Design concepts

Jesse James Garrett

The Spatial View

Everything is an object in space. Positions in space have meaning: the x and y axes map out a semantic space, generated in part by the system and in part by the user. The xy-space is not meant to be a perfect semantic visualization – semantic space is multi-dimensional, and we only have two. But two axes are enough to create a flat plane, where we can harness human spatial and visual memory to ease recognition and recall.

The z axis of the space represents time. Objects farther from the viewer are further away down the time scale. Each object occupies a unique position on the z axis, and every position on that axis is occupied by something – there are no empty spaces. Users cannot move objects along the z axis. The z-space is not linear; far-away objects move toward the user much more quickly than close objects. This allows users to visually scan large numbers of nearby objects while in motion, but keeps the distant end of the z axis from becoming too distant, creating a sub-optimal browse experience (see also Desert Bus).

Objects that have been invoked are considered active. The active set comprises the “present” end of the z axis. When objects leave the active set (when the user “closes” them) they begin to recede down the z axis. Until then, they remain fixed in the present.

Opacity and transparency communicate an object’s suitability to the current context. Every object communicates the relationship it can have with any other object. Data and presentation objects that can integrate tightly can be grouped for use as a single object. Groups called clusters are made automatically by the system to communicate strong semantic affinity. These can be edited or re-tagged by the user.

The user can also create groups of their own (semantic affinity or not) for any purpose. Users can create clusters either directly, by pushing objects closer together in the xy plane, or indirectly, by editing an object’s metadata, which will cause the object to be repositioned by the system. Either the user or the system can assign a name to a cluster, which has extra weight in any semantic analysis.

To alleviate crowding, the xy plane can be stretched or compressed by the user. Objects retain their relative size (determined by their position along the z axis), but their relative placement on the xy plane is distorted uniformly and linearly along both x and y. Stretching may also be done by the system when the user focuses on a cluster. This distortion causes the xy plane to expand beyond the viewport; therefore, users can move the viewport around on the xy plane. Signposts around the edges of the screen indicate contextually relevant clusters out of view in that direction.

Objects have gravity. Semantically similar objects are drawn toward one another, and their gravitational influence is compounded. As new objects enter the space, they are automatically placed according to these gravitational influences. For example, a webpage that is semantically similar to more than one cluster might be placed between the two clusters on the xy plane; if one cluster is significantly larger than the other, the new object would be placed closer to the larger cluster.

Classes of Objects

There are three classes of objects: people, places, and things. People objects are, at heart, a collection of contact information, but they also maintain a rich, semantic history of your communication and collaboration with that person. Some data within a person object may be live data shared by that person, either privately to specific people, in small, restricted groups, or publicly as a feed. Other data within a person object may be appended by the user.

Places are resources on the network. Anything with a URI is a place, not just webpages. For example, web applications can insert saved states into history as places. Things are, essentially, resources that are part of the user's personal data. Deciding to make a personal copy of data on the network creates a thing now separate from the original network resource, which remains a place.

There are a few special classes of things: Data objects are simple repositories of structured data. Logic objects can operate on or manipulate data objects. Presentation objects offer user interfaces or direct-manipulation visualizations for datasets and functionality. Tokens are specialized data objects used for authentication and authorization. Credit card information takes the form of a token; so does personal behavior or preference data. In either case, the user explicitly authorizes or licenses use of the data by sharing the token with a service provider.

Objects may have different metadata schemes according to their class. In other words, people have different metadata than places or things. Users can extend any object's metadata scheme and assign their own metadata to objects.

Every object can have the same set of actions performed with it, regardless of its class. Objects can be collected into a group, added to or removed from an existing group, inspected, invoked, and moved around in the xy-plane. Invoking data objects displays and/or makes editable their data. Invoking logic and presentation objects allows the user to set default parameters. Invoking a person displays that person's feed of shared data; invoking a place opens an active network connection to that resource.

Way Awesome Bar

The way awesome bar is a generic, multi-purpose text entry interface. From the way awesome bar, users can enter URLs, invoke or inspect objects by name, perform searches, use commands, jump to a point in time in the spatial view, add metadata to objects, and create text data objects. It appears as an overlay and can be invoked in any context.

The Frame

The frame appears around the edges of the screen when invoked by the user. The appearance of the frame is animated to create the sense of stepping back from the current context to get a slightly broader view of available resources.

Each corner of the screen provides access to a way to access a filtered look at the spatial view. The four corners correspond respectively to people, places, things, and the way awesome bar.

Each edge of the screen contains proxies for objects. These proxies are not objects themselves – they are only pointers to, or representations of, objects. They exist only as a means of triggering actions. The edges scale with the display: the number of proxies shown is a function of the absolute physical length of that edge.

The top edge of the frame is the shelf. This space is populated solely by the user. Any type of object (person, place, thing, group) can be placed here, and in any order. This area is intended to provide rapid access to frequently used items.

The left edge of the frame is the history stack. When the user navigates from one place to another, the earlier place goes to the top of this stack. Older items on the stack push down one position, pushing the last item off the bottom of the stack. If the user needs to see farther back in history than the stack will allow, they should search or browse the spatial view.

The right edge of the frame is the user stack. This works just like the history stack (new items always on top), but objects are placed here by the user, not by the system. This stack enables users to perform complex tasks involving multiple objects without leaving the present context.

The Wheel

The bottom edge of the frame is the wheel. The wheel cannot be invoked independent of the frame. The visible part of the wheel curves across the bottom of the screen, creating the sense of a large wheel curving away from the user, only part of which is within the user's field of view. The wheel contains proxies for all active objects. Users can inspect, open, close, and switch among objects using the wheel. There is no limit to the number of objects that can appear in the wheel.

The order of objects from left to right on the wheel can be determined by the user. By grabbing and throwing the edge of the wheel, users can quickly scan very large sets of active objects. The wheel is circular – objects that disappear from one side will reappear on the other if the user continues to spin in that direction. The most recently accessed object, or the one currently being accessed, occupies the center position along the curved edge of the wheel when the frame is invoked.

Range of Action

To manipulate or act on an object in any way other than simply moving it around the semantic space, the user pulls the object into their range of action, or, more simply, "picks it up". The range of action lies between the user and the frame, so that users can pick up objects from webpages or from the semantic space and "drop them" into the frame.

Menus, Panels, and Workspaces

Most user actions that don't involved direct manipulation of objects are accomplished through radial context menus. Some actions invoke panels: small, focused interface components that enable the user to set parameters for the action. Workspaces are groups of objects that are displayed simultaneously within a layout created by the user. Objects within a workspace can expose data and services to one another, enabling users to assemble components into an integrated display that essentially functions as a single application.

Search and Focus

Search takes place within the spatial view. The system does not perform relevance ranking, though it does privilege cluster labels over text within pages. Matching results appear opaque and/or have a distinctive visual highlight; objects that do not match the search are transparent, nearly invisible. As in the regular spatial view, every item occupies a unique point on the z axis, and no point on the z axis is unoccupied; however, in search, the objects in question are only valid results. This has the effect of compressing the z axis, so that a small result set spanning a long stretch of time does not create the Desert Bus effect.

Additionally, the xy plane initially distorts to provide a view broad enough to encompass the entire result set, but no broader. The xy plane can be further distorted by the user as usual to enable easier visual identification of individual results.

Throughout the spatial view, the user can also focus on a cluster, causing it to be presented in a manner identical to search results (effectively, performing an implicit search for all members of the selected cluster).