芽|Sprouts Working Papers on Information Systems ISSN 1535-6078

Examining Egocentric and Allocentric Frames of Reference in Virtual Space Systems

Asaf Friedman
Delft University of Technology, Netherlands

Abstract
The aim of this paper is to examine the egocentric and allocentric frames of reference, through evidence from both gesture and linguistic communication. The action of frames of reference, helps the user refer to the agent as a base for movement or to the object as a guiding point. We will show that although each uses different methods to convey movement in virtual reality, both have a distinct ability to refer to the pathâ s intrinsic value, but both handle objects with a different number of axes. This paper identifies the major components related to the mechanics of movement in virtual space. It also discusses some of the misconceptions in the field.

Keywords: Virtual Reality Egocentric, Allocentric, Navigation, Representation, Information systems

Permanent URL: http://sprouts.aisnet.org/5-16

Reference: Friedman, A. (2005). "Examining Egocentric and Allocentric Frames of Reference in Virtual Space Systems," Delft University of Technology, Netherlands . *Sprouts: Working Papers on Information Systems*, 5(16). http://sprouts.aisnet.org/5-16

Introduction

In "The City of Tomorrow" Le Corbusier argues that orthogonal, the "*orthogonal state of mind,*" expresses the spirit of the modern age. "Orthogonality," which means "intersecting or lying at right angles" or in a "linear transformation," is the correct, state of the architectural and urban "mind". Le Corbusier examines the meandering roads of current European cities and compares them to the paths traveled by pack-donkeys. According to Le Corbusier, man walks in a straight line because he has a goal and knows where he is going. The pack-donkey meanders along, meditates a little in a distracted fashion; he zigzags in order to ease the climb. Pack-donkeys generally take the path of least resistance; their early tracks have marked the layout of many cities. Le Corbusier's intuitive idea was that a linear orthogonal grid was the answer to men's means/goal approach, as differentiated from meandering roads, but is this notion cognitively correct? Is there supremacy in the orthogonal view?

The above discussion relates to the object and its properties, whether the object is thought of only in relation to other objects, or in relation to the observer. The nature of the object in 'virtual space' is a curious matter; it could be thought of as occupying a virtual space independent of its physical essence. This sort of "requirement" might even overshadow every other consideration given the present-day power of virtual space. Confusion regarding objects in virtual space stems from neglecting to think of what might be prior: "Is the egocentric view a precondition to the allocentric one? Or vise se versa?" This question goes back to the discussion of the Cartesian thought of how do we perceive the world. One of the examples that Descartes gives is the example of the spoon in a glass of water were the rays of light break the continuity of the spoon as the light hits the water. We, as humans, can understand this, and therefore think rationally as we perceive the spoon in a glass of water. One of the ways to compare the egocentric and allocentric was to consider linguistic communication. It was argued that visual communication possesses only an egocentric view while language possesses an allocentric element.

Egocentric and Allocentric Systems

The way people represent objects and the way they arrange objects, locations, and paths are, in a sense, mental constructions, especially in the constructed 'realities' of computer interface. Visual navigation represents one possible way to move about in space. Visual navigation is a sequence of arrangements of objects in a location and the paths of object action. The representation of objects capable of being manipulated generally concerns a specific choice of objects transferred from one space to another, directed in virtual space by specific gestures. The term "spatial frames of reference" has been used by specialists in several different but related areas of research; for example, in perception, cognition, neuroscience, linguistics, and information science (Jackendoff, 1983, Levinson, 1996, Campbell, 1994). Fundamentally, a spatial frame of reference is a conceptual basis for determining spatial relations between various objects. This description is applicable across situations in which person-to-object and object-to-object spatial relations are represented or described.

When Piaget and Inhelder first published their book "*The Child's Conception of Space*" in 1948, Piaget was striving to understand the development of spatial reasoning. Piaget believed that the child learned from observation; for example, a child might study a straight line and come to an understanding of projection, realizing that the three dimensions of projective space lead to the idea of a bundle of lines intersected by a plane. But the concept of the straight line itself, together with the various relationships resulting from its synthesis with the original topological relations, ultimately presumes the discovery of the part played by points of view, that is, their combined co-ordination and differentiation of egocentric and allocentric frames of reference. But to discover one's own viewpoint is to relate it to other viewpoints, to distinguish it from and co-ordinate it with the points of view of other people. To do this requires a system of true mental operations, that is, operations which are reversible and capable of being linked together. In the allocentric view (i.e. Many-to-Many relationship), objects relate to other objects in a multifaceted way. The points which represent object locations in Cartesian space relate to X and Y co-ordinates and to other objects in that space. By contrast, in the egocentric view (i.e. One-to-Many relationship), all objects relate to a single object. Frames of reference, of course, can be constructed mathematically. When spatial reasoning is introduced into various frames of reference, one is dealing with representation, visual and verbal description. The 'plan' and 'axonometric' are associated with an allocentric view, while 'perspective' is associated with an egocentric view. In order to examine how well one handles egocentric and allocentric frames of reference one examines array-rotation viewer-rotation.

