and counties}; the linking relation* ***Partition*** *applies as {(g_{22}, g_{21}), (g_{23}, g_{21}), (g_{24}, g_{22}), (g_{24}, g_{23})};*
- *GS_3 is fetched from Geonames,: Granularities {g_{31}=Administrative divisions of countries, g_{32}=Populated places, g_{33}=Roads and Rail}; the linking relation* ***FinerThan*** *applies as {(g_{32}, g_{31}), (g_{33}, g_{32})};*

- *GroupsInto also applies to granularities across these GSs as* $\{(g_{21}, g_{11}), (g_{22}, g_{12}), (g_{12}, g_{21})\}$, *and FinerThan applies as* $\{(g_{32}, g_{23}), (g_{33}, g_{24})\}$.

3.3 Granularity Conversion

A granularity systems defined above regulates the uniform order of granularities. If we scan from G_1 to G_0 monotonously, the resolution that the granularities express transforms uniformly. The left indeterminacy, whether refining or merging such transformation causes, is clarified as below.

Definition 3.5 (*Granularity Order*): *Given two granularities* $G: N \rightarrow P(S)$, $H: N \rightarrow P(S)$ *and a linking relation* $\preceq$, *s.t.* $G \preceq H$

- *Refine order: we say G, H has refine order (G refines H) under relation* $\preceq$, *denoted as* $(G \prec H)_{\preceq}$, *if , for any subgranularity of G, say G', let H' be any subgranularity of H s.t.* $G' \preceq H'$, $|G'| \geqslant |H'|$ *always holds.*
- *Merge order: inversely, we say G, H has merge order (G merges into H) under the relation* $\preceq$, *denoted as* $(G \succ H)_{\preceq}$, *if, for above G' and H',* $|H'| \geqslant |G'|$ *always holds.*

This concept classifies granularity relations into two groups, say *refining relations*: {FinerThan, GroupsInto, Partition, CoveredBy, GroupsPeriodicallyInto, GroupsUniformlyInto}, and *merging relations*: {CoarseThan, GroupedBy, PartitionedBy, Covers}. Relatively, we address granularity systems with a linking relation from the former group as *refining systems*, those with one from the latter group as *merging systems*. As a merging system can be transformed to a conjugative refining system by inverting its linking relation, we discuss refining systems hereafter *without loss of generality*.

A granularity conversion, which shifts the granular data across resolutions, is defined as follow.

Definition 3.6 (*Granularity Conversion*): *A granularity conversion is a function* $Conv_{H \rightarrow G}(H')_{\preceq}$ *to convert a subgranularity H' (i.e. subset) of granularity H to granularity G, where G,H satisfy* $G \preceq H$, *and* $\preceq$ *is a linking relation. For* $\preceq$ *s.t. either* $(G \prec H)_{\preceq}$ *or* $(G \succ H)_{\preceq}$ *holds, either of the following conversion is allowed.*

- *Refine-conversion: If* $(G \prec H)_{\preceq}$, *let G' be the subgranularity of G s.t. no other* $G^* \supseteq G'$ *satisfies* $G^* \preceq H'$, *then H' is refined to G as G', denoted as* $Conv_{H \rightarrow G}(H')_{\preceq} = G'$, *which is a total function.*
- *Merge-conversion: If* $(G \succ H)_{\preceq}$, *and there exists the subgranularity G' of G s.t. no other* $G^* \subseteq G'$ *satisfies* $G^* \preceq H'$ *and no other* $H^* \supseteq H'$ *satisfies* $G' \preceq H^*$, *then H' is merged to G as G', denoted as* $Conv_{H \rightarrow G}(H')_{\preceq} = G'$, *which is a partial function.*

In any granularity system, a conversion across $G \preceq H$ preserve the specific semantics related to $\preceq$. E.g., let $\preceq$ be Partition, then $Conv_{H \rightarrow G}(\{H(1),H(2)\})_{\preceq}$, would return the granules at G whose extent exactly equals to H(1),H(2). But we only get granules roughly covered by H(1), H(2), let $\preceq$ be FinerThan.

In-system granularity conversions are compositional, because they are defined with the same linking relation.

Property 3.2 (*Compositionality*): *Given a linking relation* $\preceq$, *if* $G \preceq H \preceq I$, *then* $Conv_{H \rightarrow G}(Conv_{I \rightarrow H}(I')_{\preceq})_{\preceq} = Conv_{I \rightarrow G}(I')_{\preceq}$

Specifically, if $G \preceq H$ and no other granularity I exists s.t. $G \preceq I \preceq H$, we say a conversion across G and H $Conv_{H \rightarrow G}(H')_{\preceq}$, is an *atom conversion*, since it cannot be decomposed to a union of conversions. Otherwise we say it's a *composed conversion*.

3.4 Weighted Granularity Graph

A granularity system can be transformed to a weighted granularity graph (WG), defined in Definition 3.7, with which we can reduce many problems incl. multi-system combination, reasoning of semantic constraints of granularity conversions, quantization of geometric or statistic imprecision in granularity conversion and granular comparison on graphs.

Definition 3.7 *(Weighted Granularity Graph)*: *A weighted granularity graph* $WG(V, E, W, L, Mv, V_0, V_1, R, M_S)$ *for a granularity system* $GS(D, \{G\}, \preceq, G_0, G_1)$ *is an acyclic digraph defined as follows:*

- *V is the set of vertexes representing granularities,* $Mv: V \rightarrow \{G\}$ *is a bijection from vertexes to granularities,* $V_0 = Mv(G_0)$, $V_1 = M_V(G_1)$.
- E is the set of edges, which is the subset of $V \times V$. For each $e \in E$, $Mv(t(e)) \leq Mv(s(e))$ and no other $G \in \{G\}$ exists s.t. $Mv(t(e)) \leq G$ and $G \leq Mv(s(e))$
- *W(e) is edge-weight to denote the gain of a conversion from Mv(s(e)) to Mv(t(e)). Its range is set as the real number in [0,1].*
- *L is the label function of edges,* $L(e) = (W(e), \preceq)$.
- *R is the label function of linking relation,* $R(WG) = \preceq$.
- M_S *is the bijection from WG to GS* ($M_S(WG) = GS$).

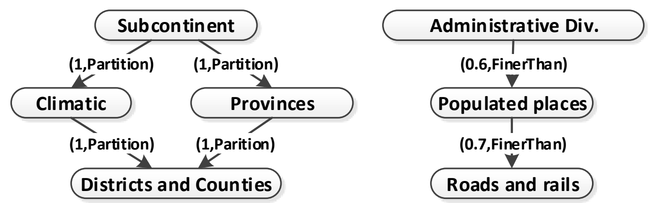


Fig. 2. **WGs for GS_2 and GS_3 of Example 3.1**

The WG has similar property on the creation of edges with a Hasse diagram of a partial-order set. It is acyclic, and no edge is created for relations obtained with transitivity. Specially, it considers the semantics of conversion between any pair of adjacent vectors. Besides, its edge-weight will be used to represent the uncertainty of conversion, which we define in Section 5. Two exemplary WGs for GS_2 and GS_3 of Example 3.1 are shown in Fig. 2, where the weight is annotated with numbers in [0,1] to denote the expected precision of atom granularity conversions.

