

Cognitively-Inspired Direction Giving

by

Gary Wai Keung Look

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

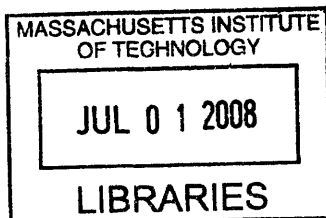
June 2008

© Massachusetts Institute of Technology 2008. All rights reserved.

Author
Department of Electrical Engineering and Computer Science
May 23, 2008

Certified by
Howard E. Shrobe
Principal Research Scientist
Thesis Supervisor

Accepted by
Terry P. Orlando
Chair, Department Committee on Graduate Students



ARCHIVES

Cognitively-Inspired Direction Giving

by

Gary Wai Keung Look

Submitted to the Department of Electrical Engineering and Computer Science
on May 23, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Computer Science and Engineering

Abstract

Online mapping services and portable GPS units make it easy to get very detailed driving directions. While these directions are sufficient for an automaton to follow, they do not present a big picture description of the route. As a result, while people can follow these detailed turn-by-turn directions, it can be difficult for them to actually comprehend where they are going.

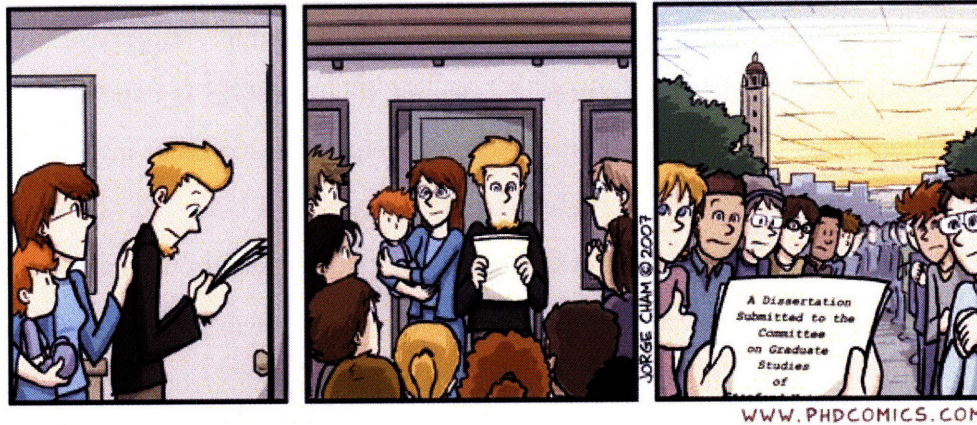
Our goal is to make such directions more comprehensible. Our approach is to apply findings from human spatial cognition, the study of how people conceptualize and organize their knowledge of large-scale space, to create a system that generates written route overviews. Route overviews provide a big picture description of a route, and are intended to supplement the information in turn-by-turn directions. Our route overviews are based on cognitively-inspired design criteria such as: the use of spatial hierarchy, goal-directed descriptions, selective suppression of detail, and the use of the trunk segments and cognitive anchor points along the route.

In our experiments, we show that we can make directions more comprehensible — independent of the particular places a person knows — by using what we know about how people think about space to structure the way we present spatial information.

Thesis Supervisor: Howard E. Shrobe

Title: Principal Research Scientist

Acknowledgments



I've been looking forward to writing this part of my thesis for a *long* time and I was excited to find the above comic to open this section. A small army of people helped me complete my thesis and it's my pleasure to acknowledge them here.

I am greatly indebted to my thesis advisor, Howard Shrobe. He was the one who first got me thinking about this problem of making driving directions more comprehensible and he provided encouragement and guidance when I would get lost. I also want to thank the other members of my thesis committee, Patrick Winston and Randy Davis. Patrick forced me to hone my ideas and Randy gave me invaluable feedback as I was writing this thesis. Randy's comments were incredibly incisive and they helped me present my ideas in a clear and convincing manner.

When I set out to implement my ideas, I received a lot of support and feedback from friends in different parts of MIT. My friends in Project AIRE — Aaron Adler, Harold Fox, Krzysztof Gajos, Tyler Horton, Kimberle Koile, Alice Oh, Stephen Peters, Kevin Quigley and Max van Kleek — helped me get through the day-to-day grind of being in grad school. I want to especially recognize Stephen for all the technical help he provided and Kimberle for her advice on getting through grad school; in this sense, Kimberle was like a second advisor. TIG, the infrastructure group at MIT CSAIL, provided invaluable help with my computing needs. In particular, Frank Tilley and Anthony Zolnik were always willing to bring more magic to bear when the situation called for it.

The cross-discipline nature of my work led me to work with a lot of people outside of MIT CSAIL. I want to thank two wonderful people at MIT's GIS Lab, Daniel Sheehan and Lisa Sweeney, for their help with ArcGIS. Daniel's expertise saved me countless amounts of time and stress in learning what ArcGIS is capable of doing. David Singer, from Brain and Cognitive Science, helped me structure my user studies in such a way that I could meaningfully analyze the data I collected. Nealia Khan, from the Harvard-MIT Data Center, cheerfully helped me grind out the statistics.

Once all the work was complete and written up, the following people read through this thesis with a set of fresh eyes to catch final proofreading errors: Joshua Chang, Emily Cheng, Erin Hung, Livia King, and Christina Lee. Christina, it would have been a fun experience if we could have written our theses at the same time.

A large number of other people have checked-in with me and provided welcome company during my time in grad school. I won't be able to name them all here but I will make special mention of a few of them now. Elliot Hui provided valuable moral and prayer support during the last nine months of the thesis writing process and I was glad to listen to his perspective, as someone who has gone through the PhD process before. My Bible study from Seattle — Jimmy Wong, Tom Young, Abe Wong, Ming Ma, Stone Chin, and Dennis Lam — provided encouragement during my entire time here at MIT and this was especially welcome during the last month leading up to my defense.

My neighbors Annie Kuo and Joy Huang would take time to bring me food when I was sick or frantically working to meet a big deadline. I looked forward to the quasi-regular routine of commiserating with Jen Chu over dinner and then going back to the thesis writing while she studied for her MCATs. Meagan Chan is, quite simply, m-chan-tastic. Rooting for (and when it suited my personal interests, against) her sports teams made life much more enjoyable when neither Seattle sports teams nor my research work was going as well as I would have liked.

Mom and Dad, I'm glad you, Amy, Michael, and Jennifer will be able to see me graduate. This is a trip seven years in the making. Thank you for all the love and support along the way.

Finally, I thank God for the grace He extends day-to-day and the constant reminders of His presence through all the adventures of the past seven years.

You have made known to me the path of life; you will fill me with joy in your presence, with eternal pleasures at your right hand.

Psalm 16:11

Contents

1	Introduction	19
1.1	Conveying Route Information	19
1.2	Route Overviews	23
1.3	Cognitively-Inspired Direction Giving	25
1.4	LAIR, CAP LOC, and ROVER	29
1.5	Contributions	32
1.6	Thesis Overview	33
2	Human Spatial Cognition	35
2.1	The Image of the City	35
2.2	The Frame for the Image of the City	39
2.3	Simplifying the Image of the City	45
2.4	Routes	49
2.5	Tying It All Together	51
3	The Design and Implementation of LAIR, CAP LOC, and ROVER	53
3.1	The LAIR Spatial Representation	54
3.1.1	Places and Paths	56
3.1.2	Regions	57
3.1.3	Routes	58
3.1.4	Cognitive Design Guidelines and LAIR	59
3.2	LAIR's Implementation in GIS	59
3.3	Previous Work That Maps Cognitive Anchor Points	62

3.4	Deriving a Mental Map of Boston	65
3.5	Analysis of the Derived Mental Map	68
3.6	Populating LAIR with CAP LOC	70
3.7	ROVER	72
3.7.1	Bringing Google Earth into LAIR	74
3.7.2	Identifying the Trunk Segments of a Route	75
3.7.3	Modelling What Places a Person Knows	76
3.7.4	Structuring Route Overviews	78
3.7.5	Modifying Turn-by-Turn Directions	83
3.8	Summary	85
4	Evaluating ROVER's Output	89
4.1	Experimental Design	90
4.2	Results and Analysis	99
4.2.1	ROVER's Output Compared to Google Earth's	100
4.2.2	ROVER's Performance Across Routes and Across Different Sub- populations	105
4.2.3	ROVER's Performance for Subjects in Other Knowledge Profiles	108
4.3	General Discussion	111
5	Future Work	115
5.1	Route Overviews Revisited and Re-examined	115
5.2	Other Evaluations of Route Overviews	117
5.3	Extending Route Overviews	118
5.4	Building up LAIR	119
6	Related Work	121
6.1	Knowledge Representations of Space	121
6.2	Identifying Important Places and Regions	124
6.3	Describing Routes	125
6.4	Finding Better Routes	127

7 Contributions	131
Bibliography	133

List of Figures

1-1	Map from Apartment A in Cambridge, MA to the Peach Farm restaurant in Boston, MA.	21
1-2	This route is used to illustrate how the route overviews generated by ROVER for a particular route differ based on which streets on the route a person already knows.	30
2-1	A map of Boston, depicted using the five spatial elements Lynch identified. Image taken from Kevin Lynch’s <i>Image of the City</i> [46].	37
2-2	The New Yorker’s View of the World. This figure illustrates a “fisheye” effect in which places closer to the reference frame’s origin are seen in greater detail than those further from the origin.	42
2-3	Reno, Nevada is actually to the <i>west</i> of San Diego, California	43
2-4	The subway map on the left emphasizes the connectivity of the different subway lines. The street map on the right accurately represents the distances between stops along the Red Line. In order to emphasize connectivity, the subway map sacrifices accurate distances. Note, for example, that although the subway map suggests that stops along the Red Line are equally spaced, the distance between the Park Street and Downtown Crossing stops is actually much less than what the subway map would suggest.	46

2-5	As this world map shows, the Americas are not nearly as aligned with Europe and Africa as most people think. Most people think the Americas are located further north than they actually are (e.g., it would surprise many people to learn that Boston, MA and Rome, Italy are located at approximately the same line of latitude). Image taken from Barbara Tversky’s <i>Distortions in Memory for Maps</i> [84].	47
2-6	The subway map on the left has aligned Manhattan Island to run more north-south than it actually is. The map on the right accurately shows Manhattan as running more northeast-to-southwest.	48
2-7	A birds-eye view of open space (left) along with its representation as an axial map (right). This figure is taken from Young Ook Kim’s <i>Linking the Spatial Syntax of Cognitive Maps to the Spatial Syntax of the Environment</i> [31].	49
3-1	Participants used ArcGIS to outline the locations of prominent places in their mental map of Boston.	66
3-2	Our aggregate map of prominent places in Boston.	68
3-3	The Regions CAP LOC identifies. Regions that are not colored green are commercial regions. Some of these Regions (namely, Newbury Street and Chinatown) were selected by CAP LOC because they are commercial regions that meet CAP LOC’s second and third search criteria. Because of this, they show up on the map in two different colors. . . .	72
3-4	A comparison of the Regions CAP LOC identifies compared to the prominent places identified in our user study. CAP LOC does not identify Beacon Hill or the North End, Regions that are defined by their unique geographies, a feature CAP LOC does not consider in its search algorithm.	73
3-5	The overall process ROVER follows to produce its output.	74
3-6	Google Earth screenshot.	75
3-7	This route is used to illustrate how a person’s knowledge profile influences the route overview generated by ROVER.	79

3-8	A map of the route used to illustrate how ROVER modifies Google Earth's turn-by-turn directions.	84
4-1	Map of Route 1.	92
4-2	Map of Route 2.	94
4-3	Map of Route 3.	97
6-1	Example of a LineDrive route map (left) compared to the corresponding standard route map (right). This route goes from Cambridge, Massachusetts to Providence, Rhode Island. Note how the LineDrive map renders different parts of the route at different scales and how LineDrive straightens out the roads on the route.	126

List of Tables

3.1	Summary of the fields in each element of LAIR	60
3.2	Prominent places listed by participants from each of our four sub- groups, along with the distance to the nearest subway stop, to the nearest 10 meters. Places that overlapped the location of a subway stop have a distance of zero.	67
3.3	Route overviews for the same route, generated using different knowl- edge profiles. Significant differences in the structure of the route overviews are highlighted.	82
4.1	ROVER's and Google Earth's directions for Route 1.	93
4.2	ROVER's and Google Earth's directions for Route 2.	95
4.3	ROVER's and Google Earth's directions for Route 3.	98
4.4	Differences in mean Likert scores between ROVER and Google Earth, for each route and each statement. The differences are based on data from study participants who fit the predicted knowledge profile used to generate the route overviews. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER.	101

4.5	The statistical significance of the difference in mean Likert scores for each Statement is determined by examining the effect of direction type (ROVER vs Google Earth) on Likert scores across all routes using ANOVA. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER. The ANOVA was based on data from subjects who fit the predicted knowledge profile for all three routes (n=12).	102
4.6	This table shows which type of directions were preferred among participants who fit the predicted knowledge profile used to generate a route overview and who had a preference between the two types of directions they saw. A majority of participants preferred ROVER's output to Google Earth's. We performed a sign test to determine if this majority was statistically significant.	103
4.7	Instances where deltas (the size of ROVER's effect) differed between different sub-populations in a statistically significant manner ($p < 0.05$). Statistical significance was determined by performing independent samples t-tests on data from subjects who fit the predicted knowledge profile used to generate the route overview.	108
4.8	Differences in mean Likert scores between ROVER and Google Earth, for each route and each statement. The differences are based on data from study participants who did not fit the predicted knowledge profile used to generate the route overviews. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER.	109
4.9	The structure and chunking of information provided in ROVER's overview for Route 3 makes it easier for people to parse the directions that are presented to them. We believe this structuring is a reason why the increase in route comprehensibility is independent of people's familiarity with the specific places mentioned in the overview.	110

Chapter 1

Introduction

I can't see the forest for the trees.

1.1 Conveying Route Information

Imagine getting into your car to drive to a restaurant that you have been meaning to eat at for a while. Since you aren't certain how to get to this restaurant, you enter the restaurant's address into the car's onboard global positioning system (GPS). The GPS displays a map and then instead of giving you a list of turn-by-turn directions, the GPS does something current GPS units currently do not do: it first tells you the neighborhood this restaurant is in and the major roads you'll be travelling on to get there. Additionally, the GPS mentions a few of the other neighborhoods you will pass through on the way to the restaurant. You then go on your way, for most of the route not listening very closely to the subsequent turn-by-turn directions from the GPS because the initial overview it gave you provided you with a good sense of how to get to dinner.

In this dissertation, we discuss the challenge of communicating spatial information to people. Specifically, we look at what can be done to help people better understand the driving directions produced by GPS units and mapping services such as MapQuest and Google Maps. The most common route-travelling situation people encounter is when they are travelling to new places within the greater metropolitan area in which

they live and work. These are the situations this work is intended to address. Our approach is to apply findings from cognitive psychology to make a given route between two places easier to understand. Our goal is not to design algorithms to find a better route between those two places (in whatever sense one route might be better than another), although the principles discussed in this dissertation could be applied to finding routes on which people are less likely to get lost because these routes are more comprehensible and thus easier to follow.

As an example of how current computer-generated driving directions leave something to be desired, consider the following set of directions from Apartment A in Cambridge, Massachusetts to the Peach Farm restaurant in Boston, Massachusetts. A map is also included for reference in Figure 1-1, on the next page.

- (1) Head southeast on Pilgrim St toward Brookline Pl, 49 ft
- (2) Turn left at Brookline Pl, 469 ft
- (3) Turn right at Franklin St, 308 ft
- (4) Turn left at Sidney St, 240 ft
- (5) Turn right at Green St, 180 ft
- (6) Turn left at Blanche St, 269 ft
- (7) Turn right at Massachusetts Ave/RT-2A, 0.3 mi, 1 min
- (8) Turn left at Vassar St, 0.3 mi, 1 min
- (9) Turn right at Main St, 0.6 mi, 3 mins
- (10) Continue on Longfellow Bridge/RT-3 S. Continue to follow RT-3 S, 0.4 mi, 1 min
- (11) Turn right at Charles St, 0.3 mi, 1 min
- (12) Turn left at Beacon St, 0.3 mi, 1 min
- (13) Turn right at Park St, 0.1 mi
- (14) Turn right at Tremont St, 0.4 mi, 2 mins
- (15) Turn left at Stuart St, 0.1 mi
- (16) Continue on Kneeland St, 0.1 mi
- (17) Turn left at Tyler St, 318 ft

After reading directions like these, it is not uncommon to feel either confusion or frustration (perhaps both). Confusion can arise because a person may not be familiar with much of the route, and even after re-reading the directions several times, may not have a clear understanding of exactly where the route takes them. Frustration may then set in. Frustration can also arise for people who know the route very well, and whom, after reading the long set of directions (17 steps!), realize they spent a lot of time and effort interpreting a set of directions with which they are already mostly



Figure 1-1: Map from Apartment A in Cambridge, MA to the Peach Farm restaurant in Boston, MA.

familiar. A common reaction from people in this category is, “Oh, I already know how to get most of the way to Peach Farm; why can’t the directions take that into account when telling me how to get there?”

Note that GPS units, which can be very useful to people when they are in the car navigating this route, are not that useful to people before they start driving, when they are trying to make sense of the route. GPS units tout the fact they have millions of landmarks¹, also known as points of interest. These landmarks include places such as shopping malls, restaurants, and entertainment venues. However, these landmarks would not be as helpful in either of the scenarios mentioned above.

In both of these scenarios, providing more information in the detailed turn-by-turn directions would only complicate the process of developing an overall sense of the route. The challenge in both these scenarios lies in describing the route using appropriate information based on what people are trying to accomplish at the time. For a person who is already familiar with most of the route, the added information would be redundant (and would only add to his frustration). For the person who is not familiar with the route, adding landmarks at each turn would be overwhelming when he is looking at all the directions at once and trying to build an overall understanding

¹Three million landmarks is considered “not luxurious, but very passable” [52].

of where the route takes him.

Even though the directions are detailed enough for an automaton to follow, it is this very same precision that can make the directions difficult or frustrating for a person to interpret. Ironically, what is necessary information for an automaton may not be necessary for a person, and what is necessary information for a person is not necessary for an automaton. When presenting route information to people, detailed turn-by-turn directions do not take into account the limitations of human memory or the organizational schemas people use to make sense of large scale space.

In this dissertation, we discuss what these limitations and organizational schemas are, and demonstrate how they can be used to better structure the presentation of written route directions. We use findings from cognitive psychology and urban planning to guide our design of a computational model for direction giving. Our primary focus will be on how these organizational schemas, which are independent of the specific places a person knows, can be used to improve route comprehensibility. A secondary focus that we will discuss is how our model is used to represent what particular places a person knows and how this specific knowledge affects direction giving and route comprehensibility.

To make detailed turn-by-turn directions easier to understand, we preface these directions with a general overview of the route. Route overviews are not intended to get you to your destination by themselves and as such, they are not meant to replace turn-by-turn directions. The primary intent of route overviews is to provide a different perspective on driving directions and thus, they should be regarded as a complement to turn-by-turn directions. We will, however, discuss how the information in route overviews can be incorporated into the actual turn-by-turn directions in Chapter 3. These overviews are similar to how some people would describe a route. For example, consider the following route overviews, given by two different people who were asked to give directions from Apartment A to Peach Farm:

Person 1: Longfellow, right on Charles street, circle around Common.

Person 2: Take the Longfellow bridge into Boston, and go down Charles St, and then wrap around the Common to get to Tremont.

The overview our system produces for the route reads as follows:

The Peach Farm Restaurant is in Chinatown. To get there, take Massachusetts Avenue to Main Street and cross over the Longfellow Bridge. Then go down Charles Street toward the Boston Common, and around the Common to Tremont Street. Then take Tremont to Kneeland to get into Chinatown.

In producing these route overviews, the approach we take is that we can provide more understanding by prefacing turn-by-turn directions with something that has less information (compared to those directions). Route overviews have less detail and present the route at a coarser level of granularity than the turn-by-turn directions produced by MapQuest and Google Maps. Paradoxically, while overviews such as these include fewer elements in the route description, they are able to provide a clearer mental image of the route as a whole, thus complementing the turn-by-turn directions and making them easier to understand.

1.2 Route Overviews

Let's take a closer look at the route overview presented in the previous section and compare it to its corresponding turn-by-turn directions. What are the differences, and how do these differences affect our understanding of the route?

One notable difference between the route overview and the turn-by-turn directions is the route overview's use of neighborhoods and prominent places² to situate both the location of the restaurant (as being in Chinatown), as well as to anchor a certain portion of the route. The use of neighborhoods and prominent places provides a structure not found in standard turn-by-turn directions. This structure is important because if we regard travel as a plan, then travel to these places can be regarded as important subplans in a means-ends-analysis decomposition of that plan.

This structure is particularly useful to people familiar with Boston. These people will think about these places in the context of other places. When they read the route

²We will explain the decision to use the term "prominent place" instead of "landmark" momentarily.

overview, they can conjure mental images of Chinatown and the Boston Common in isolation, but they can also think about the two places in relation to one another — they can think about the streets leading to, from, and connecting the two places, as well as other adjacent neighborhoods. Although there are only two specific places mentioned in the route overview, these two places invite the person who is familiar with the area to bring to mind a number of other places and paths to form a more detailed mental map.

On the surface, a route overview may convey less information compared to its corresponding turn-by-turn instructions. However, as the above example illustrates, a judicious choice of streets and places to include in the route overview will leverage the implied knowledge that a traveller has. In this work, we created a cognitively-inspired computational model of the spatial knowledge a person has and how this knowledge is organized. We then demonstrate the power this model can bring to the problem of making driving directions more comprehensible to people.

In the evaluation of our route overviews, we first wanted to determine if, controlling for what places a person is familiar with, our route overviews would make driving directions more comprehensible. The rationale in these situations is that we can increase comprehensibility by appealing to the person's familiarity with the area. We found that this was the case, but in addition to that, we demonstrate a more important benefit. Since our route overviews are based on general organizational schemas people use to make sense of large scale space, even those who are not familiar with the area through which a route passes benefit from the structure provided by the route overview.

Now let's turn from what information the overview has to what information it doesn't have.

1. Route overviews do not include distances. Instead, to indicate how far a person would travel, route overviews break the route into chunks, delineated by cross-streets and neighborhoods. Travel is described not in terms of distance, but as movement between episodes of space.

2. Route overviews may not include turn directions. When turning from one street to another, the direction in which to turn may not be mentioned. Instead, in keeping with the idea that travel is between episodic chunks, a route overview uses well-known places and paths as intermediate subgoals that a person heads towards. Turns are viewed not only as transitions from one street to another, but also as transitions from one episodic waypoint to the next.
3. Route overviews do not include every street to take or turn to make. In particular, the turn-by-turn details near the end of the route, where it is most important to have clear and accurate information [2], are noticeably absent. Instead, only a handful of major streets and prominent places are used to guide travel, with the final destination being situated in a neighborhood.

1.3 Cognitively-Inspired Direction Giving

Our approach in designing route overviews is grounded in principles of human spatial cognition. This section briefly describes what cognitive psychology tells us about the way people think about space. We then describe how these principles are incorporated in our route overviews and explain how these overviews improve the comprehensibility of a route, given what we know about human spatial cognition.

The principles of human spatial cognition that we incorporate into our route overviews are briefly described in the following list. They are explained in further detail in Chapter 2.

- People use a spatial hierarchy to organize their knowledge of space. People will move up and down this hierarchy as they think about space at different scales.
- Two major types of elements in people’s mental map of a city are cognitive anchor points and a network of major roads. Cognitive anchor points are not just centers of activity or distinct landmarks that are useful when giving directions. Rather, as the name suggests, they are the quintessential places that define a city and give it its identity.

- People perceive routes as having a skeleton made up of trunk segments. These trunk segments are a subset of a city’s major road network.
- People use many types of simplifications in the way they represent and communicate spatial information.

These principles provide us with the following design guidelines we used to determine what information to include (and not include) in our route overviews.

1. *Use spatial hierarchy and goal-directed descriptions.* Route overviews should appeal to people’s sense of spatial hierarchy and be structured in such a way that individual steps in the turn-by-turn directions can be understood in the context of some larger goal.
2. *Selectively suppress detail.* Specific details do not need to be presented in a route overview because they can be inferred by people’s knowledge of the environment or read directly from the detailed turn-by-turn directions when a person is actually travelling along the route³.
3. *Identify road skeletons and cognitive anchor points.* Explicitly identifying the major trunk segments that make up a route’s skeleton and the cognitive anchor points along that route will make driving directions easier to understand.

We now describe how route overviews that follow the above design guidelines improve the comprehensibility of driving directions.

The usefulness of route overviews comes as much from the way the information is presented as from the information itself. Consider the related case of furniture assembly instructions: the first thing the manufacturer will say is to read through the directions completely before starting. One reason for this is that before people can follow assembly instructions from start to finish, they have to understand what they are trying to accomplish in the first place. In this regard, the best assembly instructions provide an overview section that describe how all the pieces fit together.

³In Chapter 3, we will describe how this idea of selective suppression can also be used in modifying turn-by-turn directions.

The assembly instructions can then be regarded as a plan and the overview breaks down the larger plan into subplans that each make a contribution to the overall plan. The value of the overview becomes apparent when you actually follow the step-by-step assembly instructions. Because of the overview, you now have a larger context against which to situate the immediate assembly task at hand.

This same principle applies to following instructions that direct you from one place to another. If we regard the entire set of route instructions as a plan, a route overview decomposes that plan into smaller subplans. This establishes a framework against which people can situate themselves as they read the turn-by-turn directions and as they make sense of those directions. Exact turn details are not found at this scale because that information is not needed at this scale. The purpose of viewing the route at this scale is to get a sense of the route as a whole.

One of the principle tenets in this work is that to make it easier for people to understand route directions, it is necessary to provide different views of the route so that people can have the right information at the right time to suit the task at hand. If the task is initially orienting and making sense of the route (as in the vignette at the start of this chapter), then a high-level description is required. Another tenet is that we present route information in a way that mirrors how people make sense of it, both in terms of how this information is organized as well as in the choice of elements mentioned in the overview. By doing so, people will have a better sense of orientation when they read the detailed turn-by-turn directions because they will be able to situate these low-level specifics against their higher-level understanding of the route.

Our approach to presenting route directions thus exhibits certain principles from means-ends-analysis. We take advantage of the fact that people think about space at different scales and that these scales serve different purposes. Much of the focus on generating driving directions has been on the smallest scale: the actual act of travelling from one point to another. The work in this dissertation focuses on the larger scale at which a route is no longer just a sequence of go-to and turn directions. Rather, at the larger scale, a person looks at a route as a whole and situates the route

in the context of all the other places he knows. The route is viewed as a sequence of places he travels through and it is this view that forms the basis for episodic memory.

At this larger scale, one simplification used to reduce cognitive load is to view routes as a skeleton comprised of a few major trunk segments and punctuated with a handful of places that “anchor” the route. These cognitive anchor points may be well defined places but they can also be more nebulously defined neighborhoods. While this imprecision may be problematic for automatons who try to use a route overview in the same way as a set of turn-by-turn directions, this imprecision in route overviews do not pose a problem for people.

The use of nebulously defined neighborhoods is an interesting observation for two reasons. First, one person’s boundary for a neighborhood such as Harvard Square may be fuzzy, and second, two people may have slightly different definitions for what makes up Harvard Square. Yet despite this imprecision, people are able to communicate using these terms without much confusion. The reason for this is that while people’s definitions for neighborhood boundaries may be imprecise and may not align, the differences are so fine-grained that they become immaterial at the relatively coarse grain at which people are communicating. We can take advantage of this fact by creating a reasonable approximation for a given neighborhood and appealing to people’s inherent tolerances for imprecision when giving a general overview of a route.

At this point, it is worthwhile to point out the fact that we have referred to the places we include in route overviews as cognitive anchor points, and not as landmarks or points of interest. This is an important distinction. The purpose of this work is to reduce the confusion people experience when reading turn-by-turn directions by providing them with a greater understanding of the route as a whole. It is not to produce turn-by-turn directions with better landmarks (although we will describe how aspects of this work can be applied to improving turn-by-turn directions as well). Including recognizable landmarks in turn-by-turn directions (such as “turn right at the Starbucks”) helps with local disambiguation when a person is traversing the route. However, most of these landmarks are context-specific to that particular part of the route, and their usefulness in helping people orient themselves is limited

(e.g., a person may know many Starbucks coffee shops). Using cognitive anchor points in route overviews makes it easier for people to orient a larger portion of the route against fewer places, thus making it easier to mentally see and physically travel the route as a whole.