In the 1960th Piaget experimented with the limits of children's abilities to transform spatial information. Piaget and Inhelder attempted to discover the age by which children can switch from an egocentric to an allocentric frame. That is, the egocentric frame was considered to be innate, while the allocentric was considered to be acquired. Piaget and Inhelder presented children with a model of three colored mountains and asked them to indicate how it would look to an observer who viewed it from a different position by showing them different alternatives to the mountain view. Until 9-10 years of age, children tended to make what were thought to be egocentric errors when shown a representation of the array, which depicted a variety of elevations. According to Huttenlocher (1979), Piaget had chosen difficult problems of representation when the children had to choose among pictures of the mountains from differing perspectives. In an experiment conducted with a small model of a farm, the children where able to answer correctly what a new viewer would see. According to Huttenlocher (1979), array-rotation problems are much easier to solve than viewer-rotation problems. Huttenlocher proposed that in solving these problems, subjects interpret the instructions literally, recoding the position of the viewer vis-à-vis the array for viewer-rotation problems and recoding the array with respect to its spatial framework for array rotation problems. The results show that the viewer is fixed vis-à-vis the spatial context rather than fixed vis-à-vis the array.

Campbell (1993) adds to some distinctions between ways of thinking that involve an explicit or implicit dependence upon an observer and those that have no such dependence. Campbell's suggestion is that the resultant system is egocentric only if its significance can be given solely by a reference to the subject's own capacities for perception and action in terms of what he calls "causally indexical" relations. The causal significance – the judgment made about objects standing in various spatial relations – is essentially given in terms of its consequences for the subject's perception or action: casual indexical. It will be allocentric if, and only if, this significance can be given without appealing to the subject's perceptual and active abilities, causally non-indexical, in terms that give no single object or person a privileged position, which treats all the world's objects (of a given kind) on par with respect to their physical interactions.

Panoramic Communicative View

To understand the relationship between an observer and his or her surroundings one must understand the communication relationship that exists regarding that observer and others who might wish to learn about such relationships. Effective representation requires making as many distinctions as necessary and sufficient to communicate the spatial relations between an object and an observer. The representation for ordering information that restricts the location of a point with respect to some reference point is given by the particular panorama. A panorama is defined as a continuous image created from lateral motion, movement from one side to another, from which features are extracted. The examination of how one describes what one sees while moving can identify the speaker's underlying cognitive states. In order to achieve spatial cognition, i.e., the ability to represent the environment and act upon this representation to form a decision (e.g., where to go and what to see), one has to be able to distinguish between the object in a scene and its background. The English spatial predicate marks a location, an operation that designates it as one to be remembered. The spatial predicate also marks the referent and relatum position in space and arranges its parts to be accessed. The English spatial predicate takes the form of a predicate, a referent and a relatum: (1) **Referent** – the object; (2) **Relatum** – the reference object in the background; (3) **Predicate** – the spatial relationship between the referent and the relatum (Talmy 1983).

In order to be able to understand the difference between description and depiction, between encompassing view and individual point of view, we must first consider spatial reasoning. Spatial reasoning is the engagement in representation of motion in terms of the geometric properties of points in space. The most elementary properties are prepositions, which specify the main relation between objects in space; I call such prepositions *spatial prepositions*. Examples of spatial prepositions are *in* – city (x) or *next to* – house (x). Typically, the set of points where a predicate is true form a single compact *region* of space, and spatial reasoning amounts to detecting intersecting relations among combinations of regions, called environments. The relation between points in space and the observer comprise the basic elements for reasoning about space.

On the phenomenon level the event that one is looking at is an event depicted as an action sequential stream; one is in a given state and commands a change to a new state. What we are interested in is the relationship between the avatar's previous position and subsequent position relative to either the object or the avatar. We could ask ourselves: what directions are equivalent to the relationship of the subject's transformed position? A problem arises regarding the fuzzy nature of event classes. The principle of phrase structure is a **homomorphic** relationship between what is said and what is there. A homomorphism connects every point from system A to system B, without, however, connecting every point of structure B to structure A. An **isomorphism** is a symmetrical relationship; it connects every point from system A to every point of system B and vice-versa. In our case a description is inevitably a homomorphic relationship. For any event there will be subjects that are clear instances of that event type, those that clearly are not instances, and those whose membership in that class of events is unclear.

