Frames of Reference, Positional Information and Navigational Assistance

Jörg Baus and **Christian Kray**
Universität des Saarlandes, P.O. Box 15 11 50, 66041 Saarbrücken, Germany,
tel. +49 (681) 302 4137, fax +49 (681) 302 4136, baus@cs.uni-sb.de
DFKI GmbH, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany,
tel. +49 (681) 302 2865, fax +49 (681) 302 4136, kray@dfki.de

Abstract

We present an analysis of different types of frames of reference and what they are needed for in a setting of a system providing navigational assistance. Since the position of a user of such a system is a main factor in the establishment of different frames of reference, we review different technologies for measuring positional information, and present a qualitative classification for this information. Based on this and on the previous analysis, we then examine the influence of positional information and frames of reference on typical tasks that arise in the context of navigational assistance.

Introduction

Designing adaptive mobile navigation aids that extend beyond those used in modern cars, is a great challenge. Typically, different means of transportation have to be combined in order to reach a destination. Additionally, the user's situation is changing constantly, e. g., her position and orientation. A mobile navigational assistant should therefore adapt to the user's changing situation and generate navigational instructions that are easily understood, or provide the user with further information about her current environment. There are several typical tasks that have to be addressed in the context of navigational assistance, such as the identification of buildings or objects in the user's vicinity, or providing background information about these. Furthermore, a mobile assistant should be able to localize arbitrary objects, which is a necessary precondition for the generation of navigational instructions. All these task rely on some fundamental spatial tasks, e.g., the establishment of a frame of reference, and the computation of spatial relations. In addition, a system has to cope with varying positional and orientation information, which may vary greatly in terms of quality and availability in a real world setting.

This paper reviews the connections between tasks, positional information, and frames of reference in a scenario of a mobile navigational assistant. In the following section we present our notion of frames of reference. We then examine what types of positional information can be distinguished before we apply the results to the tasks arising in a navigational system. Finally, a short conclusion summarizes the main points and presents an outlook on future research.

Frames of Reference

Reasoning about space is a fundamental human ability enabling us, e. g., to generate verbal or graphical route descriptions that help an addressee to construct a cognitive map of the environment or to reach a given goal. A central component in most spatial reasoning processes is the consideration of spatial relationships between different objects. In order to unambiguously specify the location or direction of objects, a frame of reference (FOR) is required that relates the location of an object to some other location or direction. Following (Frank 1998) a reference system or a frame of reference is specified by three characteristics: the origin of the coordinate system (which is independent of the kind of coordinate system used), its orientation and its handedness (which establishes the relation between the axes).

Locations and directions can be specified either in allocentric, intrinsic or egocentric frames of reference, which are either global or local according to the relevant space. Allocentric frames of reference use a fixed coordinate system imposed by external factors such as geographic space, and are independent from the observer's or addressee's position. In navigational instructions, they are mainly used to transfer survey knowledge, e. g., information about regions or the structure of the environment. In an allocentric frame of reference, elements of the scene are referred to according to a survey perspective of the environment, i.e. "Go east through the hallway" (see also (Tversky & Lee 1998; 1999; Werner *et al.* 1997)).

Intrinsic frames of reference rely on inherent properties of an anchor object (reference object), e. g., topology, size, or shape in order to determine the coordinate system. Buildings or building sections, such as shops or offices, often have an intrinsic or prominent front, which is frequently defined by the entrance. Egocentric frames of reference are used to specify locations or directions according to the location and perspective of the observer or addressee. They are defined with respect to the body's intrinsic axis.

Verbal descriptions of a motion path through an environment most frequently rely on an egocentric frame of reference. They encode the path to follow using landmarks as well as spatial relations between objects in the current environment. The addressee undertakes a mental journey during which objects in the environment are localized in relation to her current position or to each other from an egocentric point of view. This view, sometimes called route perspective (cf. (Tversky 1993)) or field perspective (cf. (Schweizer *et al.* 1998)), helps to convey knowledge about path segments and landmarks (route knowledge). The agent's egocentric frame of reference is a special case of an intrinsic frame of reference, where the agent defines its origin (see (Klatzky 1998)).

In addition, frames of reference can be induced: the addressee is asked to reorientate, turn around, perform a (mental) relocation, or combination of these actions. Induced frames of reference can be applied either in an egocentric or allocentric setting. An egocentric use might be, e. g., "Turn around until you see the church; then, the book shop is to your left", whereas an allocentric use might result in a description such as: "Turn around until you see the book store; the public telephone is in front of the shop." In the later case we assume that the agent is in front of a building and that the building's intrinsic front provides an allocentric frame of reference with its own reference direction or orientation. At least in the case of an egocentric frame of reference (but also partially in the case of induced frames of reference), the current location and orientation of the user is of central importance. The following section therefore reviews different means for obtaining positional information.

