

KatBot

B0: Ashika Koganti, Abha Agrawal, Jade Traiger





Use Case

Storytelling robot that interacts with people to aid in language and reading comprehension

- Merging AI with educational tools
- Target Audience: elementary school age children
- Child-friendly user experience



Storytelling Companion Robot by the MIT Personal Robots Group

<https://robotic.media.mit.edu/portfolio/storytelling-companion/>



Requirements

Analog & Circuits

- Microphone input and speaker output circuits
- Audio filtering and gain stages
- ADC for audio input

Software Systems

- Template creation & story generation
- User input part of speech tagging
- Synonym detection

Signal Processing

- Audio data processing (speech processing and text to speech)

Robotics

- 2 DOF Robot Arm Motion Planning
- Face Display for robot's eyes
- Text display to follow along with story (reading comprehension development)



Challenges and Approaches

- **Finding data to train a story generator**
 - Story datasets
 - Part of speech and synonym detection datasets
- **Create a cohesive story**
 - Create story framework prior to use
 - Algorithm fills in slots in accordance with user input
- **Collecting clear audio from the user**
 - Ideally, microphone on robot
- **Implementing a speech recognition system for simple phrases**
 - Selecting an appropriate speech recognition package
- **Robotic expressions and gestures that compliment the story**
 - Pre-programmed based on story order
- **“Real time”**



Solution Approach

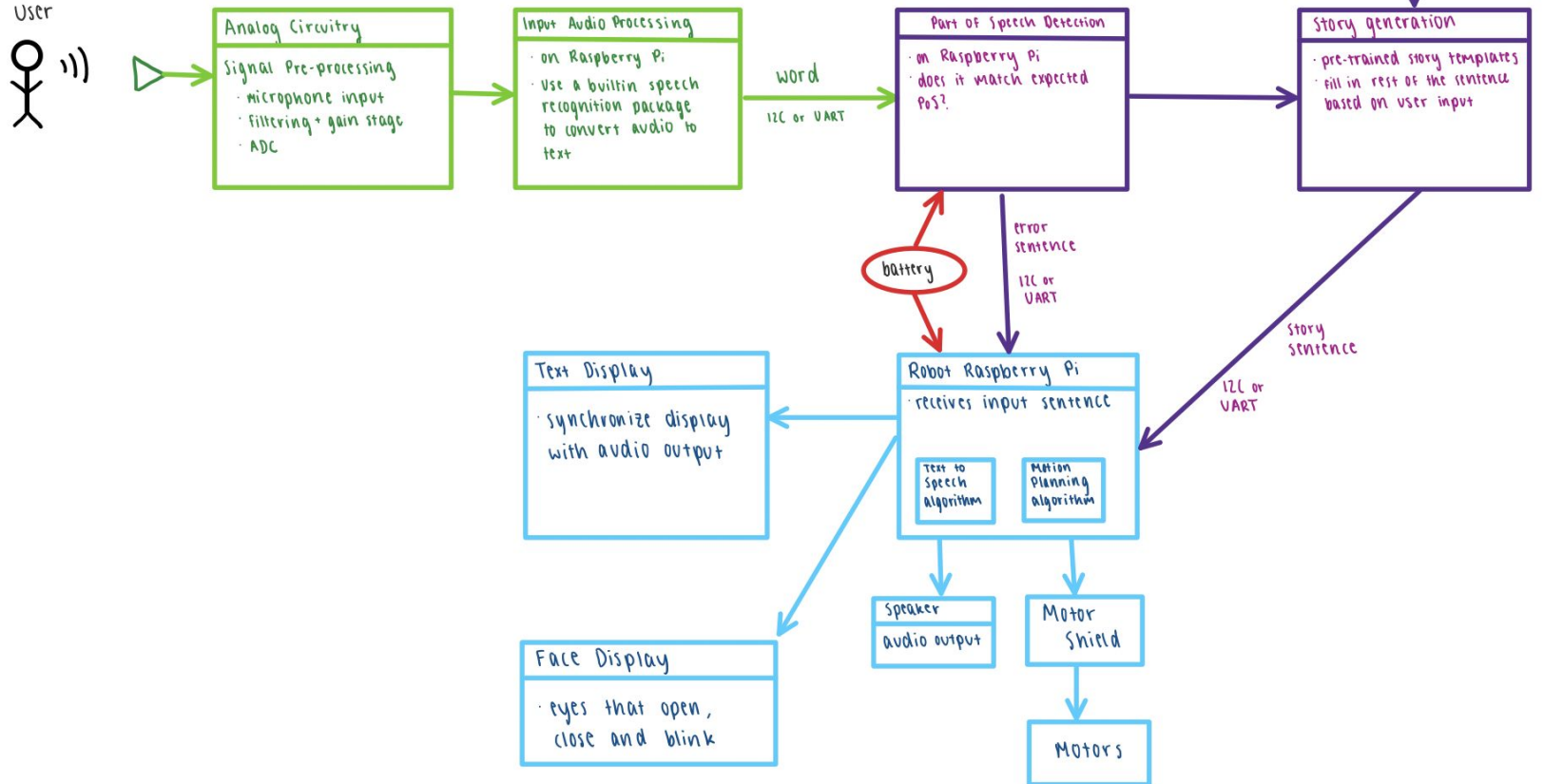
- Audio Signal Processing
 - Analog circuits developed with speech processing in mind
 - ADC output compatible with RPi input
- Speech Processing Package
 - Offline speech processing on RPi
 - Low latency between user input and word decoding
- Machine Learning Story Generation
 - BERT - bidirectionally trained transformer
 - Part of speech tagging, other entities (e.g. person, date)
 - fill in the blank, MadLib guessing
 - LSTMs for word prediction
 - Synonym with supervised learning