At this point, we have provided a general model of granularity systems, which allows heterogeneous granularity systems to coexist. Based on this, we can combine multi-systems so as to extend inter-system conversion in the next section.

4. COMBINATION OF GRANULARITY SYSTEMS

The purpose of multi-system combination is to merge multiple lattice-based systems from $\mathcal{E}_D$ into a single lattice, so as to extend original in-system functionalities of granular conversion and granular comparison across multiple systems. However, due to the heterogeneity of granularity systems, combination is necessarily restricted by the semantics of granularity conversion and feasibility of granular comparison across original systems.

We hereby discuss the property *combinability*, which guarantees significant pre-conditions of inter-system granularity conversion, i.e. semantic preservation and semantic consistency, as well as the support of granular comparison. After that, we address the approach of multi-system combination.

4.1 Semantic Preservation and Consistency

The semantic preservation and semantic consistency of composed atom conversions are defined as follows.

Definition 4.1 (*Semantic Preservation*): *Let $G_1..G_n$ be n ($n>2$) granularities, and $\leq_k$ be the linking relations s.t. $\forall k \in [1,n-1]$, $G_k \leq_k G_{k+1}$. Let G' be a subgranularity of G_1, the composed conversion from G_1 to G_n is semantic preserved if $Conv^{n-1}_{G1\rightarrow\ldots\rightarrow Gn}(G')_{\leq 1}=Conv_{G1\rightarrow Gn}(G')_{\leq 1}$. I.e., the semantics of the first atom conversion is preserved in the rest atom conversions.*

Definition 4.2 (*Semantic Consistency*): *Let $G_1..G_n$ be n ($n>2$) granularities, and $\leq_k$ be the linking relations s.t. $\forall k \in [1,n-1]$, $G_k \leq_k G_{k+1}$. Let G' be a subgranularity of G_1, the composed conversion from G_1 to G_n is semantic consistent if $\exists j \in [1,n-1]$ s.t. $Conv^{n-1}_{G1\rightarrow\ldots\rightarrow Gn}(G')_{\leq j}=Conv_{G1\rightarrow Gn}(G')_{\leq j}$. I.e., the uniform semantics can be given according with one atom conversion within the composed conversions.*

Remark *Across multiple systems, semantic preservation directly extends the conversion from a granularity in the original system to another granularity in another system within its reach with the same semantics. While semantic consistency decides a uniform semantic for a composed conversion, although it may lose the semantics in the original system. E.g., if we regard a refine-conversion of granule {g} in a GroupsInto system as to fetching all granules in a certain refined granularity that groups into {g}, then this illustration still applies to a semantic preserved inter-system conversion of {g}, but may not apply to a semantic consistent inter-system conversion.*

Thereof, across systems, the former is the requisite for direct inheriting of the conversion operation in the original systems, as well as their compositionality, even each atom conversion follows the original linking relations. The latter is a weaker constraint, which is the minimum requirement for the compositionality of conversions w.r.t. a universal relation if no other relation that hasn't originally existed in the composition is introduced.

It is noteworthy that, compositionality of conversions guaranteed by semantic consistency is signification, as it is a precondition to: (1) conformation of any inter-system conversion $Conv_{G1\rightarrow Gn}(G')_{\leq j}$ with Definition 3.6, and topological consistency in conversions [9]; (2) usability of other conversion-based aggregation functions, e.g. Sum, Avg, Max [9], preserving their linearity w.r.t. conversions; (3) transitivity and path-independence of geometric/statistic precision which enables analysis of these quantities and their usability as weight of WG for solving OCRG problem (Sect. 5.1); (4) value correctness of composed refine conversions as well as efficiency gained by the avoid of multiple atom conversions as well as inter-granularity indices.

The uniformity of linking relation has ensured these constraints for in-system conversions. However, when we extend such conversion to inter-system, semantic preservation and consistency do not always hold. Let's consider Example 4.1:

Example 4.1 *Given granularities G,H in one system, and I in another, as well as I' as a group of granules in I.*

1. *Let GroupsInto(H,I) and GroupsUniformlyInto(G,H) be true.*
2. Let GroupsUniformlyInto(H,I) and GroupsInto(G,H) be true, but GroupsUniformlyInto(G,H) be false.
3. Let FinerThan(H,I) and GroupsInto(G,H) be true, but Partition(H,I) and Partition(G,H) be false.

In the first case, the composed conversion from I to G is semantic preserved and consistent, because GroupsUniformly-Into(G,H)→ GroupsInto(G,H) holds. In the second case, the composed conversion doesn't preserve the periodicity semantic in conversion from I to H, so that $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{Grou\text{-}psUniformlyInto},G)_{GroupsInto} =Conv_{I\rightarrow G}(I')_{GroupsUniformlyInto}$ doesn't hold. But we can still deduce the semantic consistency w.r.t. GroupsInto from $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{GroupsUniformlyInto},G)_{GroupsInto}=Conv_{I\rightarrow G}(I')_{GroupsInto}$ acc. to the inference GroupsUniformlyInto(H,I)⟶GroupsInto (H,I). As for the third, neither of the two semantic constraints holds, since FinerThan and GroupsInto cannot form any mutual deduction acc. to Property 3.1.

In scenes where granularity systems are used for precise multi-resolution representation, we have to preserve the semantics of conversion in the original system so as not to lose certain properties of granular data (such as geometric congruity, periodicity, etc.), by following Property 4.1.

Property 4.1 (*Semantic Preserved Compositionality*): *Given two linking relations $\leq$, $\leq'$, we denote $\forall G,G': G\leq G' \vdash G\leq^* G$ as $\leq \rightarrow \leq^*$. Given granularities G,H,I s.t. $G\leq H\leq^* I$, then $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq}=Conv_{I\rightarrow G}(I')_{\leq^*}$ iff $\leq \rightarrow \leq^*$.*

Proof *(Sufficiency) If $\leq \rightarrow \leq^*$, then $G\leq^* H$ is deduced from $G\leq H$. Thus the second conversion in $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq}$ equals to that in $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq^*}$ acc. to Definition 3.6. As $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq^*}=Conv_{I\rightarrow G}(I')_{\leq^*}$ (property 3.2), we have $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq}=Conv_{I\rightarrow G}(I')_{\leq^*}$.*

(Necessity) If $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*},G)_{\leq}=Conv_{I\rightarrow G}(I')_{\leq^*}$ but $\leq \rightarrow \leq^*$ doesn't hold. Then $\leq^*$ doesn't apply between $Conv_{I\rightarrow G}(I')_{\leq^*}$ and $Conv_{I\rightarrow H}(I')_{\leq^*}$. Thus the second conversion in $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{\leq^*}, G)_{\leq}$ conflicts with semantic preservation. In that way, $\leq \rightarrow \leq^*$ must hold.

