Converting Spatiotemporal Data Among Heterogeneous Granularity Systems

Muhao Chen[1], Shi Gao[1], X. Sean Wang[2]

Department Of Computer Science, University Of California Los Angeles[1]

School Of Computer Science, Fudan University[2]

Spatiotemporal Data

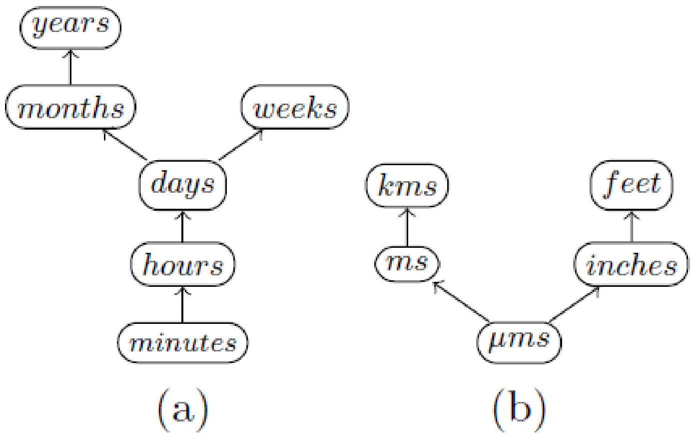
Time & Space：The inherent attributes of any existing object and event.

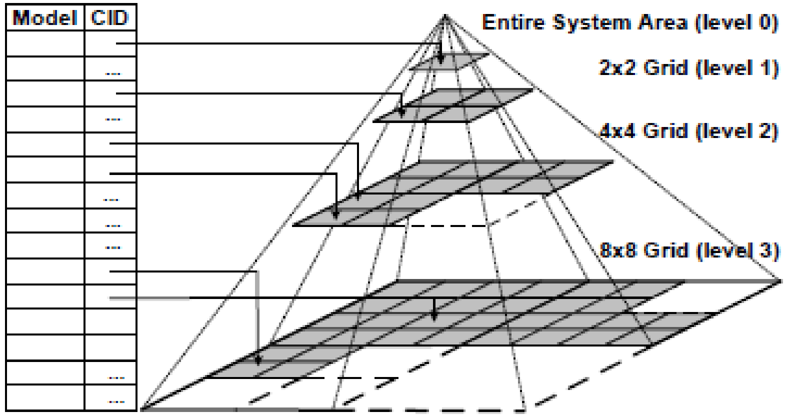
Features：

- Multi-resolution representation
- Different units of measurement
- Uncertainty (Vagueness and fuzziness)

Granularity and Granularity System (GS)

- Granularity: divides space / time into granules
- A GS: a partial-order lattice ({G}, ≤) which manages several granularities with a partial-order relation (E.g., FinerThan system)
- Two operations on GS:
  - **Granularity conversion**: convert a granular object to its "equivalence" or another granularity
  - **Granular comparison**: convert two granules to a same granularity and compare them





Granularity Relation

A topological relation between two granularities

Granularity Relations (Spatial/Temporal)

Partial-order relations

Relation	Description	Converse
GroupsInto(G,H)	*Each granule of H is equal to the union of a set of granules of G.*	***GroupedBy (H,G)***
FinerThan(G,H)	*Each granule of G is contained in one granule of H.*	***CoarserThan(H,G)***
Partition(G,H)	*G groups into and is finer than H.*	***PartitionedBy (H,G)***
CoveredBy(G,H)	*Each granule of G is covered by some granules of H.*	***Covers(H,G)***
SubGranularity(G,H)	*For each granule of G, there exists a granule in H with the same extent.*	

Symetric relations

Relation	Description
Disjoint(G,H)	*Any granule of G is disjoint with any granule of H.*
Overlap(G,H)	*Some granules of G and H overlap.*

