Converting Spatiotemporal Data Among Heterogeneous Granularity Systems

Muhao Chen, Shi Gao
University of California, Los Angeles
{muhaochen, gaoshi}@cs.ucla.edu

X. Sean Wang
Fudan University
xywangcs@fudan.edu.cn

***Abstract*—Spatiotemporal data are often expressed in terms of granularities 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 enable granular comparison and data conversion within the system. However, if data from multiple granularity systems needs 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 the existence of a CRG and unifying conversions of incongruous semantics, and supports the approach to integrate multiple systems into one so as to process 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.**

***Keywords*—spatiotemporal data; multi-granularity; granularity conversion; granular comparison; system combination**

I. Introduction

In this decade where over 80% datasets have spatial features and are usually with temporal components [2], the notion of granularity has become significant for expressing and exchanging spatiotemporal data under specific units of measurement.

It's a common practice in literatures to organize a group of granularities in a partial-order set or a lattice [2-4, 6, 8, 10, 14], where granularities are linked with a partial-order topological relation (hereafter *linking relation*) into a hierarchical set. Two major functions are normally associated with such a set, namely *granularity conversion* and *granular comparison* [3, 8]. The former enables the expression of data in different measurement units, while the latter supports the topological or statistical analysis on spatially or temporally qualified information. These functionalities can be used in applications like multi-scaled information retrieval, dynamic knowledge extraction [3, 8, 19], and in multi-dimensional databases to enhance various types of queries from big spatiotemporal datasets [5, 13, 17]. 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 up into one, we will encounter a scenario where several granularity systems are simultaneously used in the same spatial or temporal domain. Such coexistence of granularity systems usually results from different representation standards as well as separate data maintenance realms, causing *heterogeneity in granularities*. E.g., time can be recorded in different calendars like the solar calendar and the lunar calendar. The government may collect civilian data bound with locations, managed separately by systems formed with {climatic regions, provinces, direct municipalities, streets} or {economic regions, counties, cities}. Datasets can also be indexed by different multi-granular structures in different clusters of a distributed system [13]. In addition, *heterogeneity in linking relations* is reflected by the existing instances in the literature [2, 6, 8, 10, 13, 14], such as FinerThan, GroupsInto, Partition, and CoarserThan.

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 exchanging 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 respectively in their original granularity systems. A general framework that supports such data transformation in multiple granularity systems will benefit many applications such as knowledge integration and data analysis. When integrating several knowledge bases (e.g. Wikidata, GeoNames, and TGN [19]), it's indispensable to provide the transformation of knowledge among ontologies that adopt different time granularities in knowledge representation (e.g. eons, eras, dynasties, different calendars, and time zones), and even more comprehensive spatial ones. Or, consider analyzing multiple spatial or temporal datasets which have their own representation standards, by enabling conversion of granular quantities across these standards, we can finally merge these datasets and unify the global knowledge extraction in different levels of details.

However, this is a non-trivial problem which brings several new challenges to multi-granular modeling. Besides adapting the heterogeneity in a model, interoperation of data across systems unavoidably requires extending the original in-system granularity conversion and granular comparison [3, 8] to their inter-system equivalents. However, heterogeneity of granularity systems often implies incongruous semantics of conversion, and indeterminacy of the existence of a common refined granularity (CRG) to allow correct 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 uncertainty to the conversion of data across granularity systems. Such uncertainty, which is not reflected by models in current

978-1-5090-0626-7/16/$31.00

literatures, 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 heterogeneous granularity systems. This framework first generalizes coexisting granularity systems to support their heterogeneity, and defines a graph model to represent their structures, and the semantics as well as the uncertainty of granularity conversions. It then introduces two constraints for composing inter-system granularity conversions, namely *semantic preservation* and *semantic consistency*. We show that granularity systems can be combined, or they have *combinabiliy,* only if they are semantically preserved or semantically consistent in addition to globally having a CRG. A novel approach is proposed to combine multi-systems to a single lattice, where inter-system conversion and comparison can be processed transparently just like in one granularity system. Quantification of uncertainty in granularity conversions is also introduced so as to understand the precision of granular comparison.

The rest of the paper is organized as follows. In section 2, we state the background with related work. In section 3, we model the granularity systems and related concepts. Granularity system combination and granularity conversion are discussed in section 4. Section 5 focuses on the uncertainty and CRG search problem in granular comparison. In section 6, we provide a conceptual evaluation. And we conclude with section 7.

II. Related Work

In computer science community, there are several definitions of granularities for different modeling purposes. One common definition is the partition of a domain [6, 8, 14]. Some others add graph features to facilitate the reasoning of granular data inside a granularity [2, 9]. Despite how they are defined, most literatures order multiple granularities in a lattice. For example, the temporal granularity lattice was proposed by Bettini et al in [4], which has later been transplanted to organize spatial granularities [3, 6, 8, 10, 14], as any such structures benefit with a communal finest unit of representation (i.e. the zero element) and clarifies ordering of granularities w.r.t. their fine or coarse degree. Such hierarchical structure is also a support for scalable retrieval in stratified spatial or spatiotemporal datasets, like the pyramid structure in LARS [13]. Any such a hierarchical set is a granularity system we discuss.