1.4 LAIR, CAP LOC, and ROVER

Using these findings from cognitive psychology as a guide, we created LAIR, a **L**ocation **A**ware **I**nformation **R**epresentation. LAIR is populated using a system called CAP LOC (**C**ognitive **A**nchor **P**oint **L**ocator). LAIR and CAP LOC were implemented using a geographic information system (GIS). GISs are large spatial databases typically used to analyze spatially distributed phenomenon. For example, a GIS can be used by city planners to identify potential areas for urban renewal, by epidemiologists to study the spread of disease, or by business owners to find potential areas to open new stores.

Our use of GIS is novel in that the GIS isn't being used as an aid to help people make a decision by analyzing spatial patterns, but rather as a model of what places and paths a person knows. Using an approach common in AI research, we build up a knowledge base of elements and then use those elements and the relationships between those elements to address our problem of interest. In our specific case, we use the basic elements of the GIS to mirror the cognitive elements people use to organize their knowledge of space. Each entry in the GIS's spatial database therefore corresponds to some actual geographic entity, such as a place, a path, or a region. In addition to representing the cognitive elements people use, LAIR uses the GIS to group spatial elements to mirror the way people hierarchically organize all the different places, streets, and neighborhoods they know. LAIR also uses the GIS's spatial reasoning algorithms to determine such relationships as containment, and left and right orientation.

LAIR is used by ROVER, a program we created that uses the design guidelines listed in Section 1.3 to produce route **o**verviews. The route overview produced by ROVER both situates the route's destination and gives a high-level description of how to get

to that destination. Based on the contents of the LAIR model of what places and paths a person knows, the route overview ROVER generates will be different. This emulates the fact that people give slightly different directions to accommodate for differences in the assumed knowledge of the person who is asking for the directions. ROVER also modifies turn-by-turn directions to reflect the information presented in the route overview.

For example, consider the route shown below in Figure 1-2.

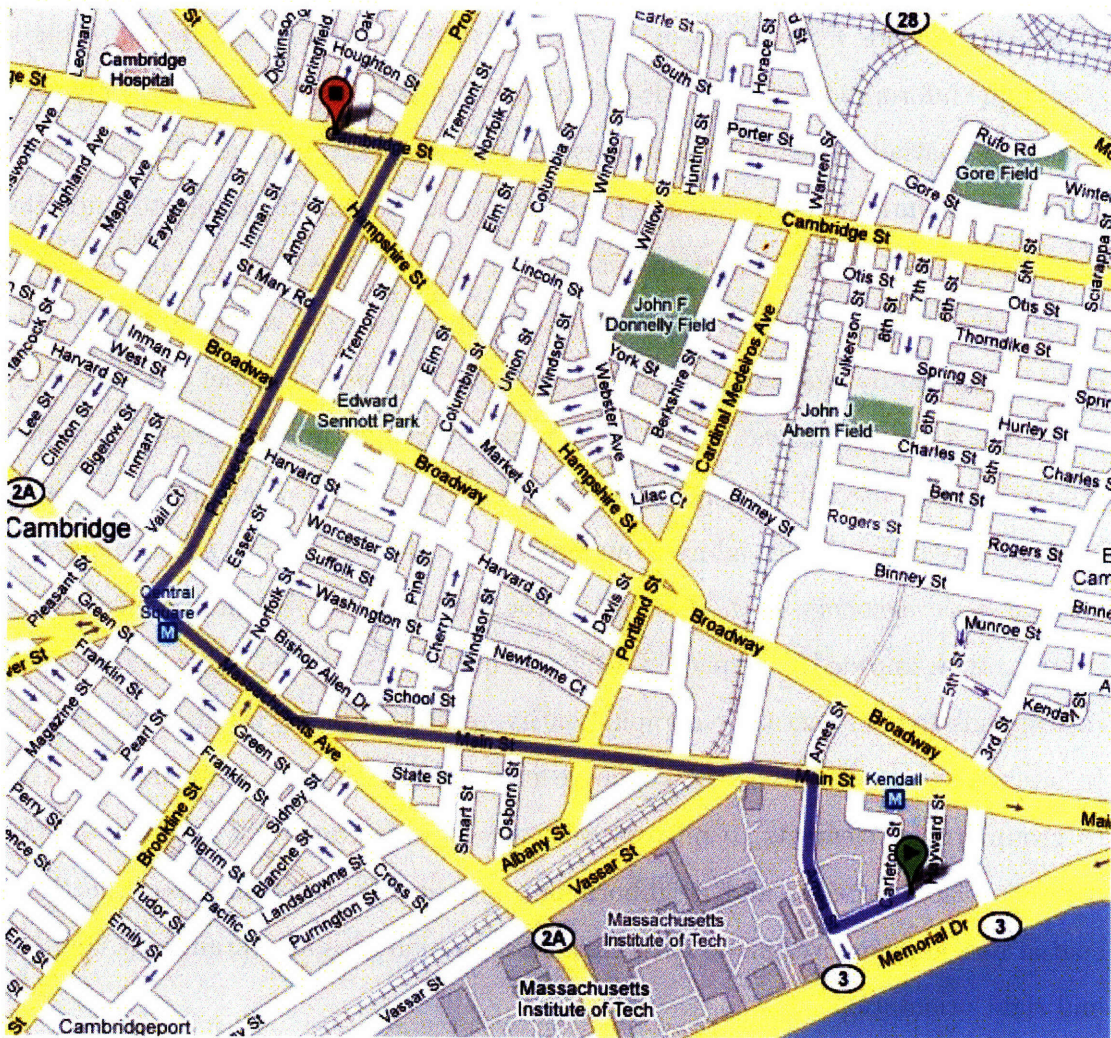


Figure 1-2: This route is used to illustrate how the route overviews generated by ROVER for a particular route differ based on which streets on the route a person already knows.

Here is an example of a route overview for this route with the assumption that a person is not familiar with the parts of the route including and after Prospect Street:

Your destination is on Cambridge St, which is a major road.

On this route, you will be travelling along 4 major roads. The first 2 major roads you will be travelling on are Main St to Mass Ave/Massachusetts Ave/RT-2A. After that, the other major roads you will be travelling on are Prospect St to Cambridge St. Along this route, you'll pass Central Square.

On the other hand, here is what a route overview would be like if the person is familiar with everything about the route except for the fact that Prospect Street connects Massachusetts Avenue and Cambridge Street:

Your destination is in Inman Square.

On this route, you will be travelling to Mass Ave/Massachusetts Ave/RT-2A, then going from Mass Ave/Massachusetts Ave/RT-2A to Cambridge St by Inman Square. The route then goes from Cambridge St to your destination. Along this route, you'll pass Central Square.

LAIR, CAP LOC, and ROVER's use of GIS will be described in greater detail in the following chapters, but one further point will be made here. There is a wealth of GIS datasets available. However, the same challenges that arise when interpreting driving directions that have too much information also appear when trying to choose the right dataset for use in generating route overviews — there is both too much information and not enough. In particular, systematically identifying neighborhoods that do not have official boundaries but are only informally defined is a challenge. This challenge led us to a second way in which our use of GIS was novel: our cognitively-inspired model of people's view of large-scale space helped us to judiciously select a small set of features from GIS datasets to make quantitative approximations of inherently qualitative spatial elements.

To validate our claim that our cognitively-inspired computational model for direction giving helps people gain a better overall understanding of a route, we conducted

a user study in which we asked volunteers to compare a set of ROVER-generated route overviews and turn-by-turn directions to a set of turn-by-turn directions generated by a traditional mapping service. Analysis of the results shows that in most situations, people rated their understanding of a route higher, in a statistically significant manner, when presented with ROVER’s output than when presented with traditional turn-by-turn directions. These results were seen in people who fit the predicted knowledge profile ROVER used to generate the route overview as well as in people who did not fit this profile. We discuss these results and the lessons learned from those situations in which the addition of a route overview was not overwhelmingly preferred in Chapters 4.

1.5 Contributions

The primary contribution of this work is the perspective it brings to the problem of making driving directions easier to understand. In this thesis, we use what cognitive psychology tells us about how people think about space to shape the lens through which we look at the problem of making driving directions more understandable. Our perspective is summarized and applied as follows:

1. Travel is viewed not as a sequence of turns which have little to do with one another, but as something that has a larger organization, with a greater sense than what is currently presented in turn-by-turn directions. This greater organization is framed by a few major trunk segments and punctuated with a handful of cognitive anchor points. In the larger view of things, people don’t go from one turn to the next; travel is conceptualized as movement from one cognitive anchor point to the next along these major trunks.
2. We apply this perspective by using the quantitative, analytic functionality of GIS to identify qualitative properties of a route such as the route’s trunk segments and the cognitive anchor points along that route.
3. We appeal to a person’s higher-level view of travel by providing a route overview

that has the elements and techniques people use to conceptualize and organize their knowledge of large-scale space. In doing so, we improve a person’s overall sense of the route and make current turn-by-turn directions easier to understand and more useful to people.

1.6 Thesis Overview

The remainder of this dissertation is organized as follows. Chapter 2 describes the principles that form the basis of this work. There, we discuss what studies from cognitive psychology and urban planning tell us about how people think about and make use of their knowledge of large-scale space. In Chapter 3, we discuss the implementation details of LAIR, CAP LOC, and ROVER. We emphasize how the principles from Chapter 2 guided our design rationale. We describe how the cognitive principles identified in Chapter 2 were used by CAP LOC to identify the cognitive anchor points in a city. We describe how these principles influenced the choice of features ROVER includes in its route overviews and how the structure of ROVER’s route overviews change based on the spatial information LAIR includes in its model of a person’s familiarity with a route. We also discuss how related work from computer science influenced our implementation decisions.

Chapter 4 describes our evaluation and analysis of the route overviews and turn-by-turn directions ROVER produces. We discuss how the results of our user study demonstrate how taking a cognitively-inspired design approach can make directions easier to understand, regardless of the level of familiarity a person has of the area through which a route passes. We also discuss how apparently “bad” results are also consistent with our cognitively-inspired design principles, and how future experiments can evaluate the effect of accounting for familiarity in route overviews.

In Chapters 5, 6, and 7, we take a look forward and a look back. Chapter 5 describes future work inspired by the analysis in Chapter 4. Chapter 6 describes other route-related work in computer science and situates this work in that context. The dissertation ends in Chapter 7 with a closing discussion of how this work makes

driving directions easier for people to conceptualize and how this contribution was made possible by combining different aspects from cognitive psychology, GIS, and computer science.

Chapter 2

Human Spatial Cognition

You don't know where you're going unless you know where you've been.

In this chapter we discuss human spatial cognition — the processes people use to conceptualize and organize their knowledge of large-scale space. The purpose of this chapter is to present the guiding cognitive *principles* upon which the work in this dissertation is based. We survey work from a number of different fields such as cognitive psychology, urban planning, and cartography. Related work from computer science that was influential in the *implementation* of LAIR will be discussed in Chapter 3, and other computer science-related work in the area of making directions more comprehensible will be discussed in Chapter 6. In this chapter, we focus on the basis for our approach in designing route overviews, and this basis stems from work in cognitive psychology, urban planning, and cartography. We start with a discussion of the types of elements people have in their spatial knowledge base and how these elements are organized. We discuss how people perceive routes, then close with a discussion of how these concepts can be used to structure route overviews.

2.1 The Image of the City

This section describes the elements people use when they think about space; these are the elements used to populate a person's mental map. Identifying what these elements are and their effects on our perception of space gives us insight on how to

structure a route overview. Numerous papers describe the elements that make up a person's mental map [24, 46, 54, 77, 85]. These papers exhibit a common set of themes, described below.

While most of the work presented in this chapter comes from cognitive psychology, one of the earliest and most influential works in the field of human spatial cognition was done by an urban planner in the 1950s and 1960s. Kevin Lynch's seminal work *The Image of the City* [46] changed the way urban planners viewed the relationship between people and the city they live in. In *Image*, Lynch looked at what sorts of mental impressions people had of the physical space around them, in order to determine whether these views had any impact on the way residents interacted with the city, and if so, how urban planners could influence these views to better meet their original design goals.

This was important work because up until that time, urban design was viewed as a one-way process. Urban planners had particular notions of form and function in mind when laying out a city, but did not consider the actual impression their designs left on inhabitants once a city was built. Lynch introduced this added dimension to the design process. Lynch's claim was that "legible" cityscapes — those whose parts can be easily recognized and placed in an overall picture — enhance the lives of their inhabitants. Ease of navigation is the most obvious benefit of legible cityscapes, but there is also a psychological benefit: "This is the obverse of the fear that comes with disorientation; it means that the sweet sense of home is strongest when home is not only familiar but distinctive as well (*The Image of the City*, pp. 4-5)."

To develop a theory of what elements people include in their image of a city, Lynch carried out extensive case studies in three different cities: Boston, Los Angeles, and Jersey City. Lynch asked participants to describe their city using words and sketch-maps. Despite large differences in how the three cities are laid out, Lynch found similarities in the elements people used to describe their mental map of their city. These elements are nodes, paths, edges, districts, and landmarks.

- *Nodes* are major centers of activity.

- *Paths* are used to travel from one place to another.
- *Edges* are linear elements that aren't paths. They serve as boundaries between areas. One example of this is the elevated expressway that once separated Boston's North End from the rest of the city.
- *Districts* are the neighborhoods in a city. Districts are usually defined by various informal criteria rather than formal municipal boundaries.
- *Landmarks* are prominent reference points.

A map of Boston, depicted in terms of nodes, paths, edges, districts, and landmarks is shown below in Figure 2-1.

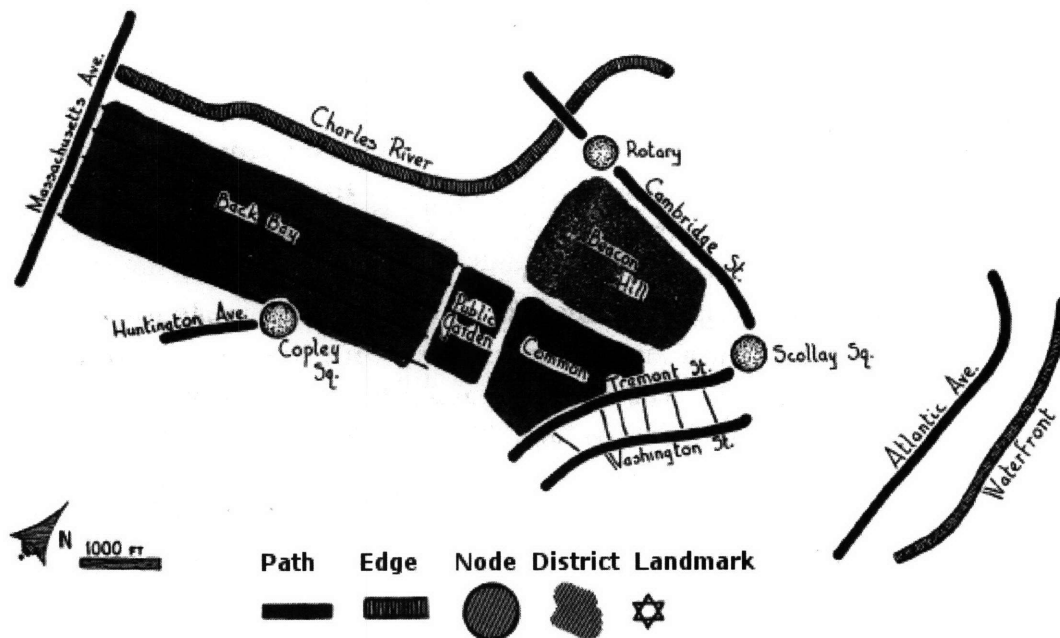


Figure 2-1: A map of Boston, depicted using the five spatial elements Lynch identified. Image taken from Kevin Lynch's *Image of the City* [46].

Of these elements, Lynch found that most people structured their image of the city around paths and districts. However, for a city to be legible, all five of these elements must be well defined and clearly integrated with one another. For example, the presence of nodes and landmarks along paths should provide a smooth sense of

transition from one district to another, as opposed to making adjacent districts seem disconnected from one another.

The impact of Lynch's work is cross-disciplinary. In describing what factors make up a legible cityscape, Lynch is not only giving guidelines to urban planners for good city form, he is also proposing a cognitive model of how people perceive their city.

Golledge's cognitive counterparts theory [24] is similar in spirit to Lynch's work. In addition to the nodes and paths that appear in Lynch's work, Golledge's cognitive counterparts theory proposes that people's mental models of space are organized using contiguity (the order and spacing between places in the mental map, a key factor that influences the episodic nature of memory¹) and a spatial hierarchy of the different sets of places people know. Section 2.2 discusses experimental results from the psychology literature that give insight on how contiguity and hierarchy influence people's perception of space.

Golledge furthers the importance of nodes and landmarks in his theory of cognitive anchor points [23]. Cognitive anchor points are not just centers of activity or distinct reference points useful when giving directions but, as the name suggests, they are the quintessential places that define a city and give it its identity. So important are cognitive anchor points in shaping a person's visceral notion of the city that they actually "pull" other less prominent places towards them, forming local reference frames centered about these anchor points.

A number of studies have demonstrated this phenomenon by showing a distance asymmetry between cognitive anchor points and other places. In these studies, college students were asked to judge the distance from a campus location considered to be a cognitive anchor point to another that was not and vice-versa. The distance estimates were collected using different methods. In some studies [71], participants were primed by having one place placed in the center of a page. They were then asked to draw the location of the other place using some standardized scale (e.g., the distance between two well-known building on campus might correspond to one

¹When people travel along a route, they move through both time and space. The impression of this route is encoded as both what we saw as well as when we saw it, giving rise to the idea of episodic memory [74].

inch). In other studies [55], participants were primed by first asking if they knew the location of a particular campus building. After that, they were then asked to estimate the location from that building to another building.

The results of these studies all demonstrated that anchor points had a sort of “cognitive gravity.” In situations where people were first primed with the name of the non-anchor point (the place they were coming from) the estimated distance to an anchor point was less than in the corresponding situations in which people were first primed with the name of the anchor point and then asked to estimate the distance to the non-anchor point. Thus, cognitive anchor points have a prominence in the mind’s eye that makes people perceive them as places that are easy to go to.

In this section, we discussed the important elements that appear in people’s mental maps of a city and which, by extension, are important for us to consider in our development of route overviews. Due to the role cognitive anchor points have in shaping a person’s visceral impression of a city, we regard these anchor points as the places that punctuate a person’s episodic memory of a city and consider them to be one of two major types of components that contribute to a person’s impression of a route. Including cognitive anchor points in route overviews makes new routes seem more familiar to people thus making driving directions easier to understand. We describe the influence of roads, the other major type of component in a person’s impression of a route in Section 2.4.

2.2 The Frame for the Image of the City

In this section, we discuss how people organize their knowledge of large-scale space. We look at how different reference frames, scale, and spatial hierarchies are used by people to solve various navigation-related tasks. As people mentally traverse a route, they move up and down their spatial hierarchy, shifting not only among different views of the route, but among different *types* of views as well. The abilities of these techniques to conjure up rich (and often metric-distorting) images in the mind’s eye provide a more effective mechanism for people to understand where a route is going

compared to detailed turn-by-turn directions, which do not appeal to these techniques. Turn-by-turn directions present a flat view of a route, and do not appeal to the mental machinery people make use of. We discuss how we can use these insights on how space is organized to produce effective route overviews.

The perspective through which spatial information is visualized in the mind's eye has received a lot of attention in the cognitive psychology literature. Spatial information can be viewed as either ego-centric, in which case you are viewing space from a first-hand perspective, as if you were walking along at ground-level, or as exo-centric, in which you are viewing space from a bird's-eye view. A commonly-cited result in different experiments [16, 39, 53, 56] is that women prefer an ego-centric view of space whereas men prefer an exo-centric view. These results appear both in indoor navigation tasks as well as outdoor tasks [40]: men were more likely to report using strategies that involved cardinal directions whereas women reported strategies relying on landmarks along a route. In one study, Allen found that this preference in describing routes a particular way had an actual effect on how well women (but not men) followed directions [3]: women made fewer navigational mistakes if directions were provided using an ego-centric perspective.

This is not to say that men and women rely exclusively on one perspective or the other [88]. Rather, the immediate wayfinding task at hand and level of familiarity with the environment has a large influence on what mental representation is used. Sholl demonstrated this phenomenon by asking college students to point to various places [76]. When asked to point to well-known places on their college campus, the students were faster at pointing to places that were located in front of them, regardless of which way they were facing. When asked to point to cities many miles from their university, the students were able to answer quicker when they themselves were facing north. These results indicate that people switch between ego-centric and exo-centric reference frames depending on the wayfinding task, and that in each of these reference frames, there is a preferred orientation in that reference frame (e.g., "in front" or north-up).

In the previous section, we mentioned that cognitive anchor points have a tendency

to affect distance perceptions. In addition to this, reference frames centered around cognitive anchor points can also distort the sense of direction. Lloyd investigated this phenomenon when he asked residents from different neighborhoods in Columbia, South Carolina to estimate distances and directions from their neighborhood to other places in the city [43]. Lloyd's results corroborated the effect cognitive anchor points had on causing people to over-estimate the distance from their neighborhood (the cognitive anchor point) to other places. In addition to this, Lloyd found that there was a systematic angular distortion in residents' directional estimations. It appeared that each neighborhood had its own reference frame in which major transportation axes were mistakenly assumed to be aligned with the cardinal directions, and that it was this local reference frame that was used to make directional assessments. Accounting for this rotation corrected the orientation biases.

Reference frames not only bias distance estimates from the cognitive anchor point around which the reference frame is based, but, as Holyoak and Mah discovered, they also bias distance estimates between other pairs of places viewed in the reference frame [27]. This is attributed to the reference frame giving focus to places closer to the reference frame's cognitive anchor point, thus producing a "fisheye" effect and giving rise to mental images similar to those cartoons illustrating the New Yorker's view of the world (see Figure 2-2 on the following page). In these cartoons, different scales are used in such a way that places closer to the reference frame's origin are seen in greater detail than those further from the origin, which are "packed" closer together using a much coarser scale. This has the effect of causing distance estimates between pairs of places closer to the origin to be over-estimated, compared to distances further from the origin because the places closer to the origin appear spaced further apart in the mental image. Moreover, this fisheye effect is flexible and occurs regardless of the reference frame chosen. In the Holyoak and Mah study, when participants were asked to imagine themselves on the East Coast of the United States, their estimates of the distance between New York and Pittsburgh were larger than participants who were asked to imagine themselves on the West Coast of the U.S. On the other hand, the East Coast group's estimates of the distance between Salt Lake City and San

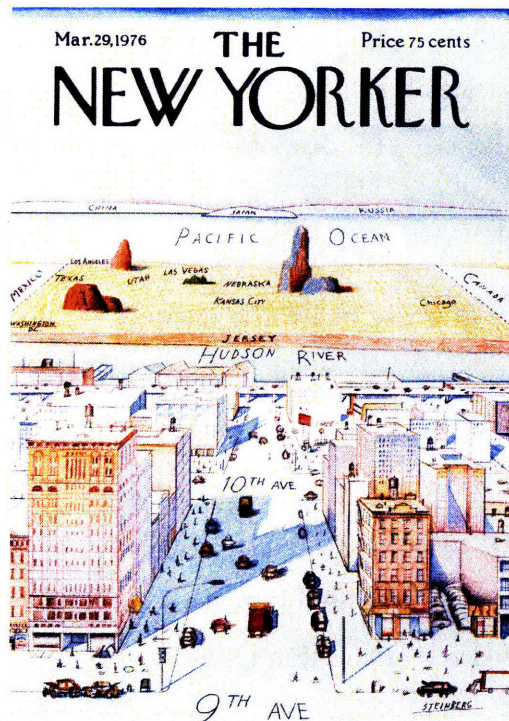


Figure 2-2: The New Yorker's View of the World. This figure illustrates a "fisheye" effect in which places closer to the reference frame's origin are seen in greater detail than those further from the origin.

Francisco were smaller than the West Coast group's.

People also use hierarchy to reason about space. Much of the evidence that supports this comes from the errors and slowdowns in reaction times people exhibit when making certain spatial judgements. One of the most compelling pieces of evidence to support the use of hierarchy was provided in a study conducted by Stevens and Coupe [81]. In this study, participants from San Diego were asked to give the direction from one North American city to another. One of the pairs used in the study was San Diego, California and Reno, Nevada. Most of the students in the study incorrectly stated that San Diego was south and west of Reno (San Diego is actually south and east of Reno, as illustrated in Figure 2-3 on the next page)². Stevens and Coupe explained this phenomenon as being due to the fact that since most students don't have a mental map that contains both Reno and San Diego, they jump up in their

²This anomaly has become so eyebrow-raising that it is included as a question in the board game Trivial Pursuit [87].



Figure 2-3: Reno, Nevada is actually to the *west* of San Diego, California

spatial hierarchy and consider the relationships between California and Nevada. Since California is generally west of Nevada, students made the (erroneous) generalization that San Diego is west of Reno.

Wilton found similar evidence for the use of spatial hierarchy when he tested geography students' knowledge about the United Kingdom [90]. Wilton's subjects were students at a Scottish university and all had decent familiarity with different cities in the U.K. In his experiment, he asked subjects to answer yes/no questions regarding whether one U.K. city was in a certain direction with respect to another. Wilton found that for cities that were in different parts of the U.K. (England versus Scotland), subjects were able to respond much quicker than when cities were in the same general area. The implication here being the dual to the one given in Steven's and Coupe's experiment: in wayfinding tasks, spatial relationships between entities higher in the hierarchy are consulted first and if these relationships are too coarse grained to provide an answer, people then go down further in the hierarchy.

Another example of the use of hierarchy is Chase's study of taxi drivers [12]. In his study, Chase found that when drivers were asked to list neighborhoods in Pittsburgh,

they usually listed them in groups based on spatial locality. Moreover, the average pause time between neighborhoods was less when neighborhoods were within the same geographic region of Pittsburgh (the North Side, South Side, or East Ends of the city) compared to when the next neighborhood mentioned was in a different geographic area than the previous neighborhood. Chase also asked the drivers to plan, in his lab, routes between 10 different pairs of places. He later conducted a field study in which he asked the drivers to take him along a route between each of those pairs of places. Chase found that in over a fifth of the routes, drivers ended up taking a shorter route than the one described in the lab. These routes used the same major arterials to go between neighborhoods as the routes described in the lab, but the in-field routes used different local roads when travelling within neighborhoods. Based on this result, Chase suggests taxi drivers use means-ends analysis to produce a general plan for a route. Drivers first create a high-level plan in which they think of themselves as travelling from one neighborhood to another along major arterials (in terms of means-ends analysis, this resolves major differences between the current state and the goal state), and then they fill in the specific details of navigating within a neighborhood (of resolving the smaller differences in the means-ends-analysis) when they are actually driving about.

In this section, we have discussed how people use different reference frames and hierarchy to structure their image of the city and solve different wayfinding tasks. Making use of these organizational schemes is a key insight in the construction of route overviews. By including cognitive anchor points in our route overviews, we provide added hints to help people better visualize a route, both from an exo-centered, at-a-glance bird's-eye perspective and from an ego-centric perspective with anchor-point-specific reference frames. Turn-by-turn directions provide information at the lowest level of the spatial hierarchy. A route overview provides a context in which to interpret detailed turn-by-turn directions; instead of being regarded as individual instructions to be carried out without context, the instructions in a set of turn-by-turn directions now are viewed as collective steps in a plan in which each step has some larger purpose.

2.3 Simplifying the Image of the City

The previous section mentioned examples of non-Euclidean distortions that arise as a result of the spatial organizational schemes people use. In this section, we provide further examples of distortions. As opposed to being a side-effect of how people organize their knowledge of space, the distortions described here arise mainly from the desire to reduce cognitive load by providing the minimal amount of information required for wayfinding. These distortions illustrate an important concept in our design of route overviews: precise details are not important if their exclusion does not affect the navigational task at hand. We can selectively suppress details when providing the gist is sufficient.

