Notes on Accessibility
In Search of a Unified Theory

Invited
Paper 2

Michael Batty
University College London
m.batty@ucl.ac.uk

Accessibility is a concept that has become central to physical planning during the last fifty years. My efforts to trace its origins reveal that the term was first occasionally used in the 1920s in location theory and regional planning, becoming important once transport planning began, mainly in North America where it came to be associated with transport networks and trip distribution patterns (Mitchell and Rapkin, 1954). By the time Hansen (1959) wrote his classic paper "How accessibility shapes land use", the term was being used casually but routinely for measures of the relative nearness or proximity of one place or person to all other places and persons, based on the notion of potential in physics, which had been introduced into geographic systems by Stewart (1948).

A well defined form of this index associates some measure of an opportunity at a place with the cost of actually realising that opportunity; in its early form, increased accessibility of a place with respect to some fixed location was assumed to vary directly with a measure of the size of the opportunity at some other place or location and to vary inversely with the distance or time taken to access the opportunities between the two places in question. This is a measure of inter-zonal accessibility from one place to another, but the usual form is to produce a composite index of total accessibility from one place or zone to all others, which gives a measure of how easy or difficult it is to realise all these opportunities from the place (zone) in question. The measure itself does not usually have any absolute meaning in terms of costs or benefits in monetary or activity values and thus it is often normalised over a certain range and interpreted in purely relative terms. This is because the quantities used are dimensionally inconsistent, or at least absolutely different from one another; the classic example is 'population potential' at a specific location which is defined as the sum of the populations of each place divided by the distances to each of those places from the location in question.

There have been many variants of this measure using different variables for spatial separation, distance, travel time, composite travel cost, and so on, as well as intervening opportunities, that is, the number of opportunities encountered as the traveller moves from the location in question to the destination place at which the ultimate set of opportunities are located. In this sense, the link to physical distance is broken for the measure is then simply some index of impedance between one place and another, which may have little or nothing to do with the actual measure of physical distance. In fact, accessibility is often seen as a measure of the cost of getting from one place to another, traded off against the benefits received once the place is reached. In this sense it may simply be a composite benefit-cost ratio for all places or a composite benefit-cost difference. The measure may then have some actual meaning in monetary terms, especially if it is computed as a difference, where costs and benefits are measured in comparable and absolute units. In a residential context, a good example might be where benefits are measured in terms of wages earned at a destination less the costs of getting to that destination, possibly meshed with the costs of living at the origin location. In this way, such measures can be linked to the behaviour and budgets of individuals and households in the urban economy. A further extension of these ideas involving consumer surplus can be consistently derived from more formal models of spatial interaction, on the basis of discrete-choice theory.

There are two key issues that make this discussion of accessibility problematic; the first is the scale at which such measures are defined, and the second is the difference between measuring the accessibility of how individuals behave with respect to economic and social activities and how individuals implicitly react to physical infrastructure. So far, the definition of accessibility from transportation pertains to relatively large scales at which clusters and agglomerations of activity are meaningful with the spatial impedance measure largely related to how individuals move and interact between places. In contrast, accessibility can be defined at the level of physical infrastructure measured in terms of how close one piece of infrastructure is to another, and, in this context, the role of individual behaviour is at best implicit. Physical infrastructure is often conceived in terms of networks of streets and related routes and clusters of land parcels or even buildings. There is of course a merging across scales and also from activities to infrastructure, in that movements, interactions, and activities ultimately take place using some physical plant.

We call this generic concept of accessibility which pertains to locational behaviour where physical infrastructure is implicit and the definition is one of how proximate or 'near' an individual is to "opportunities" type 1 accessibility. Another term for it might be geographical accessibility. We can demonstrate a second type of accessibility quite easily from this first type if we assume that opportunities are the same everywhere and the focus is purely on the distance between one location and another. If we now define accessibility, or rather inaccessibility, as the total distance from one place to all others, then we can measure this as the sum of the shortest routes in a planar graph connecting the location in question to all other locations. The locations or zones in question could be nodes at which physical segments defining transport routes intersect, as in a street network. Total distance from one node to all others is a measure of inaccessibility; it might be computed as an average if normalised by all nodes in question and it might be converted to a measure of accessibility if the reciprocal of this total or average is used. We call this kind of measure type 2 accessibility which in contrast to type 1 is a geometric accessibility. Its principal characteristic is that it is measured on physical infrastructure but as it is still a location measure where the nodes are points in a network, then it can still be associated with how individuals relate to one another. It still has a behavioural interpretation; in contrast to type 1, it is directly measured with respect to the physical infrastructure used for travel, whereas type 1 may not be, and in this sense, it is a geometric not a geographic accessibility. This is a measure that was also first developed in the 1950s and 1960s, as graph theory came to be mainly applied to urban and building layouts, and in this sense, its scale of application has always been a little finer than type 1 accessibility, although there are examples where it has been applied to larger scales (Garrison, 1960).

