The Encyclopedia Gallica of Events (or Why Geographic Information Science is Not Like Physics)

H. Couclelis

University of California, Santa Barbara
Santa Barbara, CA 93106-4060, USA
Email: cook@geog.ucsb.edu

Abstract

As a contribution towards the further development of user-centered information systems, I present an argument for a contextual and subjective view of events and related concepts in information science that is distinct from the factual view prevalent in empirical science and everyday life. The central notion is that of 'R-Event', where 'R' stands for 'Relevant'. Drawing on representations of process such as dynamic models or AI approaches to problem-solving and planning, R-Events would improve the precision and value of search results by foregrounding and ranking information about events of high relevance to the user. There is no suggestion here that this proposal is easily implementable. It is offered at this stage as a potentially fruitful thought experiment.

1. Introduction

The title is one that Borges would have loved[1]. Had he written the story of the Encyclopedia Gallica of Events, this would have also included processes, perdurants, occurrences, and non-events. And the Encyclopedia in question would have had to be infinite in length, and self-contradictory. I will argue that in an *informational* (as opposed to an empirical) context, (a), the above concepts do not refer to intersubjectively definable individuals and (b), unlike most other abstract concepts, the things these refer to are not free-standing notions but binary relations connecting information source and observer, and depending on the specific interest the latter has in certain detectable state(s) or change(s) in the world. Thus the news of my birth signaled a momentous event for my family (and a very real painful process for my mother) – but for the rest of humanity the fact of my birth was a non-event, barely registering as a blip of plus-one added to some country's population. Similarly, major flooding in south-eastern Tajikistan may be an empirical fact, but it is an event if you are in the business of importing cotton from that area, a process if you are in charge of evacuating local populations, an occurrence if you are tabulating floods in Central Asia, and noise if south-eastern Tajikistan is not among the things you care about. Similar kinds of context-dependency also hold for the endurant-perdurant distinction, and as has been argued repeatedly over the years, also in the case of fields and objects. Information systems sensitive to user-centered semantics should be able to make these determinations.

2. On Science and Metascience

It may be useful to distinguish clearly between the empirical sciences that directly measure and represent phenomena in the world, and the information sciences (which are meta-sciences) that process and present information about these phenomena in ways that meet and support the interests and purposes of information users. These are really two different epistemic layers, with different

[1] See Borges (1939) The Total Library; Borges (1943) The Library of Babel; and the "certain Chinese Encyclopedia" in Borges (1952) The Analytical Language of John Wilkins.

functions. I will argue for the importance of not conflating the two, because as information scientists we are not *doing* hydrology or *doing* forestry or *doing* urban studies, but trying to help answer questions posed by hydrologists, foresters, planners, and any others, in the most appropriate and helpful ways. There is a difference between knowledge representation on the one hand and information (re)presentation on the other, though the latter must of course draw on the former. It is no wonder that some of the smartest people in our field cannot come to an agreement on the definitions of events, processes, and so on. This is because 'it depends'. The challenge as I see it is to take data and transform them into information that is right for particular users coming to these data from particular angles.

Presenting information appropriately forces us to confront and combine three key notions: first, the changing states of the environment as measured by simple, non-cognitive sensors (the 'pre-facts'); second, the intentional users' contextual interest (or lack thereof) in particular states and changes; and third, an information science's duty to help answer questions about facts in ways appropriate for specific kinds of inquiries by these users. To fix ideas, let us consider what may fall into each of these three categories. First are the uninterpreted, uncritical raw measurements provided by sensors: temperature goes up, temperature goes down; one wave-length is sensed, another wave-length is sensed; this many cars trip the counter at time t, this many cars trip the counter at $t+n$. Second are the myriad changes in the world reflected in data at any instant, and the need to sift among them in search of relevance. Even within specific domains of interest, whether and which changes are important when is a subjective and contextual judgement. So – and third – here is the problem: how can an information modelling system be made to take data from the first group, and process them as information within appropriate categories in the second group, in a way that is logically consistent, semantically meaningful, and relevant to the user?

Galton (2015, p.7) wrote "…it is, presumably, the responsibility of the latter [the data modelling system] to extract from this processual flux those hard nuggets of salience which constitute events, and which from a human perspective represent information rather than mere data." Indeed -with the caveat that one person's hard nuggets of salience may be another person's fool's gold. In this vein, following Bateson, I will define an event as 'a difference that makes a difference'. For clarity, within an informational context I will talk about *R-Events* (where R stands for 'relevant'), to distinguish from empirical facts such as landslides or riots, where talk of events, processes, etc. as applied to specific instances is necessary for communication. R-Events are thus Galton's "hard nuggets of salience" tailored to the interests of specific observers. They emerge from (empirical) events that are likely to be significant in particular situations of interest. R-Events increase the precision of search results by picking out information of high value to the user and possibly also ranking it by significance. Thus our cotton importer would want to know right away about the SE Tajikistan flood because of likely impacts on the infrastructure and human resources of the commodity movement chain: critical roads closed, warehouses flooded, workers unable to reach loading sites. Other headlines from the region, e.g. about a toxic spill, a plane crash, or the election of a new governor would be of no direct import and are better left unstated.

3. Speculations

Elsewhere I proposed a sketch of an information system ontology that could perhaps help move such an agenda forward. It consists of three modules or interrelated parts, named after the well-known triad from linguistics: syntax, semantics, and pragmatics. This is more than mere analogy, since an information system must produce meaningful and appropriate statements about facts.

- The *Pragmatics* module reflects the *context* of the inquiry: that is, anything that shapes and constrains user interests and their appropriate mode of satisfaction. The context may be professional, educational, social, etc. and may be constrained by time, space, budget, data, etc.
- The *Syntactics* module handles the sensor and other data (pre-interpretation) and the data models (post-interpretation), following input from the *Semantics* module, and within the constraints flagged by the *Pragmatics* module.
- The *Semantics* module is the core of the system. It receives input from the *Pragmatics* module in the form of a specification of user interests and constraints, and consists of something like the 7-tier structure I proposed some years ago, which is traversed starting with the purpose of the inquiry and moving down all the way to the specification of the most appropriate spatiotemporal frame for results presentation and analysis (Couclelis 2010). At each step the *Semantics* module queries the *Syntactics* module for data relevant to the corresponding level.

As discussed earlier, there are two distinct epistemic layers, the informational and the empirical. The above three modules taken together correspond to the informational layer, while the empirical layer consists of some dynamic representation of the domain of interest. To identify R-Events, the *Semantics* and *Syntactics* modules search the empirical representation. Depending on the domain and the user's intention, as specified by the *Pragmatics* module, this could be a scientific model, a plan, or some other quantifiable pattern of activity at any appropriate level of detail.

In the next section I outline how *plans* may be used to identify changes in a domain of interest that give rise to R-Events. Often a user's intentions are actualized in a plan that may be simple or complex, implicit or explicit. Plans have been studied extensively in AI (Russel and Norvig 2009). They include goals, states, and actions, the latter meant to move states towards the plan's goals. Actions imply agents, which need not be sentient. Plans are often hierarchical, composed of goals, subgoals, targets, etc. as well as activities, tasks, subtasks, etc. In automated planning these are represented in the approach known as Hierarchical Task Networks (HTN).

In connection with plans, R-Events may be defined as changes in the world that cause a breakdown of a plan or its parts. The context may change, causing the interests of the user to change focus; the means to desired goals may change, forcing changes in strategy; the availability of relevant data may change, expanding or restricting the kinds of questions that may be asked and answered. The following kinds of changes are among those likely to give rise to R-events: those preventing planned tasks or routines to be completed; those forcing a change in strategy; those affecting the value to the user of the plan's goal itself; those necessitating a new plan and goals.

4. For Example

A sketchy illustration of the concept of R-Event in a spatial context is provided by the predicament of our hypothetical cotton importer. A successful import operation depends on the timely arrival of shipments from SE Tajikistan to the designated warehouse in the USA. There are two pieces to this operation: (a) plan of action, and (b) a corresponding spatial organization that enables and constrains the elements of that plan. The plan of action may be represented as a network the nodes that correspond to the hierarchy of goals and subgoals in the vertical dimension, while in the horizontal dimension they delimit vectors indicating sequences of activities. Each such sequence supports the goal above it. Thus both goals and activities are being systematically disaggregated as one moves down the hierarchy. In our example the network is headed by a top-level goal, to run a successful import business. Its success depends on everything functioning properly, and in particular, the two broad areas of activity: producing the cotton, and shipping it from source area to destination. Qua subgoals each of these in turn depend on further subgoals and sequences of

activities: picking, sorting, packing, loading cotton, taking it to collection points, trans-shipping it, etc. And further down: getting workers to their workplaces, ensuring appropriate supplies and transport means are available, and so on.

These activities obviously depend on a spatial organization that must function as intended. The producing areas must be physically capable of producing, and the infrastructure must be able to handle the movement of the product from source areas to destination; fields, roads, ports, and warehouses must be in an appropriate spatial configuration; the elements of that configuration must be operational and accessible as needed: e.g., all the segments on shipment routes must be traversable and all necessary warehouses and loading bays must be whole. There is thus a hierarchy of spatial granularity from the global to the very local that reflects and complements that of goals and subgoals. That is, each level of spatial features enables a corresponding level of activities, sub-activities and tasks (Howarth 2008).

Any disruption in the routines described by the plan is a potential R-Event. News of a flood in the region of interest would trigger a manual or automatic search for problems. The *Semantics* module of the information system will check for data in the *Syntactics* module in a sequence roughly as described in Couclelis (2010), which is compatible with the organization of the plan. The sequence is follows: (1) user purpose (as per the *Pragmatics* module), (2) overall function of the intended operation, (3) spatial organization supporting that function, (4) individual elements of the spatial organization, (5) relevant data to be presented to the user, and (6) appropriate spatiotemporal frame for data presentation. R-Events can appear at any level and at several different locations on the hierarchical network representing the plan. Usually the importance of R-Events diminishes as we move down the structure towards more detailed tasks and smaller areas, thus permitting the selected relevant information to also be ranked by degree of criticality.

5. Conclusion

As a contribution towards the further development of user-centered geographic information systems, this paper proposes the following points for discussion: (a), the notion of R-Event as a means of improving the precision of search results as well as the possibility of ranking these by their value to specific users; (b), the further integration of GI science with AI, in connection with the representation of user purposes and the conditions of their satisfaction; and (c), underlying (a) and (b), the potential practical utility of treating the empirical and informational aspects in information systems as distinct epistemic layers. Next steps should include refining the framework using realistic examples, and enriching it with the extensive literature on related topics.

References

Couclelis H, 2010, Ontologies of Geographic Information. *International Journal of Geographical Information Science*, 24(12):1785-1809.

Galton A, 2015, Outline of a formal theory of processes and events, and why GIScience needs one. In: Fabrikant S**,** Raubal M, Bertolotto M, Davies C**,** Freundschuh S, and Bell **S.** (eds.) *Proceedings of the 12th International Conference, COSIT 2015*, Santa Fe, NM, USA, 3-22.

Howarth J, 2008, Landscape and Purpose: Modeling the Functional and Spatial Organization of the Land. Doctoral dissertation, University of California, Santa Barbara.

Russell S and Norvig P, 2009, *Artificial Intelligence: A Modern Approach*. 3d edition, Prentice Hall, Upper Saddle River, NJ, USA