ASLBoT

Assisted Sign Language Bot Translator

Group #14

Gustavo Camero - CpE

Luis Hurtado - EE

Michael Loyd - EE

Jared Spinks - CpE

Assigned Professor

Dr. Samuel Richie

Advisor/Client

Dr. Chung Yong Chan

Motivation

- Member works at a clinic
  - Noticed a lack of efficient translation
  - Looking for standardized language interpretation
  - No support for sign language
- Classroom Setting
  - Interpretation Services are not adequate
  - Use of ASL Translators is expensive

Motivation (cont.)

- No current marketable product
  - Sign language machine translation still in research phase
  - No open-source large corpora for English to ASL
  - No standardized training framework

Goals & Objectives

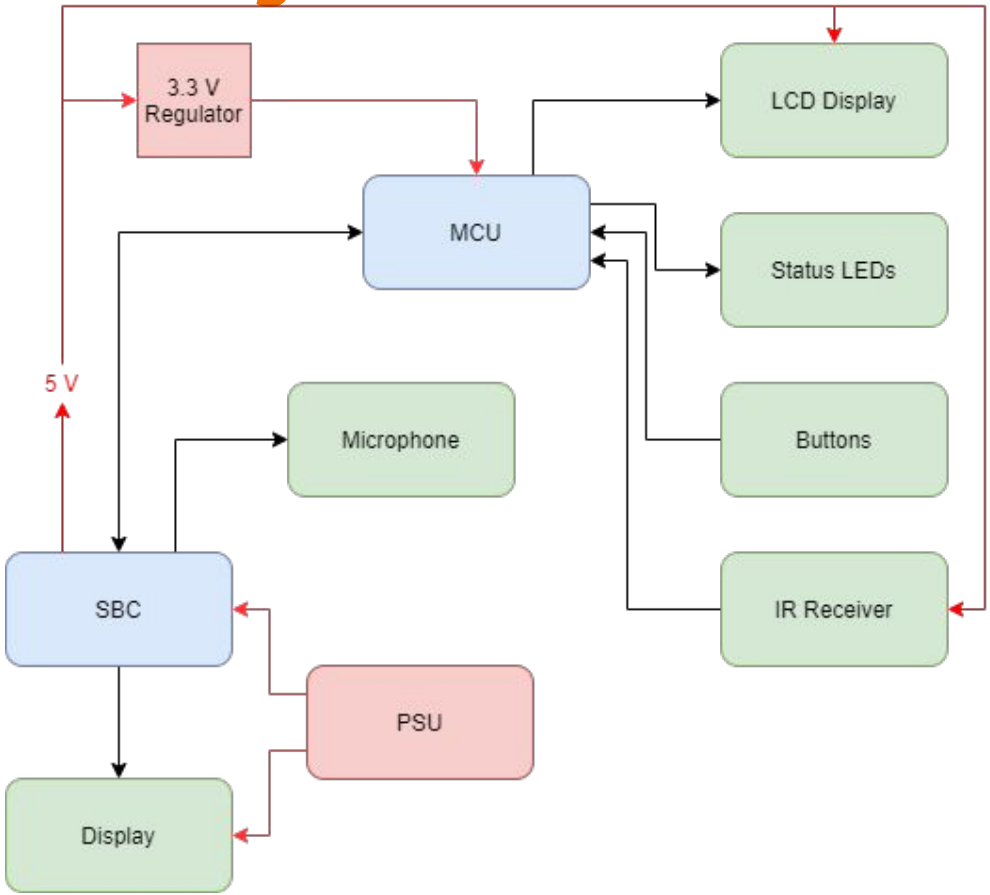
- Translate English speech into American Sign Language.
- Use small corpora to achieve ASL translations using a training framework
- Perform speech-to-text and speech-to-ASL language translations.
- Indicate various statuses of the device via LEDs.
- Capture audio at frequencies around natural human pitches that can be translated into text via speech-to-text algorithms.
- Display responsive high-quality animations of sign language gestures.
- Receive manual inputs from the user such as start/stop recording.

Requirement Specifications

Component	Parameter	Design Specification
Microphone	Frequency Response	300 Hz to 3kHz
Memory	Size	32 GB
Power Supply	Power	32 W
System	Translation Accuracy	20% (BLEU Average)
System	Cost	$500
System	Response	< 5 seconds
Vocabulary	Model Complexity	> 50 words
Neural Machine Translation	Size	< 1000 Entries

Hardware Components

Overall Block Diagram

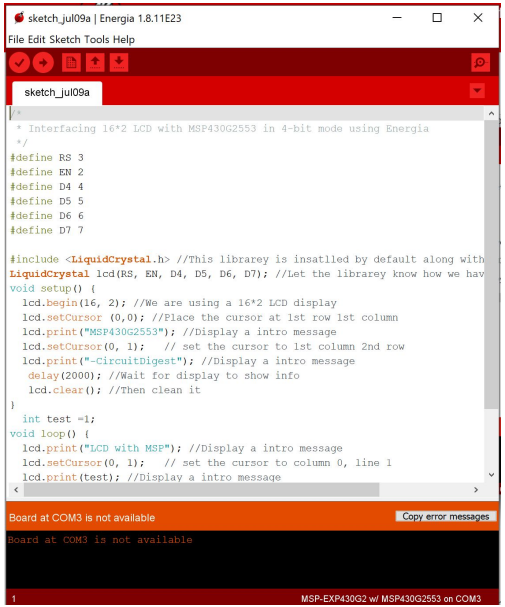


MCU Selection

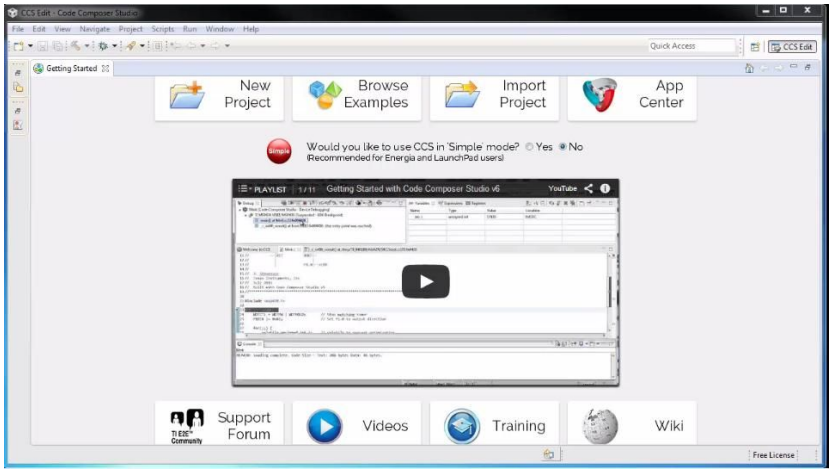
	GPIO Pins	ADCs	Communication Protocols	Cost
STM32L162ZDT6	115	12-bit/ 40 Channels	I^2C, I^2S, SPI, USART, USB	$9.80
STM32L151RDT6TR	51	12-bit/ 40 channels	I^2C, I^2S, SPI, USART, USB	$6.85
MSP430FR6922	**46**	**12-bit/ 8 channels**	**I^2C, SPI, UART**	**$5.33**
MSP430FR6989	83	12-bit/ 16 channels	I^2C, SPI, UART	$7.86