While other scenes may require to fulfill only semantic consistency among conversions so as to guarantee their compositionality, by following the condition in the next property.

Property 4.2 (*Semantic Consistent Compositionality*): *Given two linking relations $\leq$, $\leq^*$. Given granularities G,H,I s.t. $G\leq H\leq^* I$, composed conversion from I to G is semantic consistent iff any of $\leq=\leq^*$, $\leq \rightarrow \leq^*$ or $\leq^* \rightarrow \leq$ holds.*

Proof *(Sufficiency) Similar to the above proof, no matter which of $\leq=\leq^*$, $\leq \rightarrow \leq^*$ or $\leq^* \rightarrow \leq$ is given, there is always a weakest linking relation s.t. the conversions from I to G is semantic consistent w.r.t. to that relation.*

(Necessity) We can directly infer the necessity from Definition 4.2 that, if the composed conversion from I to G is semantic consistent, then one of $\leq$ and $\leq^*$ must deduce the other.

Above properties can be easily extended for conditions of three or more atom conversions via inductive method. They clarify the requisite to extend granularity conversion to inter-system regardless of the heterogeneity in $\mathcal{E}_D$.

4.2 Combinability

The precondition of multi-system combination is defined as:

Definition 4.3 (*Combinability*): *Two granularity systems from $\mathcal{E}_D$ can be combined to a single system iff*

1. *Any refine-conversion in the granularity system is semantic preserved and/or semantic consistent.*
2. *For any pair of granularities from different systems, a CRG exists in the combined system.*

Above requirement 1 enables granularity conversions in a combined system. Thereof, if we guarantee semantic preservation to any conversion, we say these systems satisfy *semantic preserved combinability*. Otherwise, inter-system conversions shall be guaranteed semantic consistency, thus these systems satisfy *semantic consistent combinability*. If requirement 1 is fulfilled, then requirement 2 enables granular comparison for any granules in a combined system.

For a pair of granularity systems GS, GS' from $\mathcal{E}_D$, sufficient-necessary (SN) condition for them to satisfy semantic preserved combinability are given as below.

Property 4.3 (*Semantic Preserved Combinability*): *Given a pair of refining granularity systems GS(D, {G}, $\leq$, G_0, G_1) and GS'(D, {G}', $\leq'$, G'_0, G'_1) from $\mathcal{E}_D$, semantic preserved combinability holds between iff one of the follows holds.*

1. $G_0=G'_0$ *($\leq=\leq'$, C1; or $\leq\neq\leq'$, C2).*
2. $\leq=\leq'$*, while* $G_0\leq G'_0$ *or* $G'_0\leq G_0$ *(C3); or* $\leq=\leq'$ *and exists a third (intermediate) granularity system* GS^* *from* $\mathcal{E}_D$ *with zero element* G^*_0 *s.t.* $G^*_0\leq G_0$ *and* $G^*_0\leq G'_0$ *(C4).*
3. $\leq\rightarrow\leq'$ *and* $G_0\leq' G'_0$*; or* $\leq'\rightarrow\leq$ *and* $G'_0\leq G_0$*; (C5) or exists a third (intermediate) system* GS^* *from* $\mathcal{E}_D$ *having linking relation* $\leq^*$*, s.t.* $\leq^*\rightarrow\leq$ *and* $\leq^*\rightarrow\leq'$*, and zero element* G^*_0 *s.t.* $G^*_0\leq G_0$ *and* $G^*_0\leq' G'_0$ *(C6).*

Proof *Since every granularity in a refining system is possible to be converted to the zero element, it's easy to know that existence of CRG and semantic preservation between any pair of granularities holds iff the existence of CRG of* G_0 *and* G'_0 *as well as semantic preservation of inter-system conversion across* G_0 *and* G'_0 *hold.*

Between a pair of granularity systems (GS, GS') on the same domain, following are all possible branches of binary relationships w.r.t. their zero elements and linking relations:

1. $\leq=\leq'$*: (a)* $G_0=G'_0$*; (b)* $G_0\neq G'_0$*;*
2. $\leq\neq\leq'$*: (a)* $G_0=G'_0$*; (b)* $G_0\neq G'_0$*;*

For above 1, semantic preservation always holds.

In the condition (a) of above 1, GS and GS' share the same zero element, thus, G_0 *provides a CRG for any granularities in the two systems. Since* $\leq=\leq'$*, acc. to property 3.2, conversion towards* G_0 *is semantic consistent. (C1)*

As for (b) of above 1, if $G_0\leq G'_0$ *or* $G'_0\leq G_0$*, then the granularity on the left of* $\leq$ *is their CRG (C3). Or,* G_0 *and* G'_0 *do not satisfy the linking relation, but in an intermediate granularity system with zero element* G^*_0*, s.t.* $G^*_0\leq G_0$ *and* $G^*_0\leq G'_0$*, then a CRG still exists in* $\mathcal{E}_D$ *(C4). Otherwise, the existence of CRG of the two zero elements is impossible under such a relationship.*

The next condition (a) of above 2 is similar to condition (a) of above 1, where G_0 *serves as a CRG, and composed conversions to* G_0 *are respectively in-system (C2).*

For last branch where $\leq\neq\leq'$ *and* $G_0\neq G'_0$ *happen, we can then refer to property 3.1. If* $\leq\rightarrow\leq'$ *and* $G_0\leq' G'_0$ *hold, then* G_0 *serves as a CRG, and conversions passing* G'_0 *to* G_0 *is semantic preserved acc. to Property 4.3. Reversely, if* $\leq'\rightarrow\leq$ *holds and* $G'_0\leq G_0$*,* G'_0 *serves as the CRG with semantic preservation similarly guaranteed. (C5) Moreover, if these factors do not holds, but in an intermediate system* GS^* *having* $\leq^*$ *as its linking relation and* G^*_0 *as its zero element, s.t.* $\leq^*\rightarrow\leq$*,* $\leq^*\rightarrow\leq'$*, and* $G^*_0\leq G_0$ *and* $G^*_0\leq' G'_0$*. We have* G^*_0 *as the CRG, to which any composed conversion is semantic preserved (C6). Otherwise the semantic preservation and existence of common refined granularity mustn't both hold in any other condition.*

Above has proved the necessity because C1~C6 have enclosed all the four possible binary relationships between GS and GS'. On the other hand, if any of C1~C6 holds, we can always provide a CRG of the zero elements of GS and GS', towards which any conversion is granted semantic preservation. Thus, they must satisfy semantic preserved combinability. Thereby sufficiency also holds.

The SN condition for the semantic consistent combinability is given as below.