The spatial predicate marks the referent and relatum position in space and arranges its parts to be accessed. **Orientation information** locates a point object in any position of the semi-straight line from the origin of the Cartesian co-ordinates with a given angle. *Orientation* information can be given by polar co-ordinates: the orientation is given by a vector – an angle and the exact position in the straight line of orientation by a distance, both measured from the origin of the Cartesian coordinates. Three spatial point objects are involved in the definition of orientation relationships by orientation model, i.e. 'a' and 'b', which define the reference system, and 'c', the object whose orientation is provided with respect to the reference system.

In investigating the hippocampus (the area of the brain thought to contain the encoding of spatial relationships), O'Keefe (1990, 1991) proposed the slope-centroid model as the way in which animals successfully navigate. This model represents the basic relations of frames by always having a reference, which is outside the simple Euclidean metric relation of trajectory vector between the current location and the desired location. O'Keefe's model contains two stages in an animal's construction of a map of its environment. In the first stage, the animal identifies a notional point in its environment, the centroid, which is a notional point in the sense in which the South Pole or the Equator are notional: there may be no distinctive physical feature at that place. It is a fixed point, in that it does not move with the animal. In the second stage, the animal also identifies a gradient for its environment, a way of giving compass directions. This is the slope of the environment; it functions like the direction east-west. The direction is fixed no matter how one moves around, and one can partially define which way one is going relative to a given object by approximating the angle that one is making in terms of the centroid. As in almost all models of mapping, we take it for granted that the animal is constructing a two-dimensional map of its environment; the third dimension is not mapped. Once the animal has identified the two stages, it can construct a map of its environment by recording the vector from the centroid to each of its targets, using the slope to define direction. Assuming that the animal has done this and now wants to know how to get to a particular target, what it must do is find the vector from itself to the centroid. Once the animal has the vector from itself to the centroid and the vector from the centroid to the target, it can find the vector from itself directly to the target.

According to O'Keefe (1990), at any point in an environment, an animal's location and direction are given by a vector to the centroid whose length is the distance to the centroid and whose angle is the deviation from the slope (360-γ). Other places (A and B) are similarly represented. This dichotomy has its roots in egocentric and allocentric frames of reference and subsequent attempts by O'Keefe (1993) to define the possibility of navigation without allocentric thinking.

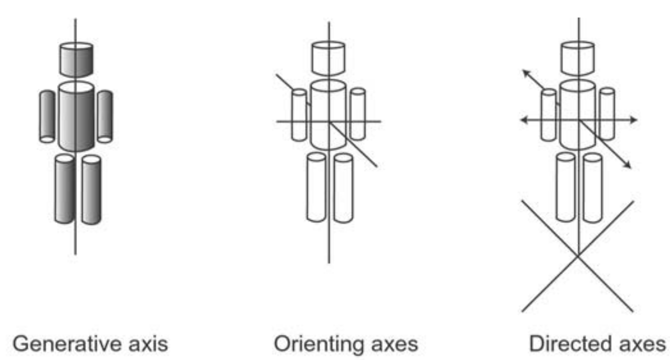


Figure 1. Three axes - object to parts construction

The co-ordinate system, centered on the viewer, seems to be based generally on the planes through the human body, giving us an up/down, back/front and left/right set of half lines. Such a system of co-ordinates can be thought of as centered on the main axis of the body and anchored by one of the body parts. Although the position of the body of the viewer may be one criterion for anchoring the co-ordinates, the direction of gaze may be another, and there is no doubt that relative systems are closely hooked into visual criteria. An axis is a locus with respect to a particular spatial position (which is defined). Landau and Jackendoff (1993)distinguish three types of axes that are required to account for linguistic terms describing aspects of an object's orientation. The **generating axis** is the object's principal axis as described by Marr (1982). In the case of a human, this axis is vertical. The **orienting axes** are secondary and orthogonal to the generating axis and to each other (e.g., corresponding to the front/back and side/side axes). The **directed axes** differentiate between the two ends of each axis, marking top vs. bottom or front vs. back.

In the TOUR model (Kuipers, 1978), the simulated robot performs two types of actions: TURN and GO - TO. The purpose of the procedural behavior is to represent a description of sensorimotor experience sufficient to allow the traveler to follow a previously experienced route despite incomplete information sensed. It is stored as sensorimotor schema of the form <goal, situation, action, result>. The "you are here" pointer describes the current position of the robot by determining its place and orientation. The **topological map** is constructed when there are enough sensorimotor schemes to allow such a construction. The topological map consists of a topological network of places (points), paths (curves), regions (areas), and topological relationships among them (connectivity order and containment). A **place** consists of an orientation reference frame, a set of paths intersecting at a particular place together with the angles of the paths relative to the orientation reference frame, and the distances and directions of other places, which are visible from this place. A **path** consists of a partial ordering of places on the path, and regions bounded by the path on the left and right. The orientation reference frame is described in terms of its orientation relative to other frames. A **district** consists of edges and paths.