MCU Development Environment

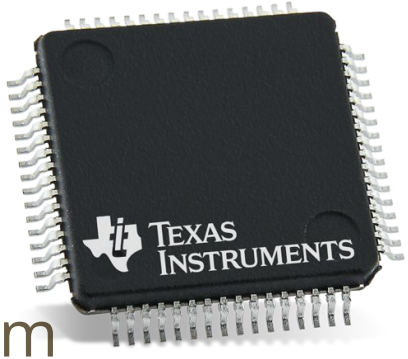
Energia



Code Composer Studio

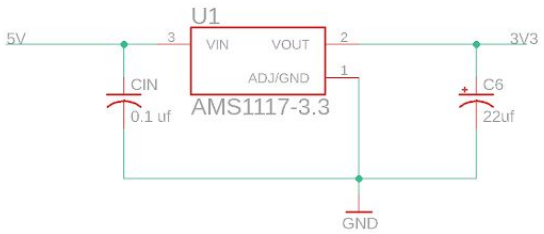


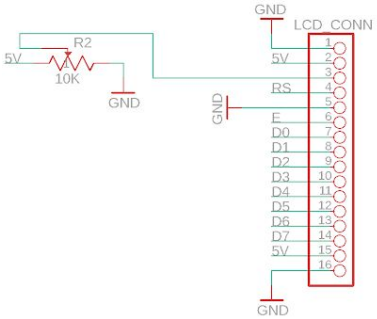
MCU Functionality

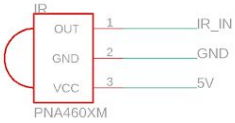


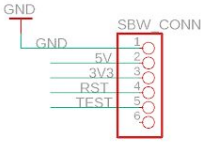
- Uses LEDs to indicate overall status of system
- Utilizes LCD to indicate recording status
- Communicates with SBC to send control signal and receive updates
- IR sensor for wake functionality
- On-board buttons to start and stop recording and reset