Usually a partial-order topological relation uniformly associates granularities in a granularity system, such as FinerThan [6,8], GroupsInto and Partition [2-4]. Based on that, the semantics and properties of granularity conversions are decided [8]. Heterogeneity of linking relations has been discussed in [3]. 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 that sets variation to semantics and properties of in-system conversion, let along challenges when conversions are extended to inter-system. We accept the heterogeneity of multi-systems with a more general model here.

Camossi et al have highlighted *granularity conversion* and *granular comparison* as the fundamental challenges in current spatiotemporal multi-granularity research [3]. In order to perform meaningful comparison, data of different granularities should be converted to a CRG [3, 8]. To enable inter-system granular comparisons, we have to verify the existence of CRG for any pair of granularities among the granularity systems.

Granularity conversion has been studied in a few works. E.g., in-system granularity conversions are respectively defined by Camossi et al [8], Moira et al [16] w.r.t. FinerThan, which preserves geometric correctness and topological consistency. Properties vary along with the conversion semantics in systems defined with other linking relations, e.g. the pyramid structure linked with CoveredBy [13]. However, these works didn't consider the conditions where linking relations are different among several systems, which bring along inconsistency of conversion semantics and incorrectness of conversion. This is solved in our work (in Section 4).

Moreover, finite precision of granules [11] and incongruity of geometric properties satisfied by different linking relations lead to uncertainty of granularity conversions, which has been tackled by Wang and Liu in [10] in a way of classification based on granular coverage. But no quantization of such uncertainty is considered in literatures, even though it is essential for precise granular comparison and related data analysis. Thus not only the geometric distortion [3] of granules is uncontrollable, but also the quantitive imprecision of granular comparison that exists among granularities cannot be evaluated. We propose a novel approach to quantify the distortion in granularity conversion, and reduce the problem for searching the optimal common refined granularity (OCRG), which has greatest expectation of geometric or statistical precision, to the LCA problem [15].

III. Modeling Granularity Systems

To set the stage, we begin with the modeling of spatial and temporal granularity systems.

A. 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 a partition on a Euclidean domain. The spatial and temporal granularities are defined as follows [8, 21].

Definition 3.1 (*Spatial Granularity*): *A spatial granularity is defined with a mapping G_S: $N \rightarrow 2^S$.*

- *$S \subseteq R^2$ is the spatial extent of the granularity in the spatial domain R^2, while R is the real number field. N is the natural number field. 2^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 non-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: $N \rightarrow 2^T$.*

- *$T \subseteq R$ is the temporal extent within the real number field R. N is the natural number field. 2^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 non-empty, then each element of $G_T(i)$ is less than all elements of $G_T(j)$. $\forall i,j,k \in N$ 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 spatial information. E.g., continents, nations, and states 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 granularities on calendars.

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

In [2, 7] have listed several topological granularity relations for spatial granularities, which are defined in [7]. They're classified into *partial-order relations* and *symmetric 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)***.
Partition(G,H): *G groups into and is finer than H. The converse is* ***PartitionedBy(H,G)***.
SubGranularity(G,H): *for each granule of G, there exists a granule in H with the same spatial extent.*
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 that of H.*
Overlap(G,H): *some granules of G and H overlap.*

These relations can be applied to temporal granularities, followed by two time-dedicated partial-orders given in [3, 4]:

GroupsPeriodicallyInto(G,H): *G groups into H. $\exists n, m \in N$ where $n<m$ and $n<|H|$, s.t. $\forall i \in N$, $\exists \{j_l\}_{l=1}^{\infty} \subseteq N$, 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,H): *G groups periodically into H, and m=1 in the above definition of* **GroupsPeriodicallyInto**.

These granularity relations are essential to manage the granularities in a granularity system. Algorithms to verify each relation can be implemented with topological relation operations in many spatial and temporal database extensions [7].

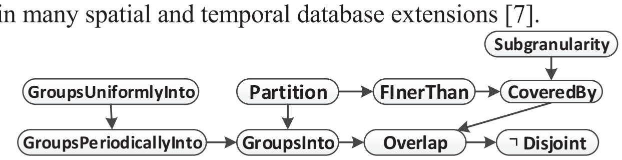

Fig. 1. The Hasse diagram of the logical inference of above granularity relations (not including the converse relations)

Property 3.1 gives the logical inference of these relations, whose transitive closure can be proved completeness. Fig.1 shows the Hasse diagram for the inference. This is a premise for verifying the semantic constraints of granularity conversion in Section 4.

