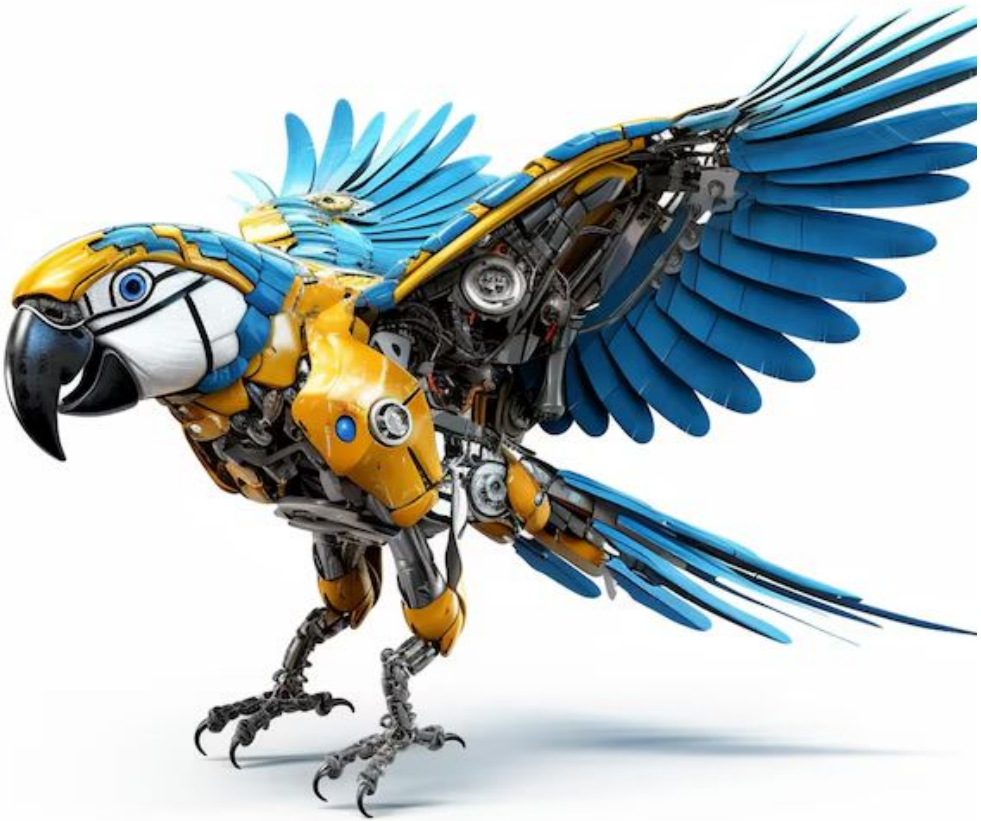


C.A.P.E.R.
Final Presentation

Group 12:
William Genevrino, Kellen Danielsen,
Paco-Jaleel Balthazar, Sarah Tse

UNIVERSITY OF CENTRAL FLORIDA
1963

MEET THE TEAM!

Paco-Jaleel Balthazar

Computer Engineering

Kellen Danielsen

Electrical Engineering

William Genevrino

Electrical Engineering

Sarah Tse

Electrical Engineering

Project Description

- C.A.P.E.R. → Compact Animated Parrot with Enhanced Responsiveness
- Life-like parrot robot with realistic body, and easy-to-use controls.
- Fundamental movements
  - Head tilt
  - Beak open and close
  - Body tilt
  - Tail (and wing) flap
- Voice response
  - AI generated responses
  - Voice detection

Motivation?

- Current robots and animatronics costs tens of thousands of dollars to produce and maintain
- Vast majority of smaller companies are not able to utilize robots for additional entertainment
- Most are massive (human size or larger!), so they are typically bolted to one spot and are unable to be moved
- Many entertainment robots are not responsive and are only there for aesthetics

Goals

- Ultimately, to create an affordable and portable entertainment robot that has smooth movement and conversational capabilities
- Affordability in comparison to existing entertainment robots and animatronics
- Ease of use
- Accessibility
- Quick response time

Objectives

Realistic, smooth motion

- Life-like motions that mimic how a real life parrot would move via solenoids
- High quality human interaction
- Pre-recorded motion; both AI response and manual push button options available

Voice activation and speech detection

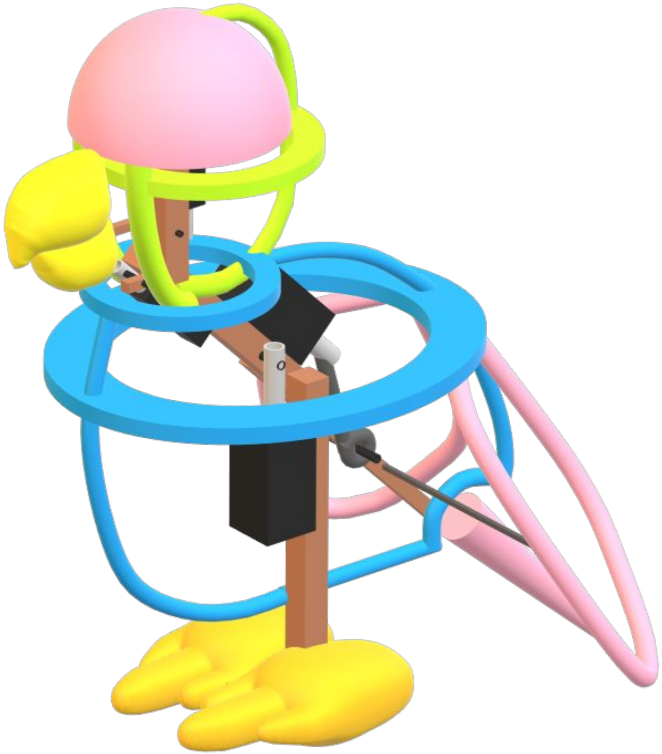
- Speech detection activates AI generated voice response
- Ability to carry a human conversation