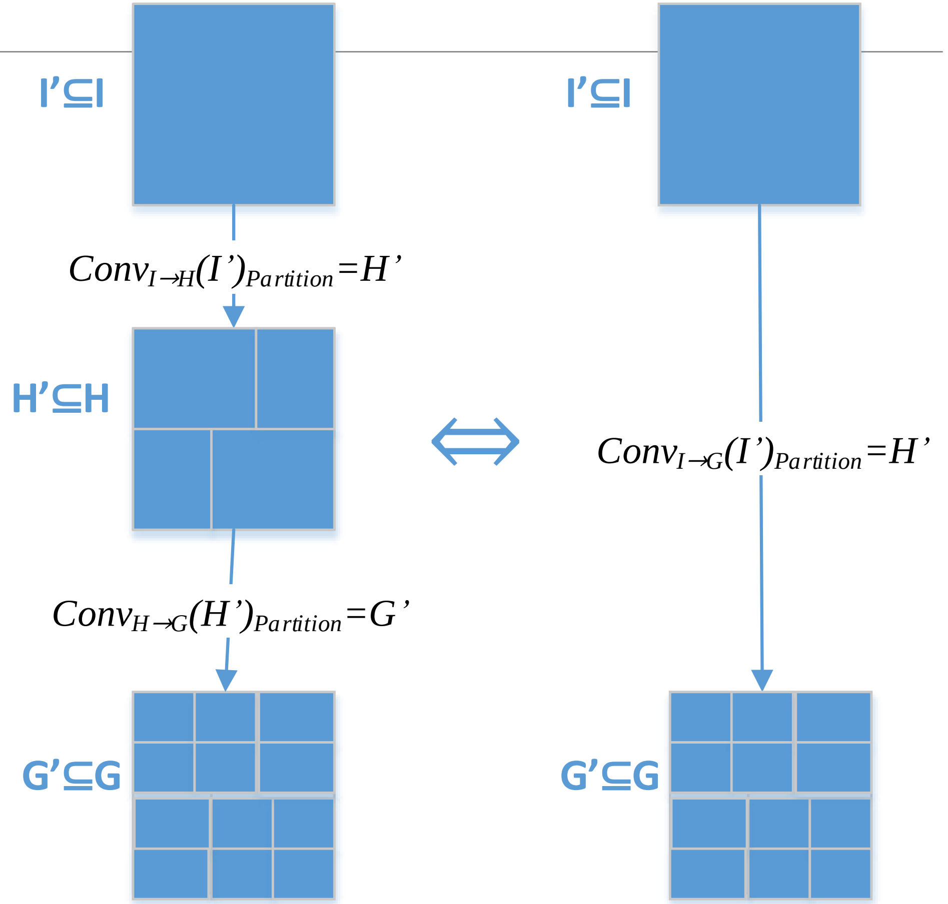
Granularity Relations (Continue)

Partial-order relations

GroupsPeriodicallyInto(G,H)	*G groups into H. ∃n, m ϵ N where n<m and n<|H|, s.t. ∀iϵN, if* $H(i)=\bigcup_{r=0}^{k} G(j+r)$ *and H(i + n) ≠ Ø then* $H(i+n)=\bigcup_{r=0}^{k} G(j+r+m)$.
GroupsUniformlyInto(G,H)	*G groups periodically into H, as well as m=1 in the above definition of* **GroupsPeriodicallyInto**.

Why A GS is a Lattice

- Compositionality of granularity conversion
  - Only one partial-order granularity relation is used
- Correctness of granular comparison
  - Existence of GLB (greatest lower bound) for any pair of granularities. (E. Camossi 2008)



Compositionality of conversions in one GS

Coexistence of Multiple Granularity Systems

Current works use only one GS to manage data

Lots of scenarios where multiple systems coexists and interacts:

- Different real-world representation standards
  - Solar/lunar calendar, history systems
  - Intl/US metrics
  - Different hierachical administrative divisions of countries
  - …
- Multiple heterogeneous GSs given respectively in literatures
- Integrate spatial/temporal knowledge bases (e.g., Wikidata, GeoNames, TGN, YAGO)

Coexistence of Multiple Granularity Systems (Continue)

Heterogeneity in Granularities:

- Inter-system granular comparison ✘ (compositionality not ensured)

Heterogeneity in Granularity Relations

- Inter-system granular comparison ✘ (GLB existence not ensured)
- Uncertainty of inter-system granular conversion ! (incongruous geom. properties)