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) *GroupedBy(G,H)⊢Overlap(G,H)*
FinerThan(G,H)⊢CoveredBy(G,H) *CoarserThan(G,H)⊢Covers(G,H)*
CoveredBy(G,H)⊢Overlap(G,H) *Covers(G,H)⊢Overlap(G,H)*
SubGranularity(G,H) ⊢CoveredBy(G,H)
Partition(G,H)⊢FinerThan(G,H)∧GroupsInto(G,H)
PartitionedBy(G,H)⊢CoarserThan(G,H)∧GroupedBy(G,H)
FinerThan(G,H)∧GroupsInto(G,H)⊢Partition(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)

B. Granularity Systems

Literatures organize a system of granularities as a lattice with a specific granularity relation (linking relation) [3, 4, 6, 8, 10] as such an algebraic system guarantees correct granular conversion and comparison in a granularity system [3]. We have addressed that, among multiple cases of these lattices, the *heterogeneity* generally lies in linking relations and granularities including identity and zero elements. We hereby generalize above features in one definition.

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

- *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 ∈{G} known as the zero element, i.e. ∀G∈{G}, G_0≤G always holds.*
- *G_1: N→P(D) is the granularity G_1 ∈{G} known as the identity element, i.e. ∀G∈{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 defined on 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 involved system is from the same $\mathcal{E}_D$. Example 3.1 forms a D-system Group with three heterogeneous granularity systems from real-world systems. We slightly modify these use cases from real-world knowledge bases to better illustrate our problem.

Example 3.1. *$\mathcal{E}_D$={GS_1, GS_2, GS_3} includes three GSs formed with granularities fetched respectively from Wikidata, GeoNames and TGN[20].*

- *GS_1 from Wikidata's Places category: Granularities {G_{11}=Continents, G_{12}=Nations}; the linking relation* ***GroupsInto*** *applies as {(G_{12}, G_{11})};*
- *GS_2 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 from Geonames: Granularities {G_{31}=Administrative divisions, G_{32}=Populated places, G_{33}=Roads and rails}; 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})}.*

C. Granularity Conversion

A granularity system defined above regulates the uniform order of granularities. If we scan through any path from G_1 to G_0, the resolution that these granularities express transforms uniformly. The remaining indeterminacy, which is whether refining or merging such transformation causes, is clarified as below.

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

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

This concept classifies granularity relations into two groups, say *refining relations*: {FinerThan, GroupsInto, Partition, CoveredBy, GroupsPeriodicallyInto, GroupsUniformlyInto}, and *merging relations*: {CoarserThan, GroupedBy, PartitionedBy, Covers}. We refer granularity systems defined with a 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 refining system by inverting its linking relation, we discuss refining systems hereafter *w.l.o.g.*

Granularity conversion, which shifts the granular data across resolutions, is defined as follows.

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

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

In any granularity system, a conversion across G≤H preserves the specific semantics related to ≤. E.g., let ≤ be Partition, then $Conv_{H\to G}(\{H(1),H(2)\})_\leq$ would return the granules at G whose extent exactly equals to H(1),H(2). But we may only get granules perhaps just roughly covered by H(1),H(2) when ≤ is FinerThan. In-system granularity conversions are compositional [8] since they are defined with the same linking relation.

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

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

D. Weighted Granularity Graph

A granularity system can be transformed to a weighted granularity graph (WG). It is defined as Definition 3.7, with which we can represent many problems including multi-system combination, reasoning of semantic constraints of granularity conversions, quantization of uncertainty in conversion and granular comparison, on graphs.

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

- *V is the set of vertexes representing granularities, Mv:V→{G} is a bijection from vertexes to granularities, where $G_0=Mv(V_0)$, $G_1=M_V(V_1)$.*
- *E is the set of edges, which is the subset of V×V. For each e(u,v)∈E, Mv(v)≤ Mv(u) and no other G∈{G} exists s.t. Mv(v)≤G and G≤Mv(u)*
- *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].*
- *R is the label function of linking relation, R(WG)=≤.*
- *M_S is the bijection from WG to GS ($M_S(WG)=GS$).*

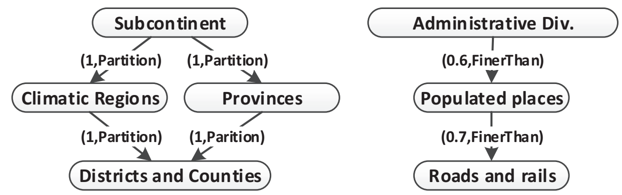

Fig. 2. WGs for GS_2 and GS_3 in Example 3.1

The WG has the similar property on edge creation to a Hasse diagram. It is acyclic, and is a transitive reduction (its edges are created for atomic relations only). 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 discuss in Section 5. Two WGs of GS_2 and GS_3 in Example 3.1 are shown in Fig. 2, where the weight is annotated with (assumed) numbers in (0,1] to denote the expected precision of atom granularity conversions.

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

IV. Combining Granularity Systems

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

We hereby discuss the property *combinability*, which guarantees essential pre-conditions of inter-system granularity conversion, i.e. semantic preservation and semantic consistency, as well as the feasibility of granular comparison. Then, we propose the approach of multi-system combination.

A. Semantic Preservation and Consistency

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

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

Remark Across multiple systems, semantic preservation directly extends the conversion from a granularity in the original system to another granularity in the second system with the same semantics. While semantic consistency decides a uniform semantics for a composed conversion, although it may lose semantics in the original system. E.g., regarding a refine-conversion of granule {g} in a GroupsInto system as fetching all granules in a certain refined granularity that groups into {g}, such illustration still applies to a semantic preserved conversion for {g}, but may not apply to a semantic consistent one for {g} as the semantics can be weaker.