The types of simplifications people make in their knowledge of space is reflected in the techniques cartographers use to produce maps [32, 47, 59]. Since maps are rendered at a scale smaller³ than 1:1, cartographers will necessarily have to emphasize some features and suppress others. To make it easier for people to solve the navigation problem for which the map was designed, cartographers will strategically call a reader's attention to different features using such visual cues as color, shape, and hue while not including other details. In some tasks, such as travelling through a subway network, the map will sacrifice accurate distances in order to emphasize connectivity (the Boston subway map, shown in Figure 2-4 on the next page is a prime example of this). The reason for this is that the global accuracy that is lost does not affect the primary purpose of the map.

Cartographers use these techniques to simplify the conceptualization of space because people often employ them effectively and without thinking. One of the most common types of simplifications people make is one of alignment. In the previous section, we described how local reference frames can cause orientation biases. Orientation biases occur in a global sense as well. For example, Tversky [84] showed students a map of the Americas and another map where South America was shifted

³When we talk about the scale used in cartographic maps, a large scale map provides more detail than a small scale map; e.g. compared to a 1:1, real-life scale, any map scale will be a smaller scale. This use has a different connotation from what is normally thought of when people say "large-scale" space vs. "small-scale" space.

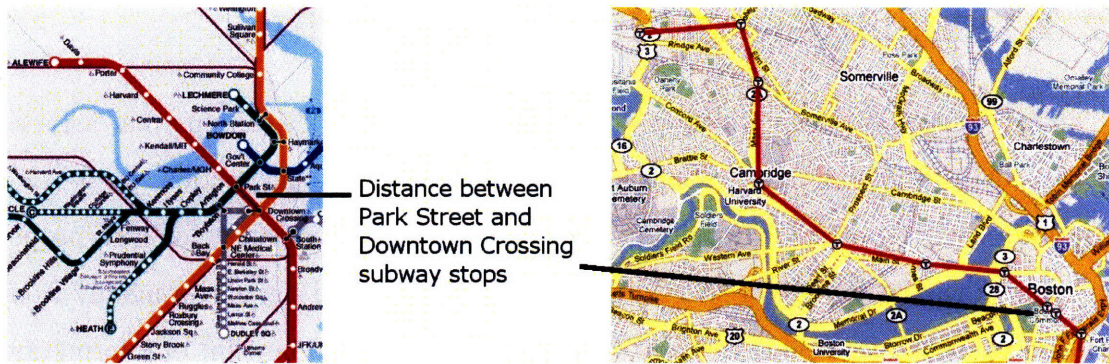


Figure 2-4: The subway map on the left emphasizes the connectivity of the different subway lines. The street map on the right accurately represents the distances between stops along the Red Line. In order to emphasize connectivity, the subway map sacrifices accurate distances. Note, for example, that although the subway map suggests that stops along the Red Line are equally spaced, the distance between the Park Street and Downtown Crossing stops is actually much less than what the subway map would suggest.

westward with respect to North America to give the appearance of a more continuous north-south line. When asked which map was the accurate one, a significant majority chose the altered map. Moreover, when shown a map in which the Americas were shifted northward to more closely align along an east-west axis with Europe and Africa, a majority of students said the altered map was the true one (see Figure 2-5). This phenomenon may be due to Gestalt laws of grouping that cause nearly-aligned objects to appear more aligned than they actually are [47].

In addition to aligning two places with respect to one another, there is also a tendency for people to rotate places so that they more closely align to the cardinal directions [42, 43]. South America (as being rotated so it is more north-south) and the 101 Highway in the San Francisco Bay Area (which some longtime residents think of as running north-south, but which actually runs east-west in some places) are two prominent examples. Manhattan Island is another prime example, as is reflected in Figure 2-6.

In addition to orientation-related simplifications, other aspects of places are simplified too. In studies where people were asked to estimate the angle of turns, their answers, whether verbal or sketched, show that people have a tendency to remember

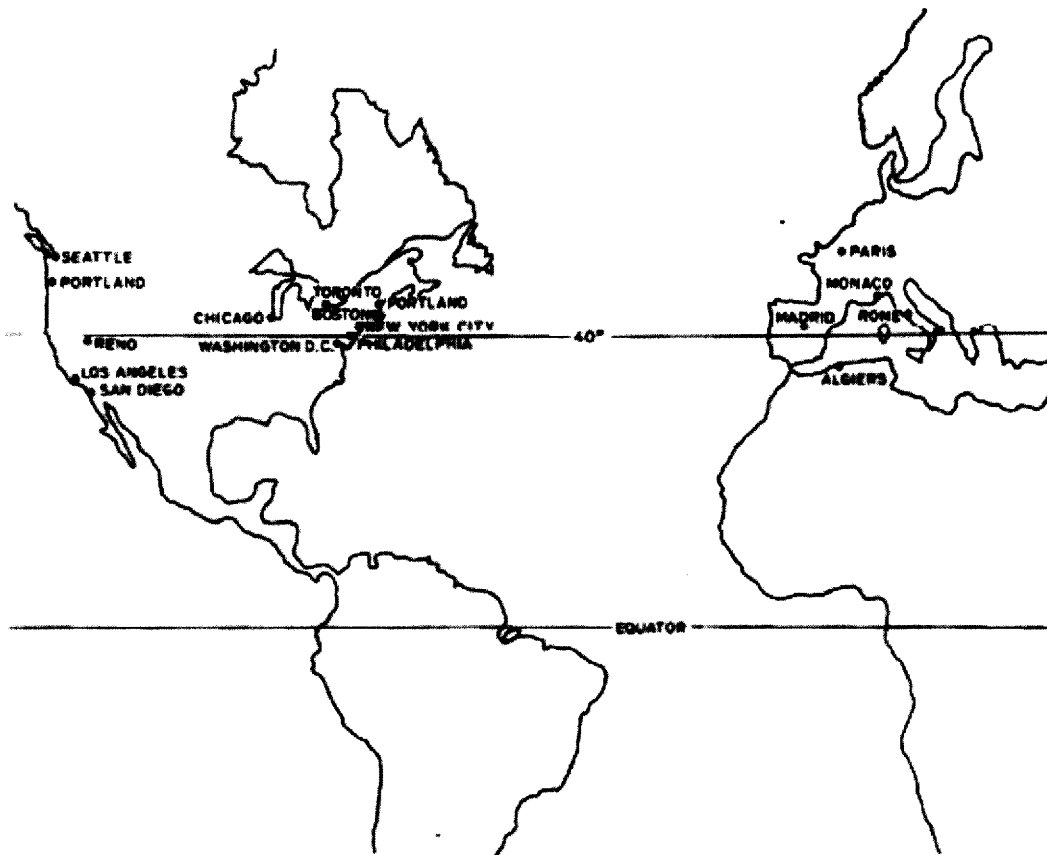


Figure 2-5: As this world map shows, the Americas are not nearly as aligned with Europe and Africa as most people think. Most people think the Americas are located further north than they actually are (e.g., it would surprise many people to learn that Boston, MA and Rome, Italy are located at approximately the same line of latitude). Image taken from Barbara Tversky's *Distortions in Memory for Maps* [84].

turns as being biased towards 90 degrees [31, 58, 72]. Streets, rivers, and municipal borders were also rendered as being straighter than they are [85].

It is also interesting to examine what imprecisions exist in the language people use when describing routes and spatial relationships. In English, there are actually only a few words (between 80 and 100 prepositional phrases) used to describe spatial relationships between two places [38]. Moreover, the spatial relationship these words describe are underspecified, unless they are qualified with additional numeric information (for example, the phrase "Boston is north of New York City" doesn't tell us how many miles north Boston is from New York City). In addition to the language

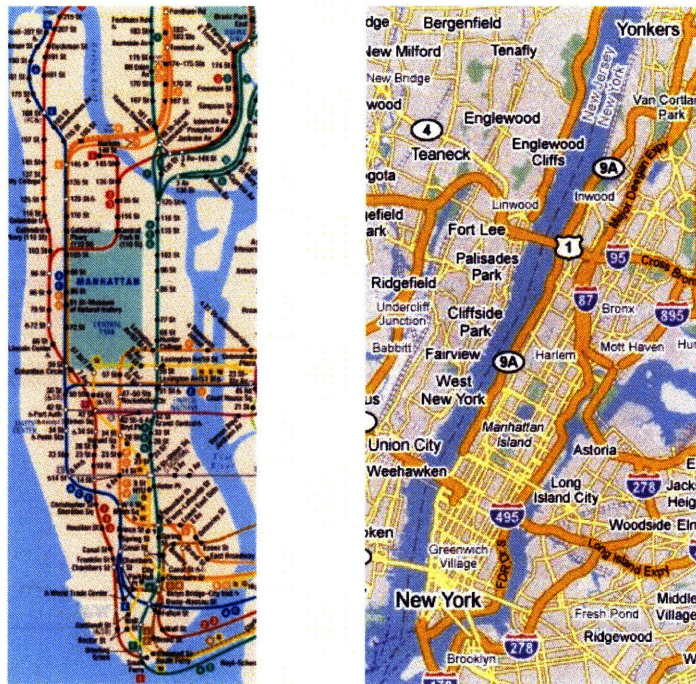


Figure 2-6: The subway map on the left has aligned Manhattan Island to run more north-south than it actually is. The map on the right accurately shows Manhattan as running more northeast-to-southwest.

used to convey route directions, people may also leave out certain turn instructions if constraints in the environment (or familiarity with the area) fill in the gaps in the directions [45]. However, these under-specifications do not prevent people from effectively communicating spatial information.

This section looked at the sorts of simplifications people make in their cognitive maps. Simplification is a powerful concept in reducing cognitive load. It demonstrates that it is not necessary to preserve the exact details in our memory because those exact details are not necessary to adequately perform a navigation task. Similarly, in producing route overviews, we do not need to describe everything along the route, but only need to provide enough information for people to have a high-level understanding of the route. With this understanding, people then have a framework against which to situate themselves as they read the detailed turn-by-turn directions ⁴.

⁴Depending on a person's familiarity, this high-level understanding may even be sufficient for a person to use environmental constraints and their own experience to infer the specific directions.

2.4 Routes

In this section, we focus on how paths impact people's impressions of the routes they travel. Here, we will describe different studies that illustrate how people regard a route as consisting of a skeleton of major trunk segments. This skeleton is the second major influence in our impression of a route. The skeleton serves as a frame that is punctuated with cognitive anchor points, the other element that influences a person's perception of a route. We discuss how this skeleton arises, is used, and its effect on people's understanding of a route.

Space syntax, a tool used in urban planning to analyze the built form of the environment [8, 91], is one tool that can be used to understand the cognitive basis for a road skeleton. One of the concepts in space syntax is the axial map, which is a way of identifying and illustrating which roads have straight, uninterrupted sightlines and highlighting the location where these lines intersect (see Figure 2-7, below). As Kim and Penn discovered in their analysis of people's sketch maps of a London suburb [31], the roads that were most often included in residents' sketch maps were the ones in the axial map of the suburb. This suggests that the overall skeleton of the city consists of roads with the longest straight-aways and the most junctions because they afford the most connectivity to other areas.

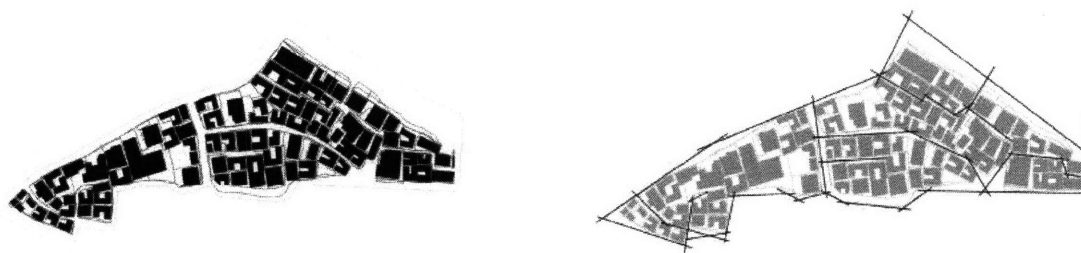


Figure 2-7: A birds-eye view of open space (left) along with its representation as an axial map (right). This figure is taken from Young Ook Kim's *Linking the Spatial Syntax of Cognitive Maps to the Spatial Syntax of the Environment* [31].

The effect of connectivity on the development of the road skeleton is further explored by Kuipers in his experiments of human wayfinding in a computer-simulated environment [34]. In this study, participants were asked to navigate from one des-

ignated location to another in a computer-simulated environment. Kuipers tracked the paths that people took to move from one place to another and found that people preferred to take paths that had many associated boundary relationships (here, a boundary relationship is defined as whether a particular place is on a path or to the right or left of that path). As people travelled along these paths, they would associate more boundary relationships with these paths, thus creating a positive feedback loop. From these explorations, a skeleton of commonly-traveled paths emerged.

Kuipers was not the first to study the use of a skeleton in route-planning tasks, but he was one of the first to describe how it could be formed. Earlier work by Pailhous [64] studied the routes taken by Parisian taxi drivers and describes how the skeleton is used by people with different levels of familiarity with Paris. Like Kuipers' work, Pailhous' work illustrates the importance of the skeleton as a route-planning tool, and consequently as a way in which people visualize the relationships between places in a city. Using a map of Paris, Pailhous designated the city's major thoroughfares as that city's skeleton and the other streets as a secondary network. He then asked his subjects to solve a detour problem. Pailhous found the skeleton was important to both novice and expert taxi drivers, but this was demonstrated in different ways. Novice taxi drivers preferred to stay on the skeleton as long as possible when finding a detour. Expert taxi drivers, on the other hand, would use roads on the secondary network in their detour so that they could more quickly return to the skeleton.

The skeleton is used for things other than route planning; as Allen demonstrated, it also influences our perception of a route [4]. In his study, Allen showed participants slides from a route and found that participants consistently used major cross streets in the route's skeleton to segment the route into episodic chunks. The segmentation of a route also has an interesting side effect on distance estimations. In the Allen study, participants demonstrated a stair-step pattern when asked to estimate distances from the start of the route to other places along the route: distances to points within the same segment increased linearly but there was a large jump between the last place in a segment and the first place in the following segment.

Cross-streets are not the only thing that have the ability to increase distance perceptions. National boundaries produce a similar effect in which estimates of distance between domestic cities are smaller than the estimated distance between a domestic city and one in another country [11]. It has been suggested that the distance distortions seen in route segmentation can be generalized to being caused by people mentally organizing things using gestalt rules of grouping [9, 29], and that when we make distance estimates between places in different groups, that distance is overestimated. Finally, there have been some studies that suggest that the greater the number of turns in a route, the longer the route is perceived to be [73]. However this finding could be a by-product of the fact that these experiments were performed using a within-subjects design [30]. In between-subjects experiments, in which a person was not able to compare routes with different amounts of turns, studies did not find that more turns increased the perception of route distance.

In this section, we have reviewed evidence for our claim that people regard a route between two places as consisting of travel along a skeleton network of roads. The impression of a route is that it is divided into episodic chunks by major cross streets, and punctuated by cognitive anchor points. This structure is the design basis upon which we structure our route overviews.

2.5 Tying It All Together

The findings presented in this chapter served as the guiding principles upon which the ideas in this thesis are based. The claim in this thesis is that people view the routes they travel as consisting of a skeleton of major roads, punctuated by a small set of cognitive anchor points. By identifying what roads make up the skeleton, which cognitive anchor points are on the route, and the neighborhood in which the destination is located, we can make turn-by-turn directions easier to understand.

In support of this claim, we noted the importance of cognitive anchor points in people's perception of the city, and how these anchor points give rise to local reference frames used when thinking about those parts of the city. We also discussed studies

that demonstrated the importance of hierarchy and the need to present the gist of a route. Finally, we looked at the tendency to break routes into episodic chunks according to the skeleton of major roads and cognitive anchor points. Together, these findings provided us with guidelines we used in determining what information to include (and not include) in our route overviews. In short, these guidelines are:

1. *Use spatial hierarchy and goal-directed descriptions.* Route overviews should appeal to people's sense of spatial hierarchy and be structured in such a way that individual steps in the turn-by-turn directions can be understood in the context of some larger goal.
2. *Selectively suppress detail.* Specific details do not need to be presented in a route overview because they can be inferred by people's knowledge of the environment or read directly from the detailed turn-by-turn directions when a person is actually travelling along the route⁵.
3. *Identify road skeletons and cognitive anchor points.* Explicitly identifying the major trunk segments that make up a route's skeleton and the cognitive anchor points along that route will make driving directions easier to understand.

In the next chapter, we discuss the specifics of how we generate route overviews using LAIR and ROVER.

⁵In Chapter 3, we will describe how this idea of selective suppression can also be used in modifying turn-by-turn directions.

Chapter 3

The Design and Implementation of LAIR, CAP LOC, and ROVER

In this chapter, we discuss the design decisions that were made in our implementation of LAIR, CAP LOC, and ROVER. The emphasis will be on how these choices are consistent with the cognitive design principles for route overviews listed at the end of Chapter 2, namely the importance of: hierarchy, goal-directed descriptions, selective suppression of detail, and the use of a route's skeleton and cognitive anchor points. We describe how following these guidelines led us to include certain features in our route overviews that are not currently available in turn-by-turn directions.

We begin by describing how LAIR models the structures used in human spatial cognition and how LAIR's representational techniques are different from other spatial representations in the computer science literature. After describing the LAIR representation and how it was implemented using a geographic information system (GIS), we then discuss insights from related work in urban planning that influenced the design of CAP LOC, a system that populates LAIR using GIS datasets. The chapter then closes with a description of how ROVER uses the information in LAIR to produce a route overview and slightly modify turn-by-turn directions in a way consistent with the cognitive guidelines listed at the end of Chapter 2. The main emphasis of this work is how the cognitive design guidelines used to create LAIR, CAP LOC, and ROVER improve the comprehensibility of driving directions, independent of the specific places

with which a person is familiar. As a secondary point, we also discuss how ROVER’s output differs based on LAIR’s model of what a person knows.

As you read this chapter, it is important to keep in mind that the focus is on the rationale behind our design decisions and the insights that led us in that direction. While the way in which we structure a route description is relevant to the natural language generation (NLG) community, the emphasis is not on NLG techniques. Moreover, while we will discuss one related project that used machine learning techniques to identify a city’s neighborhoods and the influence that work had on our approach to populating LAIR, our work does not emphasize statistical map learning algorithms. Both of these issues, and all other things described in this dissertation should be viewed through this lens: that by following a set of cognitively-inspired design principles, we can make turn-by-turn directions more comprehensible.

LAIR, CAP LOC, and ROVER are not algorithmically complex yet together they make turn-by-turn directions more comprehensible. This is a direct benefit of our representation choices and the novel way in which we apply the analytical functionality of a GIS to produce a qualitative route overview.

3.1 The LAIR Spatial Representation

When designing the LAIR location representation, we considered the differences between the problem of making driving directions more comprehensible for people to the problem that systems like TOUR [36], PLAN [13], and the Spatial Semantic Hierarchy (SSH) [37] addressed. In those latter cases, the research was intended to create a system that would be able to create a map from an agent’s experience exploring the environment. This goal is related to, but different from the problem this dissertation addresses. As such, LAIR uses many of the basic constructs used in these other representations but it also introduces other constructs that would not be used in the other representations but that make sense given the problem LAIR is addressing.

LAIR uses geometric data structures to represent the structures used in human spatial cognition. In addition to this, LAIR uses the metric precision of the street

network found in direction-giving services as a global reference frame. This global reference frame is then used to situate the experiential elements of travel (elements such as place, major road skeletons, regions and routes) found in human spatial cognition.

By situating things against a global reference frame, one immediate benefit is that it is possible to easily describe the spatial relationship between elements both in relative terms via known topological connections or in absolute geographic terms. In so doing, LAIR can be used to more fully describe where a person has been (not just the relationships between the places he has travelled). Doing this allows LAIR to describe how previously known, but topologically disconnected, places are located with respect to one another — a particularly important ability when generating route overviews. It also allows us to describe how new location information might be assimilated into the existing knowledge base before that space is actually encountered by a person. Finally, by partitioning routes into known and unknown portions and describing each portion differently, we can alter our route descriptions based on a person's *a priori* spatial knowledge.

This hybridization of cognitive structures with a global reference frame is not found in other spatial representations, and it allows LAIR to more fully model what a person knows and to more completely describe how places that are only known to be topologically connected are located in absolute terms with respect to one another. When generating turn-by-turn directions, current mapping services such as MapQuest know everything about a street network and yet know nothing about which of those streets a person is familiar with. Moreover, these directions only mention the streets on the route but do not include any other types of information to help a person situate themselves as they read through the directions or traverse the route. However, as we discussed in Chapter 2, people use a variety of elements and techniques in addition to the flat road network used by MapQuest to describe space and their movement through it. A route overview and turn-by-turn directions should include these things.

On the other hand, due to the nature of the problem they are trying to solve, computational models such as TOUR, PLAN, and the SSH can represent only those

places an agent knows. These models cannot have a global orientation frame against which they can situate the local location information the agent learns. As a result, these representations are limited in their ability to describe how two disjoint regions are related to one another.

As in TOUR, LAIR represents places, paths, regions, and routes. We describe these elements in the following subsections and will use the global reference frame defined by the street network as a two-dimensional plane to aid in describing the LAIR elements. Since LAIR was designed for use in generating route overviews, we will emphasize the characteristics of the LAIR elements that are used in this process. LAIR can also be used to simulate some of the spatial distortions described in Chapter 2. However, since that is not the focus of this dissertation, we will only briefly discuss those issues here and will return to them in Chapter 5, where we describe how having a model of spatial distortions can be incorporated into route overviews.

3.1.1 Places and Paths

Places are zero-dimensional points that lie on the plane defined by the street network. They are the point abstractions used to represent places, such as cognitive anchor points, in the real world. A Place has a two-dimensional analog called a Region, which describes the Place in two dimensions. Regions are described in Section 3.1.2.

Places lie on Paths. Paths form a network linking one Place to another. Paths can be thought of conceptually as made up of many line segments. We will refer to the individual line segments in a Path as Path segments. Since Paths are situated against a global reference frame, we can use that frame's coordinate system to define a left and right side of a Path, thus allowing us to describe the location of other things with respect to the different sides of a Path. Paths also have an attribute that describe their rank in a road hierarchy.

Since the global orientation frame defines the angles at which Paths intersect, angular distortions caused by local assumptions of street direction do not occur. However, heuristics can be used to emulate the sorts of angular simplifications and

metric distortions described in Chapter 2, such as assuming streets meet at right angles and making global assumptions about geometry based on local observations.

3.1.2 Regions

Regions are 2-dimensional, closed shapes in the plane. **Regions** are used to partition space and group different **Places** and **Paths**. Due to their geometrical nature and the use of the global reference frame, it is straightforward to determine if a **Region** contains a particular **Place** or **Path** as well as to determine containment and overlap relationships between different **Regions**.

Regions can be divided into sub-regions. The union of these sub-regions may not be a total cover of the original **Region**. Also, a given **Region** may be broken into sub-regions in more than one way. For example, a **Region** representing the state of Massachusetts may be divided into cities or major geographical areas such as the Cape, the North and South Shores and MetroWest. Cities can be subdivided into neighborhood **Regions** centered around cognitive anchor points. This last type of **Region**, neighborhood **Regions** centered around cognitive anchor points, are the ones we primarily deal with when generating the route overviews described in this dissertation.

Structuring **Regions** in this way allows us to describe a route at various levels of granularity and through different reference perspectives. This is done by first finding the smallest **Region** containing both the route's start and end points and then using that **Region**'s subregions in the route overview. For example, using the hierarchy given above, if we are travelling between two cities in Massachusetts, we can say that we either pass through certain intermediate cities or pass through certain regions of the state. If there are too few intermediary features to include in the overview (for example, if we are travelling between adjacent cities) we can step down into the hierarchy for further detail (for example, we can downward map from cities to neighborhoods and mention the neighborhoods we pass through as we travel on our route).

In addition to being able to describe the route as a whole, the regional hierarchy

can be used to describe turn-by-turn directions of different portions of the route at different levels of granularity. The turn-by-turn directions for a portion can be less detailed if a person is very familiar with that portion and more detailed if he is not.

Each **Place** has a corresponding **Region** that expands the **Place**'s zero-dimensional abstraction into two dimensions, and each **Region** has a corresponding **Place** that represents the **Region** as a point abstraction. Depending on what scale you are looking at, a given conceptual place can be regarded as either a **Region** or a **Place**. In generating route overviews, we use **Places** to describe the endpoints of a route, whereas **Regions** are used to partition the space covered by that route and give names to the episodic chunks experienced during travel along the route.

Regional boundaries are not fuzzy [60]; they are well-defined line segments. Regional boundaries can be, but don't have to be, defined by **Paths**. We will elaborate on the rationale of choosing a well-defined boundary as opposed to a fuzzy one in Section 3.3 when we describe a subsystem called **CAP LOC** that identifies different neighborhood **Regions** in Boston for use in our route overviews.

3.1.3 Routes

A **Route** between two **Places**, A and B, is made up of the **Path** segments that would be travelled to get from A to B. Some of the **Paths** on a **Route** are the trunk segments that form that **Route**'s skeleton. **Routes** are regarded as individual elements, and when viewed collectively against the global reference frame they represent a person's knowledge of large-scale space. This collection, when viewed as a network, can be used to plan new **Routes** between **Places** using heuristic means-ends analysis algorithms such as those found in **TOUR**.

Also, the collection of **Routes** can be used in a connectionist framework such as **PLAN**. The more oft-travelled **Routes** (and the **Places**, **Paths**, and **Regions** that appear on these **Routes**) can be used as a means of determining which subregions a larger **Region** is broken down into and what **Places** appear in the next level up in the hierarchy.

A summary of the elements in LAIR, along with a description of the fields in each LAIR element, is presented on the next page in Table 3.1.

3.1.4 Cognitive Design Guidelines and LAIR

In this subsection, we examine LAIR against the cognitive design guidelines outlined at the end of Chapter 2. **Regions** group space and define a spatial hierarchy. These characteristics allow **Regions** to be used in route overviews to present the individual steps in turn-by-turn directions in the context of some larger plan (namely, as episodic travel between sub-regions). The use of a global frame of reference allows us to model people's spatial knowledge. We can use this model to separate a **Route** into different portions. In those portions that a person is very familiar with, we can suppress minor details and present only the major parts of a route. In cases where a person is not as familiar with the route, we can provide a little more detail to provide the necessary scaffolding to set expectations and make turn-by-turn directions more comprehensible. Finally, the global reference frame, **Regions**, and the attributes associated with **Paths** provide a means of identifying the skeleton portions of a route and determining the cognitive anchor points along that route. In the next section, we describe how LAIR was implemented using a GIS.

3.2 LAIR's Implementation in GIS

We implemented LAIR using ArcGIS, a commercially available GIS. Typically, a GIS is used to analyze spatially distributed phenomenon. For example, a GIS can be used by city planners to identify potential areas for urban renewal, by epidemiologists to study the spread of disease, or by business owners to find potential areas to open new stores.

LAIR's use of GIS is novel in that the GIS isn't being used as an aid to help people make a decision by analyzing spatial patterns. Instead, the GIS is being used to build a model of what places and paths a person knows. Using an approach common in AI research, we build up a knowledge base of elements and then use those elements and

Summary of LAIR Elements

A **Place** is a zero-dimensional point abstraction for some place in the real-world, such as a cognitive anchor point. A **Place** has the following properties:

- Place name
- X,Y coordinates, to situate this **Place** in LAIR's global reference frame
- A pointer to a **Region** that describes this **Place** in two dimensions

Paths can be thought of as one-dimensional line segments. **Paths** form a network linking one **Place** to another. A **Path** is made up of **Path** segments, and each **Path** segment has the following properties:

- Path Name
- X,Y coordinates of the **Path** segment's two endpoints
- A score to indicate this **Path** segment's ranking in a road hierarchy.

Regions are 2-dimensional, closed shapes in the plane. **Regions** are used to partition space and group different **Places** and **Paths**. A **Region** has the following properties:

- Region name
- A set of polylines that make up this **Region**'s boundaries
- A pointer to a **Place** that serves as a zero-dimensional point abstraction for this **Region**.
- A set of pointers to other **Regions** that this **Region** can be sub-divided into.

A **Route** between two **Places**, A and B, is made up of the **Path** segments that would be travelled to get from A to B. Some of the **Paths** on the **Route** are the trunk segments that form that **Route**'s skeleton. A **Route** has the following properties:

- A pointer to the starting **Place** on the **Route**
- A pointer to the ending **Place** on the **Route**
- An ordered set of **Path** segments linking the starting **Place** to the ending **Place**.