Property 4.4 (*Semantic Consistent Combinability*): *Given a pair of refining granularity systems GS(D, {G},* $\leq$*,* G_0*,* G_1*) and GS'(D, {G}',* $\leq'$*,* G'_0*,* G'_1*) from* $\mathcal{E}_D$*, semantic consistent combinability holds between them iff one of the follows holds.*

1. $G_0=G'_0$ *($\leq=\leq'$, C1; or $\leq\neq\leq'$, C2).*
2. *Any of* $\leq=\leq'$*,* $\leq\rightarrow\leq^*$ *or* $\leq^*\rightarrow\leq$ *holds and either of these relation applies between* G_0 *or* G'_0 *(C3); or exists a third (intermediate) system* GS^* *from* $\mathcal{E}_D$ *having linking relation* $\leq^*$*, s.t. any of (*$\leq=\leq^*$*,* $\leq^*\rightarrow\leq$ *or* $\leq\rightarrow\leq^*$*) and any of (*$\leq=\leq^*$*,* $\leq^*\rightarrow\leq'$ *or* $\leq'\rightarrow\leq^*$*) hold, and zero element* G^*_0 *s.t. either of* $\leq$*,* $\leq^*$ *applies from* G^*_0 *to* G_0 *and either of* $\leq'$*,* $\leq^*$ *applies from* G^*_0 *to* G'_0*(C4).*

Proof Hint *The necessity can be proved by reasoning the coverage of C1~C4 on all conditions under all possible branches of binary relations between (GS, GS') w.r.t. their zero elements where the existence of the CRG of* G_0 *and* G'_0 *as well as conversions to that granularity to fulfill semantic consistency w.r.t. Property 4.2,. This is similar to the proof of we have given under Property 4.3. On the other hand, the sufficiency is easy to be deduced from any of C1~C4.*

Algorithms to verify either of the combinability conditions can be created as sequential procedures to verify the satisfaction of C1~C6 in Property 4.3 or C1~C4 in Property 4.4. Corresponding operations, say *SPCombinability()* and *SCCombinability()*, are within O(1) time complexity with the aid of Property 3.1 and the global granularity relation matrix introduced in Section 6.

4.3 Multi-system Combination

We can now combine multi-systems from $\mathcal{E}_D$ to a uniform system that guarantee correct granular comparison and inter-system granularity conversions.

Note that in $\mathcal{E}_D$, it's possible for several closures of combinability to coexist, each of which forms a combination. To simplify, we specify the combination algorithm in the way of finding and combining other combinable systems to a target. We process the combination in the form of granularity graphs.

Operations of lattice-based combination can be referred as *SPCombine(WG, {WG}$_D$)* (semantic preserved combination) and *SCCombine(WG, {WG}$_D$)* (semantic consistent combination). In fact, the two combination algorithms are logically similar (only different in constraints when creating edges between two systems acc. with semantic constraints). Exemplarily, we give the algorithm of semantic preserved combination.

Given a group of WGs of systems in $\mathcal{E}_D$, denote as $\{WG\}_D$, and a target graph $WG\in\{WG\}_D$, the *semantic preserved combination algorithm* is given as Algorithm 4.1, which integrates all combinable systems for WG to a single lattice.

Algorithm 4.1 *SPCombine***(***WG(V, E, W, L, Mv,* V_0*,* V_1*, R,* M_S*), {WG}*$_D$**)**

```
let Vc be the extent of Ms(WG).D 'a communal V1 for {WG}D
CreateDirectedEdge(Vc, WG.V1) 'from Vc to V1
WG.V1←Vc
for each WG'∈{WG}D do
  if WG'≠WG and SPCombinability(WG, WG',{WG}D) then
      {WG}D←{WG}D\WG'
       ClearTags(checked)
      if R(WG)=R(WG') then
```

```
9:            DFSCreateEdges(WG.V1, WG'.V1,R(WG),checked,false)
10:           DFSCreateEdges(WG'.V1, WG.V1,R(WG),checked,false)
11:       else if R(WG')→R(WG) then
12:           DFSCreateEdges(WG.V1, WG'.V1,R(WG),checked,false)
13:           R(WG)←R(WG')
14:       else if R(WG)→R(WG') then
15:           DFSCreateEdges(WG'.V1, WG.V1,R(WG'),checked,false)
16:       else for each WG*∊{WG}D do
17:           if R(WG*)→R(WG) and R(WG*)→R(WG') then
18:               {WG}D←{WG}D\WG*
19:               DFSCreateEdges(WG.V1,WG*.V1,R(WG),checked,false)
20:               DFSCreateEdges(WG'.V1,WG*.V1,R(WG'),checked,false)
21:               R(WG)←R(WG*)
22:               continue
23: for each WG'∊{WG}D do
24:     if Mv(WG.V0)=Mv(WG'.V0) do
25:         {WG}D←{WG}D\WG'
26:         MergeVertex(WG.V0, WG'.V0)
27:         CreateDirectedEdge(WG.V1,WG'.V1)
28: return WG.V1
```

Algorithm 4.2 is the *semantic consistent combination* algorithm.

Algorithm 4.2 *SCCombineWG***(***WG(V, E, W, L, Mv, Me, s, t, Gain, V_0, V_1, R, M_S), $\{WG\}_D$***)**

```
1: let Vc be the extent of Ms(WG).D 'a communal V1 for {WG}D
2: CreateDirectedEdge(Vc, WG.V1) 'from Vc to V1
3: WG.V1←Vc
4: for each WG'∊{WG}D do
5:   if WG'≠WG and SCCombinability(WG, WG',{WG}D) then
6:       {WG}D←{WG}D\WG'
7:       ClearTags(C)
8:       r←FindLow(R(WG), R(WG')) 'Find the weaker Relation
9:       if RelationApplies(WG.V1, WG'.V1,r) then
10:          DFSCreateEdges(WG.V1, WG'.V1,r,C,false)
11:          DFSCreateEdges(WG'.V1, WG.V1,r,C,false)
12:          R(WG)←r
13:      else for each WG*∊{WG}D do
14:          r←FindLow(R(WG), R(WG*))
15:          r'←FindLow(R(WG'), R(WG*))
16:          if RelationApplies(WG.V1, WG*.V1,r) and Relation-
             Applies(WG'.V1, WG*.V1,r') then
17:              {WG}D←{WG}D\WG*
18:              DFSCreateEdges(WG.V1,WG*.V1,r,C,false)
19:              DFSCreateEdges(WG'.V1,WG*.V1,r',C,false)
20:              R(WG)←FindLow(r,r')
21:              continue
22: for each WG'∊{WG}D do
23:     if Mv(WG.V0)=Mv(WG'.V0) do
24:         {WG}D←{WG}D\WG'
25:         MergeVertex (WG.V0, WG'.V0)
26:         CreateDirectedEdge(WG.V1,WG'.V1)
27: return WG.V1
```

Thereof, DFSCreateEdges given as Algorithm 4.3 links the vertexes of one system to those of the other to mark any atom relation, but acc. with the definition of the granularity graph, no edge is created for the relations gained from transitivity.