According to Escrig (1998), there are four different types of inference rules defined to manipulate knowledge embedded in this representation: (1) Rules that compare the "you are here" pointer with the topological description of the environment; (2) Rules for maintaining the current orientation with respect to the current coordinate frame; (3) Rules that detect special structural features; and (4) Rules that solve route-finding and relative-position problems.

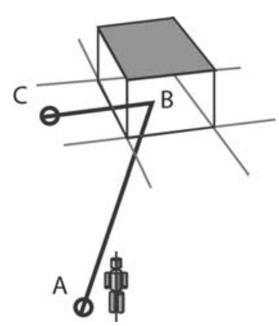


Figure 2. The reference system

The basic knowledge is represented in Zimmermann and Freksa's (1993) approach, which is the orientation of an object, c, with respect to the reference system defined by two points, a and b, that is, *c* with respect to *ab.* The vector from a to b and the perpendicular line by b define the coarse reference system (Figure 2), which divides the space into nine qualitative regions; straight-front, right-front, right, right-back, straight-back, left-back, left, left-front, identical-front.

How views Change: Axes and Frames of Reference

With the introduction of an avatar into the environment, a situation is created where the *observer* frame can be projected onto the object from a real or hypothetical observer. This frame establishes the front of the object as the side facing the observer. We might call this the 'orientation mirroring observer frame'. Alternatively, the front of the object is the side facing the same way as the observer's front. We might call this the 'orientation-preserving observer frame' (Jackendoff, 1999). To describe where something [the 'figure'] is with respect to something else [the 'ground'], we need some way of specifying angles from point of origin. In English, we achieve this either by utilizing features or axes of the ground or by utilizing angles derived from the viewer's body coordinates. The notion 'frame of reference' can be thought of as labeling distinct kinds of coordinate systems. Linguistic literature usually invokes three frames of reference: an intrinsic or object-centered frame, a relative (deictic) or observer-centered frame, and an absolute frame (see Figure 3). The frames of reference presuppose a 'view-point', and a figure and ground distinct from it, thus offering a triangulation of three points and utilizing co-ordinates fixed on the viewer to assign directions to a desired location.

Intrinsic frames of reference – the position defining loci are external to the person in question. This involves taking the inherent object-centered reference system to guide our attention, and uses an allocentric frame (Campbell, 1994, Levinson, 1996).

Relative frames of reference (deictic) – those that define a spatial position in relation to loci of the body or observer-centered. The relative frame of reference is used to identify objects' direction; this involves imposing our egocentric frame on objects (Campbell, 1994, Levinson, 1996).

Absolute frames of reference – defining the position in environment-centered, such as North, South, or East, Polar co-ordination. They use either Cartesian or Polar co-ordinates (Campbell, 1994, Levinson, 1996).