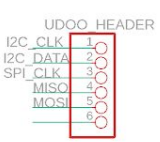
Hardware Diagram

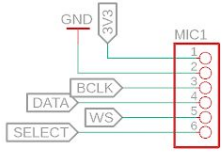


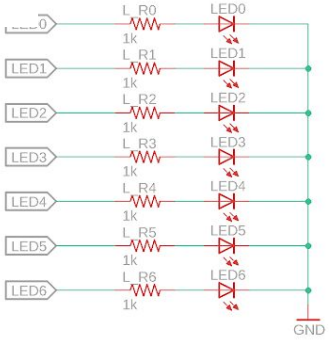


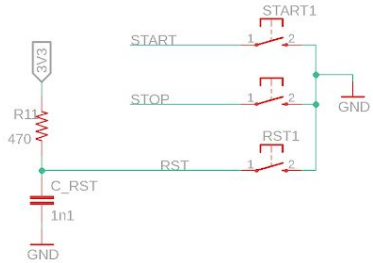


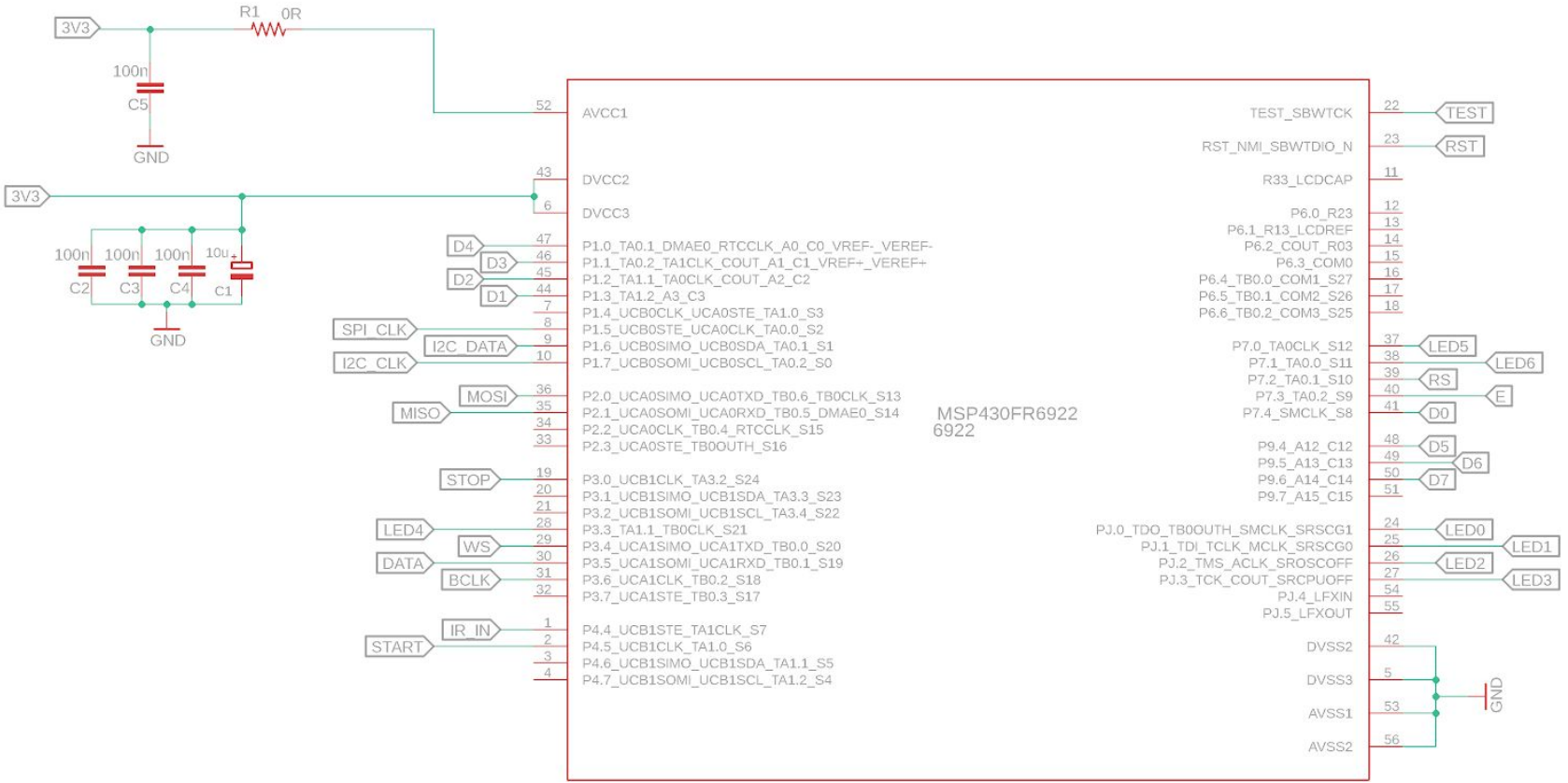




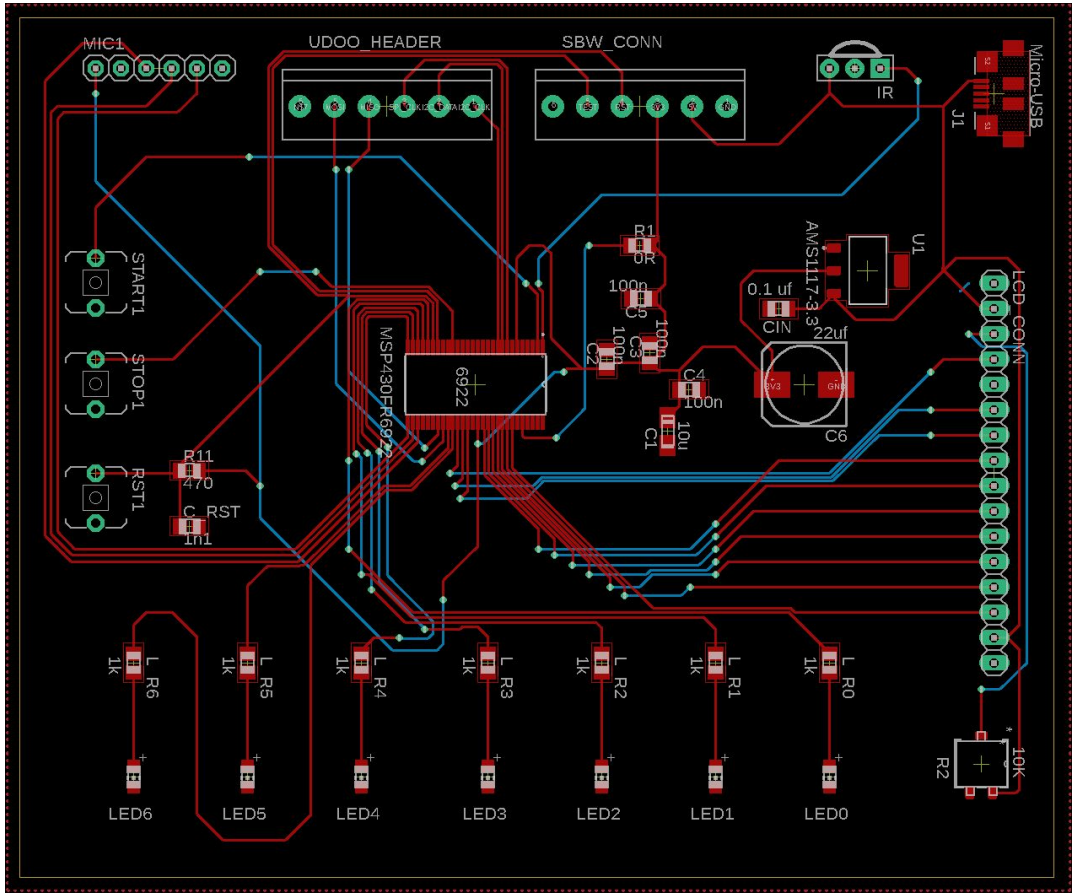







3V3
R1
0R
100n
C5
GND
3V3
100n
100n
100n
10u
C2
C3
C4
C1
GND
MSP430FR6922
6922
52
AVCC1
43
DVCC2
6
DVCC3
47
P1.0_TA0.1_DMAE0_RTCCLK_A0_C0_VREF-_VEREF-
46
P1.1_TA0.2_TA1CLK_COUT_A1_C1_VREF+_VEREF+
45
P1.2_TA1.1_TA0CLK_COUT_A2_C2
44
P1.3_TA1.2_A3_C3
7
P1.4_UCB0CLK_UCA0STE_TA1.0_S3
8
P1.5_UCB0STE_UCA0CLK_TA0.0_S2
9
P1.6_UCB0SIMO_UCB0SDA_TA0.1_S1
10
P1.7_UCB0SOMI_UCB0SCL_TA0.2_S0
36
P2.0_UCA0SIMO_UCA0TXD_TB0.6_TB0CLK_S13
35
P2.1_UCA0SOMI_UCA0RXD_TB0.5_DMAE0_S14
34
P2.2_UCA0CLK_TB0.4_RTCCLK_S15
33
P2.3_UCA0STE_TB0OUTH_S16
19
P3.0_UCB1CLK_TA3.2_S24
20
P3.1_UCB1SIMO_UCB1SDA_TA3.3_S23
21
P3.2_UCB1SOMI_UCB1SCL_TA3.4_S22
28
P3.3_TA1.1_TB0CLK_S21
29
P3.4_UCA1SIMO_UCA1TXD_TB0.0_S20
30
P3.5_UCA1SOMI_UCA1RXD_TB0.1_S19
31
P3.6_UCA1CLK_TB0.2_S18
32
P3.7_UCA1STE_TB0.3_S17
1
P4.4_UCB1STE_TA1CLK_S7
2
P4.5_UCB1CLK_TA1.0_S6
3
P4.6_UCB1SIMO_UCB1SDA_TA1.1_S5
4
P4.7_UCB1SOMI_UCB1SCL_TA1.2_S4
D4
D3
D2
D1
SPI_CLK
I2C_DATA
I2C_CLK
MOSI
MISO
STOP
LED4
WS
DATA
BCLK
IR_IN
START
22
TEST_SBWTCK
23
RST_NMI_SBWTDIO_N
11
R33_LCDCAP
12
P6.0_R23
13
P6.1_R13_LCDREF
14
P6.2_COUT_R03
15
P6.3_COM0
16
P6.4_TB0.0_COM1_S27
17
P6.5_TB0.1_COM2_S26
18
P6.6_TB0.2_COM3_S25
37
P7.0_TA0CLK_S12
38
P7.1_TA0.0_S11
39
P7.2_TA0.1_S10
40
P7.3_TA0.2_S9
41
P7.4_SMCLK_S8
48
P9.4_A12_C12
49
P9.5_A13_C13
50
P9.6_A14_C14
51
P9.7_A15_C15
24
PJ.0_TDO_TB0OUTH_SMCLK_SRSCG1
25
PJ.1_TDI_TCLK_MCLK_SRSCG0
26
PJ.2_TMS_ACLK_SROSCOFF
27
PJ.3_TCK_COUT_SRCPUOFF
54
PJ.4_LFXIN
55
PJ.5_LFXOUT
42
DVSS2
5
DVSS3
53
AVSS1
56
AVSS2
TEST
RST
LED5
LED6
RS
E
D0
D5
D6
D7
LED0
LED1
LED2
LED3
GND

