Incorporating a Virtual Presenter in a Resource Adaptive Navigational Help System

Jörg Baus, Andreas Butz, Antonio Krüger
Collaborative Research Center 378
University of Saarbrücken
PO-Box 151150
66041 Saarbrücken
Germany
[baus,butz,krueger]@cs.uni-sb.de
http://w5.cs.uni-sb.de/~[baus,butz,krueger]

Abstract

We discuss a virtual presenter in the form of a 3D avatar whose task it is to show users their way through a (3D model of a) public building. Our thesis is that when looking for way descriptions, seeing such a walkthrough guided by a virtual scout is easier to remember than a purely textual description or the kind of annotated floor plans used widely in today's public buildings. Furthermore, we discuss some technical aspects of how such a scout can be constructed and what underlying data and processes are needed for the automated generation of guided route descriptions. One of our main goals is to give the avatar the ability to react to the spatial environment in an appropriate manner, e.g., by pointing to relevant objects and following a path that can be easily memorized by the viewer. We introduce the concept of bounds objects useful for the computation of spatial relations and behavior activation of the virtual presenter. The final presentation is then performed by the virtual presenter moving along a path scouting the surroundings. The presenter is made sensitive to the objects in the environment, thus getting only the relevant information depending on its task and position. It will perform the task alone without system intervention, by querying the environment using the bounds concept to trigger behaviors and to establish an egocentric frame of reference.

Keywords: navigational help, virtual presenter, bounds objects, spatial relations

1 Introduction

In the project REAL, we investigate the design of a resource adaptive navigational help system with the capacity to adapt to various restricted resources (cognitive or technical) during the generation of multimodal route descriptions. The resource adaptive generation of graphics/animation for different classes of output devices ranging from a high-end 3D-graphics workstation down to a personal digital assistant implies dealing with limited technical resources on one hand, and taking into account a user's limited cognitive resources to decode and understand the presentation on the other hand. The top end of the scale is represented by a 3D walkthrough guided by a virtual scout, while the bottom end is represented by simple sketches of the environment and arrows indicating the direction. The 3D visualization (see Figure 1) combines an animation of a virtual presenter with accompanying text and meta-graphics in order to describe the route itself and relevant parts of the environment at the same time. In case of time pressure, the presentation will speed up and the system will reduce the amount of detail presented in the virtual walkthrough. This can be achieved for example by abstraction techniques [10] that generate visually clearer route sketches. These techniques also help to avoid giving a delayed user very long and detailed directions which are very hard to memorize and to follow correctly in the remaining time.

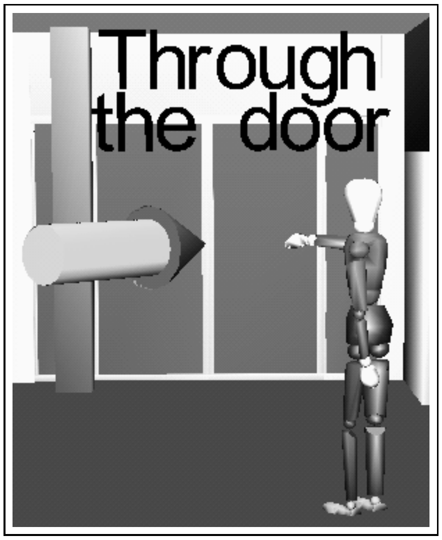

Figure 1: A 3D visualization

After a short view on the different resource restrictions and on the structure and purpose of route descriptions we will give an overview on our adaptive presentation planning approach. Then we will describe a representation of space and objects suitable for incorporating a virtual presenter in the form of a 3D persona whose task it is to show people a way through a (3D model of a) public building.

2 Resources

The construction of a presentation involves at least the following steps: determination of the information to present, determination of a presentation structure, graphical realization of the actual presentation. In addition, the information has to be presented taking the maximum benefit of particular strengths of each presentation medium, taking into account the actual content as well as the current environment/context. One goal is to generate different presentations for the same way description from the same data, i.e. a 3D-model of the domain and additional structured information (e.g., about landmarks and the route graph). As we assume that graphics always are used to communicate content between a sender and a receiver (in this case the machine and the user), the goal is to design the graphics as appropriate as possible for the given situation. Whenever graphics are presented via a certain medium to a human viewer, two different types of resources play an important role: on one hand all the *technical resources* of the machine and on the other hand the *cognitive resources* available to the user.