Specifications

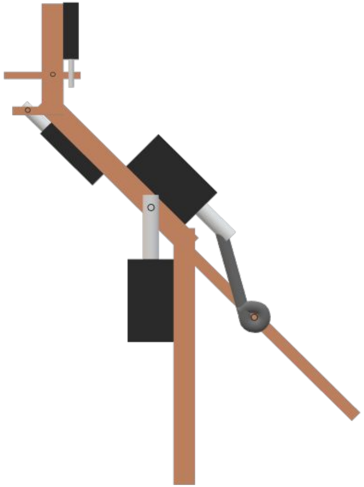
Table 2.1: Specifications

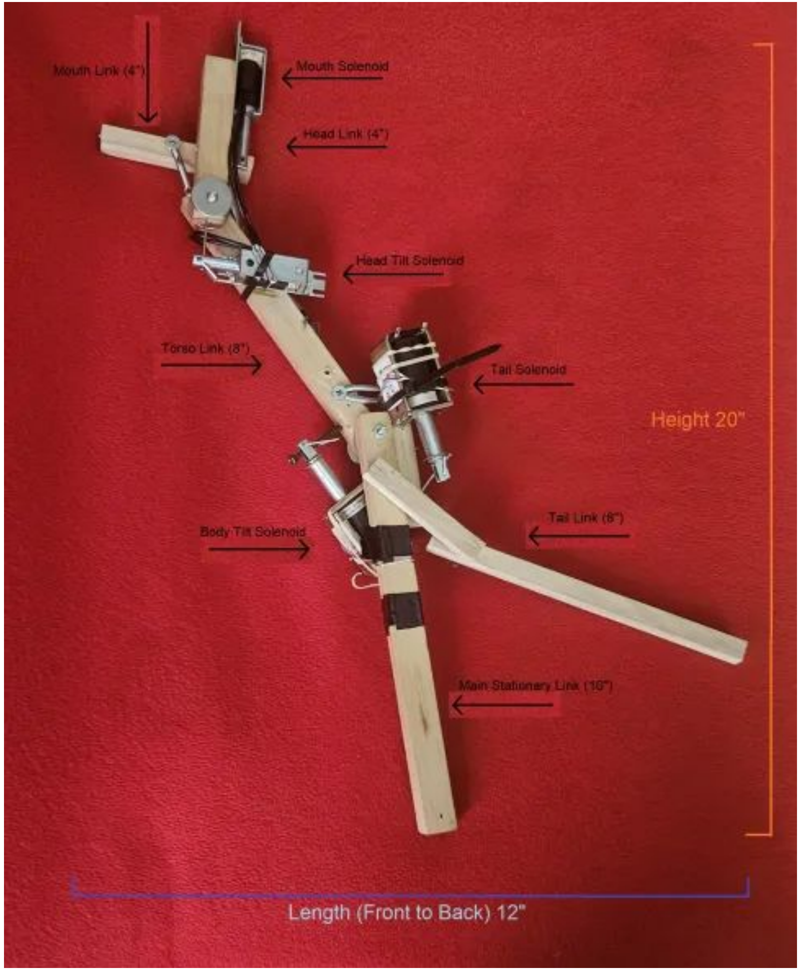
Hardware	
Movement capability	Solenoids should be able to move the joints in a way that makes the parrot life-like with fluid motions such as head tilt, mouth movements when speaking, wing flapping, etc. Head – 45 degrees Mouth – 50 degrees Body – 30 degrees Wings & Tail– 45 degrees each
Interactive modes	The selector switch will be used when the MIDI is inactive. The position of the switch will declare the mode the parrot is in. Turning the switch on would activate randomized movements for the head, body, and tail flap. Turning the switch off would activate total manual control using button inputs.
Dimensions	Around 1.5 ft tall. Maximum 2ft
Durability	Able to last at least two years with minimal maintenance
Weight	Maximum of 15 pounds.
Power Consumption	Maximum 40W
Cost Maximum	$1300
Response time to user input in push buttons	2 seconds
Software	
Voice Recognition	The parrot should be able to recognize when its name is called and respond accordingly with 80% accuracy.
Audio Response	Parrot should respond in a realistic manner with intelligible speech and 80% accuracy. Parrot should be easily understandable by the user.

Prototype Skeleton Design

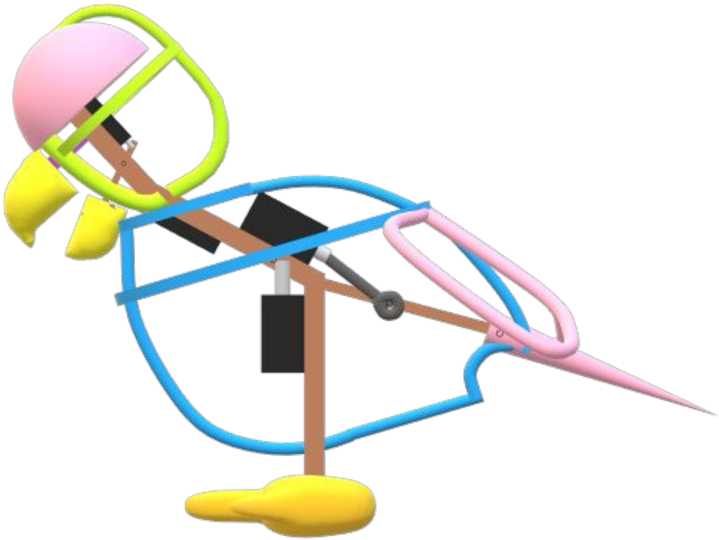


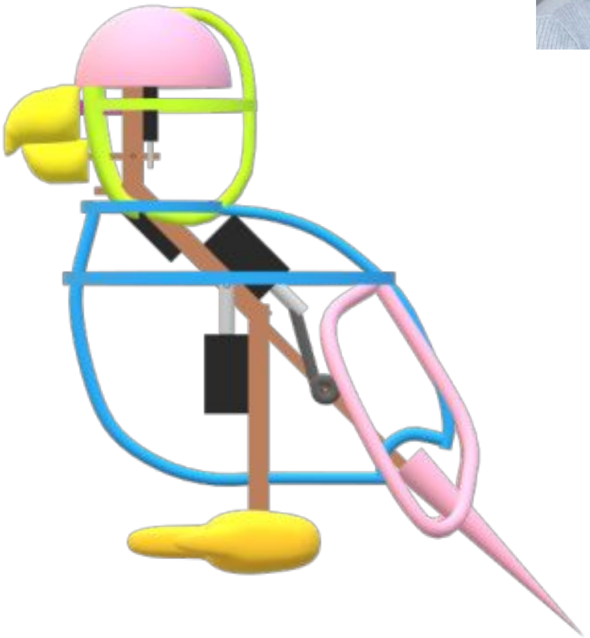
Materials:
Wood → internal frame
Solenoids → movement capability
Foam, rubber tubing, metal rods → internal shaping framework



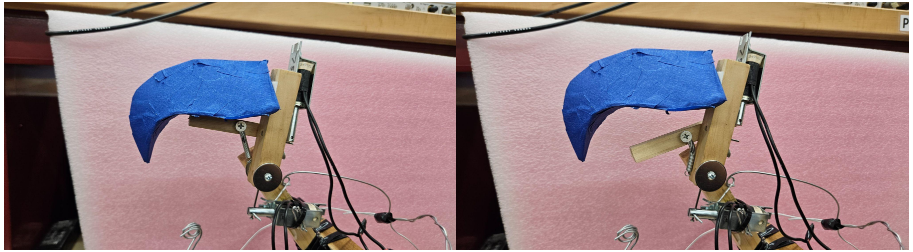

Mouth Link (4")
Mouth Solenoid
Head Link (4")
Head Tilt Solenoid
Torso Link (8")
Tail Solenoid
Height 20"
Tail Link (8")
Body Tilt Solenoid
Main Stationary Link (10")
Length (Front to Back) 12"

All four movements activated

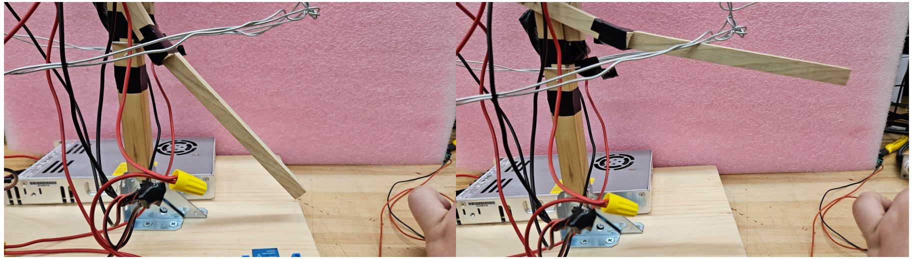




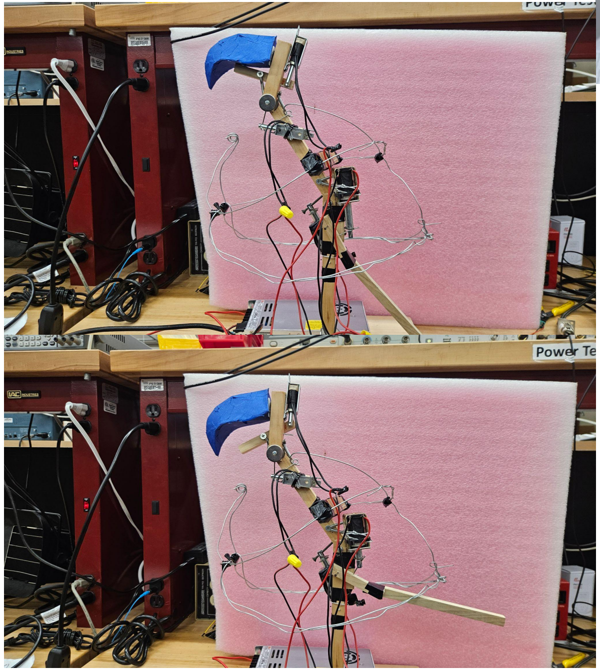
All four movements activated



Mouth Open and Close



Tail Up and Down



All four movements simultaneously

Operation Modes

Fully-manual live operation

- Push buttons

Partial-manual live operation

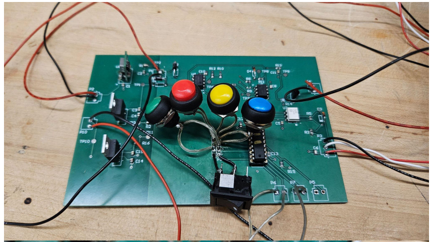
- Manual mouth control (button or microphone), with automatic asynchronous movement generation for non-mouth movements.