Table 3.1: Summary of the fields in each element of LAIR

the relationships between elements to address our problem of interest. In our specific case, we use the basic elements of the GIS to mirror the cognitive elements people use to organize their knowledge of space. These elements are then included in a route overview to make the route more comprehensible.

Implementing LAIR using a GIS was a natural decision to make for three reasons: the way the GIS organizes its data structures, the spatial algorithms the GIS has available, and the wide assortment of GIS datasets available for use.

A GIS is made up of two things: a spatial database in which each row of the database has a corresponding geometric element: either a point, polyline, or polygon; and a set of algorithms that analyze the spatial distribution of the data in the database (e.g., proximity and density calculations and cluster analysis). Each database table, with its corresponding geometric elements, can be thought of as a layer (similar to the concept of a layer in the image editing program Photoshop). In this work, since we focused on generating route overviews for travel in the greater Boston/Cambridge area, we used the state of Massachusetts' streets layer as our global reference frame. The streets layer is laid out on a coordinate system based on latitude and longitude. We situate all the other LAIR elements with respect to the streets in this layer and the coordinate system it uses.

Given the way the GIS is set up, it was very straight-forward to implement each of the LAIR knowledge elements using GIS data structures:

1. **Places** are represented in the GIS as a layer of points. In the database table, each **Place** has a name and a pointer to a **Region** that describes that **Place** in two dimensions.
2. **Paths** are represented by layers of polylines. We have one **Path** layer that represents those streets from the Massachusetts GIS streets layer that a person is familiar with. This layer contains information including the **Path**'s name, its classification according to Massachusetts' road hierarchy and a more fine-grained description of its local use (as either a principal arterial, a minor arterial, or something less significant). This layer expands as more **Routes** are travelled.

3. **Regions** are layers of polygons. **Regions** have a name, and a list of pointers to one or more layers of sub-regions. When a **Region** has multiple pointers, each pointer corresponds to a different way of dividing the **Region** into sub-regions.
4. **Routes** are layers of polylines. These layers contain a subset of the streets in the global streets layer and they list a starting and an ending **Place**.

The ability of the GIS to perform computational geometry tasks makes it easy to compute containment, overlap, and other spatial relationships. It is straightforward to determine what **Paths** and **Places** are in a given **Region**, and what **Paths** a **Place** is on. Thus, this information is not explicitly stored in any data layer, but rather inferred by the GIS when needed.

In this section we discussed how we implemented LAIR elements in a GIS. The global reference frame, **Path**, and **Route** layers were based on a GIS dataset of streets made available by the state of Massachusetts. In the next four sections, we describe how we define cognitive anchor points (a layer of **Places**) and determine the extent of these anchor points' influence (a layer of **Regions**).

3.3 Previous Work That Maps Cognitive Anchor Points

In generating route overviews, we use LAIR **Regions** to describe which cognitive anchor points punctuate the route described in the turn-by-turn directions. We produce overviews for the type of routes people usually get directions for: travel to new places within the greater metropolitan area in which people live and work. At this scale, the **Regions** we use in our overviews correspond roughly to the “Main Street” regions in Hwang’s urban planning work [28], which we will describe shortly. In this section, we describe how Hwang’s work influenced our approach in using GIS to identify where cognitive anchor points are located and the extent of the **Regions** defined by these anchor points.

There were no existing GIS datasets that had polygons we could reasonably use as **Region** boundaries. The existing datasets were either derived from formal municipal boundaries and whose polygons were too large to capture the local extent of a cognitive anchor point, or the datasets contained too many features, thus making them too fine-grained for use at the scale we are concerned with. Because of this, we decided to create LAIR **Region** boundaries by using GIS data to select **Places** that could be considered cognitive anchor points and then designating a boundary for the **Region** defined by that cognitive anchor point. The challenge in this process is that there are many possible datasets from which we can choose cognitive anchor points, each with many possible features to sift through and use in determining that anchor point's extent. The emphasis of this section is on how we framed this problem so as to produce a solution that is tractable and consistent with the results of Lynch's study on people's image of the city (see Section 2.1).

Urban planners have used GIS to find Lynch nodes to include in a Lynch-like image of the city. One of the earliest works in this area was done by Singh [78]. He used GIS to analyze land usage patterns to find geographic areas which had a large amount of diversity in the surrounding land use, and these were the areas he designated as nodes. This work was largely exploratory, and Singh's intent was to provide a sense of what was possible with the (at the time) emerging GIS technology. Singh made no conclusion as to the comprehensiveness of his method, instead positioning it as a tool to assist urban planners in their work. More recent work in this area was done by Hwang.

Hwang [28] used GIS in a mixed-initiative manner in which users of her system (urban planners, architects, and city administrators) identified canonical "Main Street" regions of a city. Hwang defined a Main Street region as the built-up commercial district where local neighborhood residents purchase goods and services. After getting user-provided examples of Main Street regions, her system, the Heuristic Noll Map (HNM), would identify similar regions using machine learning techniques on GIS datasets and then present these areas to the user for consideration. Hwang implemented the HNM using a support vector machine (SVM), and as users interacted

with the HNM, either agreeing with or disagreeing with its suggestions to either add or remove candidate Main Street regions, the HNM improved its classifier with the new training data. Hwang also analyzed which features were a large factor in the classifier's decision making.

The key insight from Hwang's work that we apply here is the way in which she approached the problem of identifying Main Street regions. Hwang's approach is to recognize that there is no one 'right' answer when identifying the locations of Main Street regions. From interviews of her intended users, she found that while in aggregate, there was enough overlap in the regions people selected for her to identify Main Street regions in different neighborhoods in Boston, each user gave a different number of Main Street regions as well as different boundaries for each Main Street region. Moreover, in her analysis of the features used by the HNM's classifier, which she interpreted as a proxy for the user, Hwang found that people used different features in varying degrees of importance when determining which areas are Main Street regions.

Hwang's work indicates that we should not be concerned with finding *the* "right" set of **Regions** to use to partition a city for use in our route overviews. So the approach we take is to use a definition for cognitive anchor point that is consistent with the one Lynch derived. Lynch described nodes as being transportation hubs or as being a concentration of public activity. Given this definition, we then search for a reasonably sized set (between 15 and 25) of **Places** consistent with this definition. Once these **Places** are identified, we then use their corresponding **Region** definition in the route overview. We search for **Places** that possess a specific set of features that are consistent with Lynch's description for a node. We will not worry about exact boundaries for the corresponding **Regions** because the granularity of the language used in route overviews renders the exact specifics of a **Region's** boundary as irrelevant (and thus the fact that some **Region** boundaries are fuzzy becomes a moot point).

Singh's and Hwang's work both emphasized commercial features in searching for Main Street-type regions. To determine if there are other types of features we should

consider in searching for cognitive anchor points, we conducted the user survey described in the next section.

3.4 Deriving a Mental Map of Boston

To understand what sorts of cognitive anchor points are prominent in a person's mental map of Boston, we conducted a survey similar to the one Lynch conducted. In our study, we recruited 30 subjects from the greater Boston area and asked them to list "prominent places in their mental map of Boston." After compiling their list, we then showed participants a street map and asked them to outline the boundaries of the places on their list. Based on feedback from people who tested early versions of this protocol, we included on the map the locations of Boston's subway stops as a spatial reference.

We were concerned that showing the subway stops on the maps we asked participants to mark up would bias the data we received. We minimized this risk by asking participants to first generate their list of prominent places and think about their locations before presenting the map to mark up. Moreover, early feedback indicated that not including the location of subway stops would have unduly hindered the mark up task, greatly outweighing the possible risks associated with showing the subway stops.

To analyze and aggregate the maps submitted by study participants, the map we asked participants to mark up was shown on a tablet PC running ArcGIS. Using ArcGIS, participants would outline on a digital map the locations of their prominent places while panning and zooming around the map using an interface similar to Google Maps (see Figure 3-1 on the next page for an example of the marked-up maps participants produced).

Although the size of our sample group was fairly modest, its composition was well varied. We had 14 men and 16 women, who ranged in age from 19 to 52, with a mean age of 30.2 and a standard deviation of 9.2 years. Participants had lived in Boston anywhere from 5 months to 47 years with a mean time in Boston of 11 years. All

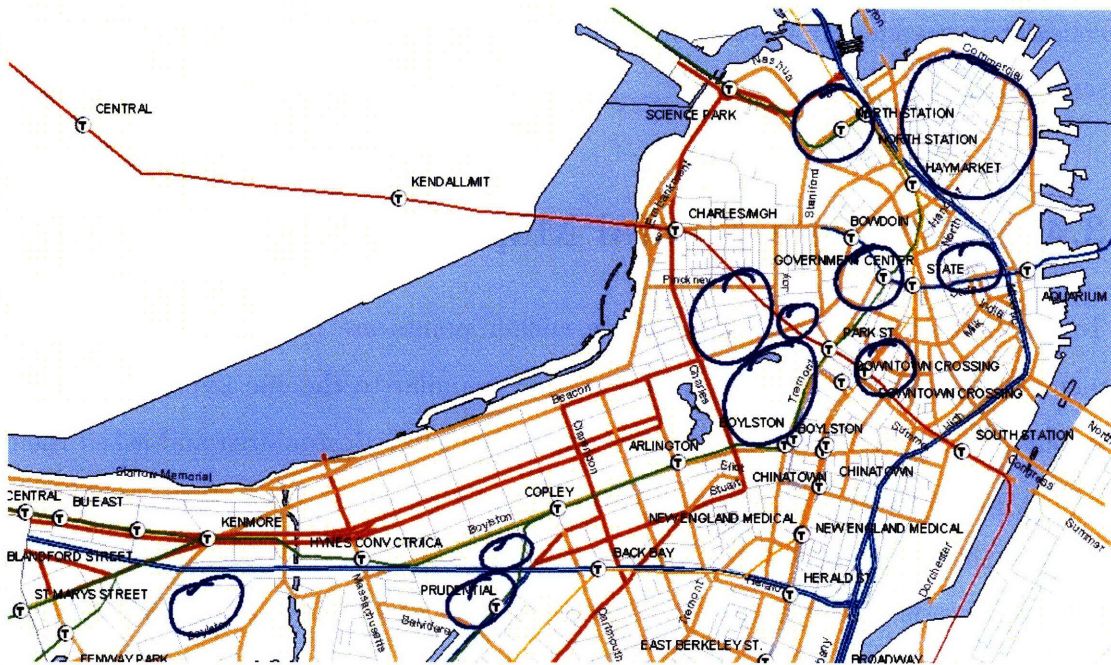


Figure 3-1: Participants used ArcGIS to outline the locations of prominent places in their mental map of Boston.

but 3 participants had lived in Boston for over 1 year. Thirteen of our subjects rode the subway more than 10 days a month, twelve subjects drove more than 10 days a month; two people appeared in both of these groups. When asked the zipcode in which their home mailing address was located, participants responded with 24 unique zipcodes. We recruited 13 participants from MIT, which is located in the nearby city of Cambridge, across the Charles River from Boston. The other 17 participants were recruited from public areas in the greater Boston area.

To combine all the maps into one aggregate basemap depicting prominent places in Boston, we first divided our sample into four subgroups and identified places that were listed by participants in all four of the subgroups. Controlling for gender and whether they were from MIT, the size of our subgroups were 7 MIT men, 5 MIT women, 6 non-MIT men, and 11 non-MIT women. There were 16 places in Boston that were named by at least 1 person in each of the subgroups. These places are listed on the next page in Table 3.2.

After identifying the names of these places, we took the corresponding shapes defined by their boundary outlines, pixelated the shapes at a resolution of 10 meters

Type of Place	Place Name	Distance from Subway Stop (m)
Large Land Parcels	Boston Common	0
	Faneuil Hall/Quincy Market	120
	Fenway Park	190
	Government Center	0
	Prudential Center	0
	Public Garden	0
	Symphony Hall	0
Commercial Regions and Areas of Cultural Importance	Charles Street	50
	Chinatown	0
	Copley Square	0
	Downtown Crossing	0
	Faneuil Hall/Quincy Market	120
	Kenmore Square	0
	Newbury Street	0
	Prudential Center	0
Theater District	10	
Neighborhoods with Defining Geographies	Beacon Hill	280
	North End	40

Table 3.2: Prominent places listed by participants from each of our four subgroups, along with the distance to the nearest subway stop, to the nearest 10 meters. Places that overlapped the location of a subway stop have a distance of zero.

by 10 meters, and laid the pixelated shapes on top of one another. We then took the pixels that appeared within the boundary sketches of people from at least 3 of the 4 subgroups. The places that appear in the resulting map are shown in Figure 3-2. As an example of how to interpret that map, for each pixel in the place labelled “Kenmore Square”, at least 1 person from 3 (possibly 4) of our 4 subgroups considered that pixel in Kenmore Square.

We chose to set our threshold at 1 person from 3 of the 4 subgroups instead of all 4 subgroups because for a few of the prominent places that we identified, there were some subgroups in which only 1 person from that subgroup named that place. This fact would be problematic in cases when either that 1 person forgot to indicate on his map the boundary to that place or if that single person’s boundary outline intersected the other outlines in a cursory manner.

Initially, we considered building a basemap of prominent places by pixelating the shapes outlined in all 30 maps, laying them on top of one another, choosing a

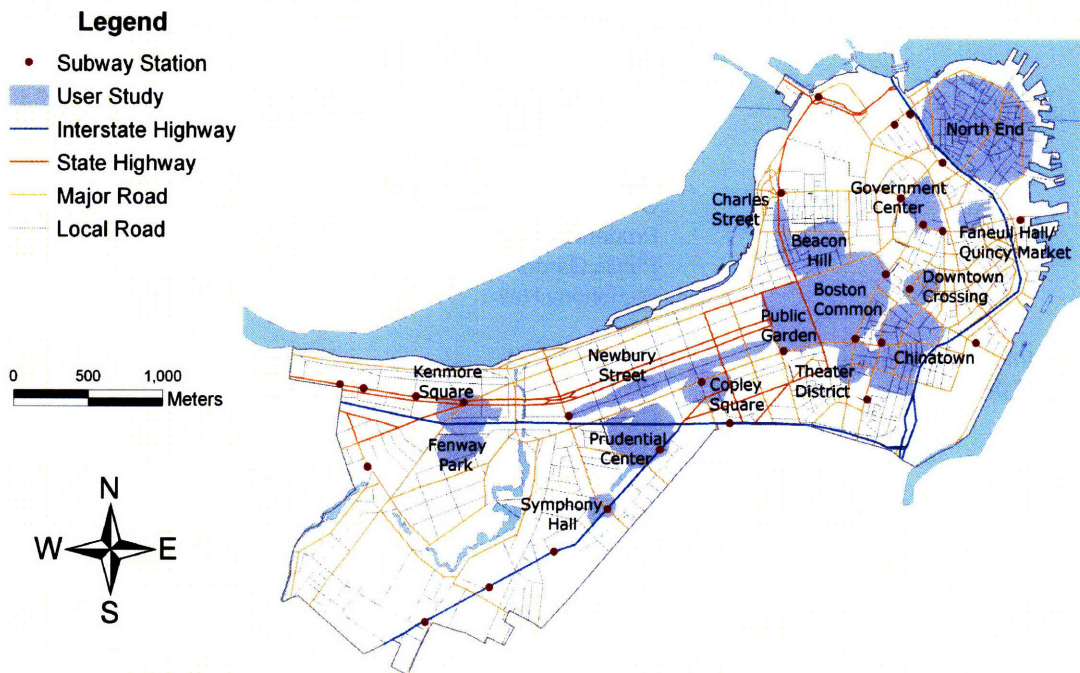


Figure 3-2: Our aggregate map of prominent places in Boston.

threshold percentage, t , and then selecting those pixels that appeared in t percent or more of the maps. We did not use this “raw count” method to generate the basemap because it does not allow us to control for any variables used to characterize study participants.

3.5 Analysis of the Derived Mental Map

Figure 3-2, above, shows our aggregate map of Boston’s prominent places. The actual city of Boston extends beyond the area shown in Figure 3-2, but we limit our analysis to the area shown, which represents the union of three municipal regions defined by the city of Boston for urban planning purposes. Compared to a commercially available map of Boston, the extent of the places shown in the figure are similar, with two exceptions. The first is Beacon Hill; the area labelled “Beacon Hill” in Figure 3-2 covers only the southern slope of the actual hill. This is where the prominent State House (capitol building) is located. The northern half of Beacon Hill is bounded by the first major road (yellow street in the figure) north of the Beacon Hill boundary

as shown. Also, in Figure 3-2, the northern border of Chinatown is drawn two or three blocks further north than on the commercial map we referenced. Note that we are only referencing a commercial map as a consistency check for boundaries. Since commercial maps often contain many tens or hundreds of “points of interest”, commercial maps would be too liberal when listing the cognitive anchors in a person’s mental map of a city.

There are two noteworthy observations regarding where prominent places are located and what types of places people included in their mental map of Boston. The first observation is that all of the places in the basemap are extremely close to subway stops. All of the places were within 280 meters of a subway stop, and in many cases, a subway stop was located within the boundary of a prominent place (indicated by a distance of zero in Table 3.2). While this data is taken from participants in Boston, it does provide empirical evidence to support the more general notion that major transit points on the scale of cities (such as subway stops in cities like New York and London) are located near important places and thus are relevant features around which people structure their mental maps.

Although this observation was based on dividing our sample by gender and whether they were affiliated with MIT, this trend still appeared when we compared prominent places listed by participants who drove more than 10 days a month (who we will refer to as frequent drivers) to those places listed by participants who rode the subway more than 10 days a month (frequent subway riders). When we selected those places that were listed by at least 30% of participants in both of these groups, we found 10 places, all of which appear in the list of places in Table 3.2. Moreover, when asked to rate how difficult it would have been to locate the places on their list if subway stations were not included for reference, 22 of our 30 participants responded that it would have been more difficult to locate their prominent places. Frequent drivers and frequent subway riders were equally represented in this group, suggesting that the location of subway stops is important in the structure of mental maps, independent of whether people primarily ride the subway or drive.

The second observation was that the prominent places people listed could be

classified into three different groups: specific, well-defined places located on large parcels of land (for 6 of the 7 places in this class, the area occupied by each of these places was more than 2 standard deviations above the average parcel size in the area of Boston shown in Figure 3-2); regions that draw many people, mostly due to commercial or cultural attractions; and neighborhoods defined by their geography (Beacon Hill's elevation sets it apart from other parts of Boston, and the North End was largely separated from the rest of Boston due to a large public construction project that was only recently completed). Table 3.2 shows how each place is classified; some commercial regions are also located on large parcels of land.

Taken together, these two observations suggest that in addition to a skeleton formed by a road network, a mental map is also framed by major transit points. Located near these transit points are noteworthy places that are prominent due to the amount of area they occupy (which may be correlated to their importance as a place of commerce, local government, or as a place of public gathering), their large commercial or cultural impact, or their defining geographical features.

3.6 Populating LAIR with CAP LOC

We created a sub-system called CAP LOC (**C**ognitive **A**nchor **P**oint **L**ocator) to populate LAIR. CAP LOC uses the results from our previous study of prominent places in Boston to identify cognitive anchor points (**P**laces) and the extent of their influence (**R**egions). CAP LOC does this by using GIS data to first identify potential cognitive anchor points at the locations of major transit points — either subway stops or street intersections. CAP LOC selects street intersections if they are formed by streets that are highly ranked in the Massachusetts road hierarchy or if they are local arterials.

After identifying these candidate cognitive anchor points, CAP LOC defines a search radius around each of these points and searches for commercial land parcels and land parcels that are much larger than their neighboring parcels (spatial outliers) in the search radius. Potential points with either a spatial outlier nearby or commercial land parcels whose total area is above a certain threshold are considered to be cognitive

anchor points (and thus form a layer of **Places**). The parcels we find define the extent of the anchor point's influence (these parcels form a layer of **Regions**). Based on our knowledge of neighborhoods in the greater Boston area, we then manually named the **Regions** CAP LOC selects. In selecting parcels, CAP LOC gives priority to spatial outliers over commercial regions, so that if we find a spatial outlier in our search radius, we select only that parcel to reflect the fact that these large parcels generally have well-known names and are the defining feature in that area.

CAP LOC uses Anselin's Local I [6] to identify spatial outliers. This is a GIS algorithm that, given a particular feature of interest, detects groups of parcels with similar values and parcels that are outliers. It does this by going from parcel to parcel to see if the value of one parcel's feature (in our case, the parcel's area) is similar or different, in a statistically significant sense, to the average values of parcels surrounding it.

The step-by-step algorithm CAP LOC uses to find cognitive anchor points and their corresponding **Regions** is presented below in pseudo-code:

1. Select subway stations that are either at subway line junctions or at intersections of high ranking roads in the Massachusetts road hierarchy. Search within a certain distance, r_1 , of these transit points for spatial outliers or commercial regions above a certain threshold size, t_1 . After experimenting with different values, we used $r_1 = 175\text{m}$ and $t_1 = 250,000\text{m}^2$.
2. For all other subway stations, search within a certain distance, r_2 , of these subway stations for spatial outliers or commercial regions above a certain threshold size, t_2 . In our experiments, we used $r_2 = 175\text{m}$ and $t_2 = 250,000\text{m}^2$.
3. Select intersections of local arterials that are more than r_{min} meters from any previously selected subway station. Search for spatial outliers or commercial regions above a certain threshold size, t_3 , within a certain distance, r_3 , of these intersections. In our experiments, we used $r_{min} = 300\text{m}$, $r_3 = 175\text{m}$ and $t_3 = 100,000\text{m}^2$.

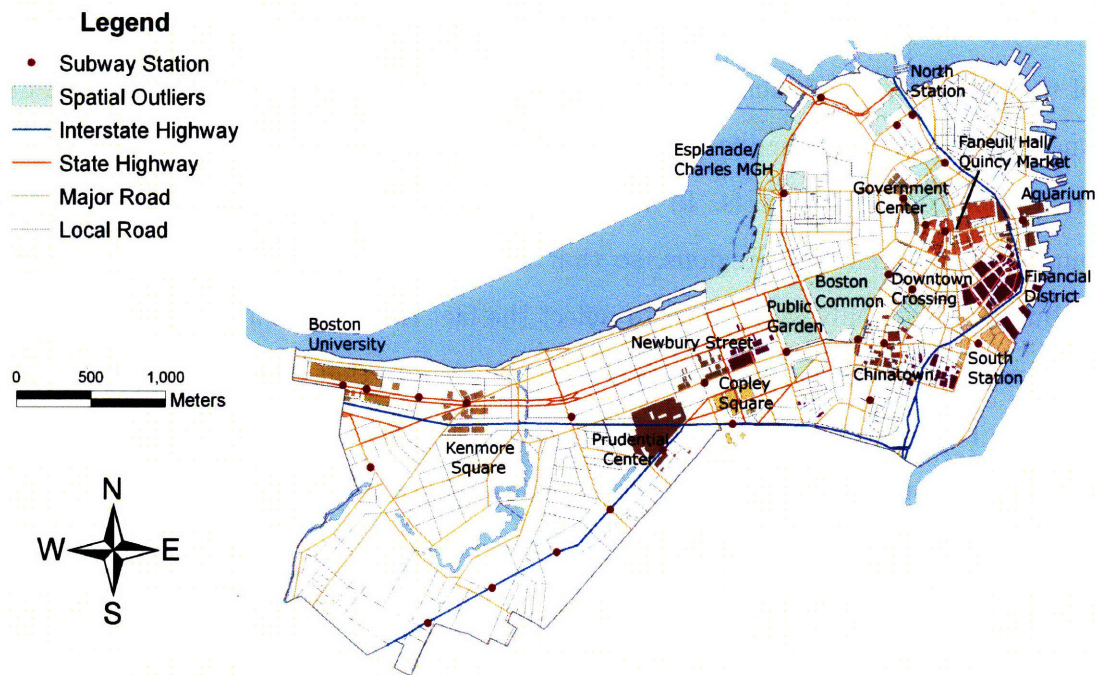


Figure 3-3: The Regions CAP LOC identifies. Regions that are not colored green are commercial regions. Some of these Regions (namely, Newbury Street and Chinatown) were selected by CAP LOC because they are commercial regions that meet CAP LOC's second and third search criteria. Because of this, they show up on the map in two different colors.

The resulting layer of Regions in Boston CAP LOC identifies is shown in Figure 3-3. For comparison purposes, Figure 3-4 superimposes the prominent places found in our user study against the Regions CAP LOC found. CAP LOC found many of the places identified in the user study. It found a few places not listed in the study, and it did not find Beacon Hill or the North End, Regions that are defined by their unique geographies, a feature CAP LOC does not consider in its search algorithm. Overall, CAP LOC's output forms a reasonable set of places to use in constructing route overviews.

3.7 ROVER

We have described the LAIR knowledge representation, its implementation, the CAP LOC system that uses different GIS datasets to populate LAIR, and the reasons why

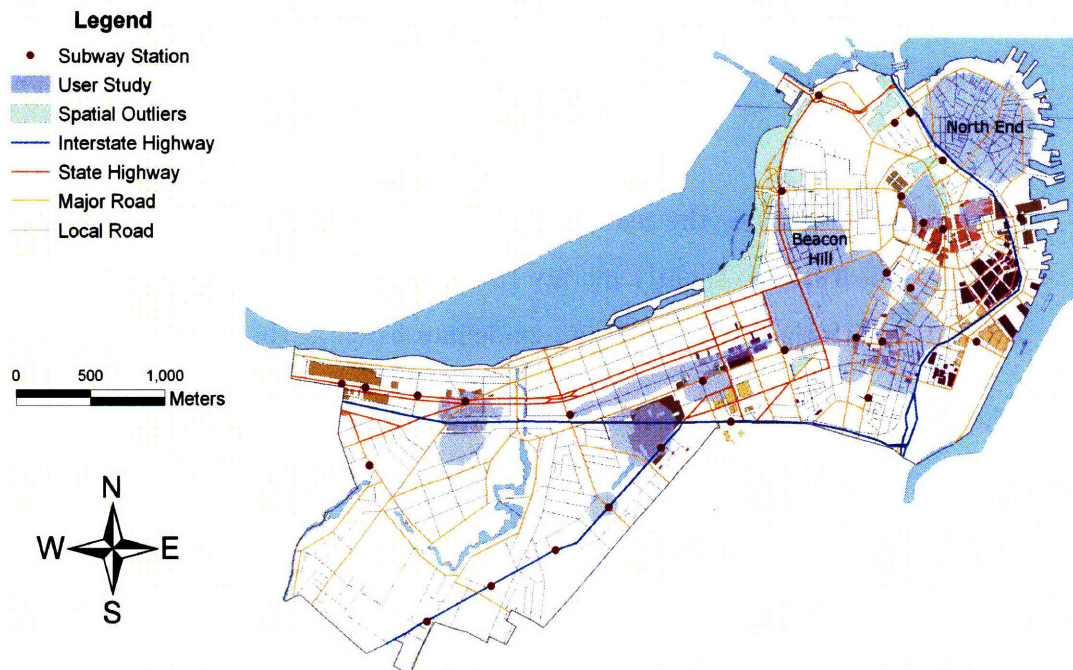


Figure 3-4: A comparison of the Regions CAP LOC identifies compared to the prominent places identified in our user study. CAP LOC does not identify Beacon Hill or the North End, Regions that are defined by their unique geographies, a feature CAP LOC does not consider in its search algorithm.

we used those datasets in the way we did. This section describes ROVER and how it uses Routes, Region boundaries, and road skeletons to generate route overviews. The overall process ROVER follows to produce its output is shown in Figure 3-5. The specific steps ROVER carries out is summarized as follows:

1. We use Google Earth as a front end to handle our route planning and generation of standard turn-by-turn directions. The route generated by Google Earth is then imported into ArcGIS as a LAIR Route.
2. ROVER examines the Route and identifies which Path segments in the Route form that Route's skeleton.
3. This Route is then added to the knowledge base of Routes a person knows and compared against the pre-existing Routes in that knowledge base to determine how much of the current Route with which a person is familiar.

4. ROVER then selects previously known **Regions** that the current **Route** passes through and situates the **Route's** destination in terms of its skeleton and known **Regions**.
5. Once ROVER identifies the features to include in the route overview, ROVER uses these features to slightly modify the turn-by-turn directions in a manner consistent with the cognitive design guidelines listed at the end of Chapter 2.
6. The features to be included in the route overview and in the modified turn-by-turn directions are then passed to a set of scripts that produce the textual route overview and modified turn-by-turn directions.