Positional Information

Our notion of positional information does not only include the absolute or relative *location* of the user, but also her *viewing direction* and *body orientation*. Additional relevant positional parameters are the *speed* and *acceleration* of the user's movements. We assume that for each of these parameters a *sensor* exists that produces a permanent stream of data. One important observation is that there is currently no technology that is capable of delivering highly accurate measurements on all parameters at any time. Therefore, applications with a more general purpose, e. g., an electronic tourist guide, have to rely on multiple sensors to determine the positional information. Fortunately, various technical solutions are available to detect positional information.

The most common system is the satellite-based *global positioning system* (GPS) that uses the runtime difference between satellite signals to determine the user's current location. The accuracy of GPS depends on the number of satellites, whose signals are received simultaneously. Since the system needs to "see" the satellites, GPS does rarely function inside buildings, and can also be unreliable during bad weather conditions or in dense vegetation. Signal deflections pose a further problem within urban areas. A similar approach uses the network cells of cellular phone companies, and does usually also reach indoor areas. The problem here is that precision depends on the network cell size, which may vary from 500 meters to several kilometers. Radio technology like *Wavelan* and more recently *Bluetooth* have much smaller cell sizes resulting in higher accuracy. However, they depend on a fully developed infrastructure of senders to work properly, which is only available in very few areas – unlike GPS or GSM, which cover most of the planet.

Other sensors for tracking locations indoor are based on *infrared transmitters* (Hartner & Hoppner 1994; Want *et al.* 1995; Butz *et al.* 2001)). Since infrared light does not pass through walls, the cells can be used to distinguish reliably between different rooms or parts of a room. Infrared cell sizes may vary from a few to several meters. In order to obtain higher accuracy, ultrasound in combination with radio signals can be used (e. g., the Cricket system (Priyantha, Chakraborty, & Balakrishnan 2000)). Unfortunately, both radio and ultrasound signals suffer from multipath problems, which can only be addressed using sophisticated correction algorithms.

Image recognition is another means to obtain positional information. It can be used to detect landmarks in the surroundings that help to determine the actual location from video images (Chiorboli & Vecchi 1993). Other direct tracking approaches rely on visual (or electromagnetic) scanning of tags or transponders. This allows, for example, to detect a person entering or leaving a certain room or area. The main drawbacks of image recognition are the high cost in terms of computational power and memory footprint as well as the need for a camera. In case of electromagnetic tags, a large number of these devices have to be installed in the environment.

Electromagnetical sensors are often used (e. g., an electronic compass) to determine the viewing direction and body orientation. The problem with these devices is their sensitivity to metallic objects in their vicinity. Some approaches therefore use infrared to determine both location and orientation (e. g., (Butz *et al.* 2001)) at a different granularity. The viewing direction can also be tracked indirectly with *accelerometers* by measuring the changes in acceleration of the user. This technique is often used in conjunction with other tracking approaches to raise the overall quality of the results. Accelerometers also facilitate the determination of the actual location by *dead reckoning*, a method that extrapolates the location from the velocity and traveling direction.

In order to handle incomplete or missing information from the sensors we assume that every sensor delivers positional information (e. g., x-, y-, and z-coordinates of the location), and a corresponding *error measure*. The error measure is a region in space that constraints the actual user location or viewing direction. The smaller the region, the more precise is the measurement. The region itself can be described as a circle, ellipse, cone or even as a polygon[1], depending on the characteristics of the sensor in use.

In Figure 1a the error of a GPS derived location is depicted as an circular area, representing the possible deviation from the current position. Darker shades correspond to higher probabilities. Infrared based tracking leads to a semicircular area (see figure 1b), reflecting the fact that infrared light is blocked by walls. The error for the viewing direction is represented by a cone. Figure 1c depicts the error for an orientation with an accuracy of 90 degrees.

This example and the list of techniques for detecting positional information illustrate that one can expect great variations in terms of quality and availability in a real world setting. Therefore, it makes sense to relate the task (e. g., localization, or navigational guidance) that requires positional

[1]For the sake of simplicity we use 2D-representations. The extension to 3D is straightforward.