Problems We Solve

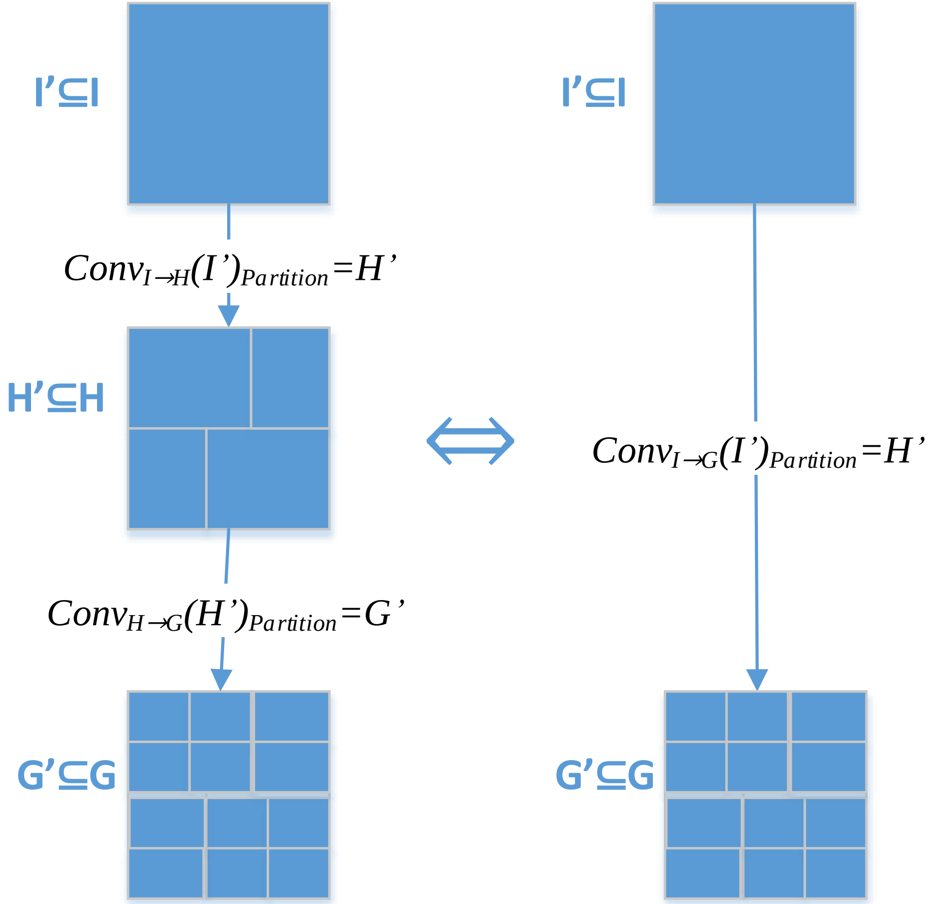
- Combine multiple heterogeneous GSs
- Extend granularity conversion and granular comparison among systems with correctness
- Model the uncertainty in inter-system conversion/comparison
- Reduce the expected uncertainty

Combining Multiple Systems

- Multiple lattices => one lattice
- Why?
  - Inter-system conversions ⇔ like in a single system
  - Inter-system granular comparison
  - Facilitate in solving the uncertainty problem later