2.1 Technical Resources

Technical resources cover all types of limitations to the presentation platform. Here two different subtypes can be identified:

- Restrictions of the output medium
- Restrictions of the generation process.

The output medium is the visual interface to the viewer. Typical output media are printed paper and computer displays, but also 3D-displays, such as stereo glasses. All these media are at least restricted by the *inner* and *outer* scale. These terms from cartography (see also [7]) describe the maximal size of the displayed graphics and its resolution (and thus the amount of detail that will appear). Both factors influence the presentation of the same graphics. For example it's not possible to infinitely reduce a graphics' size without loosing important details. If the outer scale of the display is limited, it might be better to enlarge some smaller details that are important and omit others which are not so relevant for the visualization task. Another limitation is *color*. Colors often provide important information about the depicted objects. Color can also help to focus on important objects or details (e.g. a red spot near the objects of interest) If color is not available the system has to convey this information by other means or omit it completely. The graphics generation process itself also needs resources, especially *cpu-time* and *memory*. These are often critical in realtime applications, such as the scenario mentioned in the beginning, where a complete way description has to be rendered and displayed for a hurrying passenger.

2.2 Cognitive Resources

One of the most important cognitive resources of the viewer is her *working memory*. Following [2] the working memory does consist of at least three components: an attentional component, a phonological loop and a monitoring central unit. In this work we simplify the limitations that are connected to the working memory of the user to three classes of parameters: *time restrictions*, *domain knowledge* and *familiarity* with the presentation type.

Time restrictions can be divided into two different types. The *viewing time* is the time the viewer is given to look at the graphical presentation. In contrast to this, *decoding* time describes the time the viewer needs to decode the presentation and understand it's meaning. Both times may be very limited, especially when the viewer is heavily stressed. Complicated presentations that must be decoded under time pressure often lead to vicious circles, where the sense of not understanding results in stress reducing the ability to decode the presentation and so on. The domain knowledge of the viewer also influences the decoding time in two ways. On one hand the *type* of the presented graphics may be familiar or unfamiliar to the viewer, for example whenever certain coding conventions (e.g., symbols or icons) are used, and on the other hand information might be familiar or unfamiliar, e.g. parts of the way the pedestrian has to walk.

2.3 Types of Resource Adaptivity

Graphical presentations and a system generating them can respond to limited resources in various ways. For our work, we differentiate the terms resource-*adaptive*, resource-*adapting* and resource-*adapted*. A resource-*adapted* process would, for example, make optimal use of a given medium under given circumstances. In order for a process to be called resource-*adapting*, it will have to react to known resource limitations, e.g., yielding lower quality results. A fully resource-*adaptive* process will react to upcoming resource limitations during run-time and might even totally change its strategy in order to still yield results under the new restrictions. For a more thorough discussion of these terms see [17].

3 Purpose and Structure of Route Descriptions

Route descriptions should help the addressee in constructing a cognitive map of the environment. They describe the way to follow including landmarks and regions, and mention spatial relations between objects in the current environment. Verbalization is given step by step and the addressee undertakes a virtual journey through the environment (see [13, 12]). In our case we'd like to visualize the virtual journey, i.e. a course of motion in a given environment. In the visualization we focus on this course of motion from a mainly egocentric frame of reference. Elements of the environment are localized relative to the agent's position or relative to each other from a point of view called route perspective [15]. This kind of perspective is helpful to convey route-knowledge, knowledge about path segments and landmarks. For the transfer of survey knowledge, information about regions or the structure of the environment e.g., in order to help the addressee to take a short cut, often another point of view is chosen. Elements of the scene are referred to in an allocentric frame of reference corresponding to a survey perspective of the environment (see [15, 8]). In addition to the animation of a virtual walkthrough, we want to animate the virtual presenter. The virtual guide helps users to understand and navigate the environment. A virtual human presenter is largely enriching the repertoire of available presentation options which can be used to effectively communicate information [1]. The presenter guides the user through the presentation, attracts the user's attention and conveys additional conversational and emotional signals.