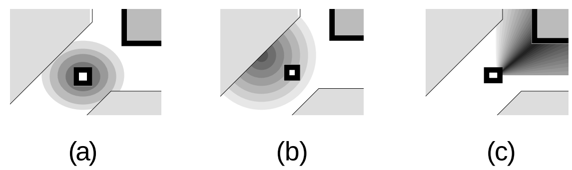

Figure 1: Error regions resulting for different sensors

information of a certain type to the information itself. This leads to four qualitative categories for positional information:

- **precise information**
  The information provided by the sensors and/or reasoning processes is sufficiently precise to solve the task at hand. In this case, no further work is needed, and the task can be solved directly.
- **unprecise information**
  Positional information of the required type is only available at a quality level (e. g., in terms of kilometers instead of meters) that is insufficient to address the current task. Consequently, the task cannot be solved. However, it may be possible to either increase the precision of the measured data (e. g., using dead reckoning), or to adapt the task to work with the available measurements (e. g., by computing coarser localizations).
- **no information**
  There may be situations when no information is available, e. g., when sensors are broken, or fail to gather any reading (see previous section). In this case, the task at hand cannot be addressed. While it may still be possible to apply reasoning mechanisms to infer positional information or to adapt the task, it will be much harder than in the case of unprecise information.
- **false information**
  Sensors or devices may return false information, e. g., electronic compasses close to large structures made of metal. If this case is not detected, the task at hand might fail without being noticed. If detected, this case is equivalent to the total lack of information.

Grouping information according to this categorization facilitates the analysis of the relationship between frames of reference, positional information, and tasks related to space, which we present in the following section.

Navigational Assistance

When applying spatial reasoning techniques in a real world setting such as a mobile navigational assistant, obviously additional factors (compared to theoretical applications) have to be taken into account. This includes factors related to the user (such as the age of the user or her familiarity with the environment) and to the context (such as weather conditions or the means of transportation). Resource consideration (e. g., the cognitive load of the user, or the available computational power (Baus *et al.* 2001)) are equally important as well as the availability and quality of positional information. The later point is of central importance since it does not only influence, which frames of reference and consequently which relational expressions are possible, but also determines the spatial context.

Position	Allocentric	Egocentric	Induced
location	$\odot$	$\otimes$	$\odot$
orientation	$\odot$	$\otimes$	$\bigcirc$
speed	$\bigcirc$	$\bigcirc$	$\bigcirc$
acceleration	$\bigcirc$	$\bigcirc$	$\bigcirc$

Table 1: Positional information required to establish a FOR; $\bigcirc$ = not required, $\otimes$ = required, $\odot$ = possibly required

FOR	Relations			
	topological	angular	distal	path
origin	$\bigcirc$	$\bigcirc$	$\otimes$	$\odot$
orientation	$\bigcirc$	$\otimes$	$\bigcirc$	$\bigcirc$
handedness	$\bigcirc$	$\otimes$	$\bigcirc$	$\bigcirc$

Table 2: FOR required by different types of relations; $\bigcirc$ = not required, $\otimes$ = required, $\odot$ = possibly required

Table 1 lists different types of positional information, and relates them to various FOR. While the establishment of a FOR usually requires the presence of an origin, its orientation, and its handedness[2], it does not per se depend on the availability of information about the user's current position. However, in the context of navigational assistance, there are connections between positional information and different frames of reference. Consider, for example, the allocentric case: its origin is determined by an external object, but it is possible to define its orientation using the user's body axis (e. g., when the anchor object does not have an obvious intrinsic front). Then, the user's orientation has to be known precisely as well as her (rough) location, in order to impose her orientation on the allocentric FOR. The establishment of an induced FOR can be achieved without any positional information, i. e., when both location and orientation are induced. If that is not the case, missing factors have to be gathered in order to successfully induce the FOR. For example, if the user's location is known but her orientation is not, then an induced FOR with a single reorientation instruction can be established ("If you turn towards the church, the fountain is to your right.") Unlike the previous cases, an egocentric FOR requires the availability of both locational and orientational information at a high degree of precision, since these factors determine the origin and orientation of the resulting FOR.

Positional information is available at varying quality in a real world setting, and it does have an impact on which frames of reference are possible in a specific situation. Table 2 lists the components of a FOR and how they are re-

[2]Handedness is not a measurable factor, and therefore has to be deduced from other sources, e. g., contextual or dialog related sources.

quired by different relations, thereby establishing a connection between positional information, its quality, and spatial relations. Topological relations, such as 'disjoint' or 'meet' (Egenhofer & Herring 1990), do not require a frame of reference, and therefore, pose no constraints on the different components of a FOR. Angular relations, such as 'left-of' (Mukerjee 1998), on the other hand depend on the presence of orientation and handedness; however, they do not necessarily need an origin as they depend on angular disparity, which can be computed between any two vectors. Distal relations, such as 'far-from' or 'close-to' (Gapp 1994), are independent of orientation, but require the presence of an origin. Path relations, such as 'along' or 'past' (Kray & Blocher 1999), in general do not depend on a FOR. However, recent results (Kray *et al.* 2001) provides evidence, that the distance between anchor and target object is of relevance. In summary, positional information and frames of reference determine the set of spatial relations that is available in a given situation.