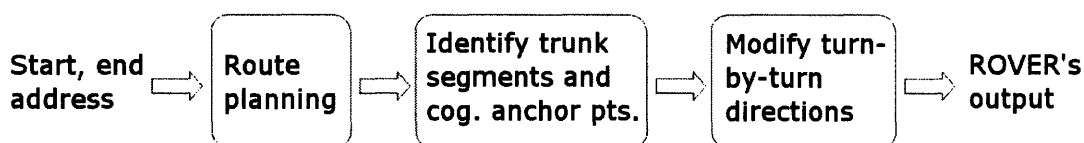


Figure 3-5: The overall process ROVER follows to produce its output.

The following subsections elaborate on this process and the rationale behind it.

3.7.1 Bringing Google Earth into LAIR

We used Google Earth as a front end because it was an off-the-shelf piece of software that made it easy to get driving directions between two places. Google Earth allows a user to specify a location either by address or by dropping a placemark on a map. A description of the streets in the route and the turn-by-turn directions that Google Earth produces is then exported to XML. We then convert this XML into a **Route** layer using a utility we created for this purpose.

One might wonder why Google Earth was used when much of this work is already based in a GIS. While ArcGIS does have the ability to do route-planning, it does not provide an API to easily identify the individual **Path** segments in the **Route** or to manipulate the detailed turn-by-turn directions produced. An added benefit to using Google Earth is that it has a user interface that is very easy to manipulate

and it invites users to pan, zoom, and rotate the display. We describe how Google Earth can be used as a visualization engine to appeal to different aspects of spatial cognition, such as hierarchy and frames of reference, in our discussion of future work in Chapter 5. Figure 3-6, below, presents a screenshot of the Google Earth interface.

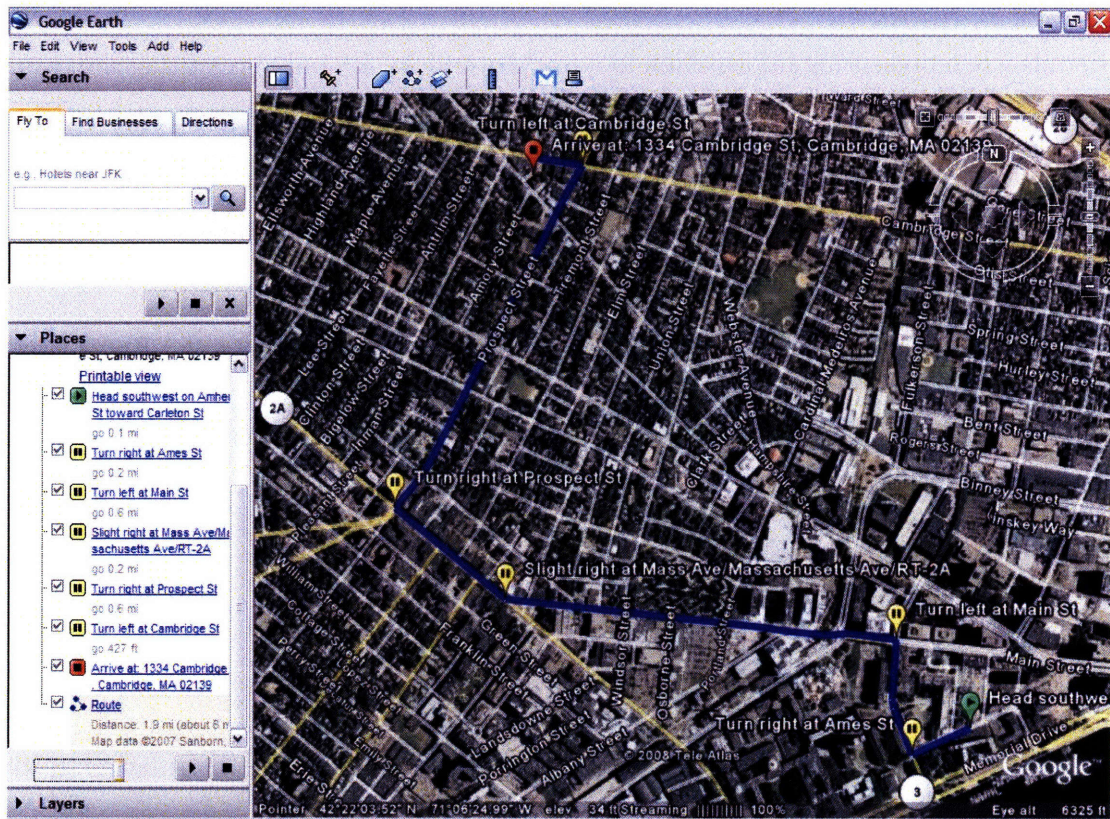


Figure 3-6: Google Earth screenshot.

3.7.2 Identifying the Trunk Segments of a Route

When Routes are imported into LAIR, we first identify which Paths make up the major trunk segments of that Route. These Paths are later included in the route overview. The trunk segments are the Paths that make up a significant percentage of the Route's length, and thus serve as a major skeleton to frame the Route. After experimenting with different threshold percentages, we selected 10% as the minimum threshold for a Path to be considered a trunk segment. We also include the last major Path (as determined by the classification system used by the state of Massachusetts) on the

Route as a trunk segment, regardless of the length of this Path. We do this because this last trunk segment is used to situate the Route's destination and is regarded as the final sub-goal to travel to in the route overview.

Note that trunks may not be contiguous. This is in line with our earlier design criteria of appealing to a person's sense of spatial hierarchy by suppressing specific details when presenting information in a route overview. By only presenting the trunk segments, we break the route into subgoals and treat movement along the route as a sequence of episodic events in which a person is travelling between major road segments and cognitive anchor points. We rely on the person's knowledge of space (and if that is lacking, the turn-by-turn directions) to fill in the details.

To summarize the criteria we use to identify the trunk segments of a Route, we:

- Select Paths that make up more than $N\%$ of the Route. In our experiments, we set $N = 10\%$.
- Select the last major Path on the Route.

3.7.3 Modelling What Places a Person Knows

Importing new Routes into LAIR is the means by which we add more information to our model of a person's knowledge of geographic space. This model consists of a layer of Places, a layer of Paths, a layer of Routes, and a layer of Regions that represent known spatial elements. This model is used to tailor route overviews to the individual, so that we only highlight those features that would best help that individual comprehend the route. (As we will discuss in the next chapter, however, the structure provided by the route overview also makes driving directions more comprehensible to people for whom the overview was not tailored.) In this subsection we describe how this model is built up.

The entries in our Paths layer represent the subset of the global streets reference frame with which a person is familiar. We seed the user's knowledge base with an initial set of Path segments taken from the area surrounding his home and work. Any Places and Regions from the set of cognitive anchor points in Boston and Cambridge

that intersect these known **Path** segments are also regarded as known by that person and are placed in their respective **Place** or **Region** layer. When a new **Route** is imported into LAIR, we add any of the **Path** segments on that **Route** not already in the known **Paths** layer to that layer. As new **Path** segments are added, we check if these segments overlap any of the **Places** or **Regions** that we defined as cognitive anchor points. If they do, then these elements are also added to their respective known layers.

Since we tailor route overviews based on people's knowledge of space, we make note of which **Path** segments on that **Route** were previously unknown. If a new **Route** includes a **Path** segment which has not been travelled before, but is part of a previously known **Path** in the same city, then that new **Path** segment is considered known. We do this to account for the case when people encounter **Paths** in disconnected pieces. In this case, knowing the name of the **Path**, even if they haven't been on this particular part of the **Path** before, carries some measure of familiarity and we want to use this to help situate people in the route overviews. The next section describes how we take into account known and unknown **Path** segments when structuring route overviews.

The following list summarizes what occurs when a new **Route** is added to LAIR:

1. ROVER identifies **Path** segments on the **Route** that were not previously in LAIR's model of **Path** segments with which a person is familiar. These **Path** segments are then added to the known **Paths** layer.
2. If the newly added **Path** segments overlap any previously unknown cognitive anchor points, ROVER adds these **Places** and **Regions** to their respective known layers.
3. ROVER then separates the **Path** segments on the **Route** into a list of known and a list of completely unknown **Paths**. A **Path** segment is considered to be known if it was previously in the known **Paths** layer or if it is part of a previously known **Path** in the same city. Otherwise, the **Path** segment is considered unknown.

3.7.4 Structuring Route Overviews

We structure route overviews in such a way that they can be read as a sequence of subgoals, and movement along the route is viewed as a sequence of episodic events in which a person is travelling between trunk segments and cognitive anchor points. We previously described how trunk segments were selected. To determine which cognitive anchor points to include in the route overview, we used ArcGIS to identify the known cognitive anchor points that overlap known *Path* segments on the *Route*. These selected anchor points are listed at the end of the route overview. We also situate the *Route*'s destination at the beginning of the overview by noting if it is located in a known *Region*. If the destination is not, then we describe it as being on either a major trunk or on a small side street just off the last trunk.

After identifying the trunk segments of a *Route*, the *Route* is then broken up into three different chunks: a beginning, a middle, and an end. We define the chunks as follows:

- The *beginning* of the *Route* consists of the first trunk segment and all *Path* segments before the first trunk.
- The *end* consists of the last trunk segment and all *Path* segments after the last trunk.
- The *middle* consists of all *Path* segments in between.
- In the case where there is just one trunk, that trunk is the middle, and all *Path* segments before it make up the beginning, and all *Path* segments after it make up the end.

Based on this chunking scheme, routes with two trunk segments will have an empty middle chunk. However, this does not affect the way in which we produce route overviews. As we will describe shortly, dividing a route into three parts is used to influence how the overall route overview is structured. This reason is different from that used in the CORAL system [17]; CORAL divided a route into three parts as a

means of segmenting turn-by-turn directions, not for generating route overviews or for accounting for how person's knowledge of the path segments in each chunk would influence how directions are given.

After we classify the Path segments in the Route as either known or unknown, we examine the pattern of known and unknown Path segments to determine how the overview should be structured. We consider three possible patterns of known and unknown Path segments. We call these patterns knowledge profiles, and describe how the knowledge profile affects how the route overview is structured. In the description of each profile, we provide an example based on a common Route, shown in Figure 3-7.

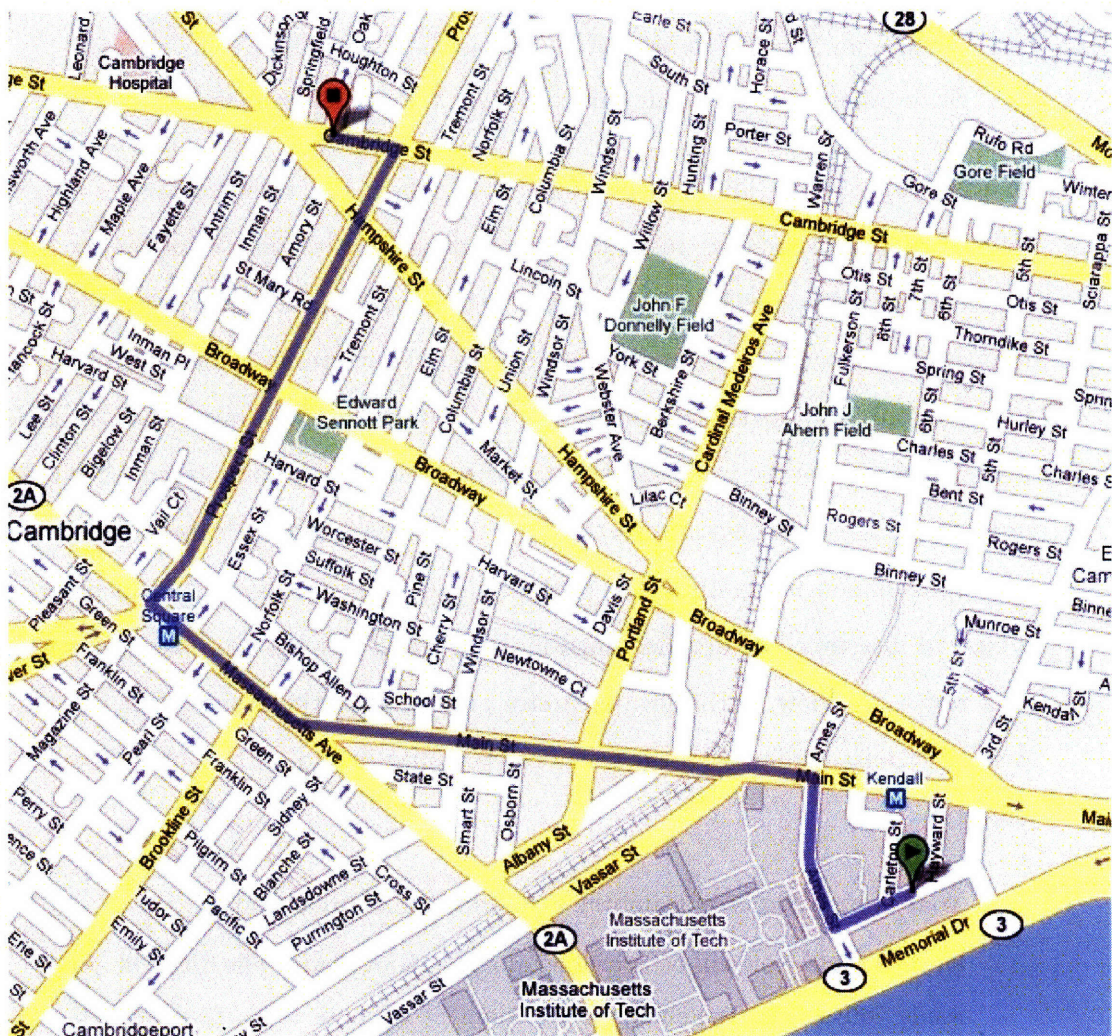


Figure 3-7: This route is used to illustrate how a person's knowledge profile influences the route overview generated by ROVER.

- In the *full-knowledge* profile, the person knows all major trunks on the route. Based on this high degree of familiarity, the route overview lists all the trunks in a straight-forward manner. The rationale for this is that the individual, with his knowledge of the major trunks and their relations to one another, can infer the specific details of the route without having to consult the detailed turn-by-turn directions. Here is an example of a route overview when a person is in the full-knowledge profile:

Your destination is in Inman Square.

On this route, you will be travelling along 4 major roads. You will be going from Main St to Mass Ave/Massachusetts Ave/RT-2A to Prospect St to Cambridge St. Along this route, you'll pass Central Square.

- In the *proper-prefix* knowledge profile, the person knows some proper subsequence of path segments in the route starting at the route's beginning (the prefix), but does not know the last trunk in the route (and possibly other trunks in between). In this case, the route overview presents the known trunks in one grouping and the unknown trunks in another grouping. The rationale behind this is that we want to explicitly distinguish the known trunks from the unknown. The unknown trunks are then presented as major episodic divisions to help a person understand the route, and to give some sort of framework in which to interpret the detailed turn-by-turns instructions. Here is an example of a route overview in the proper-prefix profile, with the assumption that the person is not familiar with the parts of the route including and after Prospect Street:

Your destination is on Cambridge St, which is a major road.

On this route, you will be travelling along 4 major roads. The first 2 major roads you will be travelling on are Main St to Mass Ave/Massachusetts Ave/RT-2A. After that, the other major roads you will be travelling on are Prospect St to Cambridge St. Along this route, you'll pass Central Square.

- In the *prefix-suffix* knowledge profile, the person is familiar with the trunks at the beginning (the prefix) and end of the route (the suffix), but does not know all the path segments in between. In this case, travel through the unknown middle creates new topological links between the prefix and suffix parts of the **Route**, essentially stitching these parts together. This situation can arise in at least two different situations. In one case, a person may be a frequent subway rider, and thus may not know the above-ground relationships between many places. Therefore, his above-ground spatial knowledge consists of disconnected islands of space. The second situation occurs when a person already knows a route between two places, but is following a new route between those places.

In this profile, we present the route as a plan that is broken into three subplans. The first subplan is travel from the start of the **Route** to the last known **Path** segment in the prefix; the last subplan is travel from the first known **Path** segment in the suffix to the destination, and the middle subplan is travel that links, or stitches together, the other two subplans. Here is an example of a route overview in the prefix-suffix profile, with the assumption that the person is unfamiliar with the fact that Prospect Street connects Massachusetts Avenue and Cambridge Street:

Your destination is in Inman Square.

On this route, you will be travelling to Mass Ave/Massachusetts Ave/RT-2A, then going from Mass Ave/Massachusetts Ave/RT-2A to Cambridge St by Inman Square. The route then goes from Cambridge St to your destination. Along this route, you'll pass Central Square.

The three different overviews for this route are listed together in Table 3.3 for comparison.

We did not consider the possible knowledge profile where a person does not know path segments at the beginning of the route but knows segments towards the middle and end because we deemed this an unlikely scenario. Given a starting point, we assume that people know the way to the nearest major trunk. Since the beginning

Full-Knowledge	Proper Prefix	Prefix-Suffix
Your destination is in Inman Square.	Your destination is on Cambridge St, which is a major road.	Your destination is in Inman Square.
On this route, you will be travelling along 4 major roads. You will be going from Main St to Mass Ave/Massachusetts Ave/RT-2A to Prospect St to Cambridge St. Along this route, you'll pass Central Square.	On this route, you will be travelling along 4 major roads. The first 2 major roads you will be travelling on are Main St to Mass Ave/Massachusetts Ave/RT-2A. After that, the other major roads you will be travelling on are Prospect St to Cambridge St. Along this route, you'll pass Central Square.	On this route, you will be travelling to Mass Ave/Massachusetts Ave/RT-2A, then going from Mass Ave/Massachusetts Ave/RT-2A to Cambridge St by Inman Square. The route then goes from Cambridge St to your destination. Along this route, you'll pass Central Square.

Table 3.3: Route overviews for the same route, generated using different knowledge profiles. Significant differences in the structure of the route overviews are highlighted.

of a route is always considered to be known, that leaves only three possible knowledge profiles because the proper-prefix profile covers both the case where a person is familiar with the middle chunk of a route but not the end, as well as the case where a person is not familiar with the middle and end chunks.

There is one final note to make about the selection of trunk segments and the cognitive anchor points we include in the route overview. Since people can only hold a limited amount of information in their short-term memory, we limit the number of trunk segments and cognitive anchor points we include in the route overview. Because people can only hold roughly 5 ± 2 items in short-term memory¹ [15], we limit the number of trunk segments we include to 5 and the number of cognitive anchor points to 3. In practice, this situation does not arise often. When it does, we take the three cognitive anchor points closest to the end of the route, and we use the following algorithm to select trunk segments:

1. Select the last trunk segment.

¹In 1956, George Miller published the influential paper *The Magical Number Seven, Plus or Minus Two* [57]. However, as Mandler and Chown point out [13, 50], Miller's use of the number 7 was as a threshold between chance and perfect performance on memory tasks (as well as a largely effective rhetorical device). More recent studies have suggested that the number of chunks people can *accurately* hold in short-term memory is closer to 5 ± 2 [15].

2. If we are in the proper-prefix or prefix-suffix profile, select the last known trunk segment before the first unknown Path segment in the route.
3. If we are in the prefix-suffix profile, select the first trunk segment after the last unknown Path segment in the route.
4. Sort the remaining trunk segments by length and then select the top N segments so that the total number of selected trunk segments is five.

3.7.5 Modifying Turn-by-Turn Directions

In addition to producing route overviews, ROVER also applies our cognitive design guidelines and uses the information included in the route overview to slightly modify the turn-by-turn directions in the following ways:

- Since we assume that people know how to travel to the first trunk segment in the route, and in keeping with the guideline of selective suppression, we replace the beginning of the turn-by-turn directions with a single instruction telling a person to go to the first trunk segment. For the given route shown in Figure 3-8, ROVER makes the following change:

Original:

1. Head southwest on Vassar St toward Mass Ave/Massachusetts Ave/RT-2A. Go 0.2 mi
2. Turn right at Mass Ave/Massachusetts Ave/RT-2A. Go 2.9 mi

ROVER's Modifications:

1. **Go to Mass Ave/Massachusetts Ave/RT-2A and turn right.** Go 2.9 mi.

- In keeping with the design decision to help people situate individual turn directives in the context of a larger goal, we describe what cognitive anchor points or major trunk segment a person is headed towards when travelling along a trunk segment. We use ArcGIS's ability to determine if spatial elements overlap to determine which known cognitive anchor points a given trunk segments passes by or heads toward. If there are no such points, we use the next trunk segment

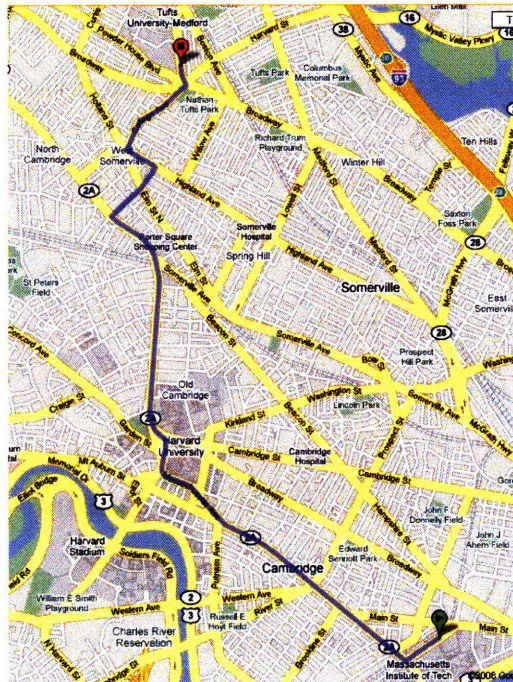


Figure 3-8: A map of the route used to illustrate how ROVER modifies Google Earth's turn-by-turn directions.

on the route, if it is known. Continuing from above, we modify the directions as follows:

Original:

1. Head southwest on Vassar St toward Mass Ave/Massachusetts Ave/RT-2A. Go 0.2 mi
2. Turn right at Mass Ave/Massachusetts Ave/RT-2A. Go 2.9 mi
3. Turn right at Russell St. Go 0.2 mi
4. Continue on Cutter Ave. Go 0.1 mi
5. Turn left at Highland Ave. Go 0.2 mi

ROVER's Modifications:

1. Go to Mass Ave/Massachusetts Ave/RT-2A, turn right, **and then head towards Central Square.** Go 2.9 mi. **As you travel on Mass Ave/Massachusetts Ave/RT-2A you will go past Central Square, Harvard Square, and Porter Square.**
2. Turn right at Russell St. Go 0.2 mi
3. Continue on Cutter Ave. Go 0.1 mi
4. Turn left at Highland Ave. Go 0.2 mi. **As you travel on Highland Ave you will go into Davis Square.**

Combining both types of changes to the turn-by-turn directions and then including the route overview, this is ROVER's output for the route shown in Figure 3-8, assuming a person is in the prefix-suffix knowledge profile:

From the Stata Center to 165 College Ave, Somerville MA

Your destination is near Tufts University.

On this route, you will be traveling on Mass Ave/Massachusetts Ave/RT-2A to get to Porter Square, then going from there to College Ave by Davis Square. The route then goes from College Ave to your destination. Along this route, you'll pass Harvard Square, Porter Square, and Davis Square.

1. Go to Mass Ave/Massachusetts Ave/RT-2A, turn right, and then head towards Central Square. Go 2.9 mi. As you travel on Mass Ave/Massachusetts Ave/RT-2A you will go past Central Square, Harvard Square, and Porter Square.
2. Turn right at Russell St. Go 0.2 mi
3. Continue on Cutter Ave. Go 0.1 mi
4. Turn left at Highland Ave. Go 0.2 mi. As you travel on Highland Ave you will go into Davis Square.
5. Turn right at College Ave and head towards Tufts University. Go 0.4 mi
6. At Powder House Square, take the 3rd exit and stay on College Ave. Go 0.1 mi
7. Arrive at: 165 College Ave, Somerville, MA 02144.

The actual production of the route overview and modified turn-by-turn directions shown above are produced by a set of scripts. As previously described in this section, ROVER uses ArcGIS to situate the Route's destination; identify the Route's trunk segments and cognitive anchor points to be included in the route overview; and denote which turn-by-turn directions should be modified. This information is then passed as input to a set of scripts written in perl. These scripts serve as templates with blank slots, and these slots are filled with the input from the GIS. A different template is used based on a person's knowledge profile.

3.8 Summary

In this chapter, we focused on the design decisions that were made in the implementation of LAIR and ROVER. We emphasized the adherence to the cognitive design guidelines from Chapter 2 as we described the elements in the LAIR representation and how they were used by ROVER to generate route overviews. The LAIR elements mirror the elements people use in their image of the city, and they possess similar organizational properties: hierarchical regions, routes, and the use of cognitive an-

chor points and road skeletons. LAIR's global reference frame, a feature not found in other location representations, allows us to model a person's spatial knowledge and more fully describe the relationships between the places that person knows. Our choice of global reference frame, the Massachusetts streets layer, was used to define the skeleton of the city.

The use of a global orientation frame was a perspective that came about because we realized the problem addressed in this dissertation is different from the problems that motivated other location representations such as TOUR and PLAN. Another instance where perspective was important was in the way we applied the cognitive principles from Chapter 2 in the process we used to identify Regions for use in our route overviews. We take advantage of the fact that when generating route overviews, the language we use is inherently underspecified, and thus it is not necessary to find *the* perfect set of Regions with which to divide space. Moreover, the granularity of language makes the specific details of exactly where a Region's boundary is located irrelevant. We rely on the importance of the road skeleton as a way of searching for important places to serve as cognitive anchor points and then determine the extent of these anchor points' influence by searching for spatial outliers and commercial areas nearby.

LAIR and ROVER were implemented using a combination of Google Earth, ArcGIS, and a set of perl scripts. Google Earth routes are imported into ArcGIS as LAIR Routes using a utility written for that purpose. The use of ArcGIS was novel in that we used its analytic capabilities to help us generate qualitative route overviews. We used ArcGIS algorithms to determine which knowledge profile a newly introduced Route falls in and which trunk segments and regions to include in a route overview. These features are then passed as input to the perl scripts. The scripts produce the route overview, along with a slightly modified set of turn-by-turn directions that accentuate the major trunk segments and cognitive anchor points encountered on the route. Both the route overviews and the changes to the turn-by-turn directions are structured in such a way as to appeal to people's use of spatial hierarchy and their view of travel as being goal-directed. Having a model of people's spatial knowledge

allows ROVER to frame — according to which knowledge profile that Route falls under — the route overview as episodic travel along a set of trunk segments and from one cognitive anchor point to another.

We accessed ArcGIS's algorithms primarily through its interactive menu system. The code to automate this process and pass input to the perl scripts has not yet been implemented, but that code would not contribute to the main thrust of this research: that by appealing to a cognitively-inspired model for direction giving, we can make driving directions more comprehensible.

In the next chapter, we discuss a user study we carried out to measure what effect our route overviews and modified turn-by-turn directions had on people's subjective measures of their understanding of a route. The findings from that study support our claim of the effectiveness of our cognitively-inspired model for direction giving. People in the knowledge profile for which the overview was intended found ROVER's output more comprehensible. More interestingly, because of the general organizational schemas used to present the route overviews, our study found that people who did not fit the intended knowledge profile also found ROVER's output more comprehensible. Furthermore, in cases where the data does not appear to support our claim, a closer analysis shows that the results are actually consistent with the cognitively-inspired design principles on which this work is based. Thus, in these latter cases, the data provides even more insight into how a cognitively-inspired set of design guidelines can be used to produce comprehensible route overviews.

Chapter 4

Evaluating ROVER's Output

To evaluate the effectiveness of our cognitively-inspired, computational model for direction-giving, we conducted a user study in which we asked participants to compare the route overviews and turn-by-turn directions ROVER produces against a standard set of turn-by-turn directions generated by Google Earth. This chapter describes the design of our study, the results, and an analysis.

The goal of our experiment was to determine if ROVER's output is more comprehensible than standard turn-by-turn directions. We wanted to determine the effect of including hierarchy, goal-directed descriptions, selective suppression of detail, and the use of a route's skeleton and cognitive anchor points in a route overview. In Section 3.7.4 we described how a person's knowledge profile, the trunk segments of a route with which a person is familiar, affects the route overview ROVER produces. In conducting our analysis, we first examined the effect route overviews had on subjects who fit the predicted knowledge profile ROVER used to generate its output. We then expanded our analysis to those who did not fit the predicted knowledge profile to determine if route overviews made driving directions more comprehensible independent of the specific places with which a person is familiar. We did not examine how individuals score different route overviews generated for the same route using different knowledge profiles, but in our analysis, we discuss what could be learned from such a follow-up study.