Board Layout



Microphone Selection

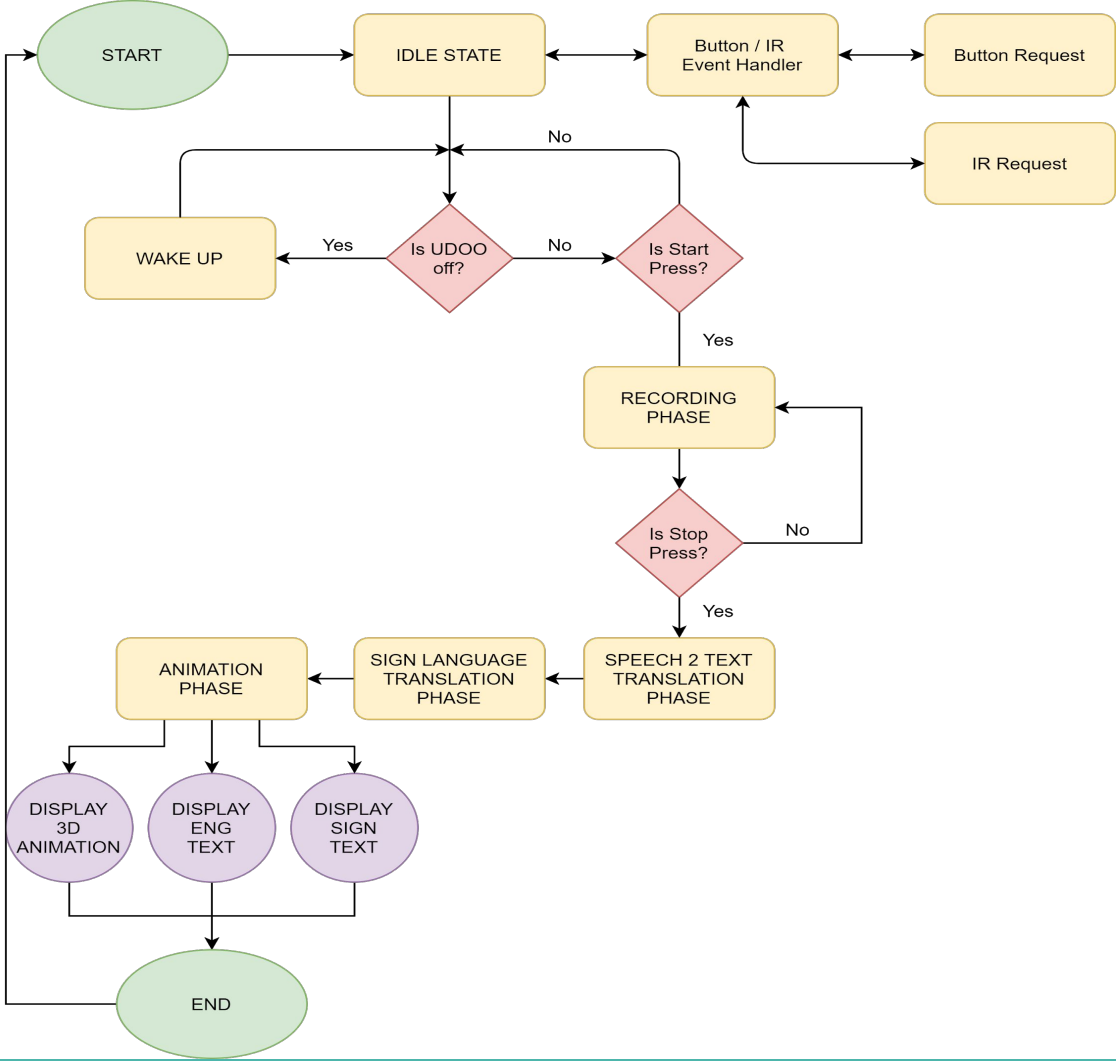
	Frequency Range	Output	Interface	Cost
Mini USB Microphone	100 Hz - 10 kHz	Digital	USB	$4.95
SparkFun MEMS Microphone (INMP401/ADMP401)	100Hz - 15 kHz	Analog	ADC	$10.95
MEMS Microphone (SPH0645LM4H)	50 Hz - 15 kHz	Digital	I^2S	$6.95
Blue Snowball	**40Hz to 18kHz**	**Digital**	**USB**	**$49.99**