Compositionality

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



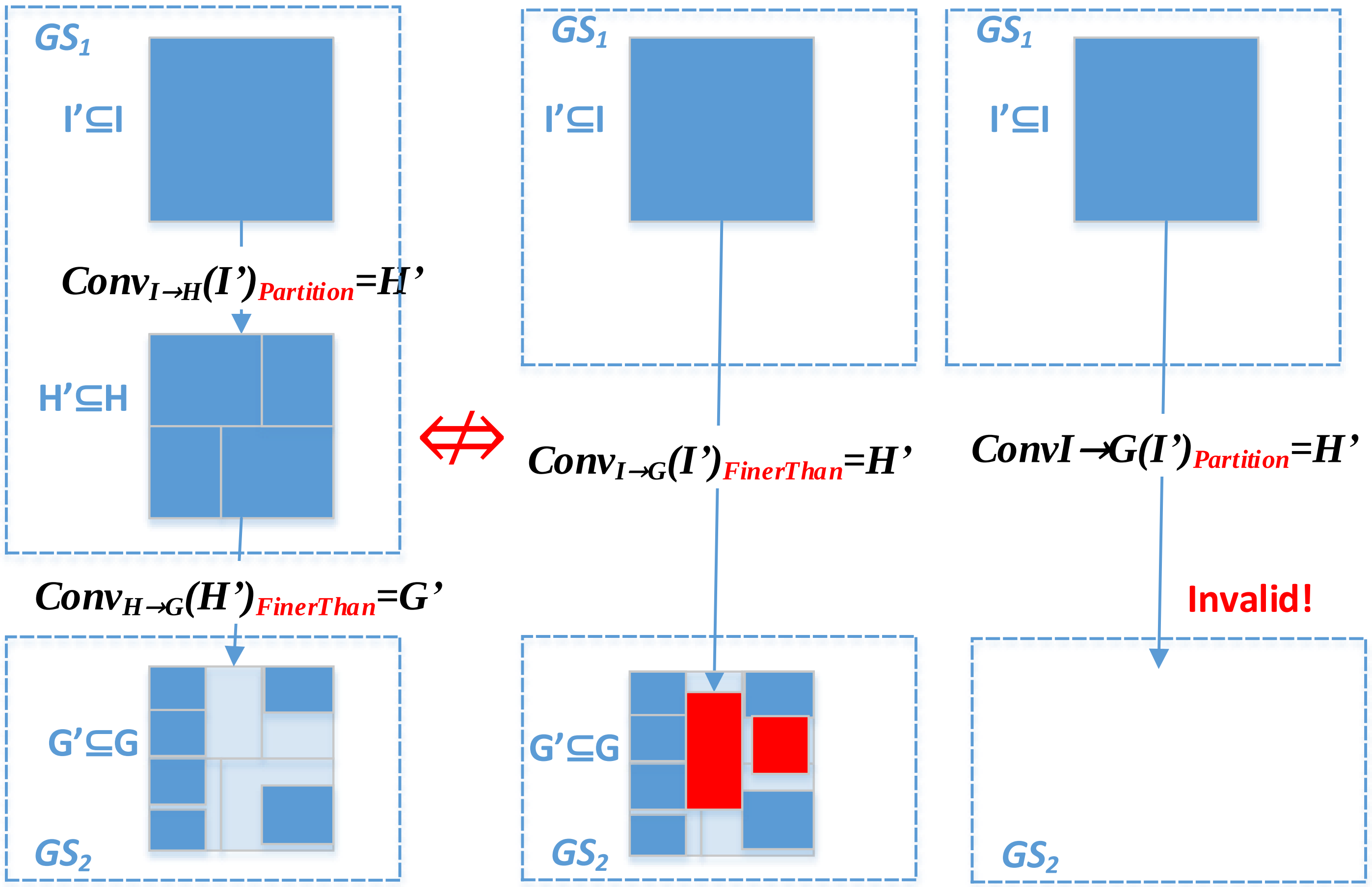
Does not necessarily hold across heterogeneous systems!

Inter-system Conversion

- Semantic inconsistency and semantic loss

  ->conversion is nondeterministic, or even invalid!

We need to find the conditions where compositionality holds across systems.



An Inference System for Granularity Relations

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)	

Two Semantic Constraints on Inter-system Conversion

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

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

Compositionality Holds for both SPC & SCC

- **Property 4.1 (*Semantic Preserved Compositionality*)**: *Given two linking relations $\leq, \leq^*$. Given granularities G,H,I s.t. $G \leq H \leq' I$, then $Conv_{H \to G}(Conv_{I \to H}(I')_{\leq^*}, G)_{\leq} = Conv_{I \to G}(I')_{\leq^*}$ iff $\leq \to \leq^*$.*
  - **The conversion semantics on a path increases monotonously.**
- **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 $\leq = \leq^*$, $\leq \to \leq^*$ or $\leq^* \to \leq$ holds.*
  - **It exists an atom conversion whose semantics is the weakest**

Combinability: Can we combine two GSs?

Definition 4.3 (*Combinability*): *Two granularity systems can be combined to a single system iff*

1. *Any refine-conversion in the combined system is semantic preserved and/or semantic consistent.*
2. *For any pair of granularities, the GLB exists in the combined system.*

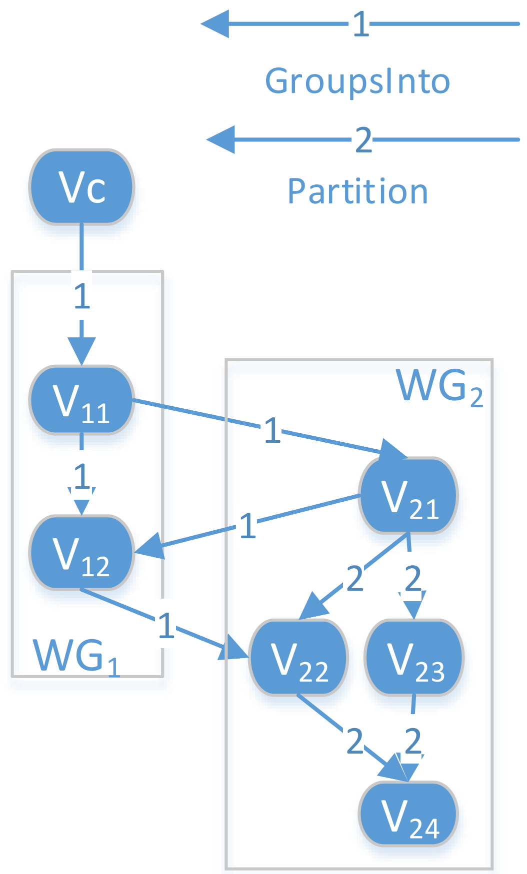
Req. 1: The S-N condition for supporting inter-system granularity conversions.

