spacebook-project.eu

Project Periodic Report

Grant Agreement number: 270019
Proposal acronym: SpaceBook
Project title: Spatial & Personal Adaptive Communication Environment: Behaviors & Objects & Operations & Knowledge
Funding Scheme: EC FP7, Call 6

Date of latest version of Annex I against which the assessment will be made:

Periodic report: 1st
Period covered: 1-March-11 - 29-Feb-12

Name, title and organization of the scientific representative of the project's coordinator:

Michael Minock
Associate Professor
Umeå University, Sweden

Tel: +46 90 786 6398
Fax: +46 90 786 6126
E-mail: mjm@cs.umu.se

Project website address: `http://spacebook-project.eu/`

1 Publishable summary

1.1 Project context and objectives

The SPACEBOOK project is prototyping a speech-driven, hands-free, eyes-free device for pedestrian navigation and exploration. SPACEBOOK is developed as an open source project and progress will be bench marked through controlled task-based experiments with real pedestrians in central Edinburgh. The SPACEBOOK project will generate concrete technical and scientific advances for eyes-free, hands-free navigation and exploration systems which will support applications in tourism. It also has potential application in rural contexts, and among user communities such as the visually impaired and the emergency services (tasks that must be unencumbered by invasive technology). While there has been recent industrial interest in voice-based, mobile city information systems (e.g. Apple's Siri system), SPACEBOOK is unique in pioneering the modelling of the pedestrian's field of view to provide situated speech-based dialogue to support both navigation and exploration. No industrial effort to our knowledge is yet attempting this.

In addition to advances in navigation and exploration systems, SPACEBOOK provides a task environment in which more fundamental scientific and technical knowledge will be generated. Specifically we seek to advance the state of the art in:

I. **Model-based approaches to plan generation and recognition.** There are essentially three approaches to building agents: 1.) they can be hand programmed; 2.) they can be learned from data; 3.) they can be *model-based*, arising from a first principles model of the agent's environment, goals, actions, sensors, etc. The model-based approach is characterized by very flexible behavior, but significant scalability challenges. The SPACEBOOK task environment is a real-world example that will help evaluate the merits of the model-based approach and drive algorithmic innovation in generation and recognition of behavior in the model-based approach.

II. **Statistical learning techniques for interaction management.** The determination of when and in what sequence to execute actions in an uncertain, partially-observable environment is a long-standing topic in Artificial Intelligence. One branch of work focuses on using reinforcement learning over *partially observable Markov decision processes* (POMDPs) to decide on a policy that determines the correct action to execute based on the agent's belief state about the world. There are three key issues that complicate this problem: 1.) determining which features of the world should map into a limited set of state variables for the POMDP; 2.) building realistic simulations to feed reinforcement learning with sufficient data; 3.) algorithmic issues that enable POMDPs to scale to larger numbers of state variables and actions. SPACEBOOK provides an ideal task environment in which to test reinforcement learning for POMDPs with actions being the next communication action of SPACEBOOK system.

III. **Machine learning of natural language understanding components.** A very active research area lately has been the learning of natural language understanding components that map natural language (e.g. English) to logical expressions in a *meaning representation language* (MRL) that can drive applications (e.g. query a database, control a DVD player, etc.). Typically the learning problem takes as input a corpus of natural language expressions paired with their corresponding MRL expressions and the output of learning is a component that can map natural language to MRL expressions. We shall drive innovation in this area by defining an MRL adequate to SPACEBOOK's complex, real world task environment, as well as create a semantic parsing corpus for the SPACEBOOK domain. This is a clearly sought after resource for the machine learning community, which is currently using only a small number of semantic parsing corpora based on simple scenarios such as booking airline flights. Moreover our approach to structural learning should extend the state of the art in learning of natural language understanding components.

While the three above areas are our declared basic research areas, there are a host of other areas in which project members seek more basic research results, including: (IV) *minimum recursion semantics (Copestake, et al.)* based meaning representation languages, (V) textual descriptions of vista space, (VI) Question Answering (QA) systems over textual resources, and more.