Solution Approach

- Text to Speech Package
 - Offline text to speech generation on a RPi
 - Voice is friendly to target user
- Robot
 - Custom-made robot with acrylic laser-cut frame and either 3-d printed or cloth shell
 - 2 Raspis for I/O to parallelize work and help system feel “real-time” to user
 - 2 2-DOF robot arms made from 2 motors and laser-cut acrylic segments
 - 2 LCD screens for Face and Text Displays

System Diagram



Testing, Verification, and Metrics



Description	Goal	Verification Method
Part of Speech Tagging	90% accuracy	SW Testing - Test Dataset
Synonym Detection	90% accuracy	SW Testing - Test Dataset
Speech Processing Accuracy	15% Word Error Rate	Measure decoding errors
Speech Processing Latency	4 - 6 sec	Time user i/p to speech o/p
Story Cohesion	Match NAP score of MadLibs	Compute NAP score of each story
Story Enjoyment	???	User Survey



Risk Management

1. Machine Learning Stories

- Template generation - use pre-programmed templates
- User input inaccuracy - minimize user input

2. Audio Processing

- Typed user input

3. Robot

- No text display
- Computer/App based



Tasks

- Audio input and output (Jade)
 - Choose package
 - Build microphone circuit
 - Speech Recognition
 - Text to speech
- Story generation algorithm (Ashika)
 - Part of speech detection
 - Synonym detection
 - Template Making
 - Fill in the blank model
- Robot (Abha)
 - Raspi Interfacing with I/O
 - Communication Protocols
 - Face and Text Displays
 - Robot arms for gestures
 - Custom robot shell and appearance
- Integration (Everyone)
 - Audio input to ML algorithm
 - Robot displays text from ML algorithm
 - Robot text to speech
 - Gesture evaluation
- Evaluation & Improvement (Everyone)
 - Whole system testing
 - Design evaluation
 - Improvements
 - Final system testing
 - Prepare for Final Presentation



Stretch

1. Sentiment analysis
 - Doable with BERT
 - Map robot gestures to emotions
2. Direction Finding
 - Microphone array processing
 - Robot will turn to face user