Event Objects for Spatial History

Karl E. Grossner

Department of Geography, University of California, Santa Barbara
Phelps Hall 3512 Santa Barbara, CA 93106-4060
Email: karlg@geog.ucsb.edu

1. Introduction

A growing number of historical scholars in social science and humanities fields are using GIS to investigate spatial questions (Gregory and Ell 2007). They require systems having greater compatibility with historiographic practices by means of temporal and semantic extensions to typical GIS data models. This research develops novel conceptual and logical models of historical knowledge representation for historical GIS (HGIS) databases, intended to support both standalone analytic applications and individually or collectively authored knowledge repositories such as the emerging genre of digital historical atlas.

Historian and political scientist William Sewell observed that societal structure is "a product of the events through which it has passed" (2005: 200), adding that "to understand and explain an event…is to specify what structural change it brings about, and to determine how [that change] was effectuated" (Ibid: 218). Geographer Doreen Massey has described place as "the meeting up of histories" (2005: 4), a poetic view that might be readily explored computationally given simple, robust event-centered data models.

Most HGIS projects to date have represented human activity in terms of *state*, i.e. conditions of some geo-referenced entities in one or several temporal snapshots. Events and historical processes are most often implicit: for example, population statistics imply birth and migration events and historical maps of transportation networks imply the various economic activities and processes related to their development.

The explicit representation of activity, events, and process—"the spatiotemporal constructs of geographic dynamics" (Yuan 2009)—has been a significant thread in GIScience research for two decades (*cf.* Hornsby and Yuan 2008). If GIScience is to discover generic principles of representing geographic dynamics, leading to systems built on general models such as proposed in the *GEM* framework (Worboys and Hornsby 2004), then intensive investigations into the dynamic aspects of natural and human phenomena for numerous domains of interest are required. The particular requirements concerning historical event data have not been addressed comprehensively and are the subject of this research.

Events are arguably the most comprehensive container for information about dynamic geo-historical phenomena. To describe an event well is to account for its purpose and results, its participant actors in roles for some interval, its location in space and time and its relations to other events. The representation of large numbers of events along those dimensions will enable faceted browsing and spatiotemporal analyses supporting the discovery of underlying social historical processes.

2. Historical cases

A general approach to geo-historical knowledge representation must meet several requirements beyond event-centeredness. All are considered in this work but a full discussion is not possible in this short paper.

Historical scholarship entails the examination, analysis and interpretation of multiple accounts of the same phenomena. The systems envisioned here will represent conflicting sets of assertions, to enable computational and visual (i.e. mapped and graphed) comparisons of their dynamic structure. Without venturing too deeply in philosophical waters, we can say it will represent multiple 'possible worlds,' in the sense that any database and its underlying ontology represent some (normally singular) view of reality. Any given study will encode those event attributes that model its author's hypotheses.

Historical data models must accommodate multi-valued logics, including probabilistic or fuzzy statements of "truth," and follow an open-world assumption (Sowa 2000) wherein null values are "unknown," not "false." Vague data is commonplace in historical studies so representing estimations of data quality will be essential.

A detailed examination of numerous print historical atlases and HGIS projects has revealed several broad classes of geo-historical phenomena that need to be modeled. Accommodating these types of data will be an effective measure of generality: i) complex events having sub-events in common, ii) theoretical processes, iii) individual space-time paths, iv) collective complex paths, and v) flows.

3. A simple conceptual model

The fundamental division between continuants and occurrents, argued convincingly in Simons (2000) and found in top-level ontologies such as DOLCE (Gangemi et al. 2002) is adopted here. Continuants include actors, information objects, geographic features and artifacts; occurrents include activity, event, process, state and period. Galton (2007) and Sowa (2000) have surveyed the varied definitions used for terms denoting basic occurrent types. As there is no empirical means for resolving definitional issues, individual researchers are left to construct useful arrangements, as I have undertaken here.