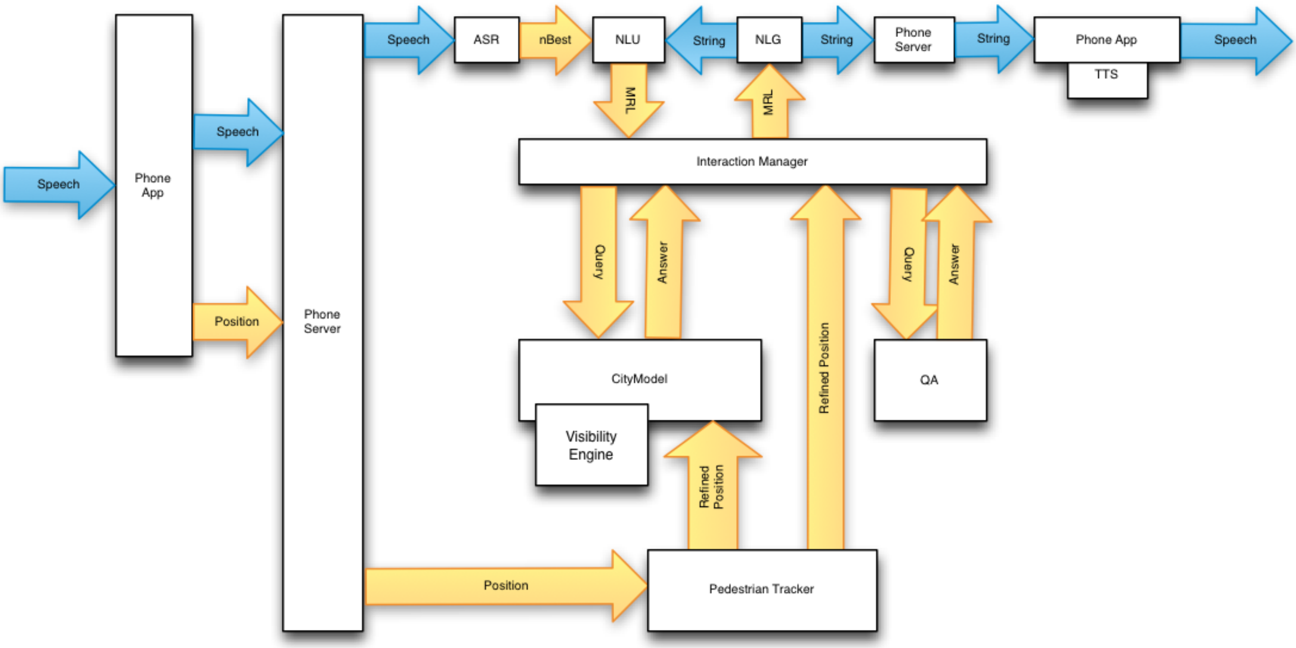

Figure 1: Initial SPACEBOOK prototype

1.2 Work performed since the project start and main results to date

SPACEBOOK has completed its 12th month (of 36) and we have met all of our deliverable goals. In short this means that all the components are in place to support the integrated SPACEBOOK prototype which is due in month 14 of the project. Figure 1 shows the components and data flows for the first SPACEBOOK prototype. Specifically:

PhoneApp runs on the pedestrian's Android mobile phone. It makes connection using the 3G network to a port on a machine running PhoneServer. PhoneApp relays speech and position data from the to the PhoneServer. Likewise it receives text from the PhoneServer to voice using Android's text-to-speech component. An alternate solution will be to use a server side text-to-speech component (e.g. Cereproc), the PhoneApp accepts audio files that it relays to the Android's speaker.

PhoneServer establishes 3G connections with the PhoneApp, the ASR component and the NLG component. Also it can relay audio files from a back-end text-to-speech component.

CityModel includes a PostGIS/PostgreSQL database that contains structured information for part of the city of Edinburgh (e.g. location and shape of buildings, monuments, streets, etc.). This database is populated with data from a variety of sources including Ordnance Survey's MasterMap, OpenStreetMap (OSM) and PointX. The city model includes attributes that describe places of interest drawn from a set of documents in *Gazetteer of Scotland* and *Wikipedia*. A full record of the pedestrian's history of positions and inferred viewing directions is also maintained in the city model.

QA a service that answers definition and factoid questions over a set of of documents from *Gazetteer of Scotland* and *Wikipedia*.

VisibilityEngine a service coupled with the CityModel which determines what can be seen from where in the city. Given any pedestrian position and viewing direction, this component can identify what entities in the city model can be seen and how well.