Following [16], a path of motion can be divided into segments. Each segment consists at least of four parts belonging to different categories: *starting point, reorientation (orientation), path/progression* and *ending point*. Since the paths we would like to visualize are continuous, the ending point of one segment can serve as the starting point for the next one. Segments are mostly separated by changes in direction, because in the presentation we have to communicate a change in direction clearly.

A segment represents a partial view of the environment by integrating different kinds of knowledge. In our proposed model, we represent knowledge about paths in a graph. Nodes of the graph correspond to turns or intersections in the real world. Edges correspond to paths between two points in the real world. Paths in the graph can be divided into segments in the aforementioned way. In order to clearly describe or visualize such a segment, we have to integrate knowledge about landmarks, especially at turning points. Here, we have to communicate the reorientation and the path progression. Depending on the next segment in the sequence we focus on objects in the direction of the path ahead of us. This is in accordance to experiments described in [11], where test persons have had problems to generate appropriate descriptions of a turn-left action when they were forced to refer to an object on the right hand side of the intersection. The 3D model used in our system stores information about object location, geometry and appearance. With this information we can render sketch-like route maps, generate animations through the environment or combine these modalities. After integrating objects of interest in such a segment, we are able to compute spatial relations between different objects in the scene. If we want to visualize a virtual walkthrough, we integrate the addressee's current position in the representation of the segments. This allows us to compute spatial relations between the virtual presenter and objects in the environment and to describe the location of objects from the agent's egocentric point of view.

Obviously, since the agent moves on its virtual journey, we have to constantly update its position and the valid spatial relations in the segmentation representation. Since construction of the segmentation structure and the integration of different pieces of knowledge can be done incrementally, we are able to cope with various time restrictions in the visualization and planning tasks.

4 Adaptive Planning of Way Descriptions

In order to generate structural descriptions of the graphics we have extended an efficient hierarchical planning approach presented in [3] for the generation of 3D animation. The main assumption here is that all generated graphical presentations can be structured in the form of a tree describing parts and subparts of the graphics to a certain depth. Each part or subpart corresponds to a node in the tree. Nodes are either terminal nodes in which case they describe portions of the graphics that will be realized by one of the graphics realization techniques or they are nonterminal nodes, in which case they specify a set of subnodes and a logical, spatial or temporal interrelationship between them.

Temporal interrelationships only apply to temporal media and include the concepts *parallel*, *sequential* and *incremental*. An example for temporally *parallel* subparts of a graphics are, for example, a camera motion and an object motion taking place over the same timespan of a 3D animation. A *sequential* interrelationship describes a sequence of subparts taking place in a temporal order, e.g., one after each other. Specifying the subparts as having an *incremental* relationship means that after a subtree is fully expanded, this subtree can be forwarded to the graphics realization component, which is not the case with every subtree in a *sequential* list. The specification of incrementally ordered subsequences allows the graphics realization process to start its work before the structure of the graphics is fully generated, and thus greatly reduces the perceived delay from the start of the whole graphics generation process to the moment the first graphical element is shown.

Spatial Interrelationships include a limited set of spatial orderings, such as *left-to-right* or *top-down* and can be used to roughly arrange visual elements such as viewports on a screen or parts of a composite graphics or diagram. In this sense they can be used within both temporal and static media and allow a rough specification of the overall spatial layout of the presentation. This is only a very coarse specification for higher levels of the presentation structure and should not to be confused with spatial arrangements of objects within a camera window or low-level elements of a diagram.

Logical interrelationships include the concepts of *alternative*, *conditional* and *additional* subtrees in the structure of a graphical presentation. Both *alternative* and *conditional* subtrees specify a list of possibilities for the realization of a certain part of the graphics. *Conditional* expansion selects one alternative from this list at planning time. In this way we can, for example, specify that a certain type of graphical presentation can be generated either as a 3D animation or as a 3D image of the scene depending on media or time restrictions. Specifying these subtrees as *alternative* postpones the decision until presentation time. The strategy here is to first expand the structurally simplest or computationally cheapest part of the tree and then – if time permits – to proceed with more complex alternatives that might be visually more appealing or clearer in the communicative sense. The resulting structure graph contains all of the various alternatives (unless planning was stopped before due to temporal restrictions) and leaves the decision which alternative is chosen to either the presentation process or even the user. An example for an *additional* relationship between subtrees of the structure tree is the labeling of a line drawing