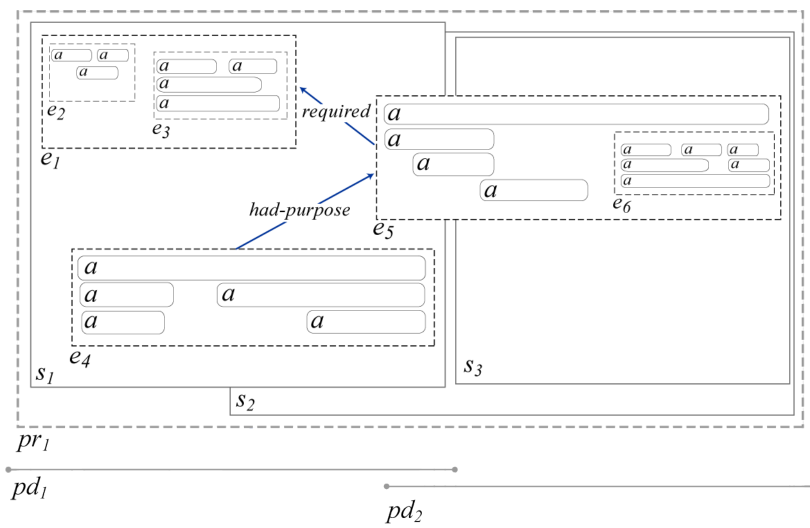

Figure 1. An asserted process pr_1 spanning periods pd_1, pd_2, having component activity instances a, events e_n, states s_n and the relations *required* (e_5, e_1); *had-purpose* (e_4, e_5).

In the simple model proposed here, events are composed of instances of activity (human or natural); compound 'macro-events' are composed of sub-event parts and in cases, non-event activity instances. Just as physical objects are composed of one or more material substance, and complex mental objects are composed of one or more concepts, discrete temporal objects are composed of activity. In this way activity can be considered 'temporal substance.' Sticks are made of wood; this walking stick, of some wood and some brass. A screenplay is composed of conceptualizations (more or less shared) of physical and mental objects stitched-together in asserted relations. A political rally might consist of instances of gathering and cheering activity and a speech event. The parts represented will depend on data resolution and the granularity of analysis.

More complex aggregations of activity instances and events, including *process* and *period* follow. For the domain of historical research, a process is usefully modeled as a 'theory of event relations,' not unlike the composite event in Worboys (2001: 134): instances of change bound "in some sort of unity or unifying principle." As depicted in Figure 1, these relations range from the topological (e_4 *overlaps-with* e_1), to the mereological (events e_2 and e_3 *part-of* e_1), to the telic (e_4 *had-purpose* e_5) or (e_5 *required* e_1). An historical period is roughly analogous to region in the spatial domain. Boundaries in either case may depend upon empirical criteria and be uncontroversial ('Reagan Era,' 'Lakes District'), or be subjective and contested ('The Enlightenment,' 'Palestine').