Algorithm 4.3 *DFSCreateEdges***(***v, u, ≤, checked[,], foundabove***)**

```
1: found⟵foundbelow⟵created⟵false
2: checked[v,u]⟵true
3: if foundabove=false and Mv(u)≤ Mv(v) then
4:     found⟵true
5: if (found∨foundabove=true) and Succ(v)≠∅ then
6:   for each v'∊Succ(v) do
7:       if checked[v,u]=false and Mv(u)≤ Mv(v') then
8:           foundbelow⟵true
9:           DFSCreateEdges(v', u, ≤, true)
10:      else for each u' ∊Succ(u)
11:              if checked[v',u']=false then
12:                  DFSCreateEdges(v', u', ≤, checked,false)
13:   if foundbelow=false then 'an atom relation is found
14:       CreateDirectedEdge(v,u) 'from v to u
15:       created⟵true
16: if created=true and Succ(v)≠∅ and Succ(u)≠∅ then
17:    for each v'∊Succ(v) do
18:       for each u' ∊Succ(u)
19:           if checked[v',u']=false then
20:               DFSCreateEdges(v', u', ≤, checked,false)
21: else if found∨foundabove=false and Succ(u)≠∅ then
22:     for each u'∊Succ(v) do
23:         if checked[v,u']=false then
24:             DFSCreateEdges(v', u, ≤, checked,false)
25: else return
```

Remark *Algorithm 4.1 scans every WG'∊$\{WG\}_D$, and checks for its semantic preserved combinability with WG. Once a combinable WG' is found, it is combined to WG through one of the possible branches of the edge creation procedure acc. to any of C1~C6 it satisfies with WG. E.g., if WG and WG' share the same semantics of conversion (i.e. R(WG)=R(WG')), edges are allowed to create both from WG to WG' and from WG' to WG. Each of such branches also decides the semantics of conversion the current WG preserves via the inference between R(WG) and R(WG'). This policy ensures any inter-system route represents a semantic preserved conversion.*

Algorithm 4.3 links the vertexes (v and its successors) from one system to those (u and its successors) from the other. It uses depth-first search. A tag array checked[,] marks if a pair of vertexes u,v have been checked. For a v, if a u s.t. u≤v is found, the algorithm will check the successors v' of v to see if there is any v' s.t. u≤v'. If no such a v' is found, an edge is created between u,v. Otherwise such check will be recurred to u and successor of v' until that condition is fulfilled. For any check where u and v don't satisfy u≤v, recursion check will be processed on their successors. However, if any pair of u,v have been checked once acc. to checked[,], no repeated check will be processed on them and their successors. This procedure ensures that only edges for relations that are not gained from transitivity are created. Combinability guarantees a V_0 after each phase of combination, and a V_C denotes the domain is created as V_1. In that way the result is still a lattice.

The last loop (23~27 of SPCombine) deals with a special case between two systems, where the linking relation of WG cannot form any logical deduction with that of any WG', but shares the equivalent V_0 with a WG'. Thus they are combined by merging V_0, but no other vertexes are linked acc. with the constraint of semantic preservation.

Overall, let $|\{WG\}_D|$ and each $|G|$ be both $O(n)$ magnitude, DFSCreateEdge take $O(n^2)$ and SPCombine take $O(n^3)$ temporal complexity. Above properties also hold for SCCombine.

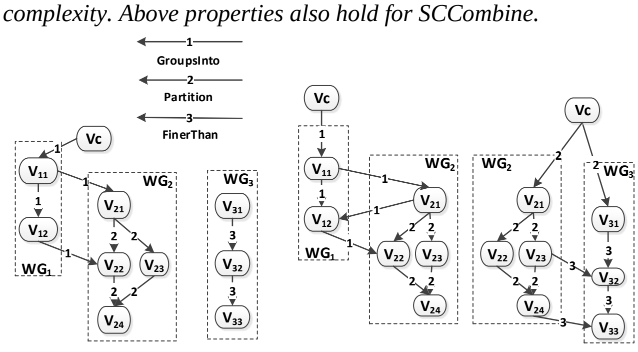


Fig. 3. **Graphic depiction for Example 4.2 (a) semantic preserved combination (b) semantic consistent combination of**

3 granularity systems

Example 4.2 *Here we combine the systems in Example 3.1. We give* $\{WG\}_D=\{WG_1, WG_2, WG_3\}$ *for* $\mathcal{E}_D$*, where* WG_1*,* WG_2 *and* WG_3 *shown in Fig.3 are respectively the WGs of* GS_1*,* GS_2 *and* GS_3*. The vertexes denotes the granularities with the same sub-scripts. Recall that across systems, GroupsInto applies as* $\{(g_{21}, g_{11}), (g_{22}, g_{12}), (g_{12}, g_{21})\}$*, and FinerThan applies as* $\{(g_{32}, g_{23}), (g_{33}, g_{24})\}$*, thus we could combine* GS_2 *to* GS_1 *or* GS_3*. The allowed SPCombine is depicted as Fig.3(a), where any inter-system conversion from a granularity in* $Ms(WG_1)$ *to another in* $Ms(WG_2)$ *preserves the semantics of GroupsInto originally in* $Ms(WG_1)$*. However,* $Ms(WG_3)$ *doesn't satisfy semantic preserved combinability with* $Ms(WG_1)$ *and* $Ms(WG_3)$*. The allowed SCCombine is depicted as Fig.3(b), where the combination process between* GS_2 *and GS is valid since conversions from* GS_2 *to* GS_3 *share the semantic consistency w.r.t. FinerThan.*

5. MULTI-GRANULAR COMPARISON

Due to the incongruity of geometric properties ensured by different linking relations, granularity conversion unavoidably contains value distortion, causing the imprecision of representation and statistic analysis of granular data.

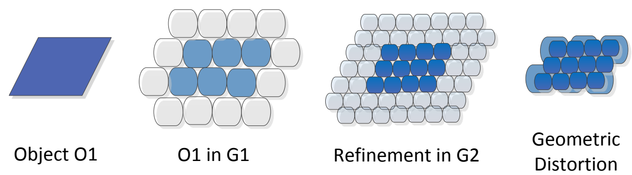


Fig. 4. **An object projection in different granularities**

On the other hand, for a pair of granularities in the combined system, more than one CRG can exist (e.g. in Fig.3(b), granularities that V_{31}, V_{12}, V_{32} and V_{34} represent are all CRGs of granularities respectively represented by V_{11} and V_{31} can compare). Selecting and converting towards the one common with least expectation of distortion in conversion can decrease related imprecision of comparison.

In this section we quantify such distortion in conversion, and pick the optimal common refined granularity (ORCG) with least expectation of distortion to improve granular comparison in multiple granularity systems.

5.1 Geometric and Statistic Distortion

Distortion of granularity conversion is reflect as in two aspects: *geometric distortion* results from the incongruity of extents by some linking relations (such as FinerThan and CoveredBy), causing a granular object become a vague object (Fig. 4); *statistic distortion* results from the loss of data aggregation in a multi-granularity system.

To quantify such distortion, we begin with the definition of geometric precision (which is the complementary of distortion) of a granule in a granularity conversion.