An additional field (besides relations), where positional information and the chosen FOR play an important role is the selection of output media and modi. For example, if the quality of positional information only allows for an induced FOR, then one or more (mental) turning or repositioning actions have to be communicated. In this case, textual or verbal output are advantageous since it is hard to realize such instructions using graphical output or animations. Although there are means to achieve this, e. g., arrows or certain camera flights, no generally accepted set of means is defined in a semantically unique way. If it is possible to establish an allocentric FOR then all media and modi are equally well suited since there is no need to encode any turning or repositioning actions beforehand. The same is true for an egocentric FOR. However, in this case a virtual flight-through (possibly including a virtual presenter) is more appropriate, as it can be aligned to the user's FOR resulting in animations that closely match the user's experience.

Typical tasks in navigation (e. g., localization, identification) define a third area, where positional information and the selected FOR have a great impact. A resource-adapting navigation system generating graphical way descriptions was implemented that takes into account these influences. The system consists of three major components. Firstly, an information booth (a 3D-graphics workstation), where a virtual walk-through through the environment can be generated including a virtual presenter, spatial utterances and meta-graphics, which complement each other. Secondly, an indoor navigation system has been build based on strong infrared transmitters and a small PDA, which can display simple sketches of the environment received via infrared. The third component is an outdoor navigation system that uses a small laptop in combination with a head-mounted display. A GPS system determines the users actual position and an electronic compass tracks the user's orientation. A detailed description of the system can be found in (Baus *et al.* 2001; Baus, Krüger, & Wahlster to appear). All graphical way descriptions that are generated by the system, are computed from scratch according to the cognitive resource limitations of the user and the technical constraints posed by the output

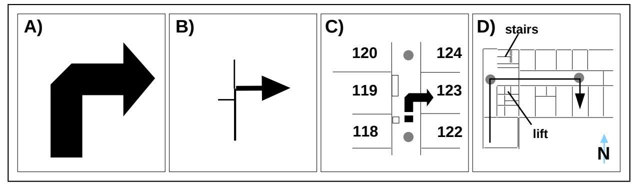

Figure 2: Four different resource adapted graphical way descriptions used by the indoor navigation component.

device. A single 3d-model of the environment is used to produce walk-throughs at the information booth and sketches for the mobile use. The system is able to adapt its services, e. g., by choosing the camera perspective, by selecting appropriate paths, by adjusting the number of landmarks depicted, and by including interactive areas in its presentation. The following example highlights how varying information about the user's location and orientation influences output behavior of the indoor navigation-systems .

When the system knows the actual position and orientation of the user precisely, it can produce a very simple instruction (i. e., an arrow as shown in Figure 2A). If the quality of the orientation information is too unprecise and the system cannot exactly tell where the user is looking at, a simple arrow could mislead the user. Therefore, the task has to be adapted and additional information has to be provided about the choices available at the decision point. Figure 2B shows such a graphical way description for an orientation resolution of +/-90 degrees. The topological diagram includes only the different choices at the current decision point, but does not show any additional landmarks. Please note that the map can still be roughly aligned to the user's walking direction to simplify her reorientation.

If the quality of orientation and the position information further deteriorates (and becomes too unprecise to address this task in the aforementioned way), landmarks can be included to compensate. Figure 2C shows a description, where the position resolution covers three potential decision points (two are indicated as grey dots). In such situations a purely topological map could cause problems and therefore an appropriately clipped area of the surrounding (here: the adjacent rooms with numbers, parts of the hallway, pillars and a locker) are displayed. By clicking on the grey dots the user can inform the system about her actual position and resolve the ambiguity of location, thus allowing the system to switch back to the topological presentation of Figure 2B (since the user did provide more precise positional information). In the worst case there is only very rough or no positional information at all and the system cannot align the map to the user's actual walking direction anymore. Now, the task has to be further adapted, and a greater portion of the map has to be chosen that may include several turns, especially those that the user has already passed (see Figure 2D). Instead of including small landmarks that are only relevant at a single decision point, global landmarks, such as stairs or elevators have to be included in the presentation. Since it is important to explain to the user that she cannot rely on the orientation of the map, the presentation contains a North Symbol to underline the allocentric frame of reference. Again, the user can communicate her position to the system by clicking on the grey dots, thereby increasing the quality of positional information and enabling the system to generate a closeup of that area of the building. But in order to align the map to the walking direction, the system has to ensure the user's correct orientation. This task can be accomplished by means of an induced frame of reference, i. e., by advising the user to reorient herself towards a landmark (e. g., "Turn around until the stairs are to your left and the lift is to your right; then walk straight on.")