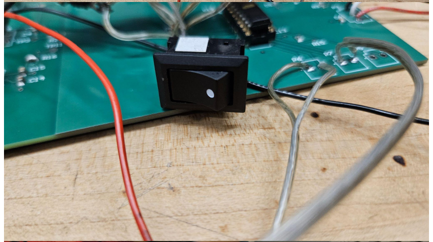
Pre-recorded sequence operation

- External playback device; digital audio workstation (DAW)
- Synchronized audio and movements

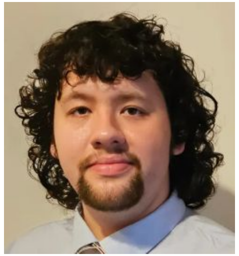
Unified conversation operation

- Detect human speech and respond synthetically





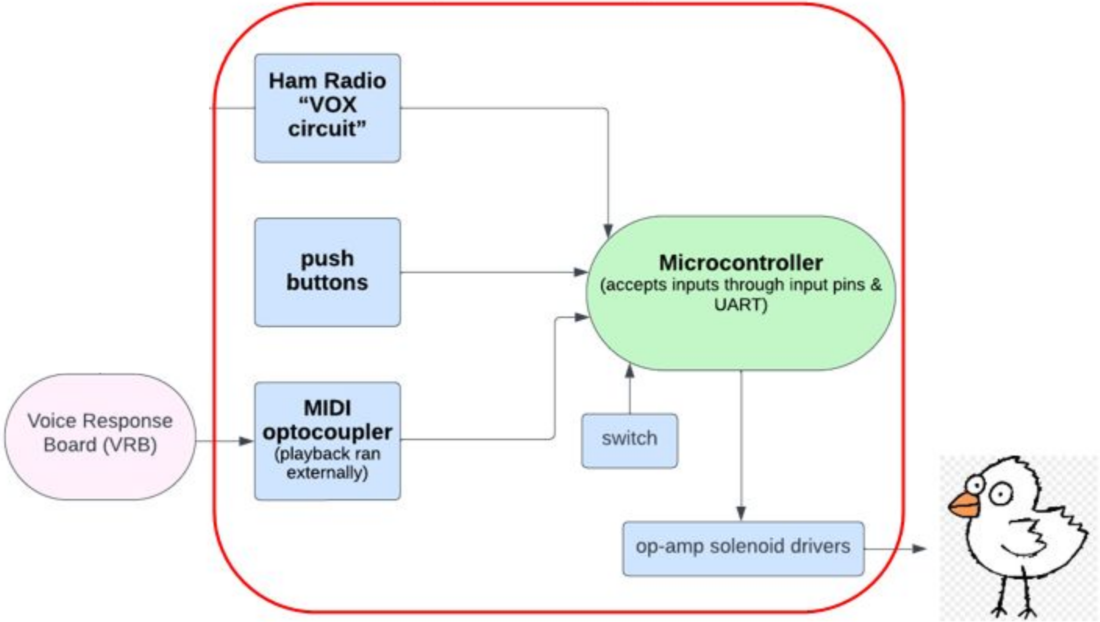
Main Block Diagram



Everything within the red outline is housed in the PCB and all require an external power supply (see individual subsection diagrams)

- Blue = peripheral circuits
- Pink = external devices
- Green = microprocessor chip

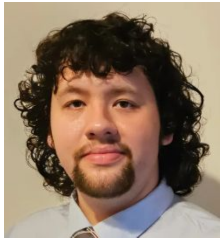
****all components are to be acquired*



IOCB

Input/Output and control board

- Microprocessor – MSP430G2553
- UART input stream processing
- Power supply for digital pins
- Peripheral circuits
  - Push buttons
  - Mode selector switch
  - MIDI optocoupler
  - Headers for external debugging
  - Power supply/regulators
  - VOX Circuit



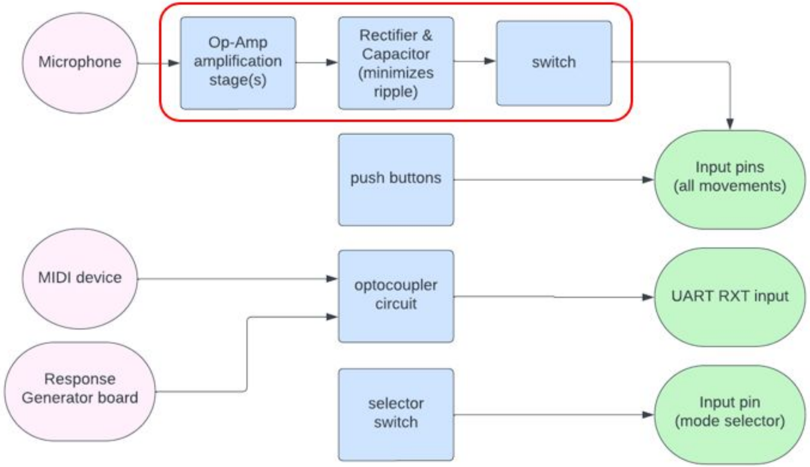
Input Control Board Block Diagram

Blue = on PCB; requires power supply

Green = microprocessor inputs

Pink = external devices (may require separate power supply)

Blocks within red outline → VOX circuit (see main diagram)

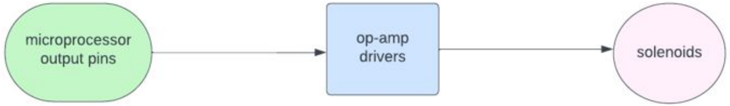


Output Control Board Block Diagram

Blue = on PCB; requires power supply

Green = microprocessor outputs

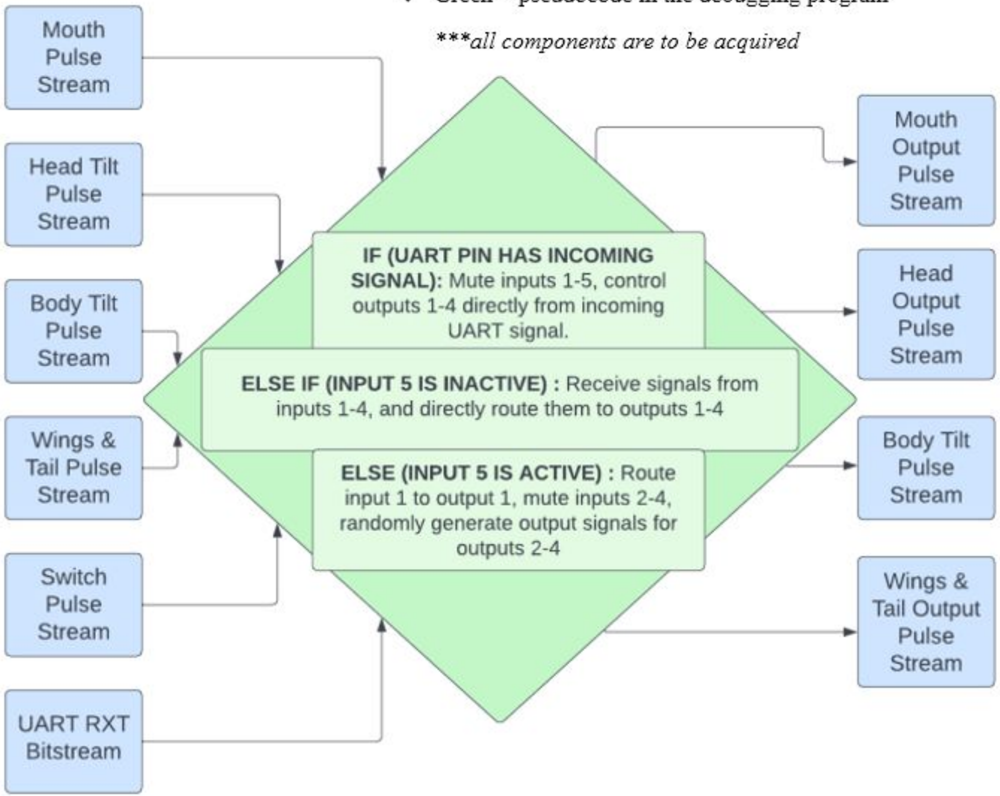
Pink = external devices (may require separate power supply)