The key distinction between these two types of accessibility is thus with respect to the relative focus on behaviour and on physical infrastructure, on activity patterns in contrast to networks, on geography in contrast to geometry. Type 1 tends to abstract the spatial system to a set of points whereas type 2 focuses more directly on the way in which these points are physically connected. In the latter, the links defining the underlying network usually have a measure of distance or travel time associated with them, but they may not, and the presence or absence of a link may be a sufficient attribute of significance in defining the measure. Context is thus all important. In the last twenty years, a third type of accessibility has emerged which is even more abstract, and which extends the physical definition on the basis of networks in order to define relationships between the physical components of the underlying network itself. This we will call type 3 accessibility and in some respects it is a dual of type 2 in a non-strict way. It is also a geometric measure, defined from the shortest routes in a network which result from the underlying planar graph of physical connections, where line segments defining the arcs of the planar graph are connected to one another if they intersect. In cases in which the network is truly planar, such a dual network is a mirror of the planar graph and little else is added, except that the focus changes to accessibility measures defined on the links of the original graph, not the nodes. If, for example, the street is the locational object in question, then this dual formulation simply gives the accessibility of a street rather than that of a node. There are problems of associating a street with a precise location, but in terms of the focus of interest, there is little doubt that streets and similar objects that are reflected in the arcs of a planar graph do have intrinsic significance for understanding proximity.

Type 3 measures only come into their own however when the underlying physical network is more than a set of line segments and arcs arranged in planar graph form. If the physical components in this network have one or more line segments, then the physical graph is no longer planar with respect to the formation of its dual. One long street, for example, may be made up of many segments which intersect once with each smaller street along its length. In this case the street in question would have many links to the smaller streets but each smaller street would have only one link to the longer street in question. This asymmetry enriches the dual network and quite substantial variations in accessibility between clusters of segments can emerge. This is the method used in space syntax (Hillier and Hanson, 1984), where the focus is on links, usually streets, of varying importance with respect to attributes associated with arcs in a planar graph that have their own integrity beyond their strict definitions as single segments between nodes. Such integrity might be based on lines of sight or on lines of unobstructed movement, but also on environments that are defined collectively as linear places. Usually in space syntax, a unit distance is associated with a link between one street and another, and thus the inaccessibility between streets is measured as a count of links that need to be traversed to move from one street to another. In this sense, the underlying graph of network links between streets is topological in that distance has not meaning and arcs exist only to define presence. There is some discussion of what such measures mean in space syntax with the integration measure being defined as the reciprocal of accessibility and variable measures of integration being defined in terms of the varying number of links at different depths from each street in question. In essence however, this type 3 accessibility measures the relative nearness of one linear physical component, such as a street, to all others.

As we define increasingly more abstract measures from the physical-spatial nexus of the system of interest, it becomes increasingly difficult to associate observable quantities with the measures in question. For example, defining a measure of physical distance between two streets, where the streets themselves reflect their own distances, requires some arbitrary definition of the street as a point. A related problem occurs with any quantity associated with a street where that quantity varies along the street, such as traffic flow. This makes it exceedingly difficult to relate type 3 measures of accessibility to point locations at which all measurement in cities and buildings takes place. As an abstract idea of how a composite physical object, such as a street, relates to another, it is important to have measures of comparability between these objects regardless of whether or not observable measures can be tagged to them. This is because the objects in question are derived from more basic physical components which tend to be less ambiguous in their definition and observation. Type 3 accessibility is thus hard to associate with traffic flow and movement, whereas type 2 is more directly applicable. Type 1 is consistent with traffic-flow theory in that flows are between point locations and not necessarily associated with any underlying physical transport. Indeed, there is a tight coupling between measures of accessibility of this type and theories of spatial interaction and traffic movement, where these accessibility measures appear as components within the models used to predict such flows.

These differences in types of accessibility simply scratch the surface of a myriad of measures that can be defined to measure nearness. Different definitions of the underlying physical system, the flow system that defines the way people interact, attributes that are associated with the determinants of travel, such as distance, costs, and so on – all these issues need to be tied together in a more unified theory that would let different professional and disciplinary perspectives talk with one another. Currently there is considerable confusion about the way that the physical structure relates to human behaviour. Moreover there are countless distance and spatial association measures that merge into questions of accessibility, on the basis of nearest neighbours, second-nearest neighbours, and so on. Indeed, spatial autocorrelation measures can in some circumstances be seen as measures of accessibility. There has been considerable progress in the last decade with respect to spatial cognition and much of this material might now be synthesised with ideas about accessibility (Montello and Freundschuh, 2006). There is a new concern for questions of mobility which relate to access to resources, while notions about accessibility in time as well as space are beginning to generate substantial interest (Miller, 1991), linking earlier ideas about space-time patterns to the burgeoning information available from local tracking.

A unified theory is urgently required. When this paper is presented, I will develop a number of examples of each of these measures that will make their relationship across the spectrum of possible measures from geography to geometry and from social and economic to physical somewhat clearer in visual terms. These is much to do to iron out ambiguities between these different perspectives and to resolve questions of what observable variables can be tied to these measures when they are used in attempting to explain how city and regional systems function in general.

References

Garrison, W. L. 1960. Connectivity of the interstate highway system. *Papers and Proceedings of the Regional Science Association*, 6: 121-137.

Hansen, W. G. 1959. How accessibility shapes land use. *Journal of the American Institute of Planning*, 25: 73-76.

Hillier, B., and Hanson, J. 1984. *The Social Logic of Space*. Cambridge, UK: Cambridge University Press.

Miller, H. J. 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, 5: 287-301.

Mitchell, R. B. and Rapkin, C. 1954. *Urban Traffic: A Function of Land Use*. New York: Columbia University Press.

Montello, D. R. and S. M Freundschuh. 2006. Cognition of geographic information. In *A Research Agenda for Geographic Information Science,* ed. R. B. McMaster and E. L. Usery. Boca Raton, FL: CRC Press, 61-91.

Stewart, J. Q. 1948. Concerning 'social physics'. *Scientific American*, 178: 20–23