Thereof, across systems, the former is the requisite for direct inheritance of the conversion operation in original systems as well as their compositionality. 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 which hasn't been originally existing in the composition is introduced.

It is noteworthy that, compositionality of conversions guaranteed by semantic consistency is the precondition to: (1) ensuring that any inter-system conversion $Conv_{G1\rightarrow Gn}(G')_{\preceq j}$ complies Definition 3.6; (2) value correctness of composed conversions, as well as efficiency in such since we avoid multiple atom conversions and can use inter-granularity indices [25]; (3) transitivity and path-independence of geometric/statistic precision which enables analysis of these quantities and their usability as weight of WG when solving the OCRG problem (Sect. 5.1); (4) usability of conversion-based aggregation functions, e.g. Sum, Avg, Max [8], preserving their linearity w.r.t. conversions.

In-system conversions have ensured these constraints with the same linking relation. However, when we extend such conversions 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 GroupsPeriodicalllyInto(G,H) be true.*
2. *Let GroupsPeriodicalllyInto(H,I) and GroupsInto(G,H) be true, but GroupsPeriodicalllyInto(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 case 1, the composed conversion from I to G is semantic preserved and consistent, because GroupsPeriodicallyInto(G,H) $\rightarrow$ *GroupsInto(G,H). In case 2, the composed conversion doesn't preserve the periodicity semantic in conversion from I to H, so* $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{GroupsPeriodicallynto}, G)_{GroupsInto} = Conv_{I\rightarrow G}(I')_{GroupsPeriodicallyInto}$ *doesn't hold. But semantic consistency still holds on GroupsInto s.t.* $Conv_{H\rightarrow G}(Conv_{I\rightarrow H}(I')_{GroupsPeriodicallyInto}, G)_{GroupsInto} = Conv_{I\rightarrow G}(I')_{GroupsInto}$ *since GroupsPeriodicallyInto(H,I)* $\rightarrow$ *GroupsInto(H,I). As for the third, neither of the two semantic constraints holds, since no inference holds from either of FinerThan and GroupsInto to the other according to Property 3.1.*

In scenarios where granularities are used for precise multi-resolution representation, we have to preserve the conversion semantics in original systems so as not to lose certain properties (such as geometric congruity, periodicity, etc.) by following Property 4.1.[1]

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

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

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

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

B. 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.*

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

Given a pair of granularity systems GS, GS' from $\mathcal{E}_D$, the sufficient-necessary (SN) conditions for them to satisfy either type of the combinability are given as below.

Theorem 4.1 (*Semantic Preserved Combinability*): *Given a pair of refining granularity systems GS(D, {G},* $\preceq$, G_0, G_1*) and GS'(D, {G}',* $\preceq'$, G'_0, G'_1*) from* $\mathcal{E}_D$, *semantic preserved combinability holds between them iff one of the follows holds. (Here C1~C6 denotes six conditions we use in proof [25])*

1. $G_0 = G'_0$ ($\preceq=\preceq'$, *C1; or* $\preceq\neq\preceq'$, *C2).*

[1] Due to space limitation, proofs to properties (Property 4.1, 4.2, 5.1, 5.2) and theorems (Theorem 4.1, 4.2, 5.1) are omitted in this version. Readers are referred to our technical report [25] which includes detailed proofs.

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. $\leq\to\leq'$ and $G_0\leq' G'_0$; or $\leq'\to\leq$ and $G'_0\leq G_0$; (C5) or exists a third (intermediate) system $GS^* \in\mathcal{E}_D$ having linking relation $\leq^*$, s.t. $\leq^*\to\leq$ and $\leq^*\to\leq'$, and zero element G^*_0 s.t. $G^*_0\leq G_0$ and $G^*_0\leq' G'_0$ (C6).

Theorem 4.2 (*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. (Here C1~C4 denotes four conditions we use in proof [25])*

1. $G_0=G'_0$ ($\leq=\leq'$, C1; or $\leq\neq\leq'$, C2).
2. Any of $\leq=\leq'$, $\leq\to\leq^*$ or $\leq^*\to\leq$ 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 ($\leq=\leq^*$, $\leq^*\to\leq$ or $\leq\to\leq^*$) and any of ($\leq'=\leq^*$, $\leq^*\to\leq'$ or $\leq'\to\leq^*$) 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).

Algorithms to verify either of the combinability conditions can be created as sequential procedures to verify the satisfaction of C1~C6 in Theorem 4.1 or C1~C4 in Theorem 4.2. 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 our techical report [25].

C. Multi-system Combination

We can now combine multiple granularity systems from $\mathcal{E}_D$ to one system that guarantees correct inter-system granular comparison and 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 algorithms in the way of finding and combining other combinable systems to a target.