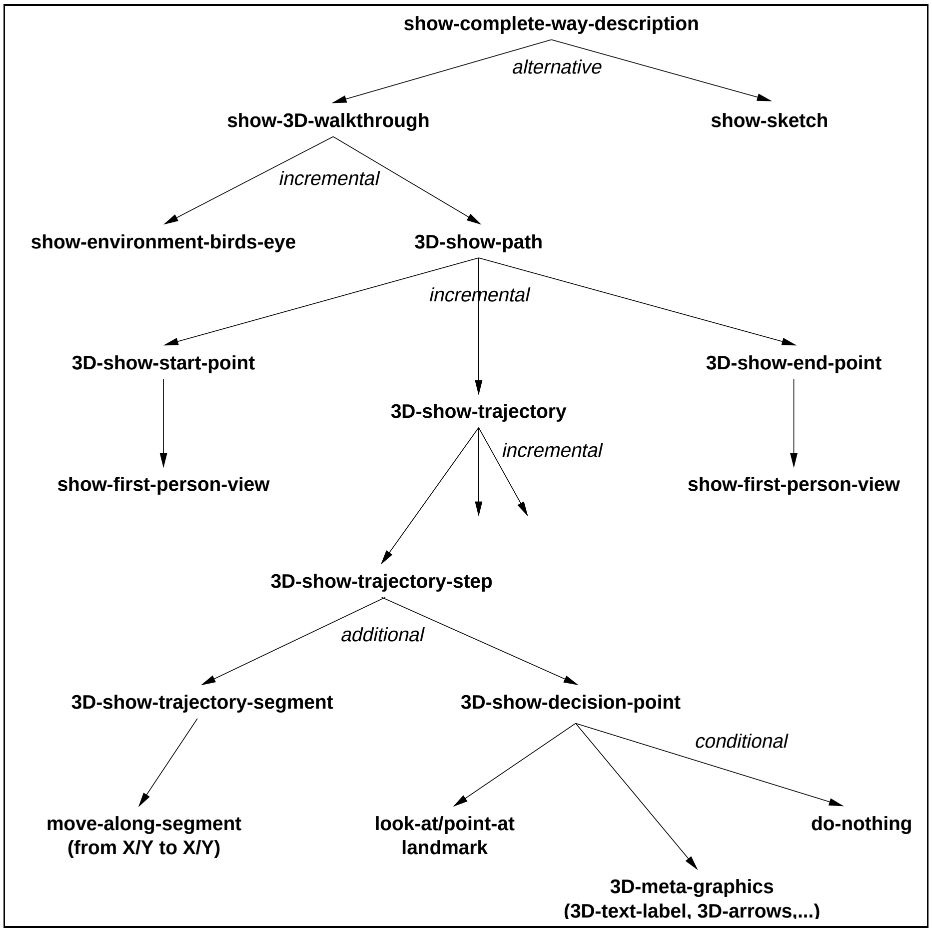

Figure 2: Simplified example of a structure tree for a graphical way description

or the creation of additional viewports for an already running 3D animation. This is assuming that already the first subtree would yield a 'working' graphical presentation, while the following subtrees contain presentation elements which will enhance the overall quality of the graphics. As it might have become obvious, the different kinds of interrelationships within the structure tree of a graphical presentation leave room for various strategies of adaptation of the generation process to limited resources either in the output medium or in the generation process. They allow the resource-*adaptive* generation of resource-*adapted* (decisions at planning time) as well as resource-*adapting* (decisions at presentation time) graphical presentations (see section 2). Figure 2 shows a simplified part of a structure tree for a graphical way description. At the root of the tree we see that the way description can be presented in the form of a 3D walkthrough or *alternatively* as a 3D sketch. One of these alternatives can be chosen at presentation time based on circumstances, media restrictions or user preferences. In the case of a 3D walkthrough a birds eye view of the environment and the actual path have to be shown. Specifying these two actions as having an *incremental* interrelationship implies that the birds eye view can already be drawn by the realization component while planning for the path visualization continues. Had we specified the two actions as being *parallel* or *sequential*, planning would have to be finished before the realization starts.