Req. 2: The S-N condition for granular comparison.

How to verify combinability?

- Semantic Preserved Combinability
  - GLB always exists + conversion is semantic preserved
- Semantic Consistent Combinability
  - GLB always exists + conversion is semantic consistant

- We proved the sufficient-necessary (S-N) conditions for both combinabities
  - Based on the relations between zero elements and granularity relations in involved GSs
  - O(1) space and time complexity



Combination Algorithms (see paper for details)

Two types of combination:

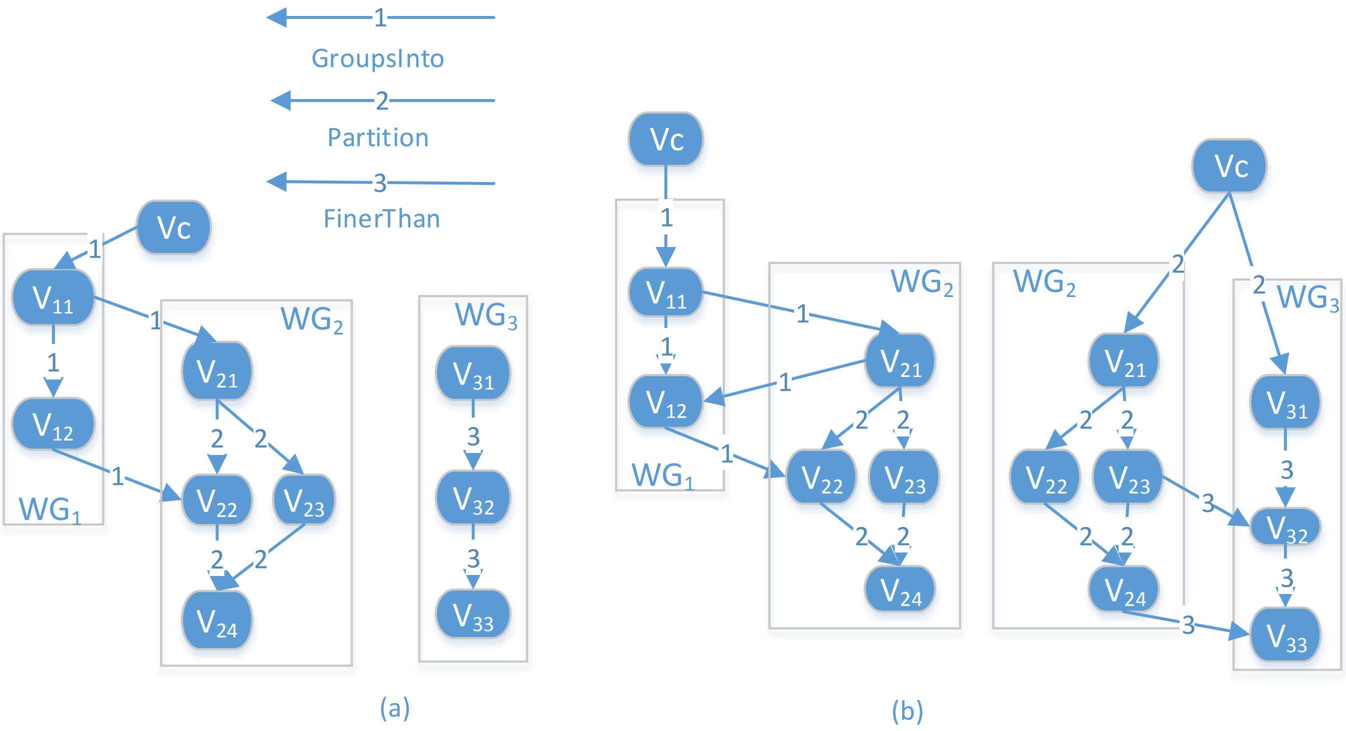
- Semantic preserved combination (SPC)
- Semantic consistent combination (SCC)