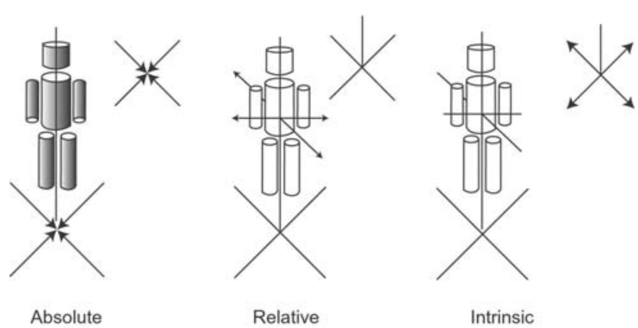


Figure 3. Three linguistic frames of reference

Let us consider an example: "the gate is in front of the house." For a manufactured artifact, the way we access or interface with the object determines its front, anchored to an already-made system of opposites: front/back, sides, and so forth. This would also be the case with any centralized symmetrical building but not for a cylindrical building. In fact, the situation is more complex. The sentence "the gate is to the left of the house" can sometimes employ a relative frame of reference that depends on knowledge of the viewer location. The statement as indicated suggests that the gate is between the viewer and the house, because the primary co-ordinates on the viewer have been rotated in the mapping onto the ground object, so that the ground object has a "left" before which the gate is situated. If one views a frame of reference as a way of determining the axes of an object, it is possible to distinguish at least eight different available frames of reference (Jackendoff 1996). Despite extensive interest in the role of frames of reference in spatial representation, there is little consensus regarding the cognitive effort associated with various reference systems and the cognitive costs (if any) involved in switching from one frame of reference to another. An experiment was conducted by Allen (2001) with regard to these issues: accuracy and response latency data were collected in a task in which observers verified the direction of turns made by a model car in a mock city in terms of four different spatial frames of reference: fixed-observer (relative-egocentric), fixed-environmental object (intrinsic-fixed), mobile object (intrinsic-mobile), and cardinal directions (absolute-global). According to Allen, (2001) the results showed that frames of reference could be differentiated on the basis of response accuracy and latency. In addition, no cognitive costs were observed in terms of accuracy or latency when the frames of reference switched between fixed-observer versus global frames of reference or between mobile object and fixed environmental object frames of reference. Instead, a distinct performance advantage was observed when frames of reference were changed.

When comparing frames of reference, a few conclusions should be noted:

- Frames of reference cannot freely "translate" into one another.
- There is common ground between visual axes and linguistic frames of reference that allows them to converge on an object, and allows one to talk about what one sees.
- Language is a process most adaptive to directing; therefore, other modalities should follow linguistic patterns (especially for tasks related to navigation).

Possible Interactions with Objects and Routes

Reaching for a nearby object requires pre-existing knowledge, knowledge of what properties define and delimit objects. When navigating in "built" or constructed environments, one employs a different strategy of grasping, that of approach and position (look) and that of reach (interaction.) According to Merleau-Ponty (1962), concrete movement is centripetal whereas abstract movement is centrifugal. The grasping frames relate to objects in two ways: **Manipulation mode** - an intention that conveys an impulse toward the object, as in the case of object use. **Observational mode** - an intention conducted away from an object, as in the case of object observation.

To design the architectural object, the architect works mostly with canonical views such as sections, plans, and elevations. The perspectives that architects use enhance the design by allowing observational comparisons of at least two sides of an object. Thus the canonical extension of axes that are most commonly used are divided into eight qualitative regions:

straight-front, right-front, right, right-back, straight-back, left-back, left, and left-front of the reference system. The canonical architectural reference system representation (see Figure 4) avoids the granularity of choice among users by setting a constraint that states that all relative distances to the avatar-object shall remain constant when the viewpoint changes, unless something contradicts the substantiality constraint that states that solid objects cannot pass one another.

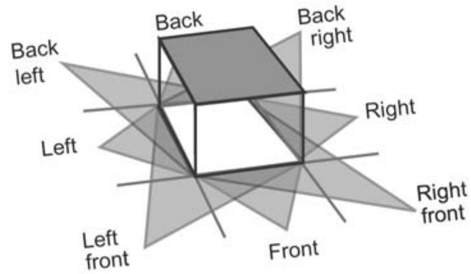
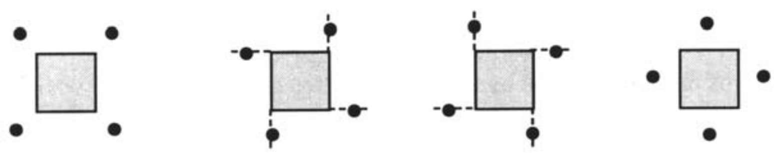


Figure 4. The architectural object reference systems, incorporating the visible area from a generating location (or convergence location of the optic rays)

In the city there is the added distinction between observation and movement. The city form dictates a physical limitation from where one can view the city. For simplification, interaction in the built environment is divided into sixteen canonical points of view to approach an object (see Figure 5). The eight points of view (of Cases 2 and 3) take on the nature of urban bound movement interaction with the avatar approaching the side of an object from a road, where only one side of an object is visible. This view is often ignored by architects when designing, and its effect on the architecture as Venturi (1977) shows is critical to the modern architect. It uses Polar coordinates. For example, approaching the building from the East, West, North, or South.



Case 1 Case 2 Case 3 Case 4

Figure 5. Sixteen different cases depending on the relative orientation of the point “a” with respect to the extended object “b”, which are grouped into four cases due to symmetries

The approach to pointing is to define a path from a to b with the position of the observer c; thus, in the panoramic model, one can point to an object and locate a new perspective relative to it. The notion of reference system can be viewed as a conceptual neighborhood with topological and linear relations (see Figure 6). Thus one can walk around to the back of the building or transport to a new position in the back of the building. There are two possibilities for manipulating an object, the first is a linear route the other is a topological conceptual system.