In our experiments, we compared ROVER's output to traditional turn-by-turn

directions using a number of different metrics. Generally speaking, when we control for the spatial knowledge a person has, study participants rated ROVER’s output as being better than traditional turn-by-turn directions when they were asked how much each helped them understand where a destination was located and the route to get there. Participants also demonstrated a preference for ROVER’s output over traditional turn-by-turn directions. While neither the type of route nor the sub-population (e.g., men vs. women) that read a route overview was generally a factor on the effectiveness of a route overview in improving route comprehensibility, we will discuss the few instances where route or type of reader did make a difference.

There were also a few instances in the study’s results in which ROVER’s output was not rated more comprehensible than or not preferred over traditional turn-by-turn directions. We discuss these instances in our analysis, and examine what insight these results provide in understanding how we should structure route overviews. Another interesting result we describe in our analysis is how effective ROVER’s output is at helping people understand a set of route instructions when these people *do not* fit the predicted knowledge profile ROVER used to generate the route overview they read. These two results are particularly worth analyzing because while we did not initially expect them, we describe how these particular results are consistent with the cognitive principles on which this work is based.

4.1 Experimental Design

We recruited 30 subjects for our study, all affiliated with MIT in various ways: undergraduates, graduate students, and staff. The average age was 28, with a standard deviation of 7.1 years. The average time a participant had lived in the greater Boston area was 6.4 years, with a standard deviation of 4.1 years. Sixteen men and fourteen women participated. On average, participants drove 4.9 days a month, with a standard deviation of 6.5 days, and they rode the subway 12.1 days a month, with a standard deviation of 10 days. Seventeen participants indicated they were more

comfortable reading maps than written directions, two participants were more comfortable reading written directions, and eleven were equally comfortable with both.

Subjects were asked to evaluate three different routes. The places at the endpoints of each route were unique, and there was minimal overlap between the routes. The study was a within-subjects design; for each of the three routes, participants saw both ROVER's output (route overview plus modified turn-by-turn directions) and the turn-by-turn directions produced by Google Earth. The presentation order of the two types of directions was counter-balanced across subjects. The routes were shown in a set order, and participants were shown all the directions of one type before seeing the directions from the other. Since we were interested in comparing the comprehensibility of ROVER's output versus traditional turn-by-turn directions, participants were provided only the textual information mentioned above. We did not show participants maps of the routes to avoid any confounding effects from the maps influencing a person's assessment of their understanding of where the route went.

Recall from Section 3.7.4 that the term knowledge profile is used to describe which of the trunk segments on a route a person is familiar with, and that the knowledge profile has an effect on the route overview ROVER produces. Since our pool of study participants was comprised of MIT students and affiliates, the LAIR spatial knowledge base ROVER used was modeled after our assumptions of what places and paths this population would know. We selected streets around the MIT campus and around the subway stops that were closest to MIT. Given this spatial knowledge base, we chose the three routes so that the corresponding route overviews were generated using a different knowledge profile. Since each person saw the same three route overviews, some of the study's participants saw route overviews that did not match their own knowledge profile. After the study, we asked participants which of the streets in the directions they were familiar with, and thereby determined which knowledge profile they actually fit into.

The three routes we used in our user study are described below. While we describe the routes here by naming their start and end points, in the directions we showed study participants, we named only the start point and gave the street address of the

destination.

1. Route 1 goes from the MIT Medical Center to S&S Deli. This route is 1.9 miles long, and is the shortest of the three routes. It covered an area we assumed most study participants would be familiar with and was in the full-knowledge profile. ROVER identified four trunk segments on this route. See Table 4.1 for ROVER's output and the turn-by-turn directions produced by Google Earth. A map is also included for reference in Figure 4-1, but again, this is shown for the reader's benefit; it was not given to study participants.

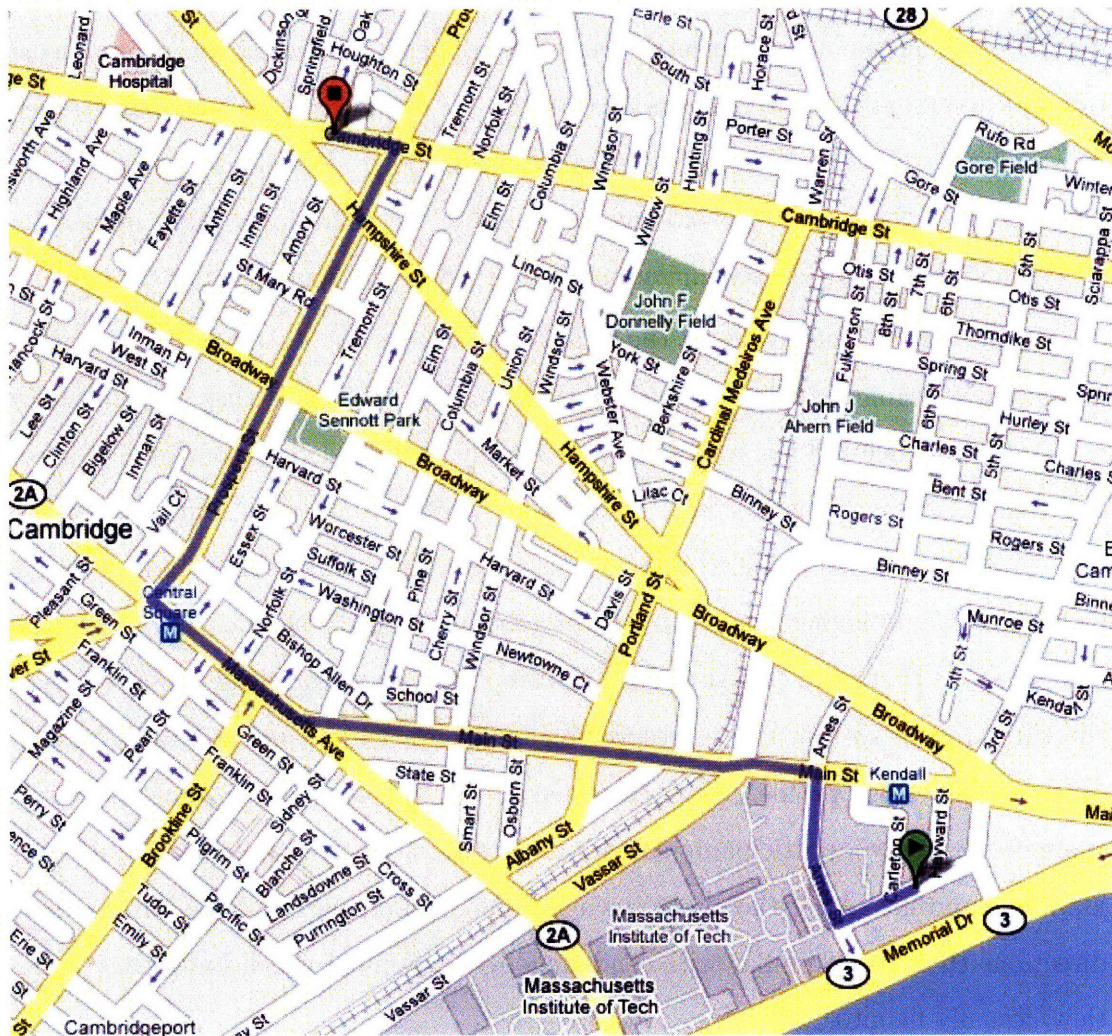


Figure 4-1: Map of Route 1.

2. Route 2 goes from Sidney Pacific Dorm to the childhood home of former pres-

ROVER's Output

From MIT Medical to 1334 Cambridge St,
Cambridge MA

Your destination is in Inman Square.

On this route, you will be traveling along 4 major roads. You will be going from Main St to Mass Ave/Massachusetts Ave/RT-2A to Prospect St to Cambridge St. Along this route, you'll pass Central Square.

Go to Main St, turn left, and then head towards Mass Ave/Massachusetts Ave/RT-2A. Go 0.6 mi

Slight right at Mass Ave/Massachusetts Ave/RT-2A and head towards Central Square. Go 0.2 mi

Turn right at Prospect St and head towards Inman Square. Go 0.6 mi

Turn left at Cambridge St. Go 427 ft

Arrive at: 1334 Cambridge St, Cambridge, MA 02139.

Google Earth's Output

From MIT Medical to 1334 Cambridge St,
Cambridge MA

Head southwest on Amherst St toward Carleton St. Go 0.1 mi

Turn right at Ames St. Go 0.2 mi

Turn left at Main St. Go 0.6 mi

Slight right at Mass Ave/Massachusetts Ave/RT-2A. Go 0.2 mi

Turn right at Prospect St. Go 0.6 mi

Turn left at Cambridge St. Go 427 ft

Arrive at: 1334 Cambridge St, Cambridge, MA 02139

Table 4.1: ROVER's and Google Earth's directions for Route 1.

ident John F. Kennedy, located in a residential neighborhood in the city of Brookline. The route is 2.4 miles long, and was in the proper-prefix profile: it starts near the MIT campus but then crosses the Charles River into Brookline, an area of Greater Boston we believed most MIT affiliates were not familiar with. ROVER identified 5 trunks, the first 3 of which were known. See Table 4.2 for ROVER's output and the Google Earth directions. Figure 4-2 shows a map of Route 2.



Figure 4-2: Map of Route 2.

3. Route 3 is the longest of the three routes, and goes from the MIT Stata Center, the home of MIT's computer science department, to an apartment 4.2 miles north of the Stata Center. This route falls in the prefix-suffix knowledge profile and is particularly interesting because it consists of two trunk segments which, according to the LAIR model of what our MIT participant would know, are

ROVER's Output

From Sidney-Pacific Dorm to 83 Beals St,
Brookline MA

Your destination is located on a local road,
and the closest major road to it is Harvard St.

On this route, you will be traveling along 5
major roads. The first 3 major roads you will
be traveling on are Sidney St to Boston Univ
Bridge/Rt-2 to Commonwealth Ave/RT-
30/US-20. After that, the other major roads
you will be traveling on are Babcock St to
Harvard St.

Go to Sidney St, turn left, and then head
towards Boston Univ Bridge/Rt-2. Go
0.5 mi
Turn right at Waverly St. Go 0.1 mi
Turn left at Brookline St. Go 141 ft
Turn right to stay on Brookline St. Go 157 ft
At the traffic circle, take the 2nd exit onto
Boston Univ Bridge/RT-2 and head
towards
Commonwealth Ave/RT-30/US-20. Go 0.3
mi
Turn right at Commonwealth Ave/RT-
30/US-20. Go 0.5 mi
Turn left at Babcock St. Go 0.6 mi
Turn right at Harvard St. Go 0.2 mi
Turn right at Beals St. Go 0.1 mi
Arrive at: 83 Beals St, Brookline, MA 02446.

Google Earth's Output

From Sidney-Pacific Dorm to 83 Beals St,
Brookline MA

Head northwest on Pacific St toward
Landsdowne St/Landsdowne Dr.
Go 351 ft
Turn left at Sidney St. Go 0.5 mi
Turn right at Waverly St. Go 0.1 mi
Turn left at Brookline St. Go 141 ft
Turn right to stay on Brookline St. Go 157 ft
At the traffic circle, take the 2nd exit onto
Boston Univ Bridge/RT-2. Go 0.3 mi
Turn right at Commonwealth Ave/RT-
30/US-20. Go 0.5 mi
Turn left at Babcock St. Go 0.6 mi
Turn right at Harvard St. Go 0.2 mi
Turn right at Beals St. Go 0.1 mi
Arrive at: 83 Beals St, Brookline, MA 02446

Table 4.2: ROVER's and Google Earth's directions for Route 2.

not connected to one another. The route passes through many neighborhoods, but in constructing the LAIR model, we assumed that some of these neighborhoods were encountered via the underground subway, and not via above ground streets. As a result, there are disconnected islands which the streets on this route connect. We allude to this disconnection and the subsequent stitching together of these islands in the route overview presented in Table 4.3. Figure 4.3 shows a map of Route 3.

Participants were shown one route at a time, and after each route, they were asked to evaluate the directions according to the following comprehensibility metrics:

- *Understanding of the route.* Using a 7-point Likert scale, participants were asked to rate their agreement with Statements about their understanding of where the destination is located and the route to get there. Participants were asked to score the following statements using a scale in which a 1 was considered “Strongly Disagree” and a 7 was considered “Strongly Agree”. The Statements are presented below and are prefaced with a short summary phrase. These phrases will be used later in this chapter as a means to refer to the full Statements.
 - Statement 1 (S1, where dest located): The directions helped me understand where the destination is located.
 - Statement 2 (S2, understanding of rt): The directions gave me a good overall understanding of the route from the starting point to the destination.
- *Information content of the directions.* Using the same 7-point Likert scale, participants were asked to rate their agreement to the following Statements regarding the content of the directions they were just presented. Note that for Statements 1-4, higher scores are better whereas in Statements 5 and 6, lower scores are better. Also, note that since Statements 4, 5, and 6 have a certain amount of overlap, we expected these scores to be correlated.

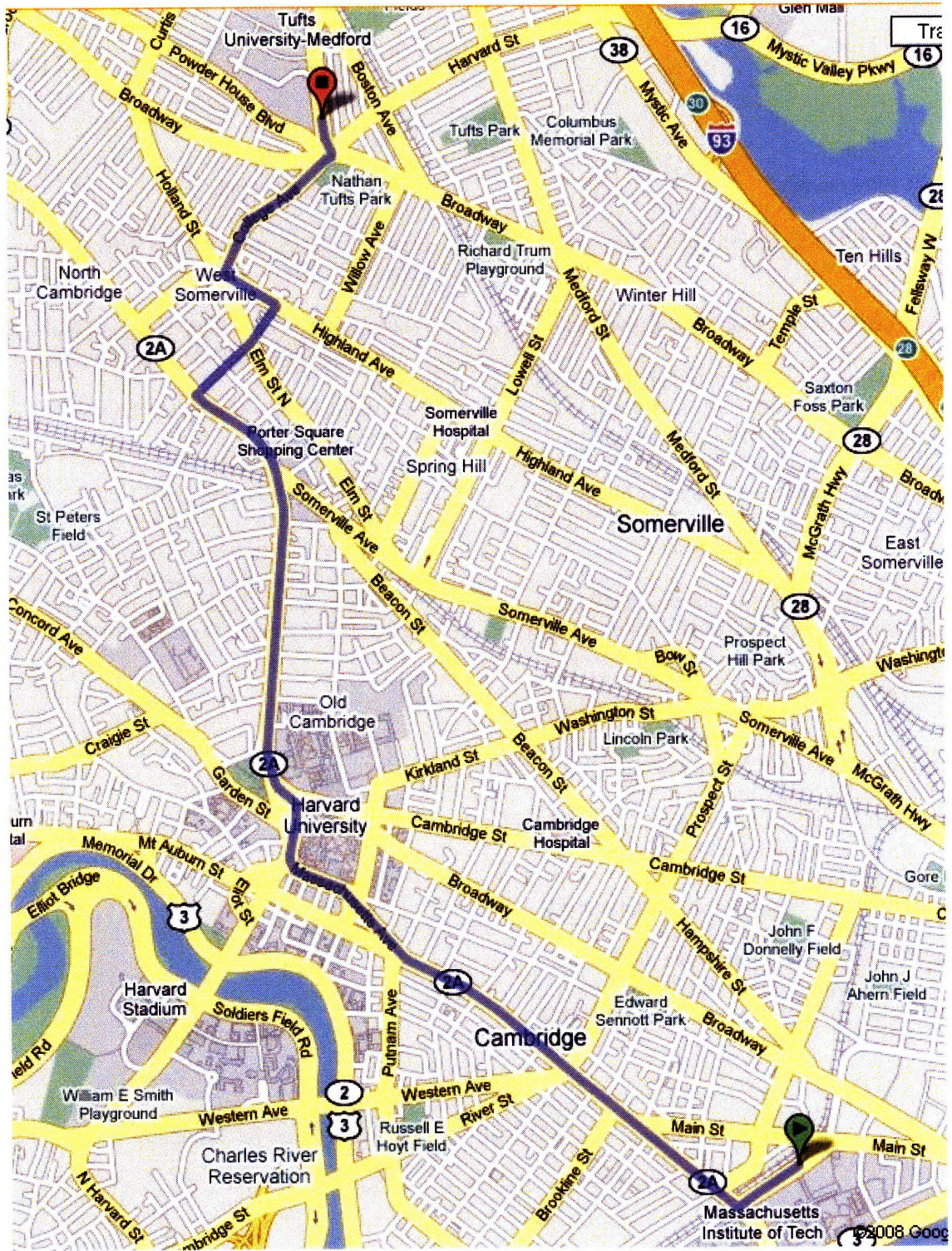


Figure 4-3: Map of Route 3.

ROVER's Output

From the Stata Center to 165 College Ave,
Somerville MA

Your destination is near Tufts University.

On this route, you will be traveling on Mass Ave/Massachusetts Ave/RT-2A to get to Porter Square, then going from there to College Ave by Davis Square. The route then goes from College Ave to your destination. Along this route, you'll pass Harvard Square, Porter Square, and Davis Square.

Go to Mass Ave/Massachusetts Ave/RT-2A, turn right, and then head towards Central Square. Go 2.9 mi. As you travel on Mass Ave/Massachusetts Ave/RT-2A you will go past Central Square, Harvard Square, and Porter Square.

Turn right at Russell St. Go 0.2 mi
Continue on Cutter Ave. Go 0.1 mi
Turn left at Highland Ave. Go 0.2 mi. As you travel on Highland Ave you will go into Davis Square.

Turn right at College Ave and head towards Tufts University. Go 0.4 mi

At Powder House Square, take the 3rd exit and stay on College Ave. Go 0.1 mi
Arrive at: 165 College Ave, Somerville, MA 02144.

Google Earth's Output

From the Stata Center to 165 College Ave,
Somerville MA

Head southwest on Vassar St toward Mass Ave/Massachusetts Ave/RT-2A. Go 0.2 mi

Turn right at Mass Ave/Massachusetts Ave/RT-2A. Go 2.9 mi

Turn right at Russell St. Go 0.2 mi

Continue on Cutter Ave. Go 0.1 mi

Turn left at Highland Ave. Go 0.2 mi

Turn right at College Ave. Go 0.4 mi

At Powder House Square, take the 3rd exit and stay on College Ave. Go 0.1 mi

Arrive at: 165 College Ave, Somerville, MA 02144

Table 4.3: ROVER's and Google Earth's directions for Route 3.

- Statement 3 (S3, useful info): The directions provided useful information to help me understand where the route goes.
 - Statement 4 (S4, right amt of info): The directions provided the right amount of information to help me understand the route.
 - Statement 5 (S5, info not needed): The directions contained information I did not need to understand the route.
 - Statement 6 (S6, not enough info): The directions did not have enough information to help me understand the route.
- *Route recall.* After a participant had scored all the statements for one route and had moved on to scoring the statements for the next route, they were asked to write down as much of the previous set of directions as they could remember. We adopted this staggered approach to see how much of the route was actually retained, and not just held in short-term memory. Participants were asked to repeat the directions for a given route once, after the first time they were presented with the directions (either from ROVER’s output or Google Earth’s). They were not asked to repeat the directions a second time after seeing the other type of directions.
 - *Direction preference.* At the end of the study, for each route, we showed participants the directions produced by Google Earth and the route overview and directions produced by ROVER. We then asked them for each route, which, if any, of the two types of directions they preferred.

4.2 Results and Analysis

We present the results of our study in subsections, according to the different questions we were interested in. In the results we present below, unless otherwise stated, all the results were based on data from people who fit the knowledge profile ROVER used to generate the route overview. When we analyze data across all three routes, we use data from those subjects who fit all three of the knowledge profiles used to produce

the route overviews. Of the different routes, 23 people fit the full-knowledge profile used to generate the route overview for Route 1, 23 people fit the proper-prefix profile used for Route 2, and 14 people fit the prefix-suffix profile used in Route 3. Twelve people fit all three of the knowledge profiles used to produce the route overviews. We used the data we collected to answer three types of questions:

1. Controlling for knowledge profile, is ROVER's output more comprehensible than Google Earth's?
2. Controlling for knowledge profile, does the effect of ROVER's output on route comprehensibility differ across individual routes, or across the different sub-populations (e.g., men vs women) who read the overviews?
3. Are route overviews created for one knowledge profile effective in helping people in other profiles comprehend a route?

4.2.1 ROVER's Output Compared to Google Earth's

In the majority of cases, ROVER's output was rated better than Google Earth's on the six Statements listed in Section 4.1.

We ran paired t-tests to compare the mean rating for ROVER's output against Google Earth's for each question and each route. Table 4.4 shows the difference between the mean scores of ROVER and Google Earth for each (Route, Statement) pair. Differences that favor ROVER in a statistically significant manner are highlighted.

For Statements 1 and 2, which deal with subjects' assessment of their understanding of where a destination is located and the route to get there, ROVER's mean score was better than Google Earth's in a statistically significant way in five of the six possible cases. There was a trend towards ROVER scoring better in the sixth case. With regards to Statements 3–6, the ones concerning the information content of the directions presented to study participants, mean scores indicate participants considered ROVER's output better at providing more useful information (Statement 3) and at providing the right amount of information (Statement 4) than Google Earth's turn-by-turn directions. This difference was statistically significant in all three routes.

	S1 (where dest located)	S2 (understanding of rt)	S3 (useful info)	S4 (right amt of info)	S5 (info not needed)	S6 (not enough info)
Route 1 (n=23)	1.26	1.00	1.17	1.00	1.74	-1.57
Route 2 (n=23)	1.00	0.52	1.00	1.04	0.74	-1.35
Route 3 (n=14)	1.07	1.07	1.43	1.21	0.64	-1.29

Differences that favor ROVER in a statistically significant manner are shaded as indicated below.

$p < 0.05$ $p < 0.01$

Table 4.4: Differences in mean Likert scores between ROVER and Google Earth, for each route and each statement. The differences are based on data from study participants who fit the predicted knowledge profile used to generate the route overviews. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER.

For Statement 5, participants on average considered ROVER’s output as having more unnecessary information than Google Earth’s turn-by-turn directions. This difference in scores for Statement 5 was statistically significant in Route 1 ($t(22) = 3.83, p = 0.0009$) but not Routes 2 or 3 ($t < 1.85, p > 0.08$). This result is not completely unexpected given that not every feature provided in a route overview may be used by a person to understand a route. In contrast to asking if the directions contained unnecessary information, Statement 6 was used to assess the complementary issue of whether or not the directions provided sufficient information to understand the route. In this case, ROVER scored better than Google Earth and this difference was statistically significant in all three routes.

The results presented in Table 4.4 show that ROVER received better scores for Statements 1-4 and 6 on each individual (Route, Statement) pairing. To verify that ROVER’s performance was generally better than Google Earth’s for those Statements across all routes, we conducted a two-way ANOVA for each Statement. In the two-way ANOVA for a particular Statement, the two independent variables were the route (Route 1, 2, or 3) and the direction type (ROVER or Google Earth). The dependent variable was the score the participant gave on the 7-point Likert scale for that Statement. In conducting the ANOVAs, we restricted the analysis to those subjects who fit the appropriate knowledge profile for all three routes. The results, shown in Table 4.5, indicate that the type of directions had a significant main effect

on the scores for Statements 1-4 and Statement 6 ($F(1, 11) = 5.48, p = 0.04$ for Statement 2, $F(1, 11) > 10.78, p < 0.007$ for Statements 1, 3, and 4). These results corroborate the earlier findings which suggested that across all routes, ROVER's output scored better than Google Earth's turn-by-turn directions on these Statements. The p-value from the ANOVA for Statement 5 was not significant. There was only one significant interaction between direction type and route, and that was for Statement 3 ($F(2, 22) = 4.41, p = 0.025$).

	Difference In Mean Likert Scores Across All Routes
S1 (where dest located)	1.11
S2 (understanding of rt)	0.92
S3 (useful info)	1.14
S4 (right amt of info)	1.42
S5 (info not needed)	0.83
S6 (not enough info)	-1.72

Differences that favor ROVER in a statistically significant manner are shaded as indicated below.

$p < 0.05$	$p < 0.01$
------------	------------

Table 4.5: The statistical significance of the difference in mean Likert scores for each Statement is determined by examining the effect of direction type (ROVER vs Google Earth) on Likert scores across all routes using ANOVA. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER. The ANOVA was based on data from subjects who fit the predicted knowledge profile for all three routes ($n=12$).

When directly asked which of the two types of directions participants preferred, for each of the three routes, a majority of participants chose ROVER's output to Google Earth's (see Table 4.6). To determine if this result was significant, we conducted a sign test to determine if we could reject the null hypothesis that participants preferred either type of directions with equal probability. Of those that had a preference, the sign test indicated that the majority was statistically significant in Routes 2 and 3, but not for Route 1.

Given these results, we see that overall, ROVER scored better than Google Earth in most of the comprehensibility metrics. We discuss two noteworthy instances where this was not the case. Table 4.4 shows that Google Earth scored better than ROVER

	ROVER	Google Earth
Route 1	14	8
Route 2	17	3
Route 3	12	2

Differences that favor ROVER in a statistically significant manner are shaded as indicated below.

p < 0.05 **p < 0.01**

Table 4.6: This table shows which type of directions were preferred among participants who fit the predicted knowledge profile used to generate a route overview and who had a preference between the two types of directions they saw. A majority of participants preferred ROVER’s output to Google Earth’s. We performed a sign test to determine if this majority was statistically significant.

for Route 1, Statement 5. Statement 5 was “The directions contained information I did not need to understand the route” and participants considered ROVER’s output as having more unnecessary information than the turn-by-turn directions. Also, when asked to compare the two directions side-by-side, although a majority of people preferred ROVER’s output to Google Earth’s for Route 1, this difference, unlike the other two routes, was not statistically significant.

Examining Route 1 and the overview ROVER generated for it in terms of the cognitive design guidelines listed in Chapter 2 provides some insight into how the above results may have come about. Route 1 was the shortest of the three routes we used, and it traversed an area that was familiar to most of the study’s participants. While ROVER identified four trunk segments in the route, one of those segments, Massachusetts Avenue (Mass Ave), is one of the most prominent roads in the road skeleton of the city of Cambridge. The route overview did not make use of this property; it just listed Mass Ave as one of the four major roads on the route.

Moreover, once ROVER modified the turn-by-turn directions to not include the small streets prior to the first trunk segment, the only streets left in the turn-by-turn directions were major trunk segments. Thus, the route overview did not have the property of highlighting a proper subset of streets from the turn-by-turn directions as major trunk segments. Instead, the route overview may have been viewed as an unnecessary restatement of the turn-by-turn directions. This redundancy could

explain the scores Route 1 received for Statement 5 and why ROVER's output for Route 1 was not preferred by more people.

This situation suggests two possible adjustments, both consistent with our cognitive design guidelines, that could be made to improve ROVER's output. One approach would be to use the road skeleton to selectively suppress certain trunk segments. ROVER could use its LAIR model of streets to determine which, if any, trunk segments are more prominent than others. If there are segments that are more prominent than others, then ROVER can emphasize these roads by suppressing any mention of trunk segments prior to the more prominent segments in the route overview. Additionally, instead of listing cognitive anchor points at the end of the overview, we could integrate the mention of cognitive anchor points with our description of the trunks in a route, similar to how they are included now in ROVER's modified turn-by-turn directions. With these adjustments, a route overview for Route 1, under a full-knowledge profile would then read like this:

Your destination is in Inman Square.

On this route, you will be travelling along 3 major roads. You will be going on Mass Ave into Central Square, to Prospect Street towards Inman Square and then to Cambridge St.

A second possible approach would be to still suppress the mention of any trunk segments prior to the prominent ones in the road skeleton in the turn-by-turn directions and then either forego producing a route overview altogether or just mention the neighborhood the destination is located in. Using this approach, ROVER's output for Route 1 would then be:

Your destination is in Inman Square.

Go to Mass Ave/Massachusetts Ave/RT-2A and head towards Central Square.