- Verification + combination: $O(n^3)$ time complexity
- $O(|\mathcal{E}_D| * |\{G\}|^2)$
  - $|\mathcal{E}_D|$: # systems on domain *D*
  - $|\{G\}|$: # granularities in each system

SPC Results

1. Result is still a lattice
2. Any path within the combined graph is semantic preserved
3. Any pair of granularities has a GLB
4. Edges are only created for atom relation (transitivity reduction)

- A similar SCCombine can be created for semantic consistent combination



Uncertainty Of Granularity Conversion

Uncertainty in granularity conversion that are not considered before:

- geometric distortion results from the incongruity of geometric properties among granularity relations
- statistic distortion results from the loss of data among granularities

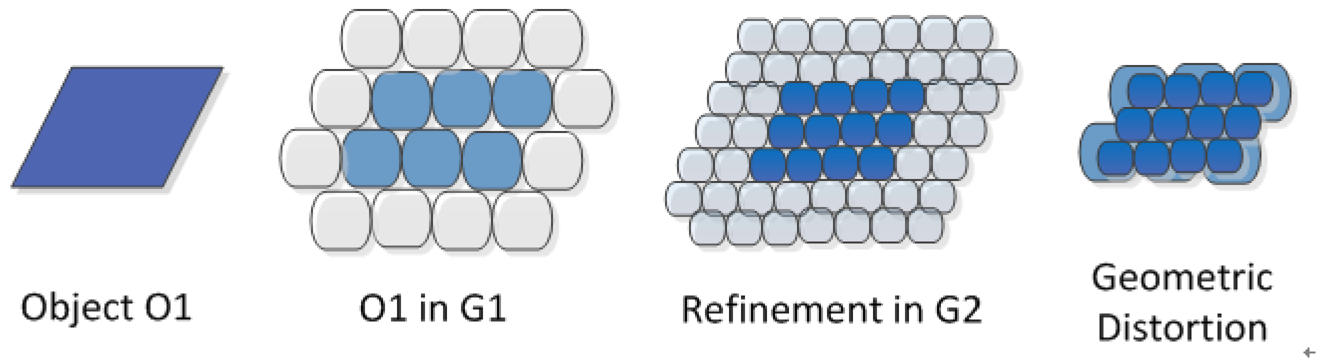


Fig. 4. An object projection in different granularities

Quantifying Uncertainty

- **Geometric precision:**

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

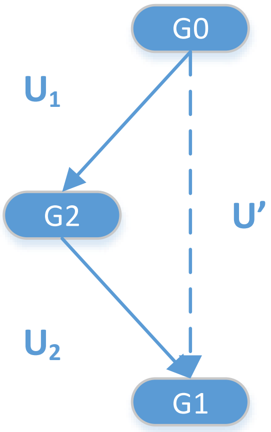
- **Statistic precision:**

$$\rho\left(C\right)=\frac{|\{e\,|\,e\epsilon E\wedge coveredBy\left(e,C\right)\}|}{C^{o}}$$

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

Properties of Uncertainty Quantification

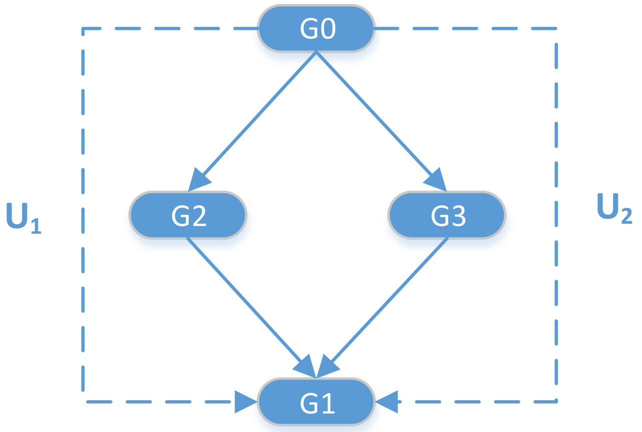
- **Property 5.1 *(Transitivity)***: *Given G,H,I s.t. $G \trianglelefteq H \trianglelefteq I$, $U(I,H) \cdot U(H,G) = U(I,G)$ and $U_\rho(I,H) \cdot U_\rho(H,G) = U_\rho(I,G)$ are always satisfied.*