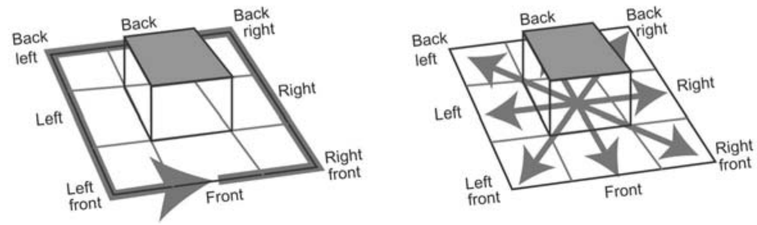
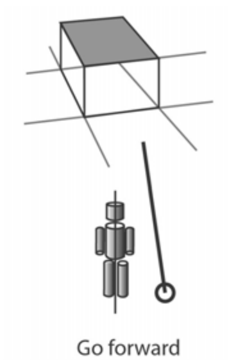


Figure 6. Topological and linear view of the conceptual neighborhood

Visual and linguistic descriptions have the ability to convey information about the path through explicit and implicit knowledge; for example, "go left" is a description, where the start point is implicit. The path can also have the end point suspended like "go towards the house" or "go into the house;" it is the equivalent of pointing visually. The converse: path can have an arrival point like "go to the left of the house." Lastly, the transverse: path can have explicit start and an end point, giving us the ability to determine the path relation to an object. The elements presented here are the full range of interaction in an immersed environment.

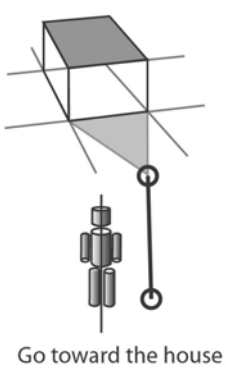


Bounded agent The agent can move in any direction desired (six degrees of freedom). It is operated by the use of correspondence of the movement of screen actions to observer purpose by combining the agent reference systems of input to output. This is an agent-entered reference system.

To move the agent, use the command GO.
Utilizes the commands: **go** → forward, backwards, left, right, up, and down.

The agent can also turn (Look) sideways.
Utilizes the commands: **Turn** to → the left/right.

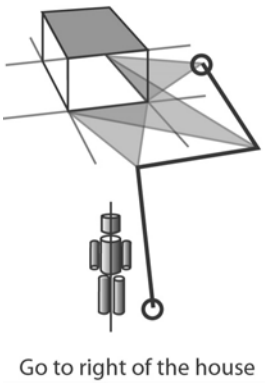
Figure 7. Bounded agent



Directed agent identifies the new position of an agent by directing it to an object; it uses the object-centred reference system.

The **go towards** command differentiates between movement and end goal. The '**go towards** an object' directed command has no ability to discern between different regions of space relative to the object reference system.

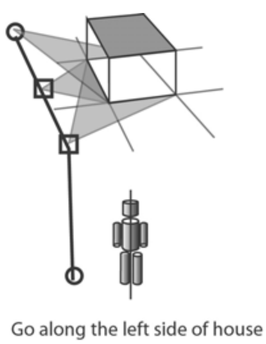
Figure 8. Direct agent



Converse path: defines a spatial path of the observer in relation to an object, object-centred. Two-Point Relations: this includes the evaluation of topological, angular, and distance-dependent relations, which share the common characteristics that they – in their most basic form – relate two objects to each other. Uses the commands: in front of, on the left/right side, behind – preference to agent role: architect, tourist, etc.
It is operated in terms of the use of identification of object and the new position directed relative to it.

The **go to** command differentiates between movement and end goal. '**Go to** an object' has the ability to discern between different regions of space relative to the object reference system.

Figure 9. Converse path



Transverse path: defines the relation of objects in relation to the path of the observer. N-Point Relations: Relations that cannot be reduced to a two-point problem, such as path relations [through/along/across].

The user operates it by identifying a target (object) and the object axis along which movement is to be performed.
Go along has the ability to discern between different path movements relative to the object.

Utilizes the commands: **Go along** → (path) and **Go around** → (an object) is also part of those commands

Figure 10. Transverse

Comparison between Egocentric and Allocentric Frames

The method of examination is usually a compromise of task analysis and information analysis. The process starts with a requirement analysis (see Figure 11). The first part of the method concentrates on user context analysis eliciting information to classify information requirements in their task context. The task requires interviews with users who have knowledge about the design process. Information analysis builds on the task model, which in our case is the possibility of action of any given spatial preposition task. In the case of information analysis, the questions that one asks at every stage of the process are:

- What input information is required for this action?
- What output information is produced by this action?
- Is any other information required to help the user complete this action?