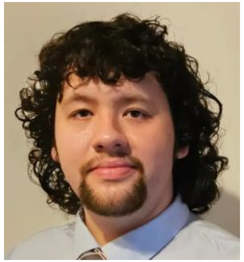


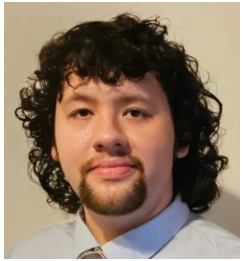
Microcontroller Software Logic

- Blue = I/O pins as seen by the **code/program**
- Orange = hardware & external devices
- Green = pseudocode in the debugging program

****all components are to be acquired*

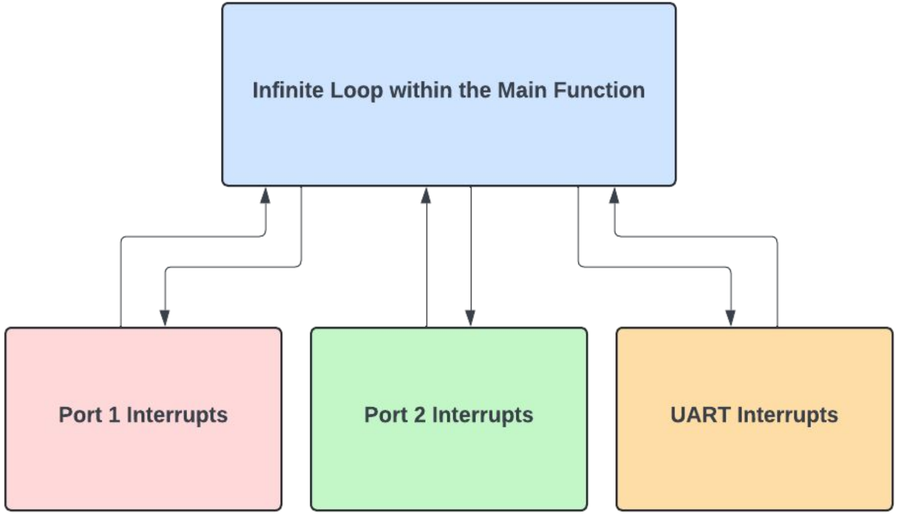






IOCB software

1. Define the “char” variables
2. Implicitly declare
3. Run the external “setUp()” function.
4. Define the four push-button input pins
5. Define the selector switch input pin
6. Set up the four outputs
7. Enable the global interrupts



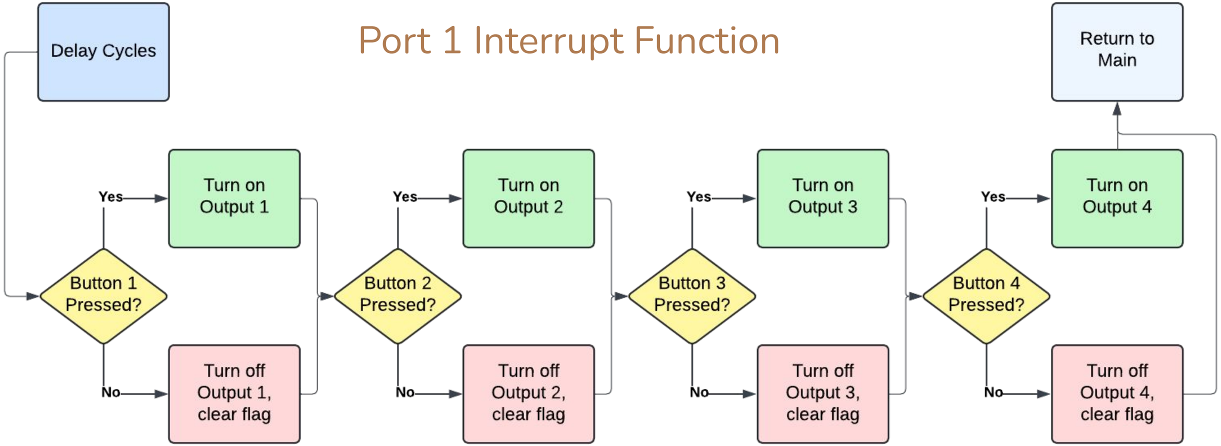
Inputs

- Mouth Pulse Stream
- Head Pulse Stream
- Body Pulse Stream
- Wings and Tail Pulse Stream
- UART MIDI Datastream
- Selector Switch Pulse Stream

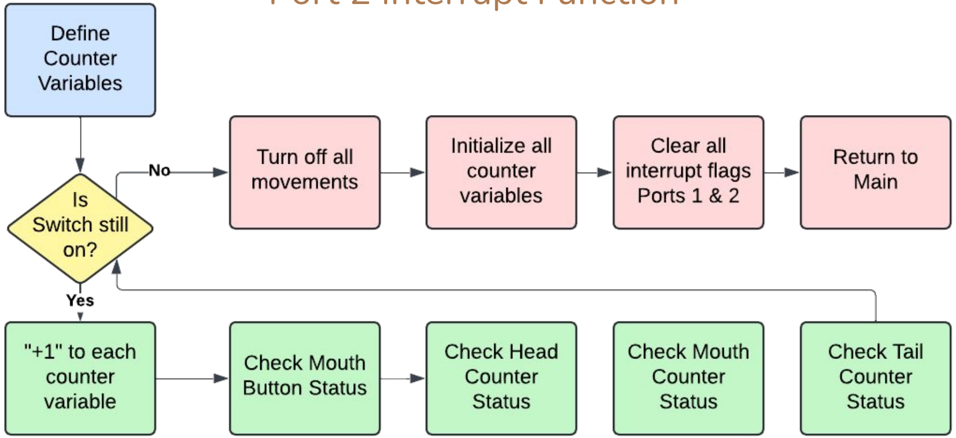
Outputs

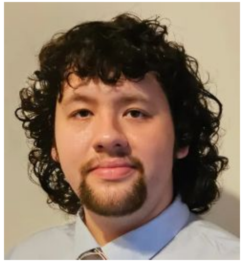
- Mouth Output Pulse Stream
- Head Output Pulse Stream
- Body Output Pulse Stream
- Wings and Tail Output Pulse Stream