We refer the combination algorithms as *SPCombine(WG, $\{WG\}_D$)* (semantic preserved combination) and *SCCombine(WG, $\{WG\}_D$)* (semantic consistent combination), which both process granularity systems in the form of their granularity graphs. In fact, the two algorithms are logically similar (only being different in the constraint of creating edges between two systems according with the semantic constraints). Exemplarily, we give the algorithm of semantic preserved combination.

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

Algorithm 4.1 *SPCombine*(*WG, {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
            DFSCreateEdges(WG.V1, WG'.V1,R(WG),checked,false)
            DFSCreateEdges(WG'.V1, WG.V1,R(WG),checked,false)
        else if R(WG')→R(WG) then
            DFSCreateEdges(WG.V1, WG'.V1,R(WG),checked,false)
            R(WG)←R(WG')
        else if R(WG)→R(WG') then
            DFSCreateEdges(WG'.V1, WG.V1,R(WG'),checked,false)
        else for each WG*∈{WG}D do
            if R(WG*)→R(WG) and R(WG*)→R(WG') then
                {WG}D←{WG}D\WG*
                DFSCreateEdges(WG.V1,WG*.V1,R(WG),checked,false)
                DFSCreateEdges(WG'.V1,WG*.V1,R(WG'),checked,false)
                R(WG)←R(WG*)
                continue
for each WG'∈{WG}D do
    if Mv(WG.V0)=Mv(WG'.V0) do
        {WG}D←{WG}D\WG'
        MergeVertex(WG.V0, WG'.V0)
        CreateDirectedEdge(WG.V1,WG'.V1)
return WG.V1
```

Thereof, DFSCreateEdges given as Algorithm 4.2 links the vertexes of one system to those of the other to mark any atom relation, while enforces transitive reduction as required in the definition of WG.

Algorithm 4.2 *DFSCreateEdges*(*v, u, ≤, checked[,], foundabove*)