The model is then analyzed in terms of the specific demands of the user and the categories of descriptions created. The **information analysis** of the user when performing a task depends on information categories and information declaration. The **information categories** are examined in terms of verb action, and **information declarations** are examined in terms of syntax and semiotics.

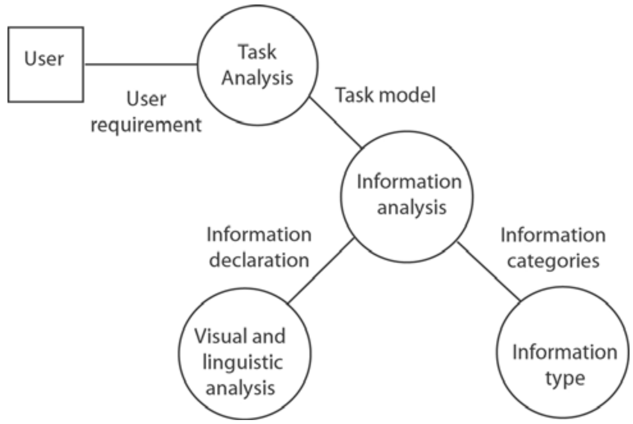


Figure 11. Method diagrammatic overview

To investigate the usability of the system one asks the following questions:

- Were there any tasks that the system was unable to represent?
- Were there any relationships between intentions and commands that the system was unable to represent?
- Is there any evidence that the use of the system would have saved the users substantial effort?
- Would use of the system create substantial new work for users?

Frames of reference have a visual component and a linguistic component. The visual components are: generative axis, orienting axes, and directed axes, while the linguistic frames of reference are: absolute, intrinsic, and relative. The linguistic system uses a combination of visual frames of reference to direct attention to an object. Thus the relative frame of reference employs directed axes, while the intrinsic frame of reference employs orienting axes, and the absolute frame of reference imposes an abstract axis on the subject. This is to argue that the linguistic system incorporates an extra set of axes compared to visual systems. According to Levinson (1996) and Campbell (1994), frames of reference already involve egocentric and allocentric thinking. Thus the major distinction between frames of reference is the way we think about them: as one-to-one relations or as one-to-many relations. According to Levinson, the three linguistic frames can be summed up in Table 1.

Intrinsic	Absolute	Relative
Origin: ≠ ego	Origin ≠ ego	Origin = ego
Object-centered	Environment-centered	Observer-centered
3-D model		2 ½ -D sketch
Allocentric	Allocentric	Egocentric
Orientation-free	Orientation-bound	Orientation-bound

Table 1. Aligning classifications of frames of reference (S. Levinson)

In order to compare the different frames, two different rotations are examined: object and array. The literature examining this phenomenon is extensive and goes back to the times of Marr (1982).

When a comparison table is drawn (see Table 1), it is relatively easy to notice that when one compares linguistic frames of reference the difference between them is a semiotic distinction, between egocentric and allocentric. This semiotic distinction is really a question of one-to-one relations between the observer and avatar, object, and polar coordinates. In the intrinsic and relative frames of reference, the observer uses an analogous process *relative* to himself or herself. In the absolute frame of reference, the observer uses an analogous process *external* to himself or herself. In any act of communication, it becomes apparent that egocentric frames obtain the sort of one-to-one relation required for effective communication. As we have shown before, this conclusion coincides well with findings in neurophysiology (Dodwell, 1982, O'Keefe, 1993).

As Levinson demonstrates (in Table 2), there is a significant difference between the various frames, but the frames can be compatible across modalities (Levinson, 1996).

F = Figure or referent with center point at volumetric center Fc
G = Ground or relatum, with volumetric center Gc, and with a surrounding region R
V = Viewpoint
A = Anchor point, to fix labeled co-ordinates
Slope = fixed-bearing system, yielding parallel lines across environment in each direction

	Intrinsic	**Absolute**	**Relative**
Relation is	Binary	Binary	Ternary
Origin on	Ground	Ground	Viewpoint V
Anchored by	A within G	Slope	A within V
Transitive	No	Yes	Yes if V constant
Constant under rotation of			
Whole array?	Yes	No	No
Viewer?	Yes	Yes	No
Ground?	No	Yes	Yes

Table 2. Summary of properties of different frames of reference (S. Levinson)

As we move in the three-dimensional world, what we see is what Marr calls 2 ½ D. That is, we see non-distal objects and the objects displayed show only the face that we look at, the isovist view; there is no natural three dimensional multiple-view representation. As we examine the different paths that an avatar can take, there emerge two properties: one is the act as a command; the other is the control over each command. When one examines the relation between the linguistic command type and the frames of reference (see Table 3), one realizes that the commands use relative and intrinsic frames of reference.