Port 1 Interrupt Function



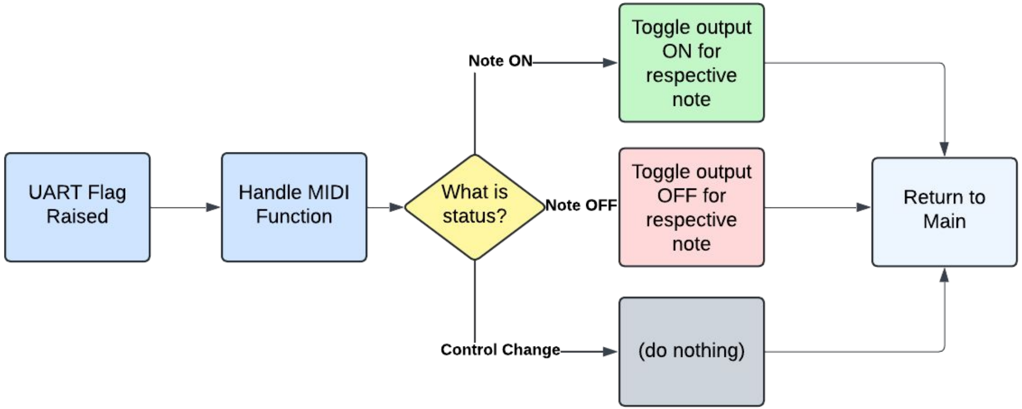
Port 2 Interrupt Function

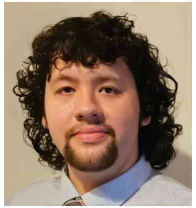




UART Interrupt Function: Pre-Recorded Sequences and Fully Live Operation

Message	Action
0x9# 0x3C 0x##	MOUTH OPEN
0x8# 0x3C 0x##	MOUTH CLOSE
0x9# 0x3E 0x##	HEAD DOWN
0x8# 0x3E 0x##	HEAD UP
0x9# 0x40 0x##	BODY DOWN
0x8# 0x40 0x##	BODY UPRIGHT
0x9# 0x41 0x##	TAIL UP
0x8# 0x41 0x##	TAIL DOWN





Motion Systems Comparison

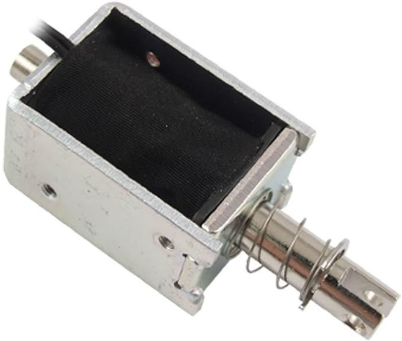
Table 3.6: Motion Systems Part Comparison

Aexit	**Features** • Within voltage and current range as specified: 14.4W @24VDC • Suitable housing dimensions: 2.6" x 1" x 0.8" • Sufficient stroke length: 1.2" • 100g mass a piece. • Pull strength of approximately 35N • Spring loaded • Within budget range: $12 to $15 a piece
Ledex	• Well within power range at under 9 watts @ 24VDC • Housing dimensions of 2" x 1.5" x 1" • Smaller stroke length of 0.69" • 192g mass a piece • 124N pull force • Not Spring loaded
Uxcell	• Rated at 72W (when fully retracted) • Dimensions 1" x 0.75" x 0.66" • 10mm stroke • 10N force • Unspecified weight (smallest size of the selections) • Spring Loaded
RS Pro Linear	• 11W to 110W consumption from 10% to 100% DC supply • $21 a piece • 2" x 0.75" x 0.75" • 12mm stroke length • Not spring loaded • 25N pull force • Corrosion resistant coating

Table 3.5: Motion systems

	Pneumatics	**Servos**	**Geared DC Motors**	**Solenoids**	**Stepper Motors**
Cost	>$300 total not including compressor or valve bank	$20.00 not including controller	<$10.00	~$15.00	$13 for smaller sized motors
Size	As small as 3 inch length	From "Micro" to "Jumbo"	Similar to servo	Similar to pneumatic	4cm wide square housing
Self-Contained (all inside CAPER)	No, requires external valve bank and air compressor	Yes, but will require additional control board	Yes	Yes	Yes, but would require additional control board
Suitable for Outdoor Use	Yes	Only the expensive ones	No	Yes	Only if outer shell is resistant (no)
Controllability	On/Off	Speed and Position	Approximate Speed and Direction	On/Off	Digital step counting, no real time position sensing
Movement	Linear Stroke	Limited or Cont. Rotation	Continuous Rotation	Linear Stroke	Continuous rotation
Power	From 200W to 4KW air compressor	~5W each, (20W total)	~0.2W each (~1W total).	~4W each (16W total)	72W each (288W total)

Motion Systems



Solenoids → Gives CAPER the capability to move

- Self contained
- Suitable for outdoor use
- On/off switch for controllability
- Power of approximately 18W each
- 100g mass a piece
- Spring loaded

Aexit	**Features** • Within voltage and current range as specified: 14.4W @24VDC • Suitable housing dimensions: 2.6” x 1” x 0.8” • Sufficient stroke length: 1.2” • 100g mass a piece. • Pull strength of approximately 35N • Spring loaded • Within budget range: $12 to $15 a piece

MCU Comparison

MSP430G2553	**Features:** • Pin layout: 20, 28 and 32 pin packages available • Package Size: 20-DIP • Low power modes: (4), can shut off all clocks and timers • Clock Modules: low frequency crystal and RC • Timers: one 16-bit timer • Serial Communication: UART, SPI, I2C, IrDA; automatic Baud-Rate detection • Analog to Digital: supports 10 bit conversions • Processing speed: up to 16MHz • RAM: 512B, 16KB Flash RAM • 16 bit architecture • Voltage: 1.8 to 3.6 VDC
MSP430FR6989	**Features:** • Pin layout: 80 or 100 pin • Package Size: LQFP (PN or PZ) • Low power modes: (3), shuts off all timers and clocks • Clock Modules: Low & high frequency crystals, DCO, MODOSC, and external • Timers; five 16-bit timers • Serial Communication: UART, SPI, I2C, IrDA; automatic Baud-Rate detection • Analog to Digital: 16 bit analog comparator, 12 bit ADC • Processing speed: 16MHz • RAM: 128KB FRAM, 2KB SRAM • 16 bit architecture • Voltage: 1.8 to 3.6 VDC • Code security: 128-Bit or 256-Bit AES Security Encryption and Decryption Coprocessor • Digital peripherals: 32-bit hardware multiplier, 3-channel internal direct memory access

Atmel ATMEGA328P-PU	**Features:** • Pin layout: 28 pin (others available, but this layout is easiest to work with) • Package Size: SPDIP • Qtouch library: 64 sense channels; capacitive touch buttons, sliders, and wheels • Clock Modules: Low frequency & full swing crystals, RC, external • Timers: two 8-bit, one 16-bit, and six PWM channels • Serial Communication: USART, SPI, and Phillips I2C • Analog to Digital: 6-channel, 10-bit (differs depending on package) • Processing speed: 20MHz (20 MIPS using single cycle instructions) • RAM: 32KB flash RAM, 2KB RAM • 8 bit architecture • Voltage: 1.8 to 5.5 VDC
ATMEGA2560	**Features:** • Pin layout: 100 pins • Package Size: TQFP or CGBA (25 pins per edge or 10x10 grid on bottom respectively) • Qtouch library: 64 sense channels; capacitive touch buttons, sliders, and wheels • Clock Modules: Low power crystal, LF crystal, Full swing crystal, internal RC, external • Timers: two 8-bit, four 16-bit, four 8-bit PWM channels & six-twelve PWM (2 to 16 bit) • Serial Communication: USART, SPI, 2-wire Byte Oriented serial interface • Six sleep modes, ADC noise reduction • Analog to Digital: 16-channel, 10-bit • Processing speed: 16 MHz (16 MIPS using single cycle instructions) • RAM: 256KB Flash RAM, 8KB RAM • 8 bit architecture • Voltage: 4.5 to 5.5VDC at 16MHz (as low as 2.7 for 8MHz) • JTAG compliance, extensive on-chip debugging support

Input Output Control Board MCU Chip

Table 3.7: Chip specifications

	MSP430G2553	**Arduino ATMEGA328P**	**Raspberry Pi SC0914 (RP2040)**
Speed	16 MHz	20 MHz	133MHz
Storage	16KB Flash	32KB Flash	264KB Flash
Pin Configurations	20 or 28 pins	21 or 28 pins	56 pins
UART	Yes	Yes	Yes
Debugging Hardware	Internal Bootstrap Loader using UART Bridge	debugWIRE	Done on PICO board
Software	TI Code Composer Studio	Arduino IDE	Multiple IDEs can work for this