The visualization of the path itself consists of showing its starting and ending point and the trajectory inbetween. Showing the trajectory in turn is nothing else than an incremental loop over all of the trajectory's segments (see section 3) and generating a 3D walkthrough for each of them. Instead of planning all actions of the virtual scout down to every detail, only an abstract plan of the presentation is prepared (e.g., reflecting the path and the resource limitations that are known in advance). The final presentation is then performed by the virtual presenter moving along a path scouting the surroundings. For this purpose, we suggest to represent the relevant domain knowledge as annotations to the objects in the domain. The presenter is made sensitive to the environment around her, thus getting only the relevant information depending on her task and position.

In the next section we focus on the representation of well designed objects and spaces which can help to perform the task by containing clues to their meaning and operations, following the ideas of agents in annotated worlds presented by [4].

5 Representation of the Environment

Our representation of the environment consists of a hierarchy of objects. Objects can be divided into two groups: mobile and static. An object representation consists of three parts: geometry, appearance and annotations. At the very least, an object must have a geometry, an appearance and one annotation, which is its unique name. The geometry and appearance representation is object centered and hierarchical. By hierarchical we mean that objects can be decomposed into parts and subparts. Geometry and appearance will be used to visualize the environment. An object might have as many annotations as we need to provide meaningful content in different tasks.

Buildings (building sections, etc.) are static objects. In addition to their obligatory representation (geometry, appearance and object name) they include the complete list of general annotations for objects. This is because buildings may be very complex and have a hierarchical structure. They may consist of many subparts, references to subparts of the building, references to rooms in the building, which may have references to their furniture and so on. They also contain different reference systems and information on how these reference systems are attached to the building, e.g., specifying the building's prominent/intrinsic front and information on how reference systems and their applicability area/influencing area[1] are scaled. This latter information depends on the object's size and the environment. For example the meaning of the word "near" describes different regions depending on the object's size. "Near the empire state building" denotes a region that is different from the one in "the electron is near the nucleus". The list of annotations may contain:

- the object's location in world coordinates, an allocentric frame of reference
- references to the object's subparts, given by their unique names
- linguistic/textual information, useful for verbal descriptions of the objects, such as: "the red house"
- information on how to depict the object in 2D graphics, e.g. an iconic representation for its visualization in a 2D graphics,

[1]For a closer look on the concepts of applicability areas or influencing areas in spatial relation computation tasks see [6], [9].

- information on how to highlight the object
- information about the different reference systems attached to the object, e.g. functions to establish the applicability area/influencing bounds used to compute different spatial relations. Objects can have different reference systems suitable for different localization tasks, such as: “stand in front of the book shelf. The book is located in the upper left corner.”
- functional annotations, e.g. functions to compute different coordinates from the object’s 3D-model, e.g. its center, 2D bounding box, 3D bounding box. In fact it could be any query to a knowledge or database needed in order to provide meaningful content

And last but not least a special *bounds* object.

- *Bounds*. Bounds specify a spatial boundary of an object used to define the applicability area of spatial relations, needed in object localization tasks or used to trigger *behaviors*, in this case called *activation bounds*. Behaviors could be attached to any kind of objects and use a *scheduling bounds*. Scheduling bounds specify a spatial boundary in which a behavior can take place. The region within this boundary is called a *scheduling region*. A behavior is not active unless an *activation volume* intersects with the behavior’s scheduling region. If there is an intersection, the behavior becomes ”alive” or enabled.

An enabled behavior can receive stimuli and respond to those stimuli performing certain actions. Stimuli are used to schedule/trigger any kind of action described in the behavior as long as the behavior is enabled. Behaviors are useful for information retrieval, interaction and animation. They provide a link between a stimulus and an action described by the behavior. Furthermore, combinations of stimuli and combinations of actions are possible. One scheduling bound can trigger different behaviors or different scheduling bounds can schedule one behavior. In our case we use the scheduling bounds of objects to activate/deactivate the virtual presenter’s behaviors.

In the next section, we describe the representation of the virtual presenter and how we use the aforementioned activation bounds to make the virtual presenter “aware” of her current environment.