4. A logical model

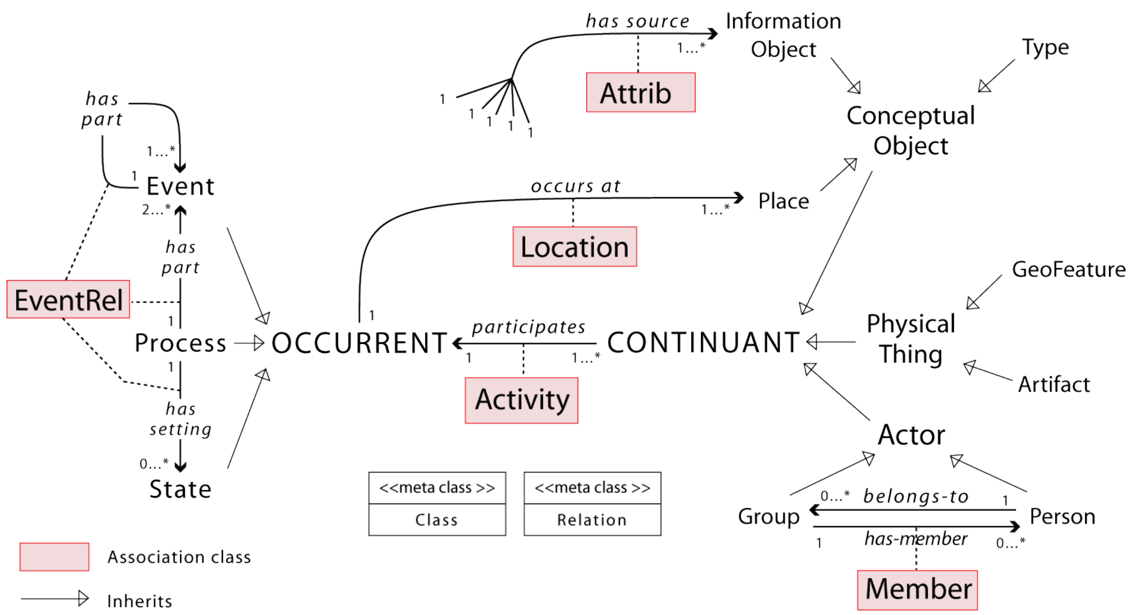

Figure 2. Principal entity classes and logical relationships constituting an ontology for spatial history

An event-centered 'information ontology'[1] for spatial history has been developed, and instantiated in a particular object-relational spatial database. Its principal logical relationships appear in the 'UML-like' diagram of Figure 2. Taxonomies of entity

[1] Use of this term follows distinctions drawn in (Frank 2003; Frank 2007)

classes and relationships can be stored in [*Class*] and [*Relation*] tables with constraining *has-domain* and *has-range* properties.
In this model, historical knowledge is held both in the definitional 'TBox' identity and compositional hierarchies (*is-a* and *part-of*, respectively), and in the 'ABox' of instances, their attributes and asserted relationships. For relational databases, a majority of instance data will be stored in association tables corresponding to core relations:

[Activity]:	has-participant (occurrent, continuant)
[Member]:	has-member (group, person)
[EventRel]:	has-influence (occurrent, (occurrent ∨ continuant))
	has-setting (process, state)
	has-part (event ∨ process, event)
	has-duration (period, interval)
	has-type (occurrent, type)
[Attrib]:	has-source ((eventrel ∨ activity ∨ member ∨ location), infobject)
[Location]:	has-location (occurrent, place)

The relational algebra of SQL can provide considerable inference capability for data in a RDBMS, given this logical structure and such features of modern systems as check constraints, triggers, stored procedures, composite data types and custom functions. A partial list of these capabilities includes class and relation propagation (*is-a*, *subclass-of*); reification (association classes); instance checking (SELECT…WHERE…AND…); domain and range (check constraints); property intersection (A *subproperty-of* B; A *subproperty-of* C → if x A y, then x B y, x C y); restrictions, e.g. all|some values from; has [default] value (SELECT WHERE…); cardinality (check constraints); unions and intersections (SQL's UNION and INTERSECT).

5. An exemplar database

Several historical datasets meeting the criteria mentioned in §2 have been stored in a PostgreSQL/PostGIS[2] object-relational spatial database. One of these, comprising detailed information about 8,088 'contentious gatherings' in Great Britain between 1758 and 1834, was originally compiled by Tilly (1995) from contemporaneous written accounts in several British periodicals. Component actions, participants and locations were encoded according to hypotheses and methods particular to the investigator and the data. Such project-specific entities as 'machine breaking' Activity, 'brawl' Events, 'blacklegs' Groups and a 'mobilization' Process are assigned a *type_id* from a project-specific taxonomy added to the ontology as subclasses of a Type class[3]. Types specialize the more general classes (e.g. Violence, Gathering, Delegation) designated by a *class_id* field. Such type ontology sub-trees are also identified as belonging to namespaces, a Semantic Web convention that will enable the filtering of 'possible worlds.'

A great many operations are made possible by holding combinations of occurrent facets constant, including spatial and temporal regions/intervals, and all attributes of occurrent participants or results, including class and type. Selected query functions appear in Table 1.

---

[2] An open-source RDBMS and spatial extensions; http://www.postgresql.org; http://postgis.refractions.net/
[3] A technique borrowed from the CIDOC-CRM (Crofts, et al 2008) ontology.

Table 1. Selected queries and corresponding database functions

Query	Function
what's happened at a place [time]	occurrence(place [, period])
who participated	participants(occurrence [(place, period)])
what occurrences, artifacts, infobjects resulted	results(occurrence [(place, period)])
what's happened in a life	activity(actor, lifespan(actor))
a person's life-path	location(lifetime) ≡ trajectory(actor)
event location (union of sub-event locations)	location(occurrence)
nature of the 'goings on'	activityType(occurrence)
what kind of person/group did what	activityType(actorType(occurrence))
eventType density during period	density_s(eventType(period))
eventType density in region	density_t(eventType(region)

6. Future work

The next stages of this work include a) formal specification of the logical model; b) testing the models' completeness with additional examples, per §2; and c) collaborative work to ensure the developing models and schema fulfill the requirements for the emerging digital historical atlas genre.

References

Crofts, N., M. Doerr, T. Gill, S. Stead, M. Stiff (Eds.) (2008). *Definition of the CIDOC Conceptual Reference Model*. Retrieved from http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.4_March2008_.pdf, 4 April 2008.

Frank, A. (2007). Data Quality Ontology: An Ontology for Imperfect Knowledge. In S.Winter et al. (Eds.): *COSIT 2007*, LNCS 4736, 406-420, Berlin: Springer-Verlag

Frank, A. (2003) Ontology for spatio-temporal databases. In M. Koubarakis (Ed.), *Spatiotemporal Databases: The Chorochronos Approach*. Lecture Notes in Computer Science 2520, pp. 9-77, Springer, Berlin.

Galton, A. (2007) Experience and History: Processes and their Relation to Events. *Journal of Logic and Computation*, 18(3):323-340

Gangemi, A., N. Guarino, C. Masolo, A. Oltramari and L. Schneider (2002). Sweetening Ontologies with DOLCE. *Lecture Notes in Artificial Intelligence*, 2473: 166-181.

Gregory, I. N. and Ell, P. S. (2007) *Historical GIS: Technologies, Methodology and Scholarship*. Cambridge, UK: Cambridge University Press

Hornsby, K. S. and M. Yuan (Eds.). (2008). *Understanding Dynamics of Geographic Domains*. Boca Raton, FL: CRC Press

Massey, D. (2005). *For space*. London: Sage

Sewell, W. H. (2005). *Logics of History*. Chicago: University of Chicago Press.

Sowa, J.F. (2000). *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Pacific Grove, CA: Brooks/Cole.

Simons, P. (2000). *Parts: A Study in Ontology*. Oxford: Oxford University Press.

Tilly, C. (1995). *Popular Contention in Great Britain, 1758-1934*. Cambridge MA: Harvard University Press

Worboys, M.F. (2001). Modelling Changes and Events in Dynamic Spatial Systems With Reference to Socio-Economic Units. In A. Frank, J. Raper and J. Cheylan (Eds.) *Life and Motion of Socio-economic Units* (pp. 129-37). London: Taylor & Francis.

Worboys, M., & K. Hornsby (2004). From objects to events: GEM, the geospatial event model. In Egenhofer, M. J.; Freksa, C.; Miller, H. J. (Eds.) *Geographic Information Science: Third International Conference*, *LNCS* 3234: 327–344

Yuan, M. (2009). Toward Knowledge Discovery about Geographic Dynamics in Spatiotemporal Databases, In H.J. Miller, J. Han (Eds.), *Geographic Data Mining and Knowledge Discovery*, Boca Raton, FL: CRC Press