Definition 5.1 (*Granular Geometric Precision*): *Given granularities G, H and a refining relation* $\le$ *s.t.* $H \le G$*. Let g be a granule of G, then the degree of granular geometric preservation between g and its conversion result at H, denote u(g,H'), is defined as below (an 'o' over an extent denotes its interior area):*

- $$u(g,H)=\frac{g^o \cap (Conv_{G\to H}(\{g\})_{\le})^o}{g^o \cup (Conv_{G\to H}(\{g\})_{\le})^o}$$

The degree of inter-granularity precision can be quantified as the expectation of granular geometric precision.

Definition 5.2 (*Inter-granularity Precision*): *Given granularities G,H s.t.* $Conv_{G\to H}(G)_{\le}=H$*, then the degree of inter-granularity precision between G,H, denote U(G,H), is the expectation of granular precision defined as below:*

- $$U(G,H)=Exp(u(G(i),H))=\frac{(\bigcup_{i\in N} G(i)^o)\cap(\bigcup_{i\in N} \mathrm{H}(i)^o)}{(\bigcup_{i\in N} G(i)^o)\cup(\bigcup_{i\in N} \mathrm{H}(i)^o)}$$

Above definition quantifies the expectation of geometric precision when a granular-represented object converts towards a different granularity.

In databases where a granularity system is used for multi-scale data management and aggregation functions [9] is available, the precision related to data density, defined as below.

Definition 5.3 (*Data density*): *Give a dataset E, and a spatial/temporal extent C, the data density ρ(C) is defined as:*

- $$\rho(C)=\frac{|\{e \mid e\in E \wedge coveredBy(e,C)\}|}{C^o}$$

The ρ-granular precision defined below denotes the ratio of statistic preservation of data bond with a granule (e.g. number of residents in a block) in granularity conversions.

Definition 5.4 (*ρ-granular precision*): *Given granularities G, H and a refining relation* $\le$ *s.t.* $H \le G$*, let g be a granule of G, and ρ be the data density, then the ρ-granular precision between g and its conversion result at H, denote* $u_\rho(g,H)$*, is defined as below:*

- $$u_\rho(g,H)=\frac{\rho(g^o \cap Conv_{G\to H}(\{g\})_{\le})^o)}{\rho(g^o \cup Conv_{G\to H}(\{g\})_{\le})^o)}$$

By computing the expectation of $u\rho$(g,H) on the extent of H, we get the ρ-precision between two granularities as below.

Definition 5.5 ***(ρ-precision)****: Given granularities G,H s.t.* $Conv_{G\to H}(G)_{\le}=H$*, and ρ is the data density function. Then the ρ-precision between G,H, denote* $U_\rho(G,H)$*, is defined as*

- $$U_\rho(G,H)=Exp(u_\rho(G(i),H))=\frac{\rho((\bigcup_{i\in N} G(i)^o)\cap(\bigcup_{i\in N} \mathrm{H}(i)^o))}{\rho((\bigcup_{i\in N} G(i)^o)\cup(\bigcup_{i\in N} \mathrm{H}(i)^o))}$$

The ρ-precision quantifies the ratio of statistic accuracy when data bound with a granule or a scaled portion of dataset (such as a fraction in the pyramid structure of LARS [14]) transform across granularities, from what the statistic bias in aggregation functions like *sum* and *count* [9] is also achieved.

Above quantities are unitized. E.g., $u(g,H)\in(0,1]$. u(g,H)=1 iff no distortion of the extent of g occurs after conversion.

For granularity systems of different utilization purposes, we can use either inter-granularity precision or ρ-precision as the weight on the edges of WG inside the scale of a system, if with their transitivity and path independence proved.

Property 5.1 ***(Transitivity)****: Given a linking relations*$\le$*, and granularities G,H,I s.t.* $G \le H \le I$*, U(I,H) U(H,G)=U(I,G) and* $U_\rho(I,H)\ U_\rho(H,G)=U_\rho(I,G)$ *are always satisfied.*

Proof *The transitivity in conditions where the extents of G,H,I are congruity is self-evident (as the U and* U_ρ *identically equal to 1). Otherwise, since* $\le$ *is partial-order, the extents of G,H,I must be either monotonically increasing or decreasing. For the former, we have* $\cup G(i)^o=(\cup G(i)^o)\cap(\cup H(i)^o)$ *and* $\cup H(i)^o=(\cup G(i)^o)\cup(\cup H(i)^o)$*, which applies similarly to H and I. As for the latter, we have* $\cup H(i)^o=(\cup G(i)^o)\cap(\cup H(i)^o)$ *and* $\cup G(i)^o=(\cup G(i)^o)\cup(\cup H(i)^o)$*, which applies similarly to H and I. By substituting them* *into the expanded formula of U(I,H) U(H,G), we can then deduce U(I,H) U(H,G)=U(I,G) through reduction. The proof for U_ρ(I, H) U_ρ(H,G)=U_ρ(I,G) is exactly the same.*

Property 5.2 ***(Path-independence)****: Given a linking relation ≤ and granularities G,H,H',I, s.t. G⊴H⊴I, G⊴H'⊴I and H≠H'. U(I,H) U(H,G)=U(I,H') U(H',G) and U_ρ(I,H) U_ρ(H,G)=U_ρ(I, H') U_ρ (H',G) always hold.*

Proof *Acc. to Property 5.1, U(I,H') U(H',G)=U(I,G)=U(I, H) U(H,G), and U_ρ(I,H') U_ρ(H',G)=U_ρ(I,G)=U_ρ(I,H) U_ρ(H,G).*

Theorem 5.1 given as below extends above properties for granularities across systems and enables *U* and U_ρ, to be used as weight on edges across systems created by Algorithm 4.2.

Theorem 5.1 *The transitivity and path-independence applies to any conversion denoted by the directed paths across systems in a combined granularity graph.*

Proof *For any path across two systems, it denotes granularities $G_1 \trianglelefteq_1 G_2 \trianglelefteq_2 \ldots \trianglelefteq_n Gn$. The semantic preservation (consistency) guarantee a uniform linking relation $\trianglelefteq_1$ ($\trianglelefteq_k$) between any adjacent pair of them, and the conversion on the path is also defined w.r.t $\trianglelefteq_1$($\trianglelefteq_k$). Thus the two properties always hold.*

5.2 Granular comparison

Conversion towards the optimal common refined granularity, defined as Definition 5.7., enables the comparison of granular data from two granularities with the least expectation of geometric distortion (U) or statistic distortion (U_ρ).

Definition 5.7 (*Optimal Common Refined Granularity*) *Given two granularities G,H, then I is the optimal common refined granularity if the geometrical mean of the gain of conversion from G to I and the gain of conversion from H to I, i.e. (Gain(G, I) Gain(H,I))$^{1/2}$ is maximum.*

Since we have combined multiple-systems as a weighted granularity graph, and quantified geometric/statistic distortion as weight, the search of the optimal common refined granularity can be reduced to the LCA problem on a weighted directed acyclic graph. An O(n) algorithm [16] for such problem, say FindOCRG(G,G'), will be easily implemented by finding the LCA with least products of edge weights via DFS.

