Abstract

The **Back Seat Driver** is a research prototype of a system to use speech synthesis as a navigational aid for an automobile equipped with localization equipment. We are evaluating the user interface by field trials. As this is work in progress, this paper will primarily give an overview of the system and describe its components. Included will be discussion of the map database, route finding algorithm, repair strategies, and the discourse generator.

Goals

The main goal of this project is to evaluate the utility of speech synthesis as the user interface to a real-time navigation system in an urban environment. Of particular concern is the discourse structure:

- how should driving acts be described?
- how can a description be generated from a route?
- how should timing considerations be applied?
- what kinds of feedback, both positive and negative, does the user require?
- what kinds of visual cues are most useful in describing an approaching location?

This information is gained from both laboratory simulations and field trials.

Our desire is to build the best possible real-time route describer. Although we believe a speech interface to the navigation unit is superior and safer than a visual interface, we do not plan to conduct direct comparison studies.

In the course of field trials to evaluate and improve our automatic direction giving, we hope to specify key components of the map database. We expect discourse behavior may need to vary with conditions (traffic, weather, day/night). It is likely that different visual cues may be useful in these situations. All must be represented in the database.

Geographic Database

Our database covers 41 square miles in the Boston area, including parts of Boston, Cambridge, Brookline, Somerville, and Watertown. It originated as a DIME (Dual Independent Map Encoding) file distributed by the United States Geological Survey[2]. A DIME file consists of a set of straight line segments, each with a name, a type, endpoints in longitude and latitude, and some additional information such as address numbers. Initially our database contained many errors. Correcting them required actually traveling most of the segments.

A DIME file alone is not sufficient for finding routes. The DIME files indicate physical connectivity, but route finding requires *legal* connectivity, i.e., one can legally drive from one segment to the next (one way streets are a simple example). We extended the data base format to explicitly represent legal connectivity. Since some streets are better than others, the database must include a measures of *quality*. We take this to be a largely subjective measure of the ease of locating and following a street. This allows the route finder to prefer to use streets of higher quality.

The generation of easy followed natural descriptions requires more extensions. We added a number of new segment types to distinguish bridges, underpasses, tunnels, rotaries, and access ramps. All these extensions were done for an earlier route finding project[1].

We are presently adding landmarks to the database. Drivers need landmarks to know how far to drive and when to turn. If the Back Seat Driver had eyes, it could simply choose landmarks as needed by looking for them in the landscape. Being blind, it must rely on landmarks coded into the map database. We have already added traffic lights to the landmark database. A main task now is to determine what else must be added.

System

Our vehicle is equipped with a localization unit built by NEC Home Electronics, Ltd., the project sponsor. It is a dead-reckoning position keeping system which uses speed and direction sensors. To compensate for error, it uses map matching on a map database stored on CD ROM. The system described more fully in [**?**].

As this is a research prototype, much of the computation is done in a base station computer laboratory (on a Symbolics Lisp Machine), rather than a computer on the vehicle. Two cellular telephones link the computer to the car. The on board navigational hardware transmits position and velocity via modem and cellular phone to the base station. The base station computer does all route planning and discourse generation. Speech synthesis is performed in a commercial text-to-speech synthesizer (Dectalk) cabled to the Lisp Machine. Synthesized instructions to the driver are relayed via the second cellular link and a speaker phone in the car. The keypad of the second phone also serves as the driver's control unit for the Back Seat Driver. Through this keypad a driver selects a destination, requests repeats of spoken information, and accesses other services of the Back Seat Driver.

A block diagram of the system appears in Figure 2, below.

Discourse Strategies

The instructions are detailed and natural, and include a rich taxonomy of driving verbs. The dialog system uses cues such as vehicle speed and difficulty of driving actions to attempt to deliver instructions at the proper pace and in a timely manner. In addition, the system can anticipate some of the driver's possible mistakes and give warnings to avoid them.

If the driver does make a wrong turn, or misses a turn, the Back Seat Driver describes the error and then incrementally calculates a new route, rather than simply back-tracking to the point of the error. Route planning includes weighting for length of the trip, difficulty of driving maneuvers (such as left turns against traffic), street quality, and complexity of the spoken directions.

As opposed to much prior work in discourse generation, the Back Seat Driver is a real-time system which must factor in a number of temporal considerations. It needs to give each stage in the directions at just the right point, in terms of the time it takes to execute the driving maneuver as well as the speed of the vehicle approaching the intersection. For safety considerations, we would rather err on the side of giving the driver plenty of warning, but a cue given too far in advance may be miscued (e.g., a turn taken at an earlier intersection). Additionally, the software must consider the length of time it will take to recite an utterance. It is better to miss a turn and plan a new route than start describing the turn at a time when it may be unsafe to execute it (i.e., already well into an intersection).

Summary

The Back Seat Driver is already working in prototype form.

Our present concerns are to determine what a spoken driving assistant should say, to understand how time and speed affect this decision, and to learn what features a map database must have to support generation of instructions.

Acknowledgments

The authors wish to gratefully acknowledge the support of NEC Home Electronics, Ltd.

References

[1] James R. Davis and Thomas F. Trobaugh. Direction assistance. Technical Report 1, MIT Media Laboratory Speech Group, Dec 1987.

[2] U.S. Department of Commerce, Data Users Services Division. *Geographic Base File GBDF/DIME: 1980 Technical Documentation*, 1980.