6 The Virtual Presenter

Following [14] a key problem posed by life-like agents that act in virtual worlds is *deictic believability*. They should in the same manner like humans refer to objects in their environment through combinations of speech, locomotion, and gestures. As mentioned the virtual presenter should be able to move through their environment, point to objects and refer to them appropriately. Deictic believability requires considering the physical properties of the presenters’s world. He must exploit its knowledge about of the positions of objects in the world, it’s relative location with respect to this objects to create deictic gestures, motions and utterances that are natural and unambiguous. To reach these goals the virtual presenter has to handle spatial deixis effectively and must be able to establish an egocentric frame of reference as described in section 3. In our scenario the virtual presenter differs from all other objects. First it is represented as a mobile object. Also, in situations in which we have to speed up the presentation of the virtual walkthrough to a level at which the animation of the presenter isn't tractable or useful anymore, we can omit its visualization, but only the visualization. The invisible virtual presenter itself stays active, still offering its abilities to scout the environment.

6.1 Establishing an Egocentric Frame of Reference

The virtual scout with its attached sensors and the influencing bounds (areas) of an object are shown in figure 3. The spheres shown around the virtual scout visualize the sensors, which

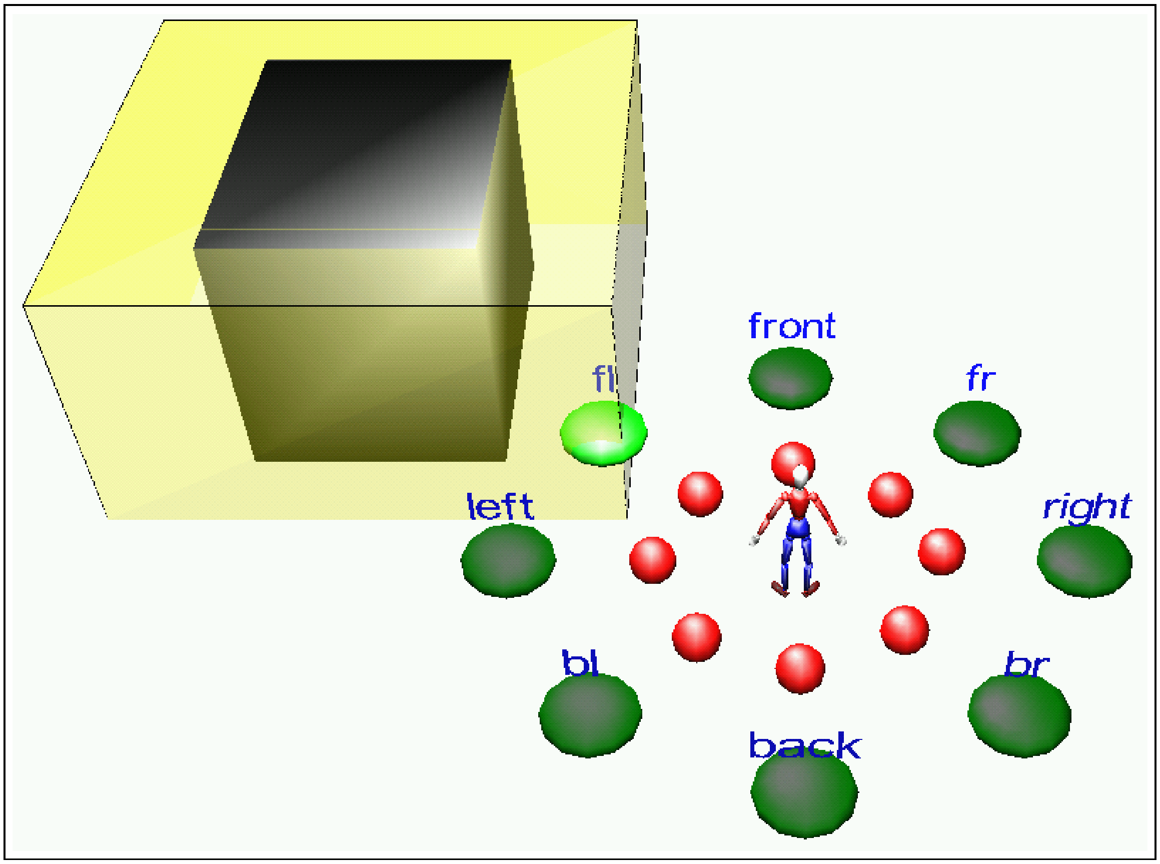

Figure 3: The virtual presenter and its sensors which establish the egocentric frame of reference. The far front-left sensor intersects with the influencing bounds of the building.