Example 5.1 *We compare the sales of product p_A in the scope of Province g and that of p_B in Municipal h. Let G=provinces and H=municipals have two common refined granularities I=districts and J=blocks in a FinerThan system or combination (blocks can cross districts so they are not FinerThan districts). Suppose ($U\rho$(G,I) $U\rho$(H,I))$^{1/2}$<($U\rho$(G,J) $U\rho$(H,J))$^{1/2}$, then the database system first process FindOCRG(Provinces, Municipals)=J, then it computes and compares Sum($Conv_{G\rightarrow J}$ (g)$_{FinerThan}$(sales_p_A)) and Sum($Conv_{G\rightarrow J}$(g)$_{FinerThan}$(sales_p_B)).*

6. CONCEPTUAL EVALUATION

In this section we present an overall conceptual evaluation of our approach. This work shows the clear goal of leveraging current multi-granularity models to a general multi-system scope, handling the heterogeneity of granularity systems in existing research work while extending the original in-system granularity conversions and granular comparison among them. This problem is with uniqueness as well as challenges.

Our framework uses a classic definition of granularity, i.e. a partition mapping of an extent. Though it is known that several literatures add graph features to their definition [2,3,10], we do not involve such features, as they benefit reasoning of granules only inside a granularity but logically cause no difference to the exchange among different granularities.

Granularity conversion and granular comparison, which we have extended from in-system to inter-system, are the major functions required to be supported by granularity systems [4]. The conversion function here is the general form of other complex functions in literatures, such as the granular aggregation functions provided in [9]. The granular comparison is the general form based on what spatially and temporally qualified granular information is compared in applications.

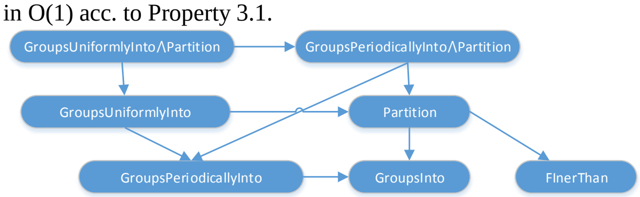
We have considered the two uncertain factors of granularity conversions among systems, i.e. incongruity of semantics and geometric uncertainty. The two semantic constraints we define, namely semantic preservation and consistency, are significant to inter-system granularity conversions w.r.t. the incongruity of conversion function semantics caused by heterogeneity of linking relations. The former preserves the inheritability of in-system conversions of original systems. The latter ensures the compositionality of conversions, which is essential for the correctness and usability of conversions and related aggregation functions as well as the weight of WG we define later.

Another key is the approach of multi-system combination. The idea of such combination stands in necessity and effectiveness. It is obvious that the requirements of combinability are exactly the requirements for the feasibility of all conversions and granular comparison between two systems. The semantic preserved and semantic consistent combinations support granularity conversions and granular comparison for multi-systems in different degrees. The former condition benefits in directly inheriting the conversion function and its semantics from the original systems as well as the auxiliary physical structures (e.g. periodic sets for GroupsPeriodicallyInto [19]), while the latter is more generalized to guarantee the feasibility of the two interoperations. Though verifying this precondition for heterogeneous systems seems complex, the SN conditions we proved for combinability require O(1) for verification. Utilizing the WG, we give one of the two similar combination algorithms to conceive the composed lattice we desire.

Meanwhile, we introduce the quantification of geometric and statistic distortion in granularity conversions. The role of this concept lies in three aspects. (1) It measures the uncertainty of conversions that semantic-related constraints don't preserve, esp. when semantic preservation are not fulfilled. (2) It gives the basis to choose the optimal of multiple CRGs for two granularities so as to improve the preciseness of granular comparison. (3) As granularity systems may be used differently either for multi-resolution representation or multi-scaled data management, *U* and U_ρ evaluate related expectation of precision in for different application requirements.

So far as we have made exploration into this fresh problem through this paper, there are still challenges remaining in the implementation of relation reasoning and granularity storage. We give some instructive strategies against them.

Granularity Relation Reasoning. Reasoning the satisfaction of a granularity relation is the most frequent operation when verifying combinability and combining multi-systems, which occupies at least O(n^2) time complexity if processed with RCC-based functions [1] in a SDBMS like PostGIS. We can use a *Global Granularity Relation Matrix* (shown in Fig. 5) to transfer such online reasoning offline. Let the lines and columns represent granularities in $\mathcal{E}_D$, each block tags the identification number of the strongest linking relation applies from its line element to its column element (null if none applies). To guarantee the soundness of inference, we import two more relations to the Hasse diagram in Fig.1, i.e. GroupsPeriodicallyInto⋀Partition and GroupsUniformlyInto⋀Partition. Then the reasoning can be deduced from the tags in O(1) acc. to Property 3.1.



	G_1	G_2	G_3	…
G_1	=	FinerThan	Partition	
G_2	CoaserThan	=	∅	
G_3	PartitionedBy	∅	=	
…				

Fig. 5. **The Extended Hasse Diagram for Refine Relations and A Global Granularity Relation Matrix**

Granularity Storage. Strategies can be used to improve conversion and granular retrieval to avoid online reasoning of the closureness constraint in a conversion. Let a WG be obtained for a combination of granularity systems, by physically mapping the stratified WG, a hierarchical index is gained where each granule points to the closure of its descendants applying the linking relation, and each granules in such a closure forms a linked list. Let the WG be a C-degree graph, this policy enables O(1) complexity for an atom conversion instead of O(n) for online reasoning. On the other side, for temporal granularity systems semantic preserved combination ensures the inheriting of relation-restricted auxiliary structures [19].

Application aspects. The issues studied here provide a baseline for a branch of emerging applications. To list some:

1) *Spatial Knowledge Integration*. The wide demands for interoperating, mapping and fusion spatial data on the Web has led to newcomer projects like YAGO2 and GeoKnow [21,24]. However such state-of-arts process integration of knowledge bases based on manually annotated semantics. In that way, to zoom-in the knowledge from one level of details (e.g. cities) to another (e.g. districts), we will have to manual-ly enumerate all the names from the latter to set the belongings to the former. This takes inconceivable labor efforts and is too hard to avoid incorrectness and missing concepts (we can miss out the Venice district of Los Angeles right?). However, the computer builds the association between the two levels with FinerThan with correct-ness and ease, and shall never miss a detail given the spatial quantities. This mean-ingful challenge, i.e. to provide implicit geographical references among spatial quanti-ties across sources of divergent representation standard, hasn't been touched yet. Once we enable the conversion and comparison among heterogeneous multi-granularity data sources and process such integration by way of combining granulari-ty systems, integrating spatial knowledge bases can be processed in an automated and computable way, where even other quantitive data values associated with these spatial quantities can be automatically exchanged through divergent levels of details. Our work now studies this non-trivial problem on its principle and provides a neces-sary logical-level solution.