```
found←foundbelow←created←false
checked[v,u]←true
if foundabove=false and Mv(u)≤ Mv(v) then
    found←true
if (found∨foundabove=true) and Succ(v)≠∅ then
    for each v'∈Succ(v) do
        if checked[v,u]=false and Mv(u)≤ Mv(v') then
            foundbelow←true
            DFSCreateEdges(v', u, ≤, true)
        else for each u' ∈Succ(u)
            if checked[v',u']=false then
                DFSCreateEdges(v', u', ≤, checked,false)
    if foundbelow=false then 'an atom relation is found
        CreateDirectedEdge(v,u) 'from v to u
        created←true
if created=true and Succ(v)≠∅ and Succ(u)≠∅ then
    for each v'∈Succ(v) do
        for each u' ∈Succ(u)
            if checked[v',u']=false then
                DFSCreateEdges(v', u', ≤, checked,false)
else if found∨foundabove=false and Succ(u)≠∅ then
    for each u'∈Succ(v) do
        if checked[v,u']=false then
            DFSCreateEdges(v', u, ≤, checked,false)
else return
```

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

Algorithm 4.2 links the vertexes (v and its successors) from one system to those (u and its successors) from the other through depth-first search. The tag set *checked* marks if a pair of vertexes u,v have been checked. For a v, if $\exists$u s.t. u≤v, the algorithm checks the successors v' of v to see if $\exists$v' s.t. u≤v'. If no such a v' is found, an edge is created between u,v. Otherwise such check is recursively applied to u and successors of v' until this condition is fulfilled. For any check where u and v don't satisfy u≤v, check is performed recursively to their successors. However, if any pair of u,v have been checked once according to checked[,], no repeated check is processed on them and their successors. This procedure ensures the transitive reduction (only create edges for atom relations). Combinability guarantees a V_0 after each phase of combination, and a V_C denoting the domain is created as V_1, making the result a lattice.

The last loop (23~27 of SPCombine) deals with a special case where the linking relation of WG cannot form any logical inference with that of any other WG', but shares the equivalent V_0 with a WG'. Therefore they are combined by merging V_0, but no other vertexes are linked between these two graphs. Overall, let $|\{WG\}_D|$ and each $|G|$ be both O(n) magnitude, DFSCreateEdge takes $O(n^2)$ such that SPCombine takes $O(n^3)$ time complexity. Above properties also hold for SCCombine.

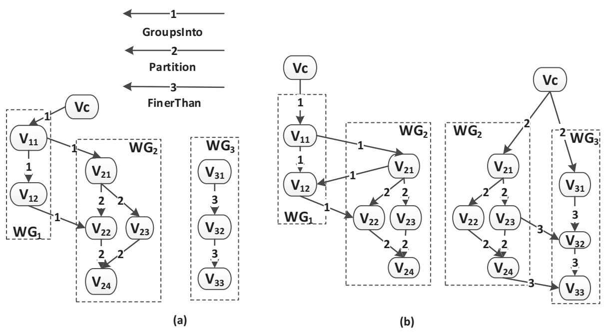

Fig. 3. Graph depiction for Example 4.2 (a) semantic preserved combination (b) semantic consistent combination of three granularity systems

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 the three WGs shown in Fig.3 are the WGs of GS_1, GS_2 and GS_3 respectively. The vertexes denote corresponding granularities with the same subscripts. Recall that across the three 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. Allowed SPCombine is depicted as Fig.3(a), where any inter-system conversion from a granularity in GS_1 to another in GS_2 preserves the semantics of GroupsInto originally in GS_1. However, GS_3 doesn't satisfy semantic preserved combinability with GS_1 and GS_3. The allowed SCCombine is depicted as Fig.3(b), where the combination process between GS_2 and GS_3 is valid since conversions from GS_2 to GS_3 share the semantic consistency w.r.t. FinerThan.*

V. Granular Comparison With Uncertainty

Since each linking relation varies in geometric properties, a granularity conversion unavoidably contains distortion that causes imprecision of representation and statistical analysis of granular data.

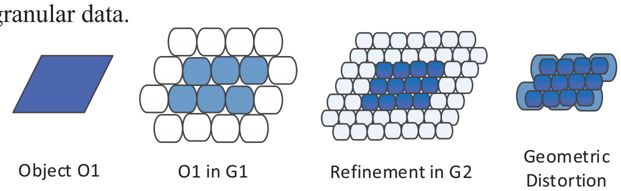

Fig. 4. An object projection in different granularities

On the other hand, in the combined system, it is very likely that more than one CRGs for a pair of granularities exist (e.g. in Fig.3(b), V_{12}, V_{21}, V_{22} and V_{24} are all CRGs of V_{11} and V_{21}; V_{32} and V_{33} are both CRGs of V_{23} and V_{31}). As numbers of CRGs may boost up along with numbers of combined systems, selecting and converting towards the one CRG with least expectation of distortion in conversions may significantly decrease related imprecision of comparison.

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

A. Geometric and Statistic Distortion

Distortion of granularity conversion is reflected in two aspects: *geometric distortion* results from the incongruity of extents by some linking relations (such as FinerThan and CoveredBy), causing a granular object to become a vague object (Fig. 4) [11]; *statistic distortion* that results from the loss of data aggregation. To quantify such distortion, we begin with the definition of geometric precision (which is the complementary of distortion) of a granule during conversion, and the granularity geometric precision which is defined as its expectation.

Definition 5.1 (*Granular and Granularity Geometric Precision*): *Given granularities G, H and a refining relation $\preceq$ s.t. $H \preceq G$. Let g be a granule of G, then the degree of granular geometric precision between g and its conversion result at H, is defined as (an 'o' over an extent denotes its interior area):* $u(g,H)=\frac{g^o\cap(Conv_{G\to H}(\{g\})_\preceq)^o}{g^o\cup(Conv_{G\to H}(\{g\})_\preceq)^o}$. *Its expectation equals to the granularity geometric precision between G,H:* $U(G,H)=E(u_\rho(G(i),H))=\frac{(\cup_{i\in N}G(i)^o)\cap(\cup_{i\in N}H(i)^o)}{(\cup_{i\in N}G(i)^o)\cup(\cup_{i\in N}H(i)^o)}$

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

In databases where a granularity system is used for multi-scale data management and aggregates [8] are available, the precision is related to data density, which is defined as below.

Definition 5.2 (*Data density*): *Give a dataset E, and a spatial/temporal extent C, the data density in C is defined as:* $\rho(C)=\frac{|\{e|e\in E\wedge coveredBy(e,C)\}|}{C^o}$

The ρ-granular precision defined below denotes the ratio of statistic preservation of data bond with a granule (e.g. the number of residents in a block) in granularity conversions. While by computing the expectation of u_ρ(g,H) on the extent of H, we also get the ρ-precision between two granularities.

Definition 5.3 (*ρ-granular and ρ-granularity Precision*): *Given granularities G, H and a refining relation $\preceq$ s.t. $H \preceq G$, let g be a granule of G, ρ be the data density, then the ρ-granular precision between g and its conversion result at H, is defined as:* $u_\rho(g,H)=\rho\left(\frac{g^o\cap(Conv_{G\to H}(\{g\})_\preceq)^o}{g^o\cup(Conv_{G\to H}(\{g\})_\preceq)^o}\right)$. *The ρ-granularity precision between G,H is defined as its expectation:* $U_\rho(G,H)=E(u_\rho(G(i),H))=\frac{\rho\left((\cup_{i\in N}G(i)^o)\cap(\cup_{i\in N}H(i)^o)\right)}{\rho\left((\cup_{i\in N}G(i)^o)\cup(\cup_{i\in N}H(i)^o)\right)}$