Conclusion

In this paper, we presented an analysis of frames of reference, positional information, and their interdependencies. We examined what types of frames of reference and positional information can be distinguished, and we reviewed their impact in a setting of navigational assistance. Three areas were highlighted in order to illustrate this influence: spatial relations, media and modi selection, and adaptation of typical tasks in this context. We then presented an example on how these considerations were applied to a resource-adaptive indoor navigation system. In the future, we plan to investigate further means to enable a system to adapt to varying conditions in terms of information quality and resource availability. One point that we intend to focus on is the adaption to incomplete or unprecise positional information by means of contextual reasoning and/or interaction with the human user.

Acknowledgments

The work presented in this paper is funded by the DFG and the Klaus Tschira Foundation, Heidelberg.

References

Baus, J.; Kray, C.; Krüger, A.; and Wahlster, W. 2001. A resource-adaptive mobile navigation system. In *Proceedings of International Workshop on Information Presentation and Natural Multimodal Dialog*.

Baus, J.; Krüger, A.; and Wahlster, W. to appear. A resource-adaptive mobile navigation system. In *Proceedings of IUI2002: International Conference on Intelligent User Interfaces 2002*. New York: ACM Press.

Butz, A.; Baus, J.; Krüger, A.; and Lohse, M. 2001. A hybrid indoor navigation system. In *IUI2001: International Conference on Intelligent User Interfaces*. New York: ACM.

Chiorboli, G., and Vecchi, G. 1993. Comments on "design of fiducials for accurate registration using machine vision". *T-PAMI* 15:1330–1332.

Egenhofer, M. J., and Herring, J. R. 1990. A mathematical framework for the definition of topological relations. In Brassel, K., and Kishimoto, H., eds., *Fourth International Symposium on Spatial Data Hadling*.

Frank, A. 1998. Formal models for cognition - taxonomy of spatial location description and frames of reference. In Freksa, C.; Habel, C.; and Wender, K., eds., *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*. Springer. 293–312.

Gapp, K.-P. 1994. Basic meanings of spatial relations: Computation and evaluation in 3d space. In *Proceedings of AAAI-94*.

Hartner, A., and Hoppner, A. 1994. A distributed location system for the active office. *IEEE Network* 8(1).

Klatzky, R. 1998. Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections. In Freksa, C.; Habel, C.; and Wender, K., eds., *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*. Springer. 1–17.

Kray, C., and Blocher, A. 1999. Modeling the basic meanings of path relations. In *Proceedings of the 16th IJCAI. Morgan Kaufmann, San Francisco, CA*, 384–389.

Kray, C.; Baus, J.; Zimmer, H.; Speiser, H.; and Krüger, A. 2001. Two path prepositions: Along and past. In Montello, D. R., ed., *Spatial Information Theory (Proceedings of COSIT 2001)*. Berlin, Heidelberg, New York: Springer.

Mukerjee, A. 1998. Neat vs scruffy: A survey of computational models for spatial expressions. In Olivier, P., and Gapp, K.-P., eds., *Computational Representation and Processing of Spatial Expressions*. Kluwer Academic Press.

Priyantha, N.; Chakraborty, A.; and Balakrishnan, H. 2000. The cricket location-support system. In *ACM/IEEE International Conference on Mobile Computing*. ACM Press.

Schweizer, K.; Herrmann, T.; Janzen, G.; and Katz, S. 1998. The route direction effect and it's constraints. In Freksa, C.; Habel, C.; and Wender, K., eds., *Spatial Cognition – An interdisciplinary approach to representation and processing of spatial knowledge*. Springer. 19–38.

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

Tversky, B., and Lee, P. U. 1999. Pictural and Verbal Tools for Conveying Routes. In Freksa, C., and Mark, D., eds., *COSIT 99: Spatial Information theory: cognitive and computational foundations of geographic information science*, 37–50. Springer.

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

Want, R.; Schilit, W.; Adams, N.; Gold, R.; Petersen, K.; Ellis, J.; and Weiser, M. 1995. An overview on the parctab ubiquitous computing experiment. *IEEE Personal Communications* 2(6):28–43.

Werner, S.; Krieg-Brückner, B.; Mallot, H.; Schweizer, K.; and Freksa, C. 1997. Spatial cognition: The role of landmark, route, and survey knowledge in human an robot navigation. *Informatik'97: Informatik als Innovationsmotor*.