Relay Module

Table 3.12: Transistors vs Relays

	BJT (2N2222)	Relay (BESTEP)
Voltage Limits	5V (emitter to base). 60V (collector to base)	30V DC (triggers at exactly 3.3V)
Current Limits	800mA	10A
Switching Time	~150 microseconds	~10 milliseconds
Cost	$0.40	~$2.00

Table 3.13: Data Voltage Control Hardware Selection

	BESTEP	HiLetgo	Teyleten
Price	<$1	$5.89	$2
Config.	1 channel	1-8 channels	1 Channel
Power	10A 250VAC or 30VDC	10A 250VAC or 30VDC	10A 250VAC or 30VDC
Trigger V.	3.3V (5V options available)	5V or 12V (no 3.3V options)	3.3V
Connections	Raw Wire	Raw Wire	Raw Wire
Norm O/C	Open or Closed	Open or Closed	Open or Closed
Switch Delay	<20ms	*unspecified	<20ms

Microphone selection (VOX Circuit)

Table 3.15: Microphone Circuit Style Comparison

Dynamic	- Dynamic microphones are the most common circuits you can find in a handheld microphone. Essentially a reverse speaker, dynamic mics use coiled wire and a permanent magnet to generate electrical waves from an incoming sound source. They **do not require an external phantom power supply**, are very lightweight and portable, and **relatively affordable** compared to other mics (anywhere from $50 to $300).
Condenser	- Condensers (the British-English term for Capacitors), rely on a pair of capacitor plates to generate soundwaves based on the gap between them. Unlike Dynamics, Condenser microphones do **require an external 70V phantom power supply** to produce signal, but are usually more articulate in the 10kHz range and higher. **Prices range from $100 to $4000+**, not including the price of the power supply.
Ribbon	- Ribbon mics are one of the oldest microphone styles you can still purchase, originating in the 1920s. Active and passive versions both exist, but they **CANNOT handle phantom power** (unless you want to instantly destroy them). In fact, they are the most fragile style of microphone on this list, and expensive too, **ranging as high as $7000 dollars.**

The VOX circuit will work with **dynamic microphones**, because they don't require phantom power and they're much more affordable than ribbon or condenser microphones.

Power Supplies

Table 3.17: Power Supply Comparison

MEAN WELL LRS-150-12	MEAN WELL HRP-150-12
• Package dimensions: 6 x 3.81 x 1.18 inches • Output voltage: 12V • Max output wattage: 150W • Working temperatures: -30~+70C • Safety approvals: TUV EN 62368-1, UL 62368-1 • Fanless design; convection cooling • Over Temperature Protection • Cost: $20-30 (typical, depending on vendor)	• Package dimensions: 6 x 3.81 x 1.5 inches • Output voltage: 12V • Max output wattage: 150V • Working temperatures: -40~+70C • Safety approvals: UL 62368-1 • Fanless design; convection cooling • Cost: $50-60 (depending on vendor)

- **NOTE:** We opted for the **LRS-350-12** to power the figure due to the high current draw from the solenoids.

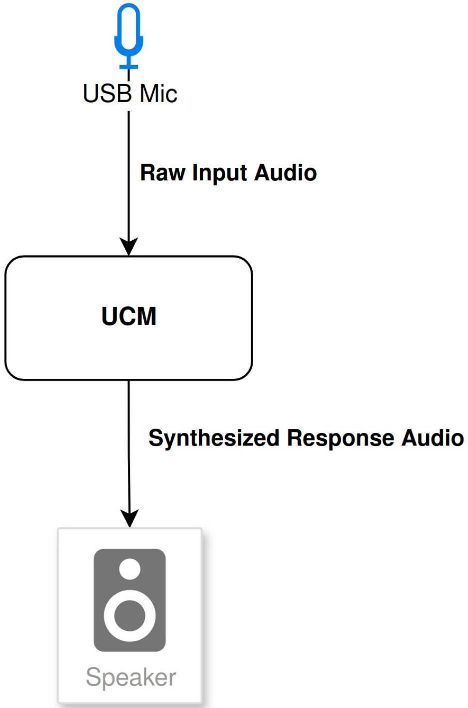
- The IOCB requires **two** power bricks for positive and negative supply voltages. Both supply **12V, 1A**, which is more than enough for the IOCB and relays.
- Power brick model: **JOVO JVN12V1ABK**

Unified Conversational Module Design (UCM)

UCM Goals:

- Listen for speech
- Respond to speech within **3 seconds**
- Apply movement to match speech

For this project, the software choice **defined** the hardware choice.



Framework Selection

	Tensorflow	Pytorch	MCT
Has toolkit for embedded conversion	✅	✅	❌
Supports CUDA	✅	✅	✅
Supports CPU	✅	✅	✅
Supports OpenCL	✅	✅	❌
Supports OpenGL	❌	✅	❌
Best Ease of Use	❌	✅	❌

Chosen Framework: **PyTorch**

Model Type Selection

Table 3.21: ASR vs LLM/SLM vs TTS

Pipeline Module	Criteria	Model Type
ASR	Highest Quality	Transformer
	Cheapest	HMM
	Best Cost-to-Performance	CNN for feat. extraction and transformer for modeling
LLM or SLM	Highest Quality	Transformer
	Cheapest	LSTM
	Best Cost-to-Performance	Transformer + LSTM
TTS	Highest Quality	Transformer
	Cheapest	GAN
	Best Cost-to-Performance	Transformer

Table 3.22: tradeoff

Name/Value	Whisper-Small	Wav2Vec2-Large-960 h
Size on disk (MB) [Minimize]	483.6	1262
Size during compute (GB) [Minimize]	0.976	1.484
WER (%) [Minimize]	3.54 (clean+dirty)	1.7 (clean) and 2.925 (clean+dirty)

Model Selection (cont.)

Model Name	Size During Compute 4-bit quantized (GB)	Context size (tokens)	MMLU Score (%)	Open-Source
Mistral 7B Instruct	7GB	8192*	60.1	✅
Meta's Llama 2 7B	7GB	4096	45.3	✅
Meta's Llama 7B	7GB	4096	35.1	✅
OpenAI's ChatGPT 3.5 Turbo	N/A	16384	70.0	❌

Whisper-Small **ASR** — Implemented via Whisper.cpp

Mistral 7B **LLM** — Implemented via LLama.cpp

SpeedySpeech **TTS** — Implemented via SpeedySpeech Python Lirbary

Computing Solution Selection

Table 3.10: Hardware vs Server Offloading

Aspect	High-Cost Edge Hardware	Server Offloading
Best performance to cost ratio	❌	✅
Best software scalability	❌	✅
Best hardware upgradability	❌	✅
Best pipeline simplicity	✅	❌
Best usage latency	✅	❌
Best educational opportunity	❌	✅

Device	Price (USD)	CPU/GPU	RAM	Storage	Performance (GFLOPs)	Language
Nvidia Jetson Nano	$149.99	Quad-core ARM A57 / 128-core NVIDIA Maxwell	2/4 GB	microSD	500	Cuda and OpenCL
Google Coral Dev Board Mini	$99.99	MediaTek 8167s SoC / IMG PowerVR GE8300	2 GB	8 GB eMMC	32 (32-bit)/64 (16-bit)	OpenCL
BeagleBone AI-64	$187.50	Texas Instruments TDA4VM	4 GB	16 GB eMMC	160 (out-of-box) 8000 (requires models bit customized)	OpenCL