are in fact invisible bounds objects used for sensing objects in the current environment. This is done by intersection tests of sensors with activation bounds/influencing bounds of objects in the environment. If a sensor intersects with an object's bounds, the virtual scout will be able to query information about the object, using the object's annotations (see section: 5) . The virtual presenter's sensors can be used to establish an egocentric frame of reference. and we are able to obtain a qualitative representation of space[2]. The reference system in figure 3, e.g., distinguishes two distance relations: *near* and *far*. It differentiates between eight primary directions. For each object referred to with respect to this coordinate system the object's distance and directions to the origin of this system will be discretized in 2 levels of distances and 8 values for direction. This is a transformation from a quantitative representation of an object's location to

[2]For a discussion of different representations of space see [5].

a qualitative representation. For example ("building–134" 6.0 45) becomes ("building–134" far front-left) which can easily be transformed into a description using building–134's linguistic annotations: "There is a white house in front of you on your left." When we transform quantitative representations into qualitative ones and verbalize them, we loose information. The localization will become vague. This loss of information can be compensated for by highlighting, labeling the object or drawing an arrow pointing to the object using the object's annotations.

6.2 The Presenter's Behaviors

Information about the location of objects in the environment from the presenter's point of view can be used to enable the presenters movement and pointing behaviors. In addition to these abilities the virtual presenter itself has to meet several requirements. According to its functional roles in the presentation, it must be "familiar" with a range of presentation gestures and rhetorical body postures (e.g. standing upright) and should adopt a reasonable and lively behavior without being distracting or obtrusive. Here we propose a high level declarative specification of the presenter's top level behaviors: *3D-show-trajectory-segment(segment), 3D-show-object(building)*. We use these top level behaviors to automatically generate the virtual presenter's animation. For a purely synthetic actor (i.e. not controlled by a human) the system must generate the right sequence and combination of parameters to create the desired movements. In order to achieve this, we propose a hierarchical structure of behaviors where our top level behaviors are assembled from other behaviors.

- 3D-show-trajectory-segment (segment):

  This is a navigational behavior which enables the presenter to follow a path computed by a search algorithm. It is scheduled by the bounds objects of streets. This behavior should make the virtual presenter's movements appear smooth, continuous and collision free. To achieve these natural movements the virtual presenter's navigational behavior generates a spline that interpolates the discretized path from the virtual presenter's current location in the environment through a list of successive control points to the target destination. The computation of control points can be done with the help of functional annotations of the street objects. The navigational behavior consists of different other hierarchically structured behaviors, for example behaviors for leg movement.

- 3D-show-objects(building):

  This behavior consists of a *look-at(building)* and/or a *point-at(building)* behavior. These behaviors will be scheduled by the object's influencing bounds. Therefore, we can assume that the presenter is near enough to the object so that he can clearly point/look at the intended object. The point-at behavior is built from different other behaviors, e.g. to move the presenter's arm, a slight rotation of the body, if necessary. The look-at behavior consists of a rotational movement of the presenter's head and eventually there is also a need for a body rotation. These two behaviors enable the scout to point and/or glance at an object. The point to look and/or point at can be computed using the appropriate function from the object's functional annotations.

It should be stated that all the actions constituting the different behaviors could be done in parallel and at different speeds. Knowing minimum and maximum speeds for the different behaviors/actions and the locations of objects in the environment from the presenter's point of view allows us to specifiy only an abstract presentation plan (e.g. 3D-show-trajectory-segment (segment)) and the scout will perform the task alone without system intervention, by querying the environment using the bounds concept to establish an egocentric frame of reference and to trigger behaviors.

7 Conclusion and Future Work

In this paper we have introduced the concept of bounds objects to specify certain spatial areas around objects in the environment suitable to model applicability areas or influence areas in the computation of spatial relations. In addition bounds objects can be used as sensors in localization tasks and to trigger the virtual scout's behaviors. The methods used to establish the objects' frames of reference and to derive information about simple spatial prepositions are based on former empirical work that has been carried out in this field (see e.g. [18]) . In order to apply more complex prepositions (e.g. path relations, such as "along") to our scenario, we have recently carried out psychological experiments that are still under evaluation. Hopefully, the results will help to improve the selection of the presenters's path, the computed spatial relations, so that a viewer can memorize the shown information more easily.