$U_1 * U_2 = U'$

- **Property 5.2 *(Path-independence)***: *Given G,H,H',I, s.t. $G \trianglelefteq H \trianglelefteq I$, $G \trianglelefteq H' \trianglelefteq I$ and $H \neq H'$. $U(I,H) \cdot U(H,G) = U(I,H') \cdot U(H',G)$ and $U_\rho(I,H) \cdot U_\rho(H,G) = U_\rho(I,H') \cdot U_\rho(H',G)$ always hold.*

Applies to any conversion denoted by the directed paths in a combined granularity graph.



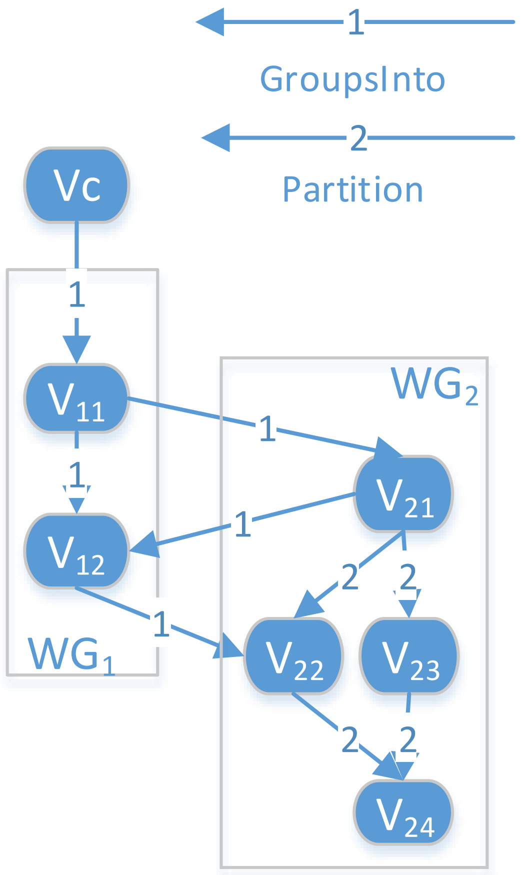
$U_1 = U_2$

The Optimal Lower Bound Problem

- To compare g ⊆ G and h ⊆ H, find the GLB with the highest expectation of precision. (i.e. $(U(G, I) \cdot U(H,I))^{1/2}$ is maximal.)

Reduce the Optimal Lower bound problem to the LCA problem on weighted DAG

O(n) solution (O(|{G}|))



The Optimal Common Refined Granularity (OCRG) Problem

Algorithm 5.1 *FindOCRG(u,v)*

let w[] **be** the cumulative gain on vertexes initialized as 0
DFSCumulate(u,w)
orcg⟵NIL
maxGain⟵0
DFSFind(v,e,ocrg,maxGain)
return (ocrg,maxGain)

Algorithm 5.2 *DFSCumulate(u,w[])*

if succ(v)=∅ **then return**
for each vϵsucc(u) **do**
    if w[v]=0 **then**
        w[v]⟵w[v]·W(E(u,v))
        DFSCumulate(v,w)

Algorithm 5.3 *DFSFind(v,w[],ocrg,maxGain)*

for each uϵsucc(v) **do**
    if w[u]=0 **then**
        w[u]⟵w[v]·W(E(v,u))
        DFSFind(u,w,ocrg,maxW)
    else totalGain⟵w[u]·w[v]
        if totalGain>maxGain **then**
            ocrg⟵u
            maxGain⟵totalGain

Remaining Discussion of the Paper

Optimization techniques:

- Using *Registration Matrix* to reduce the verification of granularity relations from $O(n^2)$ to $O(1)$
- Creating indices to reduce the operation of atomic conversion from $O(n)$ to $O(1)$

How our method may be applied to real-world applications:

- Unified spatio-temporal analysis
- Creating indices to reduce the operation of atomic conversion from $O(n)$ to $O(1)$

Reference

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

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

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

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

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

Reference

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

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

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)

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

S. Wang, D. Liu: Spatio-temporal Database with Multi-granularities. *WAIM'04*, pp.137-146. Springer-Verlag, Berlin Heidelberg (2004)

Reference

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

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

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

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

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

Reference

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

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

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

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

Thank You!