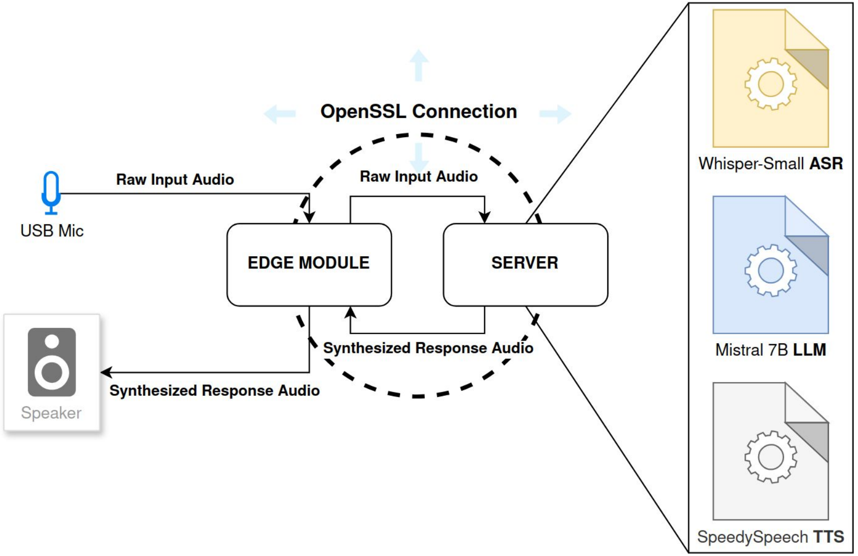
Computing Solution Selection (cont.)

Compute Heavy AI pipeline is run on Desktop OpenSSL Server.

Server Specs:

- OS: Ubuntu 22.04.4 LTS x86_64
- CPU: AMD Ryzen 7 PRO 3700 (16) @ 3.6
- GPU: NVIDIA GeForce RTX 3060 Lite 12GB

Jetson Nano is used exclusively for communication to IOCB



UCM Communication Protocols

Table 3.18: Communication Protocol

	UART	I2C	SPI
Synchronization	Asynchronous; requires transmitter and receiver to operate on same baud rate.	Synchronous; controlled by master serial clock, shared with slave.	Synchronous; master clock signal sent to slaves on separate lines.
Duplex	Half or Full-duplex	Full-duplex	Full-duplex
Slaves per Master	2 transceiving devices	128 slaves for 7 bit addressing, 1024 for 10 bits.	2-3 slaves per master
Directionality	Bidirectional; Tx to Rx for both devices	Bidirectional, two serial lines	Bidirectional; MOSI to MISO between master and slave(s)
Wires per Channel	1 per direction (2 in total)	Two serial lines; all slaves connected	7 lines going into master, 4 going into slaves, all interconnected
Speed	460kbps	100kbps	>100MHz
Error Detection	Parity bits (parroty bits in this case 😀)	ACK/NACK bit	All hardware-based

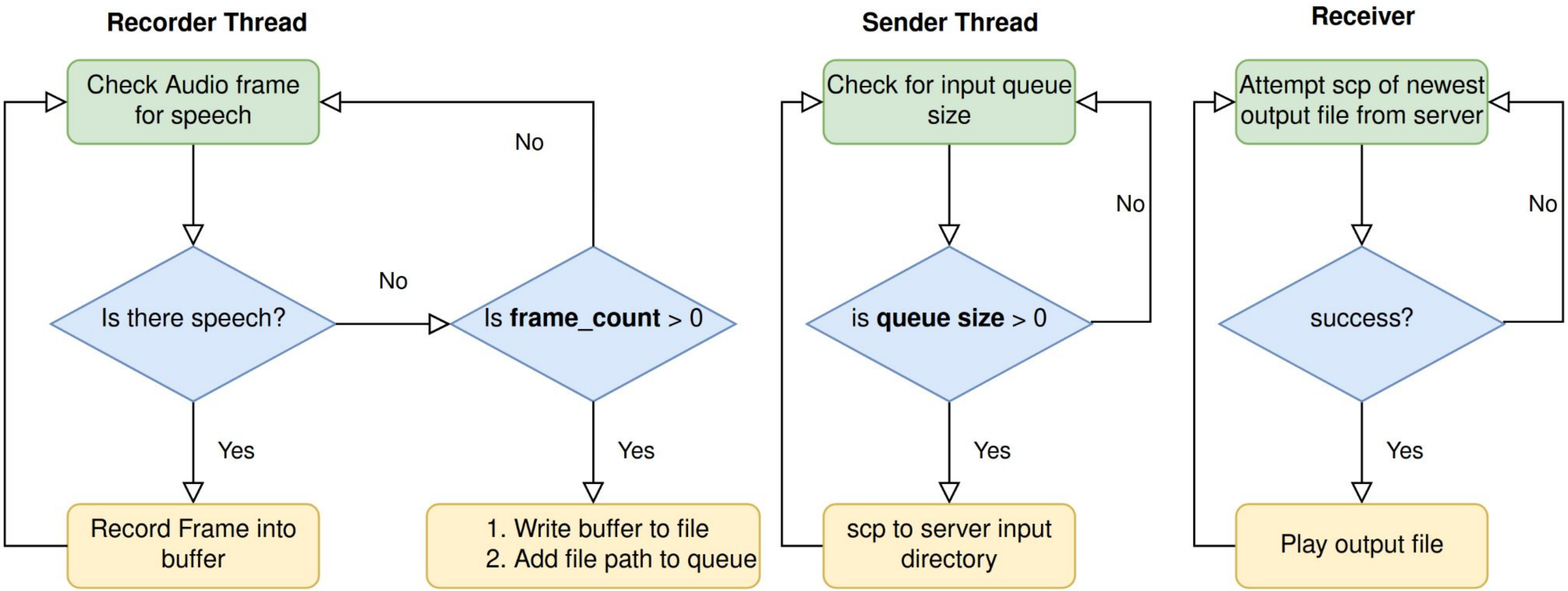
Chosen Protocol For Board-to-Board: **UART**

- Robot Body Control
- MIC Input
- Audio Output

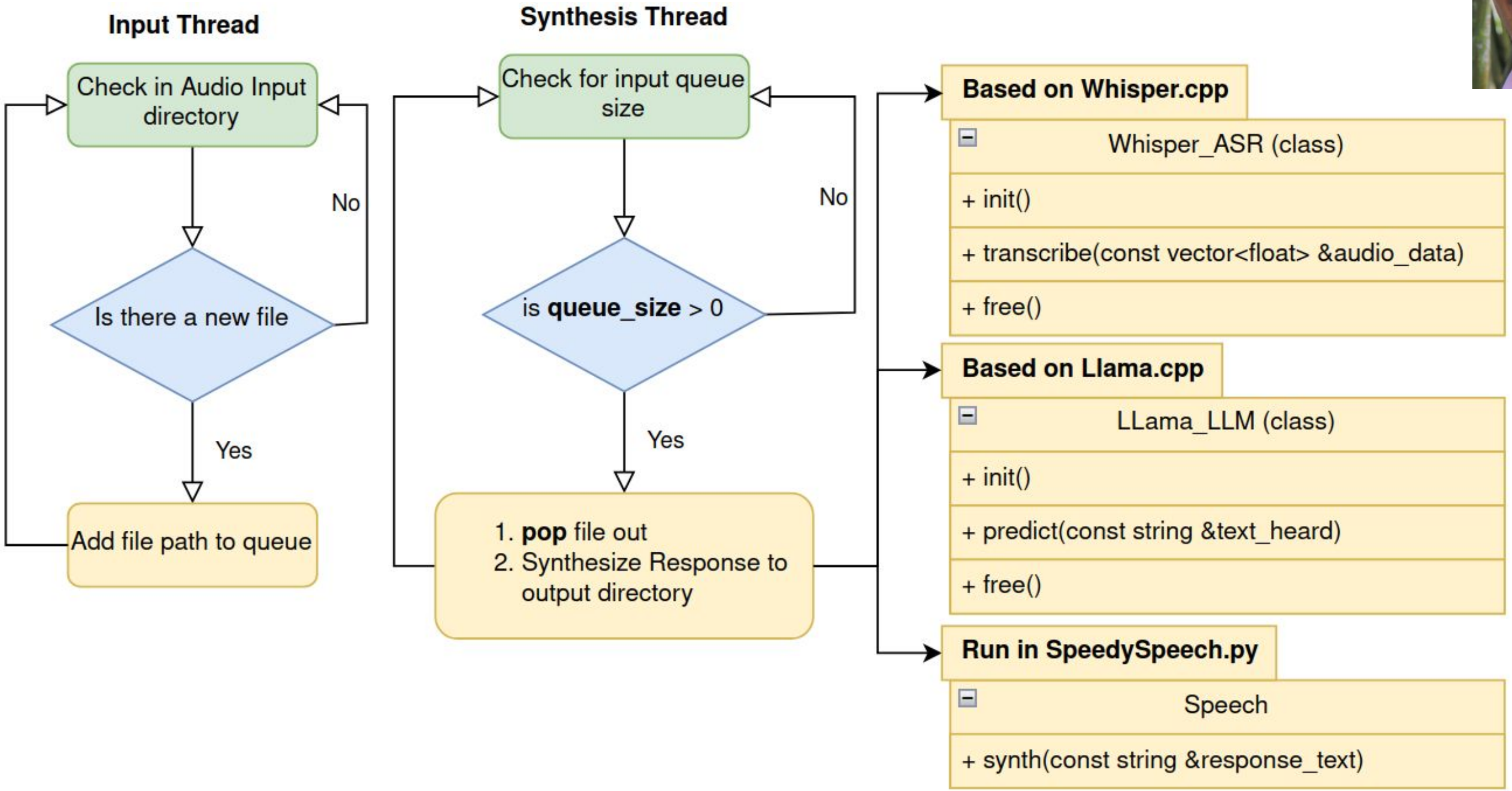
Chosen Protocol For Inter-module Communication: **OpenSSL**