2) *High-divergent Time Conversion*. Consider the diversity of time expression here. A scientist, intelligence analyst, historian, or archaeologist may encounter vast temporally qualified information of high-divergent time systems, e.g., geological peri-od systems (involving eons, eras, periods, and epochs), chronicles of different history origins, different calendars [19], and different time zones. It's not difficult to see that a unified time conversion system is equivalent to the combination of temporal granu-larity systems with heterogeneous linking relations, i.e. GroupsInto systems as geolog-ical period systems and history chronicle systems, GroupsUniformlyInto systems as each calendar, GroupsPeriodicallyInto systems as time zones, and among calendars.

3) *Integrated Data Analysis*. Scenes are not uncommon these days where sources of temporal or spatial data are joined and analyze together. For some instances, say stock prediction on multiple data sources [23], and crime data analysis [22]. Alt-hough multi-granular operations are proved to be helpful for improving the efficiency and refining the discovered knowledge in details of interest [3], heterogeneity of gran-ularities have appeared in these applications, as time representation meets already mentioned divergence, and the condition on the space goes even more complex in vast resources getting from crowdsourcing [22]. What makes things worse is that these data sources are streams that require non-blocking analysis, so that, prepro-cessing is not even available to reduce the heterogeniety. Under such circumstances it's vital to enable the automatic conversion of granular quantities across multiple granularity systems so as to meet the original needs.

4) *Querying Multiple Knowledge Bases*. When query differ-ent knowledge bases, the spatial predicate, which delimits the targets to some spatial objects, is likely to be represented in different granularity systems either due to the application's different representation criterion, knowledge base's ontologies, or user's knowledge preference (e.g. to state and county, some users may prefer regione and previncia). It is always left to front applications or users to deal with the divergent. Applying the inter-system conversion in the query plan empowers with the rule-based mapping for the spatial predicates among granularity systems, and brings along the transparency of query interpretation regardless of how the application or user wants to represent the spatial predicate.

7. CONCLUSION

In this paper we proposed a formal framework to support spatiotemporal data conversion and comparison across multiple granularity systems. Based on a general model, we have successfully dealt the heterogeneity of granularity systems reflected in literatures, and inducted the rules of semantic preservation and consistency to enable the correctness and inheritability of granularity conversions across heterogeneous systems. By studying the binary relationships of granularity systems w.r.t. linking relations and zero elements, we have deduced the SN conditions for two types of combinability, and given corresponding combination algorithms based on WG. After enabling inter-system conversion and comparison, we have quantified the uncertainty among such interoperations.

Our framework has led us to some possible future work as well as some open challenges. We plan to continue on the conceptual design and realization of our framework in SDBMS, particularly studying implementation of multiple granularity systems in distributed systems. Implementation of quantization of geometric or statistic uncertainty in databases and extending this issue for spatiotemporal granularities [2,3] are also on our list.

REFERENCES

[1] Randell, D. A., Cui, Z., Cohn, A. G.: A spatial logic based on regions and connection. In *KR'92*, pp. 165–176, Morgan Kaufmann (1992)

[2] Belussi, A., Combi, C., Pozzani, G.: Towards a Formal Framework for Spatio-Temporal Granularities. In *TIME'08*, pp. 49-53. IEEE (2008)

[3] Belussi, A., Combi, C., Pozzani, G. Formal and conceptual modeling of spatio-temporal granularities. In *IDEAS'09*, pp. 275-283. ACM (2009)

[4] Elena Camossi, Michela Bertolotto, Elisa Bertino: Multigranular spatio-temporal models: implementation challenges. In: *Procs. of the 16th SIGSPATIAL GIS*, Article No. 63. ACM (2008)

[5] C. Bettini , S. Jajodia , X. Sean Wang. Time Granularities in Databases, Data Mining, and Temporal Reasoning.. Springer (2000)

[6] M. Sester: Abstraction of GeoDatabases. In *Encyclopedia of GIS*, pp.41-45. Springer (2008)

[7] E. Camossi, M. Bertolotto, E. Bertino, et al A multigranular spatiotemporal data model. In *GIS'03*, pp. 94-101. ACM (2003)

[8] G. Pozzani, C. Combi: An inference system for relationships between spatial granularities. In *SIGSPATIAL GIS'11*, pp. 429-432. ACM (2011)

[9] E. Camossi, M. Bertolotto, E. Bertino. A multigranular Object-oriented Framework Supporting Spatio-temporal Granularity Conversions. *IJGIS*, 20(5),pp. 511-534. Taylor & Francis, (2006)

[10] M. McKenney, M. Schneider: Spatial partition graphs: A graph theoretic model of maps. In D. Papadias, D. Zhang, and G. Kollios, editors. In *SSTD'07*, volume 4605 of LNCS, pp. 167–184. Springer (2007)

[11] S. Wang, D. Liu: Spatio-temporal Database with Multi-granularities. In *WAIM'04*, pp.137-146. Spring-er-Verlag, Berlin Heidelberg (2004)

[12] A. Pauly, M. Schneider. Spatial vagueness and imprecision in databases. In *Procs. of the 23rd SAC*, pp. 875-879. ACM (2008)

[13] M. Vazirgiannis. Uncertainty handling in spatial relationships. In *Procs. of the 15th SAC*, pp. 494-500. ACM (2000)

[14] J. J. Levandoski, M. Sarwat, A. Eldawy, et al: LARS: A Location-Aware Recommender System. In *ICDE'12*, pp. 450~461. IEEE (2012)

[15] H. R. Schmidtke, W. Woo: A size-based qualitative approach to the representation of spatial granularity. In *IJCAI '07*, pp. 563–568 (2007)

[16] M. A. Bender, G. Pemmasani, S. Skiena, et al, Finding least common ancestors in directed acyclic graphs. In *12th ACM-SIAM SODA*, pp. 845-854. Society for Industrial and Applied Mathematics (2001)

[17] M. C. Norrie, M. Grossniklaus. Multi-granular Spatio-temporal Object Models: Concepts and Research Directions. In *ICOODB'09*, pp. 132-148. Springer (2009)

[18] T. Iwuchukwu, J. F. Naughton: K-anonymization as spatial indexing: toward scalable and incremental anonymization. In *VLDB'07*, pp. 746-757. ACM (2007)

[19] Bettini, S. Mascetti, X.S. Wang. Supporting temporal reasoning by mapping calendar expressions to minimal periodic sets. *JAIR*, 28: 299-348. AAAI (2007)

[20] F. Giunchiglia, V. Maltese, F. Farazi. GeoWordNet: A Resource for Geo-spatial Applications. In The Semantic Web: Research and Applications, Volume 6088, pp 121-136. Springer (2010).

[21] AG Rojas, S Athanasiou, J Lehmann et al. GeoKnow: Leveraging Geospatial Data in the Web of Data. W3C Open Data Workshop, W3C (2013)

[22] S. Shah, et al. CROWDSAFE. SIGSPATIAL GIS'11, pp. 521-524 ACM (2011)

[23] D. Wu, et al. Integrating multiple data sources for stock prediction. WISE'08, pp.77-89. Springer (2008)

[24] J. Hoffart, et al. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. IJCAI'13. AAAI (2013)