The ρ-precision quantifies the ratio of statistic accuracy when data are bond with a granule or a scaled portion of dataset (such as a "fraction" in the pyramid structure of LARS [13]) that transforms across granularities, from what the statistic bias in aggregate functions like *sum* and *count* [8] is also achieved. Above quantities are all unified, e.g.. $u(g,H)\in(0,1]$ and $u_\rho(g,H)\in(0,1]$. Corresponding precision equals 1 iff no distortion occurs after conversion.

For granularity systems of different utilization purposes, we can use either geometric granularity precision or ρ-granularity precision as the weight on the edges of WG for the system, if with their transitivity and path independence satisfied.

Property 5.1 ***(Transitivity)***: *Given a linking relations≤, and granularities G,H,I s.t. G⊴H⊴I. Then U(I,H)·U(H,G)=U(I,G) and* $U_\rho(I,H)\cdot U_\rho(H,G)=U_\rho(I,G)$ *always hold.*

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'. Then U(I,H)·U(H,G)=U(I,H')·U(H',G) and* $U_\rho(I,H)\cdot U_\rho(H,G) =U_\rho(I, H')\cdot U_\rho (H',G)$ *always hold.*

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

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

B. Granular comparison

Conversion towards the OCRG, defined as below, enables the comparison of granular data from two granularities with the greatest expectation of geometric precision (U) or statistic precision (U_ρ). Thereof, the *Gain* function can use either U or U_ρ.

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

As we have combined multiple-systems as a weighted granularity graph, and assigned quantified distortion as weight, the search of the OCRG can be reduced to the Least Common Ancestors (LCA) problem on a weighted acyclic digraph. An O(n) algorithm [15] for such problem, say FindOCRG(G,G'), is easily implemented by finding the LCA with least products of edge weights through two-way DFS.

Example 5.1 *We compare the sales of product* p_A *in the scope of Province g and that of* p_B *in the scope of Municipal h. Suppose G=provinces and H=municipals have two CRGs I=districts and J=blocks (blocks can cross districts, so they are not FinerThan districts). Suppose* $(U\rho(G,I)\cdot U\rho(H,I))^{1/2}<(U\rho (G,J)\cdot U\rho(H,J))^{1/2}$, *then the database system first processes FindOCRG(Provinces, Municipals)=J, then it computes and compares* $Sum(Conv_{G\to J}(g)_{FinerThan}(sales\_p_A))$ *and* $Sum(Conv_{G\to J}(g)_{FinerThan}(sales\_p_B))$.

VI. Conceptual Evaluation and Application Aspects

The framework proposed in previous sections is not specific to a particular scenario, but it demonstrates with generality and feasibility. In this section, we present an overall conceptual evaluation of our approach. Also we will present how this approach will greatly benefit several application aspects in knowledge representation and management.

A. Conceptual Evaluation

The framework uses the classic definition of granularity, i.e. a partition mapping of a domain. It has extended granularity conversion and granular comparison, which are the major functions to be supported by granularity systems [3], from in-system to inter-system. We have considered the uncertainty factors of inter-system conversions, i.e. incongruity of semantics and geometric uncertainty. Two semantic constraints we define are essential to inter-system granularity conversions w.r.t. the incongruity of conversion semantics caused by heterogeneity of linking relations. The former preserves the inheritability of conversions from original systems. The latter ensures the compositionality of conversions, which is essential for their correctness and usability and related uncertainty quantization.

Another key is the approach of multi-system combination. It's obvious that the requirements of combinability are exactly the requirements for all conversions and granular comparison between two systems. The two types of combinability support granularity conversions and granular comparison in different degrees. The former condition benefits in directly inheriting the semantics of conversion function and corresponding auxiliary structures (e.g. periodic sets [18]) from the original systems. The latter is more generalized to guarantee the feasibility of the two interoperations. Though verifying combinability seems complex, the SN conditions enables solutions of O(1) complexity. Utilizing the WG, we introduce the combination algorithms to conceive the composed lattice we desire.

Meanwhile, we've introduced the quantification of geometric and statistic distortion in granularity conversion. Roles of this concept lie in three aspects. (1) It measures the uncertainty of conversion which semantic-related constraints don't preserve, especially when semantic preservation is not fulfilled. (2) It gives the basis to choose the optimal one from multiple CRGs so as to improve the preciseness of granular comparison. (3) *U* and U_ρ evaluate related expectation of precision for different application requirements of data.

B. Remaining Technical Details

Moreover, we explore some instructive strategies against the remaining challenges (for details please refer to our technical report [25]). 1) Verifying granularity relations is the most frequent operation in our approach, which needs at least $O(n^2)$ time complexity to process with RCC-based functions [7]. A *Global Granularity Relation Matrix* is an offline auxiliary to optimize such reasoning in O(1) time using Property 3.1. 2) Online reasoning of the closure constraint in conversions can be avoided by physically mapping the WG into a hierarchical index, where each granule is linked to the closure of its descendants applying corresponding linking relation. This structure reduces the cost of an atom conversion from O(n) to O(1). Also, for temporal granularity systems, semantic preserved combination inherits relation-specific auxiliary structures [18].

C. Application Aspects

We have demonstrated the power of our framework in systematically tackling the data manipulation among heterogeneous granularity systems. The approach and optimizations can be easily extended for following applications:

1) Spatial Knowledge Integration. The wide demands for interoperating, mapping and fusion of spatial information on the Web has led to newcomer projects like YAGO [12] and GeoKnow [20]. 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 detail (e.g. states) to another (e.g. cities), we will have to manually enumerate all the names from the latter to set the belongings to the former. This process, which is usually done by human efforts, is extremely time-consuming and error-prone. However, letting computers build the association between such two levels on FinerThan is with correctness and ease, and shall never miss a detail given the spatial quantities. This challenge, i.e. to provide implicit geographical references among spatial quantities across sources of divergent representation standard, hasn't been touched yet. Once the data conversion and granularity system integration are enabled with the help of our framework, integrating spatial knowledge bases can be processed in an automated and computable way so that the spatiotemporal knowledge can be automatically exchanged through divergent levels of details. We are investigating the extensions on knowledge management systems to support efficient data conversions in knowledge integration.

2) High-divergent Time Conversion. Consider the diversity of time expression here [1]. A scientist, intelligence analyst, historian, or archaeologist may encounter vast temporally qualified information of high-divergent time systems, e.g., geological time scale (GTS) systems (including eons, eras, periods, and epochs), chronicles of different history origins, different calendars [18], and different time zones. It's not difficult to see that a unified time conversion system is equivalent to the combination of temporal granularity systems with heterogeneous relations, e.g. GroupsInto in GTS systems and history chronicle systems, GroupsUniformlyInto in each calendar, GroupsPeriodicallyInto in time zones and among calendars.

3) Integrated Data Analysis. In many scenarios temporal or spatial data are joined and analyzed together, such as stock prediction based on multiple data sources [23] and crime data analysis [22]. Although multi-granular operations are proved to be helpful for improving the efficiency of data mining algorithms [3], heterogeneity, however, appears easily in these scenarios. As time representation meets already-known divergence, while the spatial side goes even more complex in vast data obtained from crowdsourcing [22]. What makes things worse is that, these data sources are data streams which require non-blocking analysis [24], therefore preprocessing to reduce the heterogeneity is not even available. Under such circumstances it's vital to enable the automatic granular quantity conversions across granularity systems so as to meet the original needs.

4) Querying Linked Data. The expression of query conditions that involve spatial/temporal entities or properties may vary in different knowledge bases since they use different granularity systems. Even within the same knowledge base, it exists multiple standards for the same property due to the application's diverse representation criterion, or user's knowledge preference (e.g. to state and county, some users may prefer regione and provincia). It is always left to users or front applications to deal with the heterogeneity. Applying the inter-system conversion in the query plan empowers with the rule-based mapping for such conditions among granularity systems, and brings along the transparency of query interpretation regardless of how application or user wants to represent the conditions.

VII. 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 with the heterogeneity of granularity systems reflected in literatures, and introduced 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. After enabling inter-system conversion and comparison, we have quantified the uncertainty in such interoperations.

Our framework has led us to some possible future work as well as open challenges. We plan to continue on the conceptual design and realization of our framework, particularly the implementation of our approach in knowledge integration. Implementing quantization of geometric or statistic uncertainty in databases and extending this model for spatiotemporal granularities [2] are also on our list.

REFERENCE

[1] G. Zhao, et al. Increasing temporal diversity with purchase intervals. *SIGIR'12*, pp. 165-174. ACM (2012)

[2] A. Belussi, et al. Formal and conceptual modeling of spatio-temporal granularities. *IDEAS'09*, pp.275-283. ACM (2009)

[3] E. Camossi, et al. Multigranular spatio-temporal model.: implementation challenges. *SIGSPATIAL GIS'08*, Article No. 63. ACM (2008)

[4] C. Bettini, et al. Time Granularities in Databases, Data Mining, and Temporal Reasoning. Springer (2000)

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

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

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

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

[9] M. McKenney, et al. Spatial partition graphs: A graph theoretic model of maps. *SSTD'07*, pp. 167–184. Springer (2007)

[10] S. Wang, D. Liu. Spatio-temporal Database with Multi-granularities. *WAIM'04*, pp.137-146. Springer (2004)

[11] A. Pauly, M. Schneider. Spatial vagueness and imprecision in databases. *SAC'08*, pp. 875-879. ACM (2008)

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

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

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

[15] M. A. Bender, et al, Finding least common ancestors in directed acyclic graphs. *SODA'01*, pp. 845-854. SIAM (2001)

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

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

[18] C. Bettini, et al. Supporting temporal reasoning by mapping calendar expressions to minimal periodic sets. *JAIR*, 28: 299-348. AAAI (2007)

[19] F. Giunchiglia, et al. GeoWordNet: A Resource for Geo-spatial Applications. *The Semantic Web: Research and Applications*, Vol. 6088, pp 121-136. Springer (2010)

[20] AG Rojas, et al. GeoKnow: Leveraging Geospatial Data in the Web of Data. *W3C Open Data Workshop*, W3C (2013)

[21] M. Chen, et al. Converting Spatiotemporal Data Among Multiple Granularity systems (Poster). *SAC'16*. ACM

[22] S. Shah, et al. CROWDSAFE: crowd sourcing of crime incidents and safe routing on mobile devices. *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] B. Babcock, et al. Models and issues in data stream systems. *PODS'02*, pp. 1-16. ACM, 2002.

[25] M. Chen. Converting Spatiotemporal Data Among Multiple Granularity Systems. *UCLA CSD Technical Report 150001*. (2015)