- TCP Based Audio Transfer
- Secure
- Fast

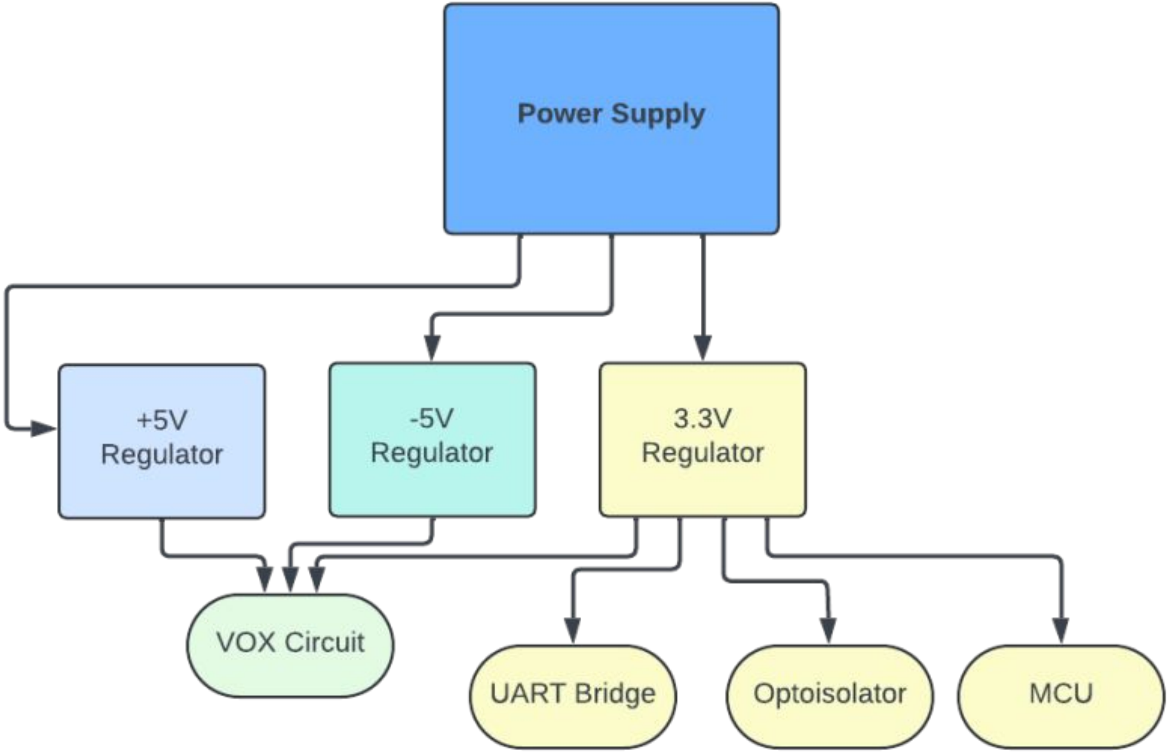
Edge-Side Threading

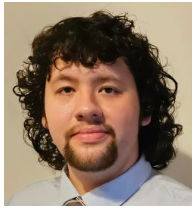


Serve-Side Threading

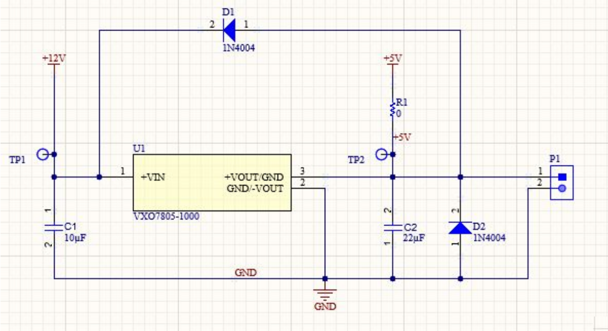


Switching Power Supply – 12V



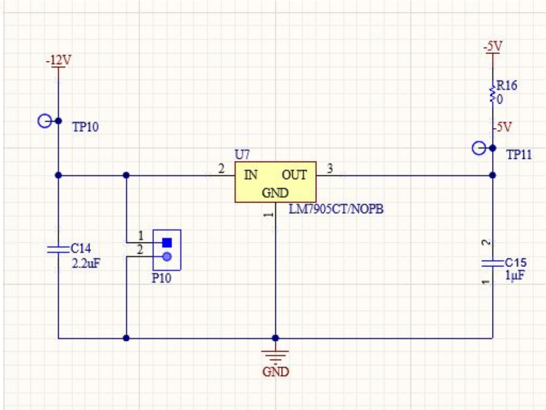


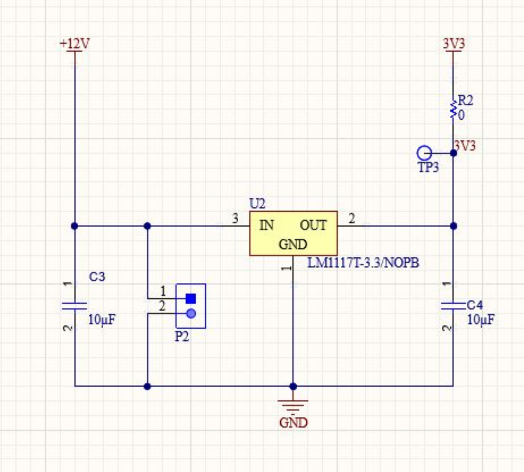
Voltage Regulator Circuits



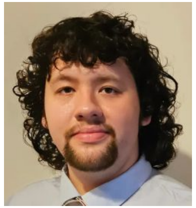
+5V regulator using the VXO7805

-5V regulator using the 7905





+3.3V regulator using the LM1117

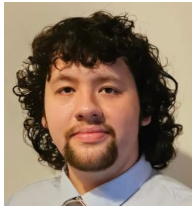


Voltage Regulator

Capacitor Values

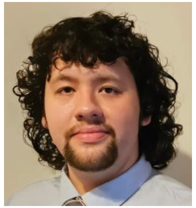
Chip	Input Shunt Capacitor	Output Shunt Capacitor
LM1117T-3.3	10uF	10uF
VXO7805	10uF	22uF
7905	2.2uF	1uF