8 Acknowledgments

The authors are members of the Collaborative Research Center 378 "Resource adaptive cognitive processes" partially funded by the DFG (Deutsche Forschungsgemeinschaft).

References

[1] E. André and T. Rist. Coping with temporal constraints in multimedia presentation planning. In *Proceedings of the 13th National Confrence on Artificial Intelligence*, pages 142–147, 1996.

[2] A. D. Baddeley. *Working Memory*. Oxford University Press, Oxford, 1986.

[3] Andreas Butz. Anymation with CATHI. In *Proceedings of AAAI/IAAI '97*, pages 957–962. AAAI Press, 1997.

[4] P. Doyle and B. Hayes-Roth. Agents in Annotated Worlds. Technical Report KSL 97-09, Kowledge Systems Laboratory, Department of Computer Science Stanford University, 1997.

[5] C. Freksa, C. Habel, and K.F. Wender, editors. *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*. Springer, 1998.

[6] K.-P. Gapp. Objektlokalisation: Ein System zur sprachlichen Raumbeschreibung. Studien zur Kognitionswissenschaft. Deutscher Universitätsverlag, Wiesbaden, 1997.

[7] Günter Hake. *Kartographie 1*. Walter de Gruyter, Berlin, 1982.

[8] T. Herrmann, K. Schweizer, G. Janzen, and S. Katz. Routen– und Überblickswissen. Konzeptuelle Überlegungen. Bericht Nr.1 des Mannheimer Teilprojekts 'Determinanten des Richtungseffekts' im Schwerpunktprogramm 'Raumkognition', Universität Mannheim, Lehrstuhl Pschychologie III, 1997.

[9] D. Kettani and B. Moulin. A spatial model based on the notions of spatial conceptual map and object's influence areas. In C. Freksa and D.M. Mark, editors, *COSIT 99 :Spatial information theory: cognitive and computational foundations of geographic information science*, pages 401–417. Springer, 1999.

[10] Antonio Krüger. *Automatische Abstraktion in 3D-Graphiken*. PhD thesis, Universität des Saarlandes, 1999.

[11] W. Maass. *Von visuellen Daten zu inkrementellen Wegbeschreibungen in dreidimensionalen Umgebungen: Das Modell eines kognitiven Agenten*. PhD thesis, Universität des Saarlandes, 1996.

[12] W. Maass and D. Schmauks. MOSES: Ein Beispiel für die Modellierung räumlicher Leistungen durch ein Wegbeschreibungssystem. *Zeitschrift für Semiotik 20*, pages 91–103, 1998.

[13] D. Schmauks. Kognitive und semiotische Ressourcen für die Wegfindung. *Kognitionswissenschaft, Sonderheft zum Sonderforschungsbereich (SFB) 378*, 1998.

[14] S. Towns, J. Voerman, C. Callaway, and J. Lester. Coherent G estures, Locomotion, and Speech in Life-Like Pedagogical Agents. In *Proceedings of the 1998 international conference on Intelligent user interfaces*, pages 13–20, 1998.

[15] B. Tversky. Cognitive Maps, Cognitive Collages, and Spatial Mental Models. In *Spatial Information Theory. A Theroretical Basis for GIS, COSIT'93*, pages 14–24, 1993.

[16] B. Tversky and P. U. Lee. How Space Structures Language. In C. Freksa, C. Habel, and K.F. Wender, editors, *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*, pages 157–177. Springer, 1998.

[17] Wolfgang Wahlster and Werner Tack. Sfb 378: Ressourcenadaptive kognitive prozesse. In Matthias Jarke, editor, *Informatik '97: Informatik als Innovationsmotor*, pages 51–57, Berlin, 1997. Springer.

[18] H.D. Zimmer, H.R. Speiser, J. Baus, A. Blocher, and E. Stopp. The use of locative expressions in dependence of the spatial relation between target and reference object in two–dimensional layouts. In C. Freksa, C. Habel, and K.F. Wender, editors, *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*, pages 223–240. Springer, 1998.