	Bounded agent	**Directed agent**	**Converse path**	**Transverse path**
Relative Frame	X	X		X
Intrinsic Frame			X	
Absolute Frame				

Table 3. Aligning classifications of frames

Concluding Remarks

In this paper we have compared two basic methods of handling objects: gestural and linguistic communication. We have shown that although each method uses different means to convey movement, and each has a distinct ability to refer to a path's intrinsic value, both methods handle objects with the same frames and axes. When one puts side by side the visual and linguistic frames of reference, one observes that they are virtually the same; the difference between them is the number of axes they refer to.

The egocentric and allocentric cannot constitute a distinction between frames of reference, only a broad description regarding ways of thinking about frames of reference. Frames of reference concern the observer, the avatar, objects, and polar coordinates. When using intrinsic and relative frames of reference, the observer is using a process relative to himself (or herself), while in the absolute frame of reference the observer uses a process external to himself (or herself). In the case of intrinsic frames of reference, experts tend to support the use of an allocentric frame of reference. Following O'Keefe (1990, 1991), we prefer to think that when one employs an egocentric frame of reference, in a one-to-one relationship, one gives preference to vision, suggesting, ultimately, that egocentric and allocentric frames of reference are independent of linguistic frames of reference.

References

Allen, L. G. (1999). Spatial Abilities, Cognitive Maps, and Wayfinding; Bases for individual differences in spatial cognition and behavior. In Wayfinding Behavior Cognitive mapping and other Processes. Baltimore: Johns Hopkins University Press.

Campbell, J. (1994). Past, space, and self. Cambridge, Massachusetts: The MIT Press.

Dodwell, P. C. (1982). Geometric approaches to visual processing. In D. J. Ingle, Goodale, M.A., Mansfield, R.J.W. (Ed.), Analysis of Visual Behavior. Cambridge: MIT.

Escrig, M. T., and Toledo, F. (1998). Qualitative Spatial Reasoning: Theory and Practice Application to Robot Navigation (Vol. 47). Amsterdam: IOS Press.

Hazen, N. L. (Ed.). (1980). Spatial Orientation: a comparative approach. New York: Plenum Press.

Huttenlocher, J., and Presson, C. C. (1979). The coding and transformation of spatial information. Cognitive Psychology, 11, 375-394.

Jackendoff, R. (1983). Semantics and Cognition. Cambridge, Massachusetts: The MIT Press.

Jackendoff, R. (1999). The Architecture of the Linguistic - Spatial Interface. In P. Bloom, Peterson, M. A., Nadel, L., and Garrett, M. F. (Ed.), Language and Space. Cambridge, Massachusetts: The MIT Press.

Kuipers, B. (1983). The Cognitive Map: Could It Have Been Any Other Way? In J. H. L. a. A. L. P. Pick (Ed.), Spatial Orientation: Theory, Research, and Application (pp. 345-359). New York: Plenum Press.

Landau, B., Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. behavior and brain sciences, 16, 217-265.

Le Corbusier, The City of Tomorrow, trans. Frederick Etchells (Cambridge: MIT Press, 1971).

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, Peterson, M.A., Nadel, L. & Garrett, M. F. (Ed.), Language and Space. Cambridge, Massachusetts: MIT press.

Marr, D. (1982). Vision, a computational investigation into the human representation and processing of visual information. New York, New York: W. H. Freeman and Company.

Merleau-Ponty, M. (1962). Phenomenology of Perception (C. Smith, Trans.): Rutledge.

O'keefe, J. (1993). Kant and the Sea-horse: an Essay in the Neurophilosophy of Space. In N. Eilan, McCarthy, R., and Brewer, B. (Ed.), Spatial Representation Problems in Philosophy and Psychology: Oxford.

Piaget, J., & Inhelder, B. (1960). The Child's Conception of Geometry (E. A. Lunzer., Trans. 1967 ed.). New York: Rotledge and Kegan.

Rock, I., & Wheeler, D., & Tudor, L. (1989). Can we Imagine How Objects Look from Other Viewpoints? Cognitive Psychology, 21, 185-210.

Talmy, L. (1983). How Language Structures Space. In J. H. L. Pick, and Acredolo, L. P. (Ed.), Spatial Orientation: Theory, Research, and Application (pp. 225-282). New York: Plenum Press.

Venturi, R., Brown, D.S., Izenour, S. (1977). Learning from Las Vegas. Cambridge: MIT Press.

Zimmermann, K., Freska, C. (1993). Qualitative spatial reasoning using orientation, distance and path knowledge. Paper presented at the IJCAI - 93, Cambery.

芽|Sprouts

Working Papers on Information Systems | ISSN 1535-6078