Go 0.2 mi

Turn right at Prospect St and head towards Inman Square. Go 0.6 mi

Turn left at Cambridge St. Go 427 ft

Arrive at: 1334 Cambridge St, Cambridge, MA 02139.

Both of the above approaches would reduce the amount of unnecessary information presented by abstracting the beginning of the route as travel to a prominent road in the road skeleton.

We close this subsection by reporting one final result. Although ROVER's output was not intended to make turn-by-turn directions more memorable, we also ran a series of independent-samples t-tests to determine if participants who had first seen ROVER's output remembered more of the directions than participants who had first seen Google Earth's turn-by-turn directions. These t-tests revealed that on average, participants who had first seen ROVER's output remembered more of the streets and turns on a route than those who had first seen the Google Earth directions, but this difference was not statistically significant for any of the three routes.

4.2.2 ROVER's Performance Across Routes and Across Different Subpopulations

The previous subsection described how in general, ROVER's output scored better than Google Earth's turn-by-turn directions on most of the six Statements we used to assess route comprehensibility. In this subsection, we examine the *differences* in the scores individuals gave ROVER and Google Earth to determine if the size of ROVER's effect varies. We refer to these differences as deltas. In most cases, there was no evidence to suggest that these deltas (the size of ROVER's effect) were significantly different across routes or across different sub-populations, so we focus the discussion on those situations in which the deltas did differ.

For each of the Statements, we ran a one-way repeated measures ANOVA on the deltas for that particular Statement to determine if they differed across routes. These ANOVAs, as with the others described in this subsection, were performed on the data from subjects who fit the predicted knowledge profile for all three routes. With the exception of Statement 3 (S3, useful info), none of the ANOVAs suggested that the deltas differed across routes in a statistically significant way (F 's < 2.60 , $p > 0.1$).

However, in Statement 3, the mean delta for Route 2 was 0.583, for Route 1 it was

1.167, and for Route 3, it was 1.667. In this case, the route had a significant main effect on the size of the deltas: $F(2, 22) = 4.41, p = 0.0246$. A Tukey test indicates that the delta for Route 2 is different from Route 3 in a statistically significant manner. Statement 3 asked participants to rate their agreement with the following statement: “The directions provided useful information to help me understand where the route goes.” The difference in the deltas for Statement 3 can be attributed to the fact that, because of a lack of cognitive anchor points along Route 2, the route overview for Route 2 identified only that route’s major trunk segments. In the other two routes, the route overview identified both the neighborhood the destination was located in and included a number of cognitive anchor points the route passed through. Since Route 3 was the longest of the routes and passed through more cognitive anchor points, the inclusion of these additional points may have increased the size of the deltas for Route 3 on Statement 3, compared to the other routes.

The fact that Route 2 mentions the major trunks in the route but no cognitive anchor points demonstrates the benefit of explicitly identifying major trunk segments in the overview. The paired t-tests comparing the Route 2 scores for ROVER’s output against Google Earth for each of the Statements show that ROVER scored better in five of the six Statements in a statistically significant way (see Table 4.4). Identifying trunk segments helps a person comprehend where a route goes, especially in the absence of cognitive anchor points.

Including cognitive anchor points in route overviews is still an important thing to do, however. While there was no statistically significant difference between the deltas for the other Statements, Route 2 had the smallest deltas in three of the six Statements in Table 4.4, thus suggesting that the overall benefit of ROVER’s output for Route 2 was not as large as it was in Routes 1 and 3. Had there been a cognitive anchor point to include in the route overview for Route 2, these deltas may have been closer to the deltas for Route 1 and Route 3.

To determine if there was a difference in the deltas between different sub-populations for a given route, we ran a series of independent t-tests for each combination of Route, Statement, and sub-population grouping. In each t-test, we used data from people

who fit the knowledge profile used to generate the overview for that route. The different sub-population groupings we considered were:

- men versus women
- people who drove five or more days a month versus those who did not
- people who rode the subway thirteen or more days a month versus those who did not
- people who have lived in Boston for more than five years versus those who lived in Boston for five or fewer years
- people who considered themselves more comfortable reading a map versus those who were either more comfortable reading written directions or who were equally comfortable reading maps and written directions.

The t-tests found statistically significant differences in five cases. These are summarized in Table 4.7. However, these differences in the deltas were found only in those particular combinations of Route, Statement, and sub-population groupings. To determine if any of the differences listed in Table 4.7 extended across all three routes we ran a two-way repeated measures ANOVA with route number as the within-subjects factor and sub-population as the between-subjects factor. We did not find any statistically significant results (F 's < 1.72 , $p > 0.2185$).

However, when frequency of subway riding was the between-subjects factor in the two-way ANOVA for Statement 6, the test revealed a significant main effect for frequency, $F(1, 10) = 4.98$, $p = 0.0498$ (with no significant interaction effect between frequency of subway riding and route number). For Statement 6, participants rated their agreement with the statement, "The directions did not have enough information to help me understand the route." In this case, more negative deltas indicate a better relative performance by ROVER compared to Google Earth. The mean delta for Statement 6 for subjects who rode the subway twelve days/month or fewer was -0.95 compared to -2.8 for people who rode the subway thirteen or more days a month.

So, looking across all the routes, ROVER scored better on this metric with frequent subway riders than with those who did not ride the subway as often. We think this is because frequent subway riders are more likely to structure their mental map in terms of subway stops, and even though these stops are not directly used to navigate driving directions, their absence in the directions makes it more difficult for frequent subway riders to understand the driving directions.

Route	Statement	Sub-Population	Deltas
1	S4 (right amt)	Gender	Men = 1.85 Women = -0.1
1	S6 (not enough info)	Gender	Men = -2.54 Women = -0.3
1	S2 (understanding of rt)	Time in Boston	Five or fewer years = -0.08 More than five years = 2.18
2	S4 (right amt)	Frequency riding the subway	Thirteen or more days/month = 1.7 Twelve or fewer days/month = 0.54
3	S5 (info not needed)	Frequency riding the subway	Thirteen or more days/month = 2.5 Twelve or fewer days/month = -0.75

Table 4.7: Instances where deltas (the size of ROVER’s effect) differed between different sub-populations in a statistically significant manner ($p < 0.05$). Statistical significance was determined by performing independent samples t-tests on data from subjects who fit the predicted knowledge profile used to generate the route overview.

For each (Route, Statement) pair, we also ran a regression between the amount of time a person had lived in Boston and their delta score for that pair. We did not find a strong correlation to suggest that time in Boston had an effect on the degree to which participants scored ROVER differently than Google Earth (t ’s < 1.96 , $p > 0.06$).

4.2.3 ROVER’s Performance for Subjects in Other Knowledge Profiles

We now examine how people who did not fall into the predicted knowledge profile scored ROVER’s output. Our tests indicate that for these people, even though ROVER’s output was not intended for their knowledge profile, they nonetheless scored ROVER’s output as better than Google Earth’s.

To determine if there was a difference in how these people scored ROVER’s output against Google Earth’s turn-by-turn directions, we also ran paired t-tests to compare

the mean rating for ROVER’s output against Google Earth’s for each possible (Route, Statement) pair. Table 4.8 shows the difference between the mean scores of ROVER and Google Earth for each (Route, Statement) pair. The t-tests show that in general, people who did not fall into the predicted knowledge profile for a given route rated ROVER and Google Earth in a similar manner as people who did fit the knowledge profile: namely, ROVER scored better than Google Earth for Statements 1-4 and 6, and worse in Statement 5.

	S1 (where dest located)	S2 (understanding of rt)	S3 (useful info)	S4 (right amt of info)	S5 (info not needed)	S6 (not enough info)
Route 1 (n=7)	2.14	1.29	1.86	1.86	1.86	-2.57
Route 2 (n=7)	2.14	1.00	1.71	1.43	2.86	-1.71
Route 3 (n=16)	2.69	1.88	2.25	1.56	1.50	-1.29

Differences that favor ROVER in a statistically significant manner are shaded as indicated below.

$p < 0.05$ $p < 0.01$

Table 4.8: Differences in mean Likert scores between ROVER and Google Earth, for each route and each statement. The differences are based on data from study participants who did **not** fit the predicted knowledge profile used to generate the route overviews. Positive differences in Statements 1-4 and negative differences in Statements 5 and 6 favor ROVER.

Looking back at our cognitive design guidelines, these results are not unexpected since the guidelines were drawn from general principles of human spatial cognition and were independent of which particular knowledge profile a person was in. The results from looking at how non-knowledge profile subjects score ROVER’s output demonstrates that there is additional descriptive power in using what we know about how people think about space to guide how we structure directions. It is this structuring of the directions, independent on how much of that space a person is actually familiar with, that makes ROVER’s output more comprehensible.

The overview for Route 3 (reproduced in Table 4.9) is a particularly good example of this; we elaborate on this instance since the population of people who did not fit the predicted knowledge profile used to generate this route’s overview was fairly large. The overview for Route 3 contained many of the canonical elements we mentioned in our design guidelines: goal-directed descriptions, selective suppression of detail, and

ROVER's Output

From the Stata Center to 165 College Ave,
Somerville MA

Your destination is near Tufts University.

On this route, you will be traveling on Mass Ave/Massachusetts Ave/RT-2A to get to Porter Square, then going from there to College Ave by Davis Square. The route then goes from College Ave to your destination. Along this route, you'll pass Harvard Square, Porter Square, and Davis Square.

Go to Mass Ave/Massachusetts Ave/RT-2A, turn right, and then head towards Central Square. Go 2.9 mi. As you travel on Mass Ave/Massachusetts Ave/RT-2A you will go past Central Square, Harvard Square, and Porter Square.

Turn right at Russell St. Go 0.2 mi
Continue on Cutter Ave. Go 0.1 mi
Turn left at Highland Ave. Go 0.2 mi. As you travel on Highland Ave you will go into Davis Square.

Turn right at College Ave and head towards Tufts University. Go 0.4 mi
At Powder House Square, take the 3rd exit and stay on College Ave. Go 0.1 mi
Arrive at: 165 College Ave, Somerville, MA 02144.

Google Earth's Output

From the Stata Center to 165 College Ave,
Somerville MA

Head southwest on Vassar St toward Mass Ave/Massachusetts Ave/RT-2A. Go 0.2 mi
Turn right at Mass Ave/Massachusetts Ave/RT-2A. Go 2.9 mi
Turn right at Russell St. Go 0.2 mi
Continue on Cutter Ave. Go 0.1 mi
Turn left at Highland Ave. Go 0.2 mi
Turn right at College Ave. Go 0.4 mi
At Powder House Square, take the 3rd exit and stay on College Ave. Go 0.1 mi
Arrive at: 165 College Ave, Somerville, MA 02144

Table 4.9: The structure and chunking of information provided in ROVER's overview for Route 3 makes it easier for people to parse the directions that are presented to them. We believe this structuring is a reason why the increase in route comprehensibility is independent of people's familiarity with the specific places mentioned in the overview.

the use of the skeleton and cognitive anchor points. These provided a rich overview with details to help a person comprehend where the turn-by-turn directions take him. Regardless of the actual knowledge profile he is in, the overview helps to prime a person with what to expect when he travels the route; it establishes expectations of prominent things the person would make note of as he builds the episodic impression of the route. By identifying which episodic milestones to expect, a structure is imposed on the driving directions, thus improving route comprehensibility.

4.3 General Discussion

The following is a summary of the important results we described in this chapter:

1. When controlling for spatial knowledge, people who fit the knowledge profile ROVER used to generate route overviews gave ROVER better scores than Google Earth in various metrics used to assess their understanding of a route and the information content of a set of directions.
2. When asked to score the directions based on the usefulness of the information included in those directions, the amount by which ROVER's output outscored Google Earth's (the size of ROVER's effect) differed across the routes. This was the only comprehensibility metric that was affected by the choice of Route.
3. People who did not fit the knowledge profile for which ROVER's output was designed also rated the ROVER output better than the Google Earth directions.
4. Compared to participants who did not ride the subway as often, participants who frequently ride the subway scored ROVER's output much more favorably than Google Earth's when they were asked if the directions did not have enough information to help them understand a route (Statement 6). While all participants scored Google Earth's directions as providing less information than ROVER's output, the difference was much greater for those who frequently rode the subway compared to less frequent riders. We found no other difference

between sub-populations in the scores of any of the other Statements used to assess route comprehensibility.

Our results demonstrate the benefit of incorporating the general concepts of human spatial cognition into a route overview. A follow-up question to ask would be to compare how people would score a route overview tailored to their knowledge profile against overviews for the same route but produced for different knowledge profiles. A follow-up study in which we show people multiple route overviews of the same route would allow us to determine if tailoring the route overview to a person would improve comprehensibility further.

The results from such a study may indicate one of three things: scores may be better for those overviews that are tailored to the individual's knowledge profile; or there is no marked difference between route overviews for different knowledge profiles; or route overviews produced for a particular knowledge profile receive better scores, regardless of what knowledge profile a person actually is in. Since ROVER includes only cognitive anchor points it considers to be known by the person, our hypothesis is that the knowledge profile used to produce a route overview would have an effect on the score it receives, at least for routes that a person is somewhat familiar with. We think this is why the deltas for frequent subway riders in Statement 6 were greater than the deltas for those who rode the subway less often. Frequent subway riders are more likely to structure their mental map in terms of subway stops, and even though these stops are not directly used to navigate through driving directions, their absence in the directions makes it more difficult to understand the route.

In our study, we have demonstrated the benefit of incorporating principles from human spatial cognition — concepts such as hierarchy, road skeletons, and cognitive anchor points' — in making driving directions more comprehensible to people. Based on these concepts, ROVER's output outscored a traditional set of turn-by-turn directions along various metrics used to assess a person's understanding of a route, independent of the particular knowledge profile a person was actually in. Moreover, when ROVER did not receive a better score, as it did when one route overview was considered to have too much unnecessary information, this provided another example

of how the principles of human spatial cognition could be used to improve ROVER's output. Determining the effect of knowledge profile on ROVER's effectiveness is an area of future work.

Chapter 5

Future Work

This chapter discusses four different areas in which we can expand the study of human spatial cognition and how the concepts from this domain can be used to improve the comprehensibility of driving directions. The four areas are: content analysis of human-generated route overviews, alternate methods of evaluating route overviews, other approaches to presenting route overviews, and other methods of building up the LAIR spatial knowledge base.

5.1 Route Overviews Revisited and Re-examined

One interesting line of future research would be to investigate how people produce an overview of a route that is presented to them. We could conduct a study where we provide a person with a map and a route marked on it, but not provide the turn-by-turn directions for the route. The person would then be asked to produce a route overview from the map. After building up a corpus of different route overviews, we could then see if there are patterns in the route overviews people produce and then reproduce these patterns in ROVER's output. Providing this ability to describe in words what's shown in a figure would be particularly useful for current GPS systems, which have difficulty displaying maps due to limited screen sizes.

A study structured in this way, focusing on the task of generating route overviews, would be novel from a cognitive psychology standpoint. Our proposed experiment

would investigate how people interpret a specific type of spatial information presented to them. This is in contrast to the constructionist experiments seen in much of the cognitive psychology research. In much of that literature, the studies dealing with direction giving and spatial cognition involve people either planning a route and then giving detailed turn-by-turn directions for that route [18, 22] or describing spatial relationships between places previously committed to memory [86]. While there have been some experiments that have looked at how people integrate new spatial information as they travel [25, 26] and others that have looked at how people scan maps (without routes on them) [48], none of these studies have looked at the particular problem of how people interpret where the route is telling them to go.

Our cognitively-inspired design guidelines led us to focus on identifying trunk segments and cognitive anchor points in the route overviews we produced. This proposed study would allow us to see if there are other features or relationships people make note of when they make sense of a route that is presented to them. The following is a list of some of the issues we could investigate using this corpus:

- *Spatial distortions.* Are spatial distortions like those described in Chapter 2 (e.g., alignment with canonical axes or shifting places so that they are more aligned with one another) found in the route overviews people give, even though these overviews are based on a map which do not exhibit these distortions?
- *Within-person, between-route variation.* Are there significant between-route variations in the content and structure of the route overviews a particular person produces? If there is significant between-route variation in how route overviews are structured, can the route overviews be easily categorized in a small number of classes? Do the features of a route influence how the overview for that route is structured?
- *Between-person, within-route variation.* Is there significant variation in terms of content and structure in the route overviews different people give for the same route? When giving route overviews, are there patterns similar to the phenomenon of women preferring turn-by-turn directions presented from a landmark-

based, ego-centric perspective versus men who prefer a bird's-eye, exo-centric perspective [16, 56]?

5.2 Other Evaluations of Route Overviews

The issue of within-route variations in the route overviews brings up the question of how people's familiarity with an area influences route comprehensibility. In this dissertation, we have not specifically looked at this issue. Instead, we focused on how we can increase route comprehensibility by incorporating in a route overview the cognitive structures people employ to make sense of space. We then demonstrated that the benefit of using these structures is independent of the specific places and streets with which a person is familiar. However, as described in Chapter 4, another area of future work would be to examine how people's familiarity with different areas affects how they score route overviews intended for different knowledge profiles. In which situations does tailoring a route overview to an individual's knowledge profile make a route significantly more comprehensible? And if the corpus analysis reveals that different people structure their route overview for a particular route differently, are there situations where one particular way of structuring a route overview is unanimously scored higher?

Another type of evaluation of the effectiveness of route overviews would be an in-the-field evaluation. In this proposed study, people would read a set of directions with a route overview and then traverse the route. A control group would only be given the turn-by-turn directions. This study could then measure if people who are given a route overview refer to detailed turn-by-turn directions less or make fewer wrong or missed turns than those who are not given an overview. Other variations of this study could include examining the strategies people in the two groups use to find their way if they get lost or if they have to plan a detour to the route.

5.3 Extending Route Overviews

A third major line of future research would be to investigate other approaches to presenting route overviews, both in terms of the information presented in the overviews and how people interact with the route overview.

An alternative to emphasizing the trunk segments in a route would be to appeal to a person's familiarity with an area and structure the route overview by describing it in the context of previously travelled routes. In this approach, we could structure the route overview in such a way that the route is described as travel to some intermediate area followed by localized travel to zero in on the final destination (e.g., "Take Massachusetts Avenue into Harvard Square and your destination is a few blocks from the Harvard Bookstore."). The route could also be presented in terms of a previous route, with a few minor modifications ("The route is similar to the way you take to go to the gym, but instead of turning left onto Main Street, you turn right.").

Structuring the route overview in this manner puts a greater emphasis on the notion of selective suppression of detail and leaves more to the individual to interpret the things left unsaid. This is in contrast to ROVER's current approach of providing a framework that explicitly identifies the major trunk segments that resolve major geographic differences between two places. With these two alternate ways of presenting route overviews, we could compare selectively suppressing detail against providing details about a route's skeleton, and determine the relative importance of both in affecting route comprehensibility.

In addition to expanding route overviews to include knowledge of previous routes travelled, we could also expand how people use route overviews by making them more interactive. For example, in the computer desktop environment, we could enhance Google Earth's effectiveness as a route visualization tool by highlighting the major trunk segments and cognitive anchor points on a route. In an in-car GPS system, we could also make the route overviews dynamically adjust the scale at which they describe a route overview, based on how much of the route a person has travelled. For example, at the start of a trip from Manhattan to Boston, the route overview

could describe the route as a trip along major interstates to go between New York and Massachusetts, but then switch to a finer grain description using neighborhoods once the person arrives within the Boston city limits. The GPS could also supplement the dynamically shifting maps with appropriate speech output at the various points where the map scale shifts. This would have the benefit of giving a person an additional cue to adjust their mental perspective of what part of the route they are currently travelling as well as compensate for the small display on the GPS (which due to size constraints, has to limit the amount of the route that can be shown).

5.4 Building up LAIR

One final area of future work would focus on the methods used to build up the LAIR model of what places a person knows and using the knowledge of how this model is built up in a route overview. Two research problems in this area are investigating alternate methods to identify cognitive anchor points and incorporating people's canonical views of space in the route overview.

Both the rule-based approach LAIR uses and the machine-learning approach Hwang uses [28] are based on GIS datasets. An alternative to GIS-based approaches would be to use computer vision to divide an area into sub-regions based on common visual cues. Using Torralba's work in scene gisting [63], we could emulate what people do when they travel around and build up new areas based on what Lynch calls the "thematic continuity" of a neighborhood. Given the work of Google's Street View¹ team, the data to attempt this may already be in a form usable by Torralba's system. This approach would be particularly useful in dense urban areas such as Manhattan. While GIS data may be available to describe the differences in building types and usage, many neighborhoods have certain unique visual characteristics (such as the type of material used to construct buildings) that are not easily captured in GIS data but are readily apparent to the naked eye, and thus using a vision-based approach may be more effective in differentiating among closely packed neighborhoods.

¹<http://books.google.com/help/maps/streetview/>

Another area of future investigation is to incorporate knowledge of the usual routes a person travels and how this affects their canonical view of a place. As people spend more time in an area, their particular mental impression of that place evolves: a larger neighborhood may be split into two smaller areas, a person's perception of a neighborhood may expand in size as he explores more of the area, or two smaller neighborhoods could merge into one larger entity. We could include these canonical views in route overviews. For example, we could describe a route as approaching a destination from the "back" side (that is, from the direction that is the opposite of the usual direction of travel).

One other feature to include in route overviews is the ability to describe when parts of the turn-by-turn directions may run counter to a person's intuition. This could occur either because of spatial contradictions that arise due to heuristics people use to simplify their understanding of space or when new paths are introduced that optimize a route that would have only used previously known paths. In the first case, LAIR could reduce the sense of confusion that arises when the predictions produced by these heuristics are found to be false. LAIR could do this by including heuristics to simulate the spatial simplifications that people make when travelling, model spatial distortions that arise when local simplifications are extended globally, and then point out in the route overview that the actual geography is different from what the traveller would have expected. One example of this is when two streets are considered to be parallel in one part of town but then eventually meet (or even more confusingly, cross, so that Street A, which was once to the right of Street B, is now to the left; Commonwealth Ave and Beacon St in Boston is a good example of this). In the second case, when a new path is introduced, you could include in the directions a statement along the lines of "you might have thought of travelling along Street X, but Street Y is better because it is more direct and cuts through Neighborhood N instead of going around it like Street X."

Chapter 6

Related Work

Our attempt to make driving directions more comprehensible involved breaking the problem into three sub-tasks and using cognitively-inspired design guidelines to address each sub-task. The three sub-tasks were: creating an appropriate knowledge representation (LAIR), populating this representation (CAP LOC), and then using selected features from this representation to produce more comprehensible directions (by creating route overviews with ROVER). This chapter describes alternate approaches to each of these sub-tasks. We close with a look at route-planners that, instead of optimizing for travel time or distance, minimize the likelihood of making a mistake while following the turn-by-turn directions, or maximize people's preference for taking familiar paths.

6.1 Knowledge Representations of Space

LAIR uses a global metric reference frame (the street network) to situate geometric data structures that represent the structures used in human spatial cognition. A number of other spatial representations have been proposed previously, but these representations were designed for use in building a map from an agent's experiences exploring the environment. As a result, these knowledge representations do not have access to a global map.

One of the earliest knowledge representations of space is Kuipers' TOUR model [35,

36]. Lynch's work had a strong influence on the elements included in the TOUR model, which has a topological network of places and paths, routes that describe travel through this network, a hierarchy of regions, and place-specific orientation frames. When presented with go-to and turn instructions, TOUR uses inference rules to update its internal representation of its location and to add new topological information to the TOUR knowledge base.

The TOUR model was designed to accommodate the fact that the agent may have only partial knowledge of the environment (e.g., it may know where Place A and Place B are each located with respect to Place C, but it may not know how Place A is located with respect to B). TOUR incrementally builds up a global map by incorporating newly discovered spatial information with previously known information. The structure of the TOUR knowledge representation and the inference rules that operate on the representation demonstrate that a wide variety of interesting behavior can occur in the absence of complete knowledge and metric information.

The TOUR model was later followed by Kuipers' Spatial Semantic Hierarchy (SSH) [35, 37]. LAIR and TOUR demonstrate the rich capabilities of topological representations of space. The SSH had a broader scope: it incorporated multiple levels of spatial representations, with each higher level abstracting and supplementing the information from the level below it. The goal of the SSH was to provide a framework that could be used to design a robot with the ability to build a map of the area it explored. The overall design of the SSH closely follows theoretical frameworks on the development of human spatial knowledge put forth by Piaget [67, 68] and Siegel and White [77]. In these frameworks, spatial knowledge is thought to progress from knowledge about places to knowledge about routes between places, and then finally to topological and metric relationships between different places.

An alternate to using geometric data structures to represent places, paths, and regions is the use of vision to represent schematized views of these spatial elements. This is the approach taken by Chown's PLAN system [13]. In contrast to a hierarchy that featured increased sophistication in the spatial representation at higher levels as in the SSH, PLAN featured a hierarchy that emphasized schematized views which

cover a larger expanse as you go up the hierarchy. In addition to this, PLAN was a connectionist model in which spreading activation [5] was used to model the prominence of different places and paths. The spreading activation influenced both which paths are selected in route planning and which places appear in the next higher level of the spatial hierarchy.

An alternative to using vision to build up a map is to use other types of sensors such as ultrasonic or laser range finding. This is the approach used in SLAM (Simultaneous Localization And Mapping). This technique is used by autonomous robots and vehicles to construct a map of the environment while tracking the agent's location as it moves through that environment [79]. The map that is constructed does not contain the cognitive structures of the previous knowledge representations, and instead (as a result of the type of sensors used to construct the map) shows the barriers the agent has to avoid as it moves around. The challenge with SLAM is that noisy sensor data will accumulate over time, and without any correction, the map and the agent's estimate of its location will become more and more inaccurate. The issue of error-correction isn't emphasized in the previous knowledge representations, but it is critically important in SLAM. SLAM algorithms correct for accumulated error by using various statistical methods such as Kalman filters and particle filters to bring the noisy map image into focus [61].

While most approaches to SLAM are not as grounded in human spatial cognition as the work in this dissertation, the Atlas framework [10] is one exception. Atlas uses a concept similar to Lynch regions to allow SLAM algorithms to efficiently map large scale environments. The framework maintains multiple small-scale local maps which are later stitched together to form a global map. As the agent explores the environment, it builds up one local map at a time, and when the number of features in a local map exceeds a threshold, it starts creating a new local map. Using these smaller maps parallels how people think of travel in episodic chunks; as the robot agent (or person) moves from one region of space to another, each new region gets its own mini-map or canonical view (an episode in memory). These individual pieces are later stitched together into a comprehensive whole. Approaching the mapping

problem in this way increases the efficiency of SLAM algorithms.

6.2 Identifying Important Places and Regions

We have already mentioned a number of methods used to identify important places and regions. PLAN uses image data; SLAM algorithms use ultrasonic and laser range finding; and Singh [78] and Hwang's [28] work, mentioned in Chapter 3, use GIS data. In this section, we describe three other methods used to identify important places and regions.

The ubiquitous computing (ubiquitous computing) community uses location as an important factor in designing context-sensitive and mobile applications. The ubiquitous computing community is interested in identifying places that have major bearing on people's daily activities in order to provide them with timely information regarding tasks that can be opportunistically carried out, based on a person's location and what other things he is doing [19, 75, 83]. To this end, work in this field [7, 66] has used GPS data to track a person and identify places that person frequently visits in his day-to-day activities. So while these places do not serve as cognitive anchor points in the sense we use them in route overviews, they are nonetheless important to the ubiquitous computing field because they structure a person's day-to-day activities.

LAIR's approach to identifying the extent of a **Region** is to take a central point and then radiate outwards from there, in the direction of large land parcels or concentrations of commercial regions. Steinhauer [80] takes an alternate approach by identifying abstract regions in cartographic maps (such as villages in a map showing only the location of houses) using context-free grammars. These grammars build up the region by looking for the presence of the individual parts of the region and then clustering them together. To determine region boundaries, Steinhauer uses a modified convex hull algorithm that restricts the length of the edges that can be included in the hull. The resulting hull preserves some of the concavities in region boundaries, giving them a more natural-looking shape, similar to how people might outline boundaries.

Finally, instead of using a single definition to identify regional boundaries, another

approach is to appeal to the wisdom of the masses and build up a map of important places by aggregating individuals' submissions. This approach has the added benefit of tasking the people making the submissions with naming the places they identify. The web-based WikiMapia [89] and CommonCensus [14] projects take this approach.