SBC Selection

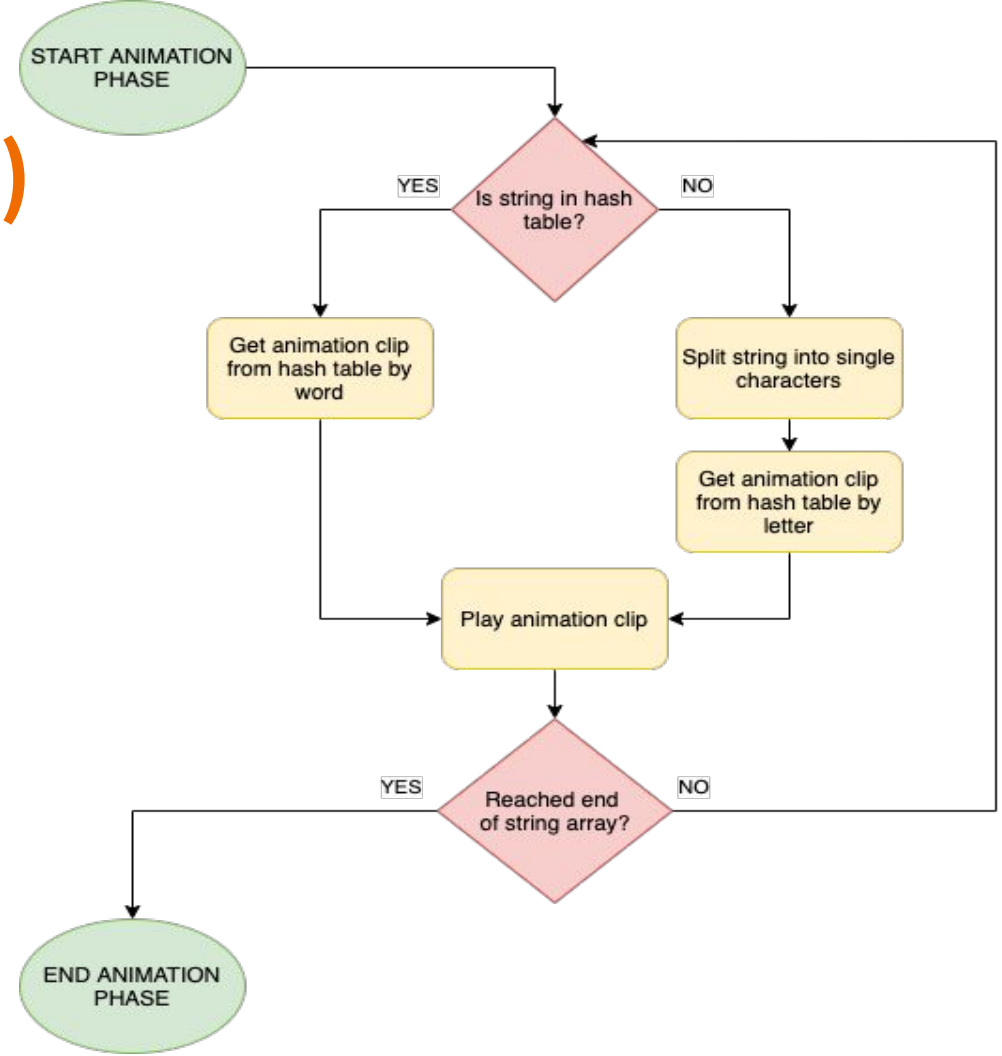
	Architecture	Communication Protocols	OS	RAM	Memory	Cost
NVIDIA Jetson Nano	ARM64	I^2C, SPI, UART	Linux	4 GB	External Only	$99.00
ASUS TinkerBoard	ARM64	I^2C, SPI, UART, I^2S	Linux	2 GB	External Only	$70.00
UDOO x86 Advanced Plus	x86_64	I^2C, SPI, UART	Linux Windows	4 GB	32 GB	$176.00

Software Components

Software Diagram



Software Diagram (Cont.)



Machine Translations

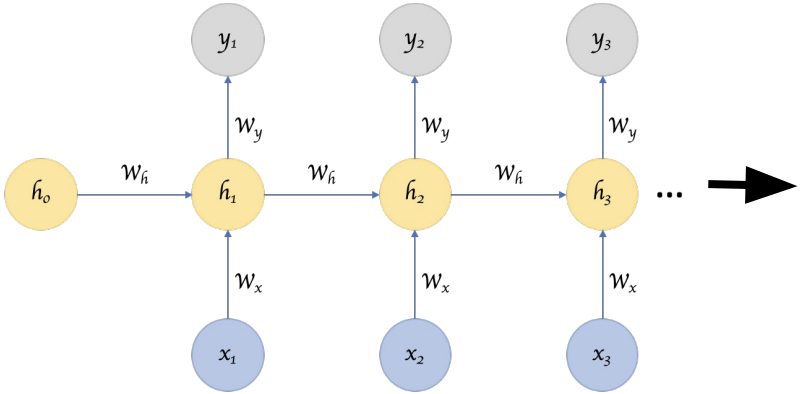
Machine Translation

Rule-Based MT (RBMT)	Statistical MT (SMT)	Neural MT (NMT)
• Large database of linguistic rules and millions of bilingual dictionaries. • For the use of RBMT in text to ASL text translation, a massive database with the rules and resources will need to be created.	• Statistical models that have parameters based on the analysis of monolingual and bilingual corpora. • Relies heavily on existing corpora, minimum of 2 million words. • CPU-intensive	• Uses neural networks with training sets to tune weights. • Examines entire sentence before translating, maintains the context of the sentence. • Training requires GPU • Large Training sets

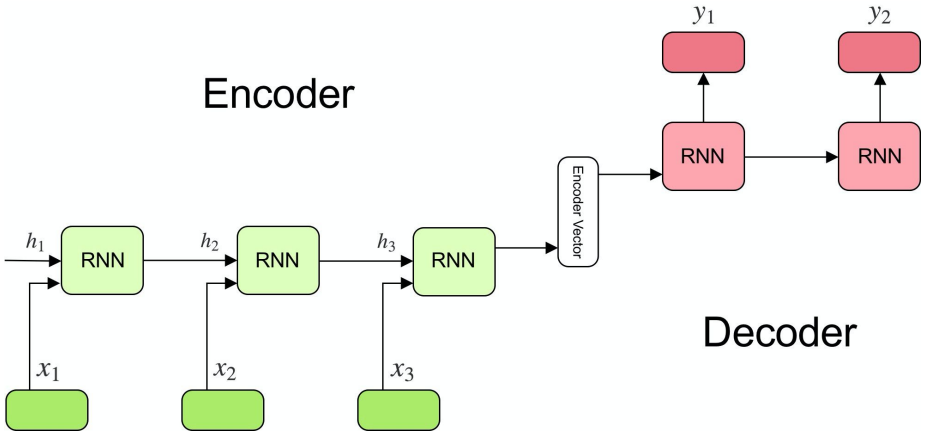
OpenNMT

Open source and user support lacked in other systems	**Provides Pytorch implementation**	**Based on seq2seq models that allow robust translation**

Recurrent Neural Network



Sequence-to-sequence



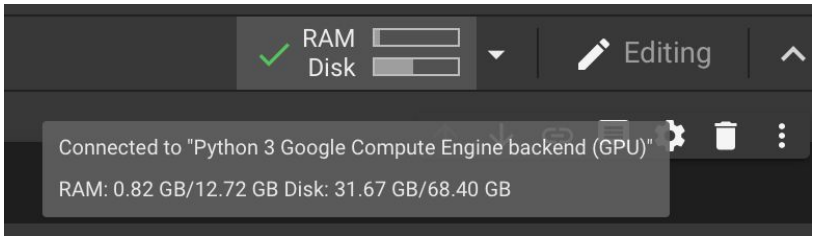
RNNshttps://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce
Seq-seq https://towardsdatascience.com/understanding-encoder-decoder-sequence-to-sequence-model-679e04af4346