PedestrianTracker a service coupled with the CityModel that filters the raw position and accelerometer data from the PhoneApp to determine likely pedestrian positions and viewing directions. See public deliverable *D3.4: A Pedestrian Position Tracker* for an in-depth description of this component.

NLU takes text from the ASR (Speech recognition) component and generates an expression in SPACEBOOK's MRL (Meaning Representation Language). NLU also takes the text generated by the NLG unit to help maintain dialog context.

InteractionManager which, based on the information and spatial state of the pedestrian, decides the next dialog move. The interaction manager can be primed into responding to the pedestrian when MRL expressions asynchronously arrive from the NLU component. Likewise events happening in the pedestrian's spatial situation can provoke proactive responses to be issued. In all cases verbal responses to the pedestrian triggered by the interaction manager sending a high-level communication request to the NLG component. Note that we have developed both a statistical interaction manager as well as a more traditional rule-based interaction manager for purposes of bench-marking progress in basic research topic II above.

NLG takes the next dialog move specified in high-level MRL from the InteractionManager. In a strategic generation phase, the NLG module first fleshes this out in more specified MRL that determines the exact content, followed by a tactical generation phase that maps the specified MRL to surface text which is relayed by the PhoneServer to the PhoneApp for voicing on the Phone TTS component or alternatively is relayed to a backend text-to-speech component to generate an audio file which is relayed via the PhoneServer to the PhoneApp to be played as an audio file.

ASR and TTS are performed by off-the-shelf components although it should be noted that we have built a grammar for the ASR component. Access to the city model is via SQL statements from the IM. Queries to the visibility engine are made through XML requests. Textual questions that engage the QA system are sourced from documents drawn from custom web based services.

Extra achievements beyond the work-plan

Use of PhoneApp and PhoneServer in WoZ studies

We decided shortly after kick-off to use a very basic version of our PhoneApp and PhoneServer to assist in Wizard-of-Oz studies. While this required the development of a Wizard-of-Oz interface, the PhoneApp and PhoneServer's use in the WoZ experiments forced them to undergo the kind of rigorous testing that systems such as SPACEBOOK require. We also confirmed that 3G network connectivity in Edinburgh is sufficient to support prototype experiments.

A system-wide meaning representation language (MRL)

Another extra effort was the development of a system-wide meaning representation language for SPACEBOOK. We decided upon adopting a syntax based on *minimum recursion semantics* and we defined a vocabulary capable of representing communication action, pedestrian mental state, the entities and relationships of the city model, spatial relationships and pedestrian physical actions. In other words all representations of meaning in SPACEBOOK are captured in a uniform syntax over what we call SB-MRL(SPACEBOOK meaning representation language).

An early pilot system

A major risk with projects like SpaceBook is that the team fails to achieve a working prototype and that experimental studies with real pedestrians are not able to be performed. In month 11 we conducted some initial pilots with a system that could guide a user to a given goal, using the PhoneApp/PhoneServer architecture.

Dissemination Activities

Project members, based on SpaceBook related activities, have published 8 conference papers (2 in IJCAI, 2 in ENLG, 1 in LBS, 1 in IWSDS, 1 in Interspeech and 1 in SIGDIAL) in 2011 and are slated to publish and present at least 2 papers in 2012. In addition there are several additional submitted papers under review. Finally our member HWU of the SpaceBook consortium proposed "The GRUVE Challenge: Generating Routes under Uncertainty in Virtual Environments" – a shared task for the natural language generation community based around generation of instructions for pedestrians navigating in open-world virtual environment.

We note that WP7, (Dissemination of SPACEBOOK results) does not start until month 16 of the project, and thus no dedicated, coordinated dissemination activity has yet commenced. In our upcoming dissemination period (month 16-18), in addition to witting a number of joint research papers, we intend to publish a series of videos on our public web-site that showcase both our initial component prototypes as well as a video of a running prototype system.

1.3 Expected final results and potential impact and use

In late Spring and early Summer of 2012 we will conduct field tests with real pedestrians in Edinburgh. As our prototype is stabilized and then modified to maximize objective navigation and exploration evaluation metrics, we have the potential to revolutionize the way in which people navigate and explore the city.

1.4 Project website

SPACEBOOK's public website address is `http://spacebook-project.eu`.