WikiMapia combines a wiki's ease of collaboration with a Google Maps interface to allow people to identify places and regions on Earth. No matter what scale WikiMapia's map is set to, people can see the largest places at that map scale. However, these places are not necessarily the most significant places at that scale. Because of the large number of WikiMapia entries, it is difficult to determine which places are relevant cognitive anchor points versus places that have significance only to the person who created the entry. Therefore, the places identified by WikiMapia are not the most suitable for generating route overviews.

Instead of trying to name all places on Earth, the goal of CommonCensus is to identify the boundary of local neighborhoods by asking people where they live and the name of the neighborhood they live in. Although CommonCensus allows for free form responses, it uses a gazetteer to generate a list of candidate neighborhoods from which people can choose. The results from each person's CommonCensus then form a spatial scatter-plot that give an indication of different neighborhood boundaries.

6.3 Describing Routes

The textual route overviews ROVER produces make turn-by-turn directions more comprehensible by identifying a route's major trunk segments and cognitive anchor points, and selectively suppressing other details. An alternative is LineDrive [1], which is a map-rendering algorithm that makes directions more comprehensible by schematizing the map of a route, making it resemble a sketch a person might produce on a cocktail napkin (see Figure 6-1, on the next page). The maps LineDrive produces show the entire route, but they reduce visual clutter and make it easier to understand all the turns in the route by varying scale: shorter roads are lengthened and longer roads

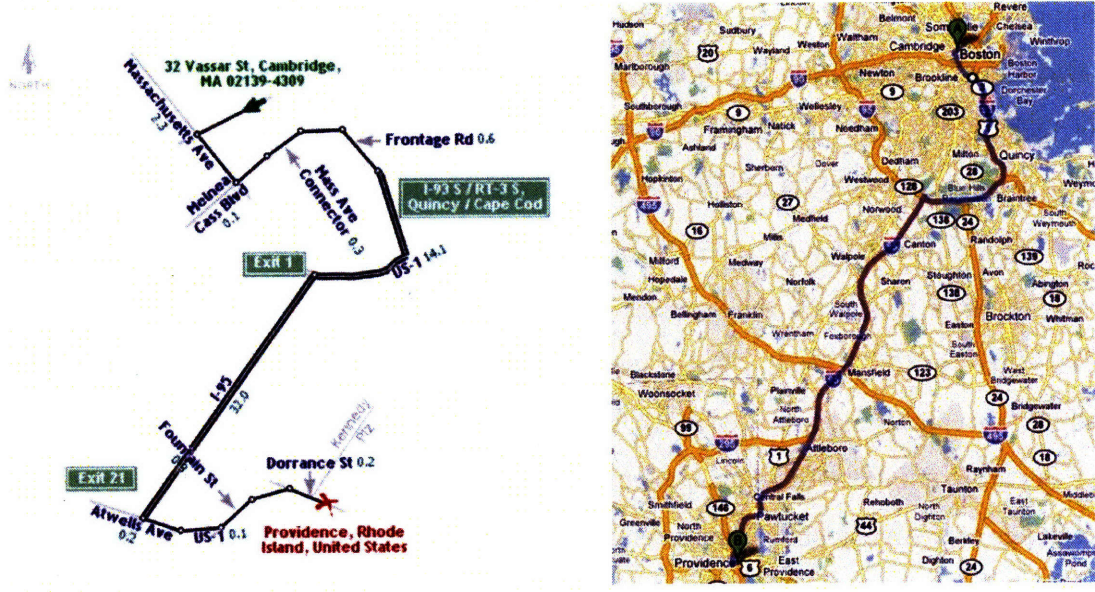


Figure 6-1: Example of a LineDrive route map (left) compared to the corresponding standard route map (right). This route goes from Cambridge, Massachusetts to Providence, Rhode Island. Note how the LineDrive map renders different parts of the route at different scales and how LineDrive straightens out the roads on the route.

are shortened. The angles at which roads intersect are also simplified and curves in a road are removed to make the map easier to read.

ROVER and LineDrive focus on improving the comprehensibility of a set of directions by helping people develop an overall understanding of a route. There is also an abundance of research that has looked into making detailed turn-by-turn directions easier to follow by identifying and including salient landmarks in the directions [21, 62, 69]. Note that these approaches regard landmarks in the sense that we defined them in Chapter 1: i.e., as places useful for local disambiguation, to identify where to make a specific turn. These landmarks are not intended to help a person globally orient themselves as cognitive anchor points do.

The work that finds salient landmarks to include in turn-by-turn directions takes each decision point along a route (road intersections), and uses GIS data to identify buildings near the decision point with high salience. A building has high salience if certain characteristics, such as its height or facade shape, are outliers compared to its neighboring buildings. There has also been research that has looked into formalizing

how the go-to and turn directives in a route can be chunked together to produce a more natural sounding set of directions [33, 70]. Richter and Klippel have proposed combining landmarks with this formalization used for chunking to present turn-by-turn directions not as a linear set of singular directives, but rather as a small handful of easily memorable chunks, with each chunk containing a small set of easily memorable go-to and turn directives [70].

Kuipers MARCO system [49] investigates the issue of interpreting the turn-by-turn directions people produce. This is a challenge given all that is left unsaid in these directions. For example, the instruction “Take the blue path to the chair” requires a person to find and get onto the blue path and then look in both directions to find the chair before he knows in which direction on the path to travel. MARCO approaches this problem by translating linguistic clauses into a set of actions to be carried out by a software agent in a virtual environment. It reasons about what is implicitly stated in the directions by identifying pre- and post-conditions for travel, and it carries out a planning stage and executes this plan to achieve the necessary pre-conditions for travel.

Finally, we note that the research in turn-by-turn directions extends to environments other than outdoor driving directions. Fontaine and Denis have compared the characteristics of directions given to navigate underground environments to those for outdoor above-ground settings [22]. Also, LAIR was used to create the Stata Walking Guide [44], an application that produced directions between different places in MIT’s computer science building, the Stata Center.

6.4 Finding Better Routes

The work described in the previous section focused on making a particular route between two places easier for people to understand. These routes are usually chosen by optimizing for shortest distance or minimal travel time. Other work has planned routes that optimize for ease of following the route. This section describes that work, which is orthogonal to ROVER’s work and nicely complements it. It would be possible

to modify ROVER to interface with these systems and analyze the routes these systems produce, identify trunk segments and cognitive anchor points in these routes, and use these trunk segments and anchor points to produce a route overview.

Duckham has proposed a simplest-path path-planning algorithm [20] that does not use any sort of distance metric and instead uses the complexity of negotiating a turn as a cost metric [51, 82]. The cost function takes into account the number of possible turns that can be made at an intersection as well as how difficult it is to recognize the intersection. This cost function can also be modified to take into account preferences for other properties of the route, such as road type (major artery versus local road). Anecdotally, Duckham found that the routes his algorithm produced looked easier to follow because they contained fewer turns. These simpler routes were, on average, 16% longer than the shortest route.

In Chapter 5, one area of future work we proposed was describing route overviews in terms of previously known routes. Patel's MyRoute system [65] accounts for previously travelled routes when it plans new routes and uses those previous routes in the turn-by-turn directions. The MyRoute system plans a route by minimizing a cost function that takes into account the number of steps in a set of turn-by-turn directions (previously travelled routes that are included in MyRoute's turn-by-turn directions are described as a "go-to Place X" instruction, and thus count as only one step), the length of the route, and the time required to travel the route. Unlike ROVER, Patel's work doesn't take into account neighborhoods.

Lastly, in Letchner's study of people's route preferences, she found that people take the fastest route only 35% of the time [41]. Instead of explicitly identifying what factors account for this, Letchner's TRIP system uses an individual's previous routes as expressions of which roads they prefer to travel along. The multiplicative factor by which the fastest route is faster than the actual route a person takes is then used by TRIP to plan new routes; TRIP multiplies previously travelled roads by this factor to discount the cost of travelling along a familiar path versus taking a path that has never been travelled. Letchner found that applying this discount, which implicitly accounts for preferences and doesn't explicitly model what factors

go into determining those preferences, is more accurate than using the fastest route as a predictor of which routes people use to travel between two places.

TRIP's assumption that, after accounting for the discount factor, people prefer to travel along roads they've taken before, can be limiting because it doesn't account for the benefit people get from learning new paths that could be used in future route planning. We discussed this same issue previously in Chapter 5, where we described how one area of future work for ROVER is to explain how taking a route with a new path would be better than taking a route using only previously known paths.

Chapter 7

Contributions

This dissertation presented a cognitively-inspired approach to make driving directions more comprehensible. By applying insights from studies of human spatial cognition, we developed the LAIR spatial representation, populated LAIR with the CAP LOC system, and built the ROVER system to produce route overviews.

Route overviews provide a high-level understanding of where a route goes. They are intended to complement rather than replace turn-by-turn directions. Studies of human spatial cognition indicate travel is viewed not as a sequence of turns which have little to do with one another, but as something that has a larger organization, a greater sense than what is currently presented in turn-by-turn directions. Route overviews account for this larger organization. The structure and content of the route overviews we created were guided by the following cognitively-inspired design principles:

1. *Use spatial hierarchy and goal-directed descriptions.* Route overviews should appeal to people's sense of spatial hierarchy and be structured in such a way that they situate the individual steps in the turn-by-turn directions in the context of some larger goal.
2. *Selectively suppress detail.* Specific details do not need to be presented in a route overview because they can be inferred by people's knowledge of the environment or read directly from the detailed turn-by-turn directions when a

person is actually travelling along the route. Thus, route overviews do not include distances or turn directions, and they do not list every street and turn. By selectively suppressing detail in this way, we make the bigger picture easier to see.

3. *Identify road skeletons and cognitive anchor points.* We suppress detail and we emphasize the major trunk segments and cognitive anchor points on the route. Cognitive anchor points are not just distinct landmarks that are useful to help people navigate turns. Rather, as the name suggests, they are the quintessential places that define a city and give it its identity. In the larger view of things, people don't see themselves as going from one turn to the next; in our episodic memory, we travel from one cognitive anchor point to the next along these major trunks.

We conducted a user study in which we asked participants to compare the route overviews and turn-by-turn directions ROVER produces against a set of standard turn-by-turn directions generated by Google Earth. Generally speaking, when we control for the spatial knowledge a person has, study participants rated ROVER's output as being more comprehensible than traditional turn-by-turn directions. Participants also demonstrated a preference for ROVER's output over traditional turn-by-turn directions.

More interestingly, people who did not fit the spatial knowledge profile for which ROVER's output was designed also rated the ROVER output better than the Google Earth directions. These results are consistent with our cognitively-inspired design guidelines, since the guidelines were drawn from general principles of human spatial cognition and were independent of which particular places a person knows. These results demonstrate the additional descriptive power in our approach of using what we know about how people think about space — independent of what particular places they know — to guide how we structure the presentation of spatial information.

Bibliography

- [1] Maneesh Agrawala and Chris Stolte. Rendering Effective Route Maps: Improving Usability Through Generalization. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pages 241–249. ACM Press, 2001.
- [2] Gary Allen. From Knowledge to Words to Wayfinding: Issues in the Production and Comprehension of Route Directions. In *Spatial Information Theory: A Theoretical Basis for GIS*, volume 1329 of *Lecture Notes in Computer Science*, pages 363–372. Springer, 1997.
- [3] Gary L. Allen. Principles and Practices for Communicating Route Knowledge. *Applied Cognitive Psychology*, 14(4):333–359, July/August 2000.
- [4] Gary L. Allen and Kathleen C. Kirasic. Effects of the Cognitive Organization of Route Knowledge on Judgments of Macrospatial Distance. *Memory and Cognition*, 13(3):218–227, 1985.
- [5] John R. Anderson. A Spreading Activation Theory of Memory. *Journal of Verbal Learning and Verbal Behavior*, 22:261–295, 1983.
- [6] Luc Anselin. Local Indicators of Spatial Association — LISA. *Geographical Analysis*, 27(2):93–115, 1995.
- [7] Daniel Ashbrook and Thad Starner. Learning Significant Locations and Predicting User Movement with GPS. In *International Symposium on Wearable Computing*, pages 101–108, Seattle, WA, 2002.
- [8] Sonit Bafna. Space Syntax: A Brief Introduction To Its Logic And Analytical Techniques. *Environment and Behavior*, 35(1):17–29, January 2003.
- [9] Bettina Berendt. *Representation and Processing of Knowledge about Distances in Environmental Spaces*. PhD thesis, University of Hamburg, 1999.
- [10] Michael Bosse, Paul Newman, John Leonard, and Seth Teller. SLAM in Large-Scale Cyclic Environments Using the Atlas Framework. *The International Journal of Robotics Research*, 23(12):1113–1139, 2004.

- [11] Christopher T. Burris and Nyla R. Branscombe. Distorted Distance Estimation Induced by a Self-relevant National Boundary. *Journal of Experimental Social Psychology*, 41:305–312, 2005.
- [12] William. G. Chase. Spatial Representations of Taxi Drivers. In *Acquisition of Symbolic Skills*, pages 391–405. Plenum Press, 1982.
- [13] Eric Chown, Stephen Kaplan, and David Kortenkamp. Prototypes, Location, and Associative Networks (PLAN): Towards a Unified Theory of Cognitive Mapping. *Cognitive Science*, 19(1):1–51, 1995.
- [14] CommonCensus. <http://www.commoncensus.org/>.
- [15] Nelson Cowan. The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences*, 24(1):87–114, February 2001.
- [16] James M. Dabbs, E-Lee Chang, Rebecca A. Strong, and Rhonda Milun. Spatial Ability, Navigation Strategy, and Geographic Knowledge Among Men and Women. *Evolution and Human Behavior*, 19(2):89–98, 1998.
- [17] Robert Dale, Sabine Geldof, and Jean-Philippe Prost. CORAL: Using Natural Language Generation for Navigational Assistance. In *ACSC '03: Proceedings of the 26th Australasian Computer Science Conference*, pages 35–44, 2003.
- [18] Michel Denis, Francesca Pazzaglia, Cesare Cornoldi, and Laura Bertolo. Spatial Discourse and Navigation: An Analysis of Route Directions in the City of Venice. *Applied Cognitive Psychology*, 13(2):145–174, 1999.
- [19] Anind K. Dey and Gregory D. Abowd. CybreMinder: A Context-Aware System for Supporting Reminders. In *Proceedings of the 2nd International Symposium on Handheld and Ubiquitous Computing*, pages 172–186. Springer-Verlag, 2000.
- [20] Matt Duckham and Lars Kulik. Simplest Paths: Automated Route Selection for Navigation. In Werner Kuhn, Michael Worboys, and Sabine Timpf, editors, *COSIT 2003*, volume 2825 of *Lecture Notes in Computer Science*, pages 169–185, Kartause Ittingen, Switzerland, September 2003. Springer.
- [21] Birgit Elias. Extracting Landmarks with Data Mining Methods. In *COSIT 2003*, volume 2825 of *Lecture Notes in Computer Science*, pages 375 – 389. Springer-Verlag, January 2003.
- [22] Sylvie Fontaine and Michel Denis. The Production of Route Instructions in Underground and Urban Environments. In Christian Freksa and David M. Mark, editors, *COSIT 1999*, volume 1661 of *Lecture Notes in Computer Science*, pages 83–94, Stade, Germany, August 1999. Springer.

- [23] Reginald Golledge. Representing, Interpreting, and Using Cognized Environments. In *Papers and Proceedings of the Regional Science Association*, volume 41, pages 169–204, 1978.
- [24] Reginald Golledge. Geographical Perspectives on Spatial Cognition. In *Behavior and Environment: Psychological and Geographical Approaches*. Elsevier Science Publishers, 1993.
- [25] Reginald G. Golledge, Valerie Dougherty, and Scott Bell. Acquiring Spatial Knowledge: Survey versus Route-Based Knowledge in Unfamiliar Environments. *Annals of the Association of American Geographers*, 85(1):134–158, March 1995.
- [26] Reginald G. Golledge, Amy J. Ruggles, James W. Pellegrino, and Nathan D. Gale. Integrating Route Knowledge in an Unfamiliar Neighborhood: Along and Across Route Experiments. *Journal of Environmental Psychology*, 13:293–307, 1993.
- [27] Keith Holyoak and Wesley A. Mah. Cognitive Reference Points in Judgements of Symbolic Magnitude. *Cognitive Psychology*, 14:328–352, 1982.
- [28] Jie-Eun Hwang. *Heuristic Noll Map: A Machine Learning Approach for Interpreting Urban Morphology*. PhD thesis, Harvard University Graduate School of Design, 2007.
- [29] Petra Jansen-Osmann and Bettina Berendt. What Makes a Route Appear Longer? An Experimental Perspective on Features, Route Segmentation, and Distance Knowledge. *The Quarterly Journal of Experimental Psychology: Section A*, 2005.
- [30] Petra Jansen-Osmann and Gunnar Wiedenbauer. The Influence of Turns on Distance Cognition: New Experimental Approaches to Clarify the Route-Angularity Effect. *Environment and Behavior*, 36(6):790–813, 2004.
- [31] Young Ook Kim and Alan Penn. Linking the Spatial Syntax of Cognitive Maps to the Spatial Syntax of the Environment. *Environment and Behavior*, 36(4):483–504, July 2004.
- [32] Alexander Klippel, Kai-Florian Richter, Thomas Barkowsky, and Christian Freksa. The Cognitive Reality of Schematic Maps. In Liqiu Meng, Alexander Zipf, and Tumasch Reichenbacher, editors, *Map-based Mobile Services - Theories, Methods and Implementations*, pages 57–74. Springer; Berlin, 2005.
- [33] Alexander Klippel, Heike Tappe, Lars Kulik, and Paul U. Lee. Wayfinding Choremes - A Language for Modeling Conceptual Route Knowledge. *Journal of Visual Languages and Computing*, 16(4):311–329, 2005.
- [34] Benjamin Kuipers. The Skeleton in the Cognitive Map. *Environment and Behavior*, 35(1):81–106, 2003.

- [35] Benjamin Kuipers. An Intellectual History of the Spatial Semantic Hierarchy. In *Robotics and Cognitive Approaches to Spatial Mapping*, volume 38, pages 243–264. Springer, 2008.
- [36] Benjamin J. Kuipers. *Representing Knowledge of Large-Scale Space*. PhD thesis, Massachusetts Institute of Technology, July 1977. MIT AI/TR 418.
- [37] Benjamin J. Kuipers and Tod S. Levitt. The Spatial Semantic Hierarchy. *Artificial Intelligence*, 119:191–233, 2000.
- [38] Barbara Landau and Ray Jackendoff. “What” and “Where” in Spatial Language and Spatial Cognition. *Behavioral and Brain Sciences*, 16(2):217–238, 1993.
- [39] Carol A. Lawton. Gender Differences in Wayfinding Strategies: Relationship to Spatial Ability and Spatial Anxiety. *Sex Roles*, 30, 1994.
- [40] Carol A. Lawton. Strategies for Indoor Wayfinding: The Role of Orientation. *Journal of Environmental Psychology*, 16:137–145, 1996.
- [41] Julia Letchner, John Krumm, and Eric Horvitz. Trip Router with Individualized Preferences (TRIP): Incorporating Personalization Into Route Planning. In *Proceedings of the Eighteenth Innovative Applications of Artificial Intelligence Conference*, Boston, MA, July 2006. AAAI.
- [42] Robert Lloyd. Cognitive Maps: Encoding and Decoding Information. *Annals of the Association of American Geographers*, 79(1):101–124, March 1989.
- [43] Robert Lloyd and Christopher Heivly. Systematic Distortions in Urban Cognitive Maps. *Annals of the Association of American Geographers*, 77(2):191–207, 1987.
- [44] Gary Look, Buddhika Kottahachchi, Robert Laddaga, and Howard Shrobe. A Location Representation for Generating Descriptive Walking Directions. In *IUI '05: Proceedings of the 10th International Conference on Intelligent User Interfaces*, pages 122–129. ACM Press, 2005.
- [45] Kristin L. Lovelace, Mary Hegarty, and Daniel R. Montello. Elements of Good Route Directions in Familiar and Unfamiliar Environments. In Christian Freksa and David M. Mark, editors, *COSIT 1999*, volume 1661 of *Lecture Notes in Computer Science*, pages 65–82, Stade, Germany, August 1999. Springer.
- [46] Kevin Lynch. *The Image of the City*. The MIT Press, 1962. Second printing.
- [47] Alan M. MacEachren. *How Maps Work*. The Guilford Press, 1995.
- [48] Alastair MacFadden, Lorin Elias, and Deborah Saucier. Males and Females Scan Maps Similarly, but Give Directions Differently. *Brain and Cognition*, 53(2):297–300, November 2003.

- [49] Matt MacMahon, Brian Stankiewicz, and Benjamin Kuipers. Walk the Talk: Connecting Language, Knowledge, and Action in Route Instructions. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*, pages 1475–1482, 2006.
- [50] George Mandler. Organization and Memory. In *The Psychology of Learning and Motivation*, volume 1. Academic Press, 1967.
- [51] David M. Mark. Automated Route Selection for Navigation. *IEEE Aerospace and Electronic Systems Magazine*, 1:2–5, 1986.
- [52] Scott Martin. How Many Points of Interest (POI) is Good Enough on a GPS? <http://www.gpslodge.com/archives/014618.php>, November 2007.
- [53] Diane McGuinness and Janet Sparks. Cognitive Style and Cognitive Maps: Sex Differences in Representations of a Familiar Terrain. *Journal of Mental Imagery*, 7(2):91–100, 1983.
- [54] Timothy P. McNamara. Memory’s View of Space. In *The Psychology of Learning and Motivation*, volume 27, pages 147–186. Academic Press, Inc., 1991.
- [55] Timothy P. McNamara and Vaibhav A. Diwadkar. Symmetry and Asymmetry of Human Spatial Memory. *Cognitive Psychology*, 34:160–190, 1997.
- [56] Pierre-Emmanuel Michon and Michel Denis. When and Why Are Visual Landmarks Used in Giving Directions? In Daniel R. Montello, editor, *COSIT 2001*, volume 2205 of *Lecture Notes in Computer Science*, pages 292–305, Morro Bay, California, September 2001. Springer.
- [57] George A. Miller. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63:81–97, 1956.
- [58] Ian Moar and Gordon Bower. Inconsistency in Spatial Knowledge. *Memory and Cognition*, 11(2):107–113, 1983.
- [59] Mark Monmonier. *How to Lie with Maps*. University of Chicago Press, 2nd edition, 1996.
- [60] Daniel R. Montello, Michael F. Goodchild, Jonathon Gottsegen, and Peter Fohl. Where’s Downtown?: Behavioral Methods for Determining Referents of Vague Spatial Queries. *Spatial Cognition and Computation*, 3:185–204, 2003.
- [61] Michael Montemerlo, Sebastian Thrun, Daphne Koller, and Ben Wegbreit. Fast-SLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, 2002.

- [62] Clemens Nothegger, Stephan Winter, and Martin Raubal. Selection of Salient Features for Route Directions. *Spatial Cognition and Computation*, 4(2):113–136, 2004.
- [63] Aude Oliva and Antonio Torralba. Building the Gist of a Scene: The Role of Global Image Features in Recognition. *Visual Perception, Progress in Brain Research*, 155:23–36, 2006.
- [64] Jean Pailhous. Representation de l’Espace Urbain et Cheminements (Representation of Urban Space and Roadways). *Le Travail Humain*, 32:239–270, 1969.
- [65] Kayur Patel, Mike Y. Chen, Ian Smith, and James A. Landay. Personalizing Routes. In *UIST ’06: Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, pages 187–190, New York, NY, USA, 2006. ACM.
- [66] Donald J. Patterson, Lin Liao, Krzysztof Gajos, Michael Collier, Nik Livic, Katherine Olson, Shiaokai Wang, Dieter Fox, and Henry Kautz. Opportunity Knocks: a System to Provide Cognitive Assistance with Transportation Services. In Nigel Davies, Elizabeth Mynatt, and Itiro Sio, editors, *Proceedings of UBIComp 2004: The Sixth International Conference on Ubiquitous Computing*, volume LNCS 3205, pages 433–450. Springer-Verlag, October 2004.
- [67] Jean Piaget and Barbel Inhelder. *The Child’s Conception of Space*. Routledge and Kegan Paul, London, 1956.
- [68] Jean Piaget, Barbel Inhelder, and Alina Szeminska. *The Child’s Conception of Geometry*. Basic Books, New York, 1960.
- [69] Martin Raubal and Stephan Winter. Enriching Wayfinding Instructions with Local Landmarks. In Max J. Egenhofer and David M. Mark, editors, *GIScience*, number 2478 in Lecture Notes in Computer Science, pages 243–259, Boulder, Colorado, September 2002. Springer.
- [70] Kai-Florian Richter and Alexander Klippel. A Model for Context-Specific Route Directions. In *Spatial Cognition*, pages 58–78, 2004.
- [71] Edward K Sadalla, W. J. Burroughs, and L. J. Staplin. Reference Points in Spatial Cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 6(5):516–28, September 1980.
- [72] Edward K. Sadalla and Daniel R. Montello. Remembering Changes in Direction. *Environment and Behavior*, 21:346–363, 1989.
- [73] Edward K. Sadalla and Lorin J. Staplin. The Perception of Traversed Distance. *Environment and Behavior*, 12(2):167–182, 1980.
- [74] Daniel L. Schacter. *Searching for Memory: The Brain, the Mind, and the Past*. Basic Books, 1996.

- [75] A.J. Sellen, G. Louie, J.E. Harris, and A.J. Wilkins. What Brings Intentions to Mind? An In Situ Study of Prospective Memory. *Memory*, 5(4):483–507, 1997.
- [76] M. Jeanne Sholl. Cognitive Maps as Orienting Schema. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13:615–628, 1987.
- [77] Alexander W. Siegel and Sheldon H. White. The Development of Spatial Representations of Large-Scale Environments. In *Advances in Child Development and Behavior*, volume 10, pages 9–55. Academic Press Inc, 1975.
- [78] Raj R. Singh. Adapting Geographic Information Systems to Sketch Planning Needs. Master’s thesis, Massachusetts Institute of Technology, 1996.
- [79] Randall Smith, Matthew Self, and Peter Cheeseman. Estimating Uncertain Spatial Relationships in Robotics. In *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag New York, Inc., New York, NY, USA, 1990.
- [80] Joe Heike Steinhauer, Tom Wiese, Christian Freksa, and Thomas Barkowsky. Recognition of Abstract Regions in Cartographic Maps. In *COSIT 2001: Proceedings of the International Conference on Spatial Information Theory*, pages 306–321. Springer-Verlag, 2001.
- [81] A. Stevens and P. Coupe. Distortions in Judged Spatial Relations. *Cognitive Psychology*, 10:422–437, 1978.
- [82] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to Tell People Where to Go: Comparing Navigational Aids. *International Journal of Man-Machine Studies*, 22(5):549–562, 1985.
- [83] Sakari Tamminen, Antti Oulasvirta, Kalle Toiskallio, and Anu Kankainen. Understanding Mobile Contexts. In Luca Chittaro, editor, *Mobile HCI 2003*, volume 2795 of *Lecture Notes in Computer Science*, pages 17–31. Springer, 2003.
- [84] Barbara Tversky. Distortions in Memory for Maps. *Cognitive Psychology*, 13(3):407–433, 1981.
- [85] Barbara Tversky. Cognitive Maps, Cognitive Collages, and Spatial Mental Models. In *Spatial Information Theory: A Theoretical Basis for GIS, Proceedings COSIT ’93*, volume 716 of *Lecture Notes of Computer Science*, pages 14–24, 1993.
- [86] Barbara Tversky. Remembering Spaces. In E. Tulving and F.I.M. Craik, editors, *Handbook of Memory*, pages 363–378. Oxford University Press, 2000.
- [87] Barbara Tversky. Navigating by Mind and by Body. In *Spatial Cognition III*, volume 2685 of *Lecture Notes in Computer Science*, 2003.
- [88] Barbara Tversky. Structures of Mental Spaces: How People Think About Space. *Environment and Behavior*, 35(1):66–80, 2003.

- [89] WikiMapia. <http://www.wikimapia.org>.
- [90] R. Wilton. Knowledge of Spatial Relations: A Specification from Information Used in Making Inferences. *Quarterly Journal of Experimental Psychology*, 31:133–146, 1979.
- [91] Craig Zimring and Ruth Conroy Dalton. Linking Objective Measures of Space to Cognition and Action. *Environment and Behavior*, 35(1):3–16, January 2003.