OpenNMT

NMT Issue	Requires GPU for training

Google Colaboratory (Colab)

- Allows users to write and execute Python code on google's servers
- Free available GPU for processing
- Zero Configuration required



colab

Training OpenNMT

Requirements:

- Source (English)
- Target (Sign language Gloss)
- Source Validation
- Target Validation

Source	Target	Source Validation	Target Validation
I know sign language	ME KNOW SIGN	I want to know sign language	ME WANT KNOW SIGN LANGUAGE

Training OpenNMT - Issue

NMT Issue	Requires large Corpora

Two approach system:
● Specific Domain ● Redundancy

Training OpenNMT - Approach 1

Approach #1 Specific Domain:

- Limits variety of words used
- Corpora based on daily school

hurry we need to go now
on campus
I need to write this paper tonight
this project is due tomorrow
my lecture starts in five minutes
I am going to college this semester
what classes are you taking
I took that class last semester
I wanted to go to the student union
did you finish the paper last night
I was absent from class yesterday
Was the homework supposed to be done individually
I am making new friends in college
I made a friend yesterday
Do you want to be friends
there are a few ways to get to that building
why is the campus so large
the engineering building is very big
I need help because I can not find this classroom
I can't find this classroom
I need help making these signs
How many buildings are there on campus
How much did you write for the paper

Training OpenNMT - Approach 1 Results

- Improvement in translations
- Still very limited translations

```
SENT 1: ['where', 'can', 'I', 'find', 'the', 'classroom', 'building']
PRED 1: ME NEED HELP MAKE SIGN SIGN
PRED SCORE: -0.0293
```

```
SENT 1: ['I', 'need', 'help', 'finding', 'the', 'classroom']
PRED 1: ME NEED HELP MAKE SIGN SIGN
PRED SCORE: -0.0002
```

Training OpenNMT - Approach 2

Approach #2 Specific Domain + Redundancy:

- Continue to limit variety of words used
- Breakdown existing sentences
- Introduce similar sentences

the engineering building is big
the engineering building
engineering building
building is big
engineering building is big

when did you finish your homework
when did you finish
did you finish your homework
finish your homework

Training OpenNMT - Approach 2 Results

Model now better suited to translate English to Gloss

```
SENT 1: ['where', 'can', 'I', 'find', 'the', 'classroom', 'building']
PRED 1: ME NEED HELP MAKE SIGN SIGN
PRED SCORE: -0.0293
```

First approach

```
SENT 1: ['where', 'can', 'I', 'find', 'the', 'classroom', 'building']
PRED 1: WHERE ME CAN FIND CLASSROOM
PRED SCORE: -0.0006
```

+Second approach

```
SENT 1: ['I', 'do', 'understand', 'the', 'homework']
PRED 1: ME NOT WANT KNOW
PRED SCORE: -0.2670
```

First approach

```
SENT 1: ['I', 'do', 'understand', 'the', 'homework']
PRED 1: ME UNDERSTAND HOMEWORK
PRED SCORE: -0.0006
```

+Second approach

Training OpenNMT - Approach 2 Results

The model may still struggle with some translations:

```
SENT 1: ['I', 'want', 'to', 'make', 'friends', 'this', 'semester']
PRED 1: ME WANT MAKE FRIEND AT COLLEGE
PRED SCORE: -0.0245
```

```
SENT 1: ['I', 'need', 'help', 'with', 'the', 'homework', 'because', 'it', 'is', 'difficult']
PRED 1: ME NEED HELP BECAUSE CANNOT FIND CLASSROOM
PRED SCORE: -0.0050
```

BLEU Score

Benefits:

- Quick calculation
- Simple score calculation
- Language does not matter
- Widely adopted (documentation available)

Problems:

- Shorter sentences (even if accurate) get close to zero scores
- Score changes depending on the reference file

```
BLEU = 13.81 62.0/30.2/14.1/1.6
```

