Programming by Voice in Scratch

Ramaraju Rudraraju, Srinivasa Datla, Avishek Banerjee, Mandar Sudame
University of Alabama at Birmingham
{ramaraju, dsnraju, avi1987, mandars}@uab.edu

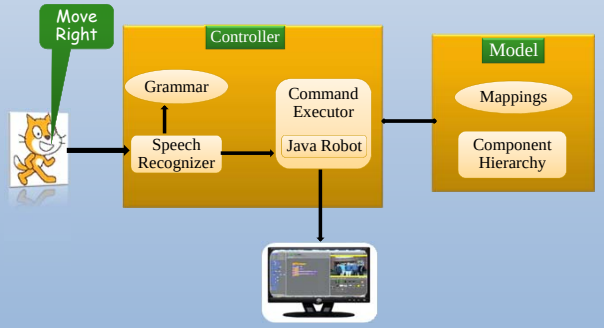
Jeff Gray
University of Alabama
gray@cs.ua.edu

This research was sponsored by a Google Research Award.

Motivation

This poster introduces Myna, which assists those with restricted limb mobility in learning programming skills through a voice-driven interface. Specifically, Myna supports programming by voice with Scratch. Although the native Scratch environment allows users to create a program by arranging graphical blocks logically, Scratch is dependent on the Windows/Icons/Mouse/Pointer interface that requires dexterity in using a mouse/keyboard (limiting those with physical disabilities). Myna processes voice commands from the user, interprets those commands according to a Scratch-based grammar, and simulates synonymous actions of a mouse/keyboard in Scratch. The result is an environment that assists those with disabilities (in particular, children) in experiencing the joy of programming.

Research Overview



Any voice driven user interface faces the challenge of mimicking the actions of keyboard input and mouse clicks without adding any significant overhead. To overcome the challenges of mapping the WIMP metaphor to a voice-driven interface, we investigated and developed several forms of interaction, which are driven by a grammar for a speech recognition engine that is translated into commands performed by an instance of the Java Robot Class (which provides programmatic control of the mouse and keyboard).

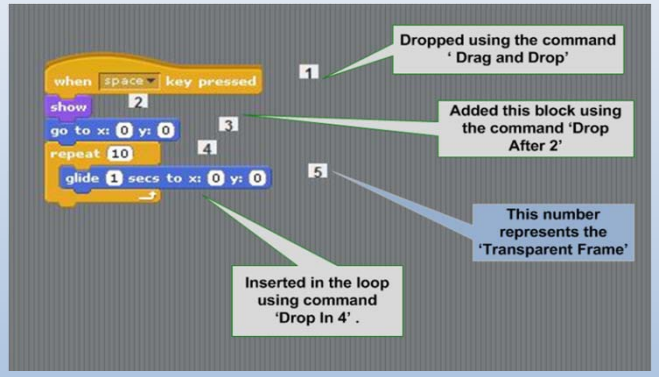
Features

Types of Navigation

- **Drag and Drop Navigation:** Mimics the sequence of click-drag-release actions of a mouse
- **Continuous Navigation:** Mimics the continuous movements of the mouse cursor
- **Using Transparent Frames:** Makes use of small labels that mark every component in the script window

Macro Commands

- **Drop After :** Enables to drop blocks after the specified block
- **Drop In:** Enables to drop the block inside the block specified
- **Drop Before:** Drops the block before the specified block

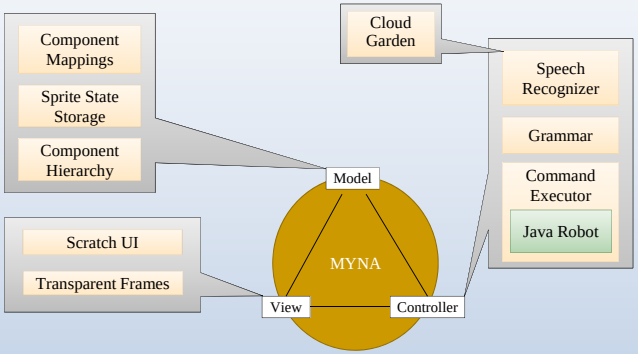


Contribution

Myna is a tool that runs alongside Scratch and provides a voice interface that can be used to specify programming tasks. The project has investigated several common interaction patterns that can be imitated through shortcuts to minimize the amount of speaking required (vocal strain is an important issue in voice-driven programming environments such that economy of expression is vital). The project unites ideas of human-computer interaction with Computer Science education to provide an assistive environment for teaching computational thinking using Scratch. In our implementation, we also provided a way to extend the support for other IPEs that may be targeted in the future. In particular, we modularized the various grammars that define the commands to be issued by voice, which will allow extensions to be added more easily.

Myna Architecture

An important design consideration was the requirement that Myna exist outside of the Scratch implementation (i.e., Myna does not require any source code modification to the Squeak implementation of Scratch). This requirement posed several challenges (e.g., adding transparent frames as visual overlays to Scratch) that influenced our design.



The design choice to build Myna outside of the Scratch source code was made to provide flexibility for adapting Myna to future versions of Scratch. Another design goal was to use an architecture that is robust enough to accommodate future changes in Scratch's user interface. To assist in separating the dependencies between the user interface and the core functionality, we applied the Model View Controller (MVC) design pattern.

Conclusion & Future Work

- Using the primary capabilities of our voice-driven application, Scratch can be used in a manner that overcomes the physical challenges of WIMP to provide a path for those with disabilities to have access to Scratch as a learning environment.
- Related work in this area has not emerged in the specific context of applying programming by voice to IPEs. Existing speech recognition systems are not tailored to tools like Scratch, which required the customizations described in this poster.
- Video examples at `http://www.youtube.com/user/Teammyna`
- The current Myna implementation has several limitations that we are working toward as future work.
- The next phase of research involves collaboration with *United Cerebral Palsy of Birmingham* (UCP) for human subject testing