The Search for Places as Emergent Aggregates

ANDREA BALLATORE
Postdoctoral Researcher
Center for Spatial Studies
University of California, Santa Barbara
Email: aballatore@spatial.ucsb.edu

Searching for places is the most popular geographic online task, in which names and categories are used as the main referents to locate places in the geographic space. By typing "hotels in Santa Barbara" in any popular search engine, the user expects a list of places matching a category ("hotel") contained within another place called "Santa Barbara." Answers to such a query can be generated by relying on a gazetteer containing some form of spatial footprint for the symbol "Santa Barbara," a database of points-of-interest categorized as "hotels" and, indeed, some strategy to compute the relevance of the potential results. This approach satisfies well-defined information needs, but fails to account for more complex, nuanced, fuzzy, and yet cognitively intuitive questions about the town. What places are similar to Santa Barbara with respect to its general atmosphere—but perhaps less expensive? What other towns in Southern California offer a comparable array of amenities? What tourist areas in Italy provide a similar combination of mountain-related and marine activities? Our current computational models of place search do not seem to provide easy answers.

Our intimate familiarity with place clashes with the difficulty of dealing with it computationally. Because of its centrality in human cognition and culture, the notion of place is unsurprisingly characterized by high polysemy, strong context-dependence, and innumerable metaphorical uses, carving social meanings from neutral, unbounded spaces (Agnew, 2011). The intellectual prominence of place has waxed and waned over time, being obscured for centuries by more abstract notions of space, and making a reappearance in recent decades (Casey, 1997). Being intensely debated in the social sciences and the humanities, place and its representations have now become an active research frontier in geographic information science (Goodchild, 2011). In this area, a central concern is that place names are often ambiguous, vague, and vernacular. Place categories are culturally-dependent, arbitrary, and inconsistently applied. More strikingly, places are implicitly assumed to have a name and to fit, at least to some degree, known categories. Current efforts focus on more sophisticated place-name interpretation in text documents (Purves, and Jones, 2011), new reference theories tailored to place (Scheider and Janowicz, 2014), and semantically more expressive gazetteers (Keßler et al., 2009).

In a complementary approach, I advocate a view of *place as an aggregate* of objects and processes that interact at a given scale, inter-locked by spatial colocation. This view of place relies on the discovery of *implicit relations*, and not on some explicit labels assigned by an observer. The approach relies on some assumptions: Places are inescapably multi-faceted (comprising diverse processes), they are socially constructed (emerging as the result of human agency and practices), relational (emerging in a context, not in a vacuum), scale-dependent (different places exist at

different scales), and they are dynamic (emerging, changing, and ultimately disappearing). Following Thrift (1999), I regard places as emergent entities in a complex, non-linear system, and they appear as assemblages of heterogeneous things that meet in space and time. Place can be fruitfully viewed through a holistic lens, emphasizing its contextuality and inherent interconnectedness, rather than as an object in isolation (Ballatore et al., 2012). In practical terms, this approach aims at supporting multi-faceted, context-dependent aggregate search, going beyond the current forms of name-based search for well-defined, individual places. In this sense, places can be searched for on the basis of their emergent distributional characteristics, rather than in an arbitrary, crisp categorization (e.g., *city* or *town*). For example, a Japanese tourist in San Francisco might search for places in which architectural landmarks co-occur with museums and galleries, or for places that, as an aggregate, present similar characteristics to Shibuya, the Tokyo shopping district she is familiar with.

	Text Information Retrieval	Place-as-aggregate Search
Vector space	Set of text documents	Geographic space
Vector	A document as a sequence of words	A place as an aggregate of spatially located objects
Dimensions	Words (high dimensionality)	Characteristics of objects (dimensionality defined by application, potentially very high)
Index	Sparse document-word matrix	Sparse object-to-object colocation matrix
Search	Weighted keyword matching, topic models, similarity	Colocation queries, query-by-place

Table 1: Place search overview

Computationally, this approach to place search can be modeled in analogy to traditional information retrieval in a vector space model, as summarized in Table 1. Given a geographic space, treated as a corpus containing a potentially infinite number of place-as-aggregate, the proposed approach has a number of opportunities and challenges to be tackled. The efficient computation of spatial colocation at a large scale—the identification of categories of objects that co-occur spatially (and temporally) in non-random patterns—is an open problem (Cromley et al., 2014). As many alternative places-as-aggregates can encompass the same entities at different scales, new heuristics are needed to construct optimal aggregates that meet user informational needs at a given scale, based on statistical measures of informational entropy. As I hope I have demonstrated, the uneasy relationship between space and place offers opportunities for unlocking new ways of searching the ocean of geo-information.

References

Agnew, J. (2011). Space and Place. In J. Agnew & D. Livingstone (Eds.), *Handbook of Geographical Knowledge* (pp. 316–330). London: Sage Publications.

Casey, E. S. (1997). *The fate of place: A philosophical history.* Berkeley, CA: University of California Press.

Cromley, R. G., Hanink, D. M., & Bentley, G. C. (2014). Geographically Weighted Colocation Quotients: Specification and Application. *The Professional Geographer* 66(1): 138–148.

Goodchild, M. F. (2011). *Formalizing Place in Geographic Information Systems*. In L. M. Burton, S. A. Matthews, M. Leung, S. P. Kemp, & D. T. Takeuchi (Eds.), *Communities, Neighborhoods, and Health* (pp. 21–33). New York: Springer.

Keßler, C., Janowicz, K., & Bishr, M. (2009). An Agenda for the Next Generation Gazetteer: Geographic Information Contribution and Retrieval. In *Proceedings of ACM GIS '09* (pp. 91–100). New York: ACM.

Purves, R., & Jones, C. (2011). Geographic information retrieval. *SIGSPATIAL Special* 3(2): 2–4.

Ballatore, A., Wilson, D. C., & Bertolotto, M. (2012). A holistic semantic similarity measure for viewports in interactive maps. In S. Di Martino, A. Peron, & T. Tezuka (Eds.), *Web and Wireless Geographical Information Systems* (pp. 151–166). Berlin: Springer.

Scheider, S., & Janowicz, K. (2014). Place reference systems. *Applied Ontology 9*: 97–127.

Thrift, N. (1999). Steps to an Ecology of Place. In D. Massey, J. Allen, & P. Sarre (Eds.), *Human Geography Today* (pp. 295–322). Cambridge, UK: Polity Press.