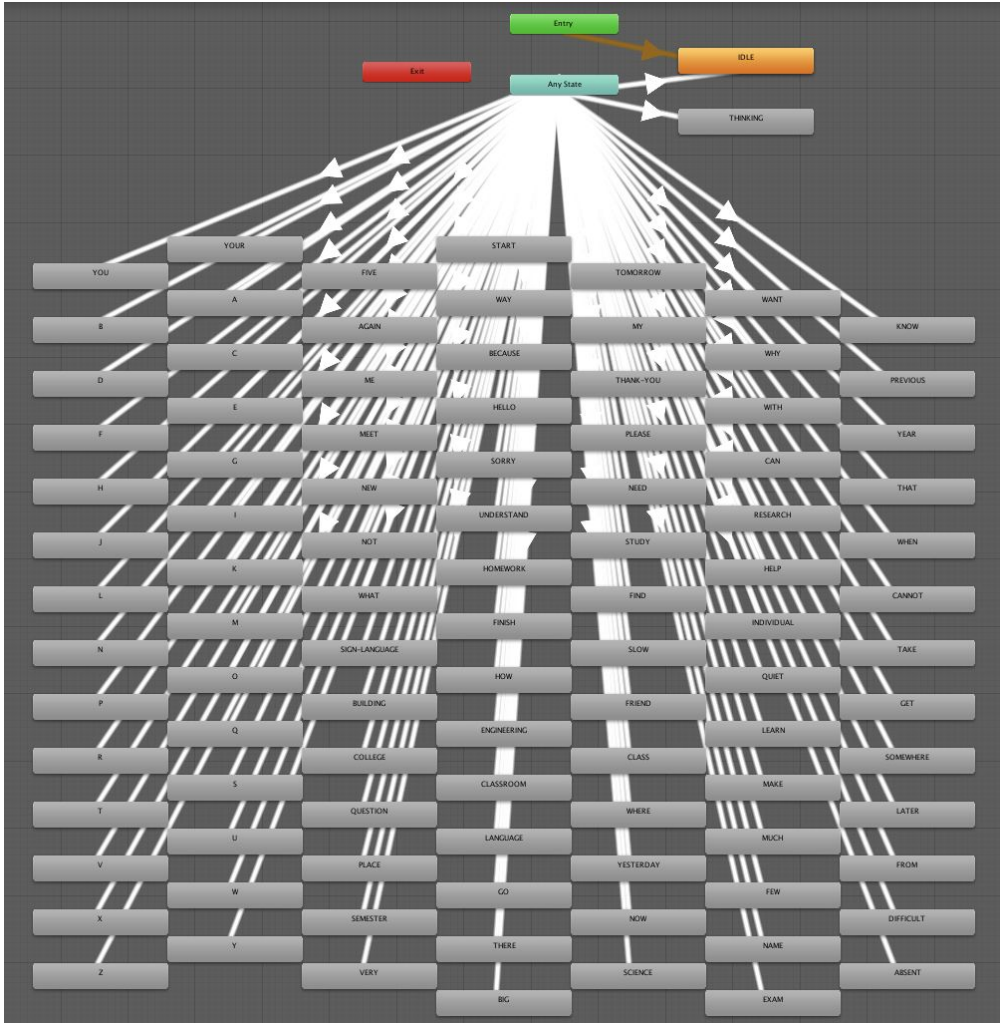
Game Engine

- Functions of the game engine:
  - Display animations
  - Integration of text-to-speech, speech-to-text, and (possible) audio-to-WAV API

	Minimum RAM	Compatible with Linux-based ARM64	Compatible with x86_64	Community Support for Programming
Unity	**None**	**No**	**Yes**	**Full**
Unreal Engine 4	8 GB	Yes	Yes	Limited
Godot	4 GB	No	Yes	None

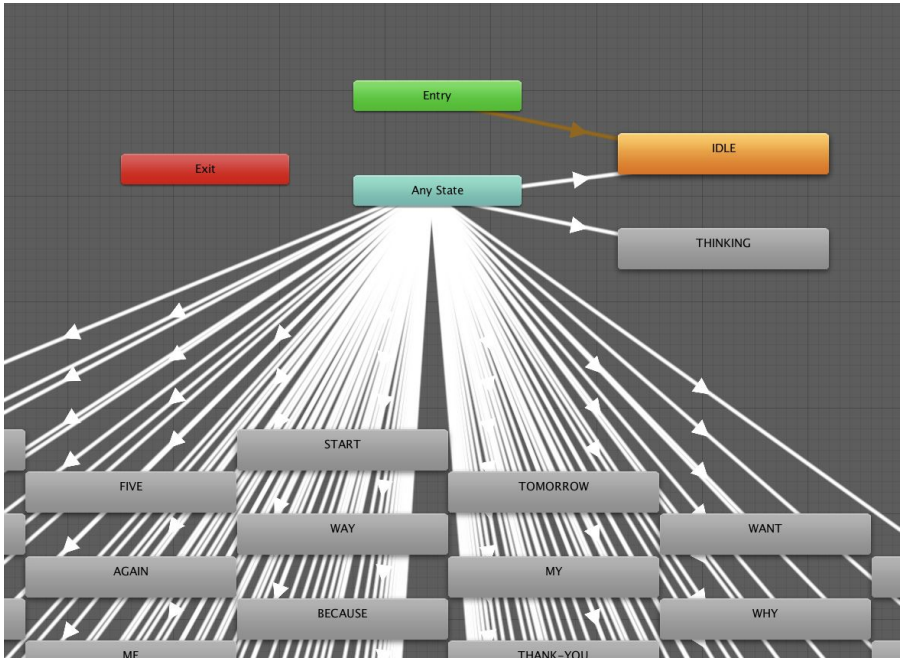
Video Rendering

- Dynamic state machine
- Hash Table used to map word as string to corresponding animation
- Trigger parameter used to start an animation



Video Rendering

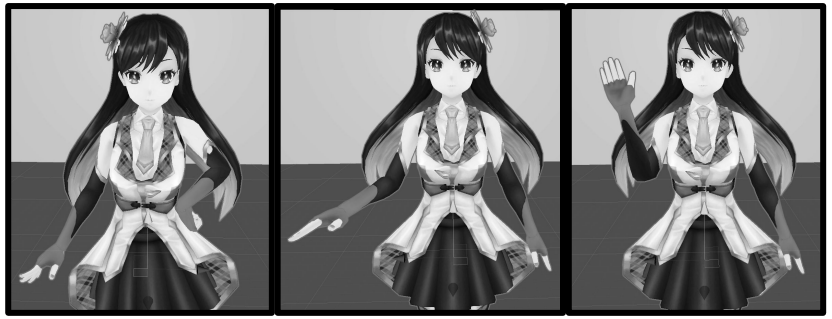
- Dynamic state machine
- Hash Table used to map word as string to corresponding animation
- Trigger parameter used to start an animation



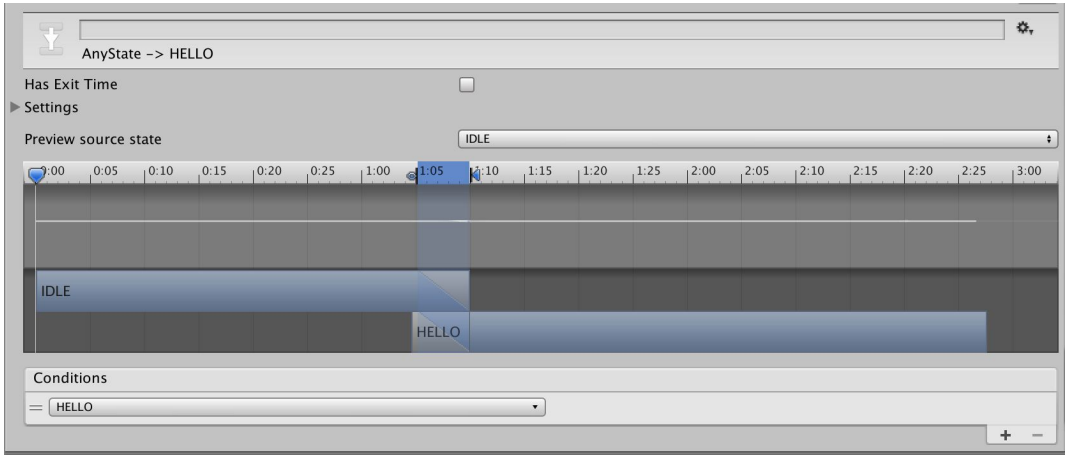
↑ Sample trigger parameter.

Video Rendering

- Dynamic state machine
- Hash Table used to map word as string to corresponding animation
- Trigger parameter used to start an animation



“IDLE” → <transition> → “HELLO”

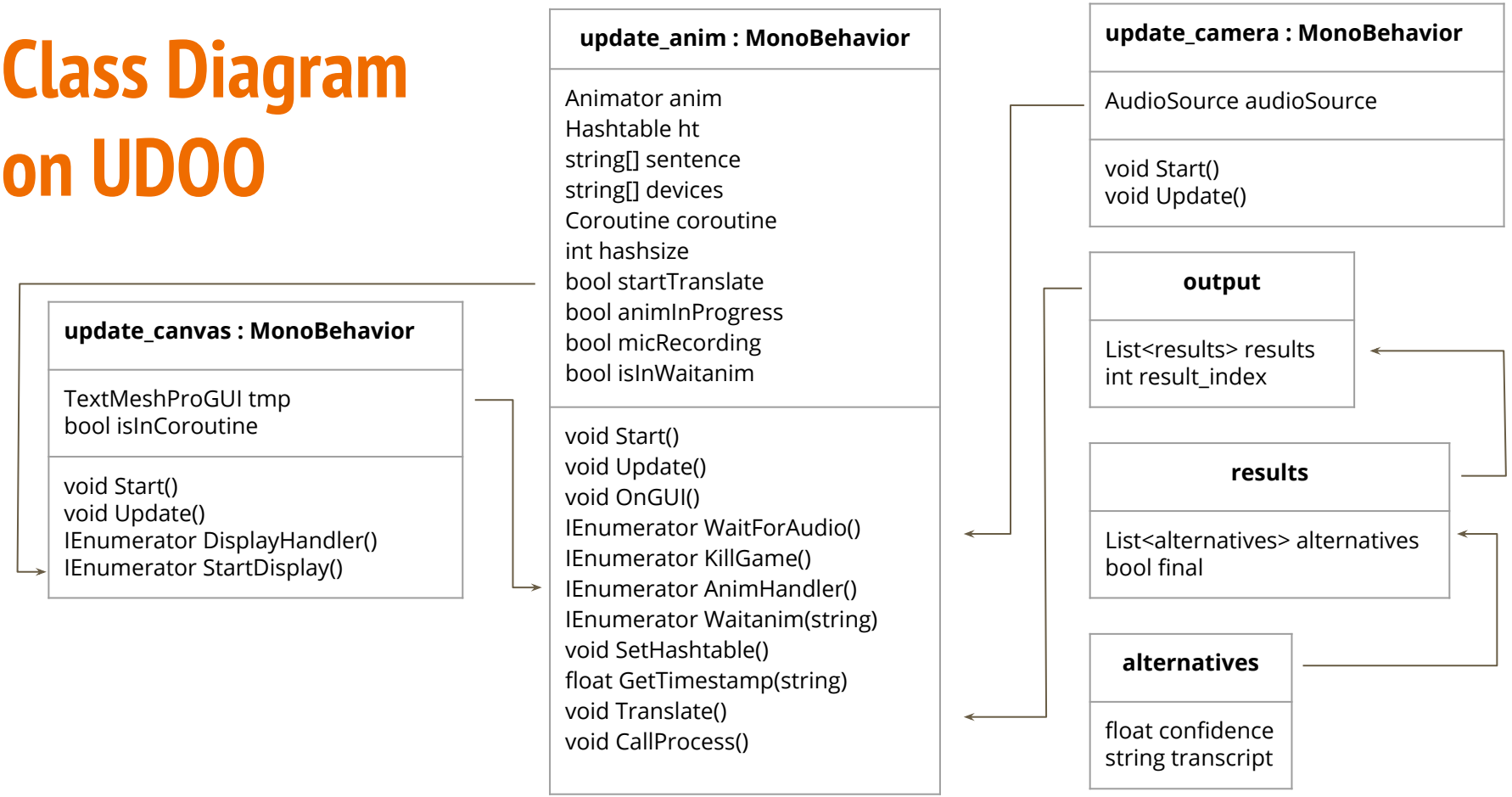


Used under license. © SapphiArt Inc.

Extra Internal Processes

- Vocabulary to Hashtable
  - To map each word to an animation.
- Microphone Module within Unity
  - To capture audio sounds from the microphone.
- Watson API request call for Speech to Text service
  - To generate a Speech to Text translation and store it in a JSON file.
- JSON Parsing
  - To transfer the translation from JSON to text file.
- OpenNMT Translation Call
  - To translate the translation from English to ASL.

Class Diagram on UDOO



Challenges & Issues - Hardware

- SBC architecture compatibility (**Solved**)
- FPGA - difficulty for soldering and programming given timeframe (**Solved**)
- USB Port (**Solved**)
- I^2S compatibility for microphone (**Solved**)
- Lack of 3D printed enclosure
- Limited disk drive size

Challenges & Issues - Software

- Granularity-based OpenNMT (trial and error) (**Solved**)
- Bleu Score 3-/4- gram calculation (short sentences)
- Serial Communication between Udoo and Arduino (**Solved**)
- Utilization of command prompt within Unity (**Solved**)
- Finding compatible microphone software to interface with Unity (**Solved**)

Administrative

Work Distribution

Member	Embedded	Machine Translation	Graphics Rendering	Integration
Gustavo Camero	Secondary	Secondary	Secondary	**Primary**
Luis Hurtado	Secondary	**Primary**		Secondary
Michael Loyd		Secondary	**Primary**	Secondary
Jared Spinks	**Primary**			Secondary

Proposed Project Budget

Part	Supplier	Unit Cost	Quantity	Cost
NVIDIA JETSON	NVIDIA	$95.96	1	$95.96
Altera Cyclone IV	UDOO	$8.90	1	$8.90
MSP430FR6989	Mouser	$7.86	1	$7.86
MEMS Microphone SPH0645LM4H	Adafruit	$6.95	1	$6.95
IR Receiver TSOP38238	Adafruit	$1.95	1	$1.95
WiFi Module	Amazon	$15.00	1	$15.00
Display Monitor	Amazon	$59.99	1	$59.99
PCB	JLCPCB	$50	1	$50
Miscellaneous	--	--	--	$200
Total	Supplier			$446.61

Actual Project Budget

Part	Supplier	Unit Cost	Quantity	Cost
UDOO x86 Advanced Plus	UDOO	$176.00	1	$176.00
Power Supply	UDOO	$8.90	1	$8.90
PCB/MCU	JLCPCB	$19.07	1	$19.07
IR Receiver	Adafruit	$1.95	1	$1.95
Display	Amazon	$59.99	1	---
Snowball Microphone	Blue Designs	$49.99	1	---
LCD	Amazon	$5.99	1	---
IR Remote	Adafruit	$4.95	1	--
Miscellaneous	---	---	---	$132.93
Total		$326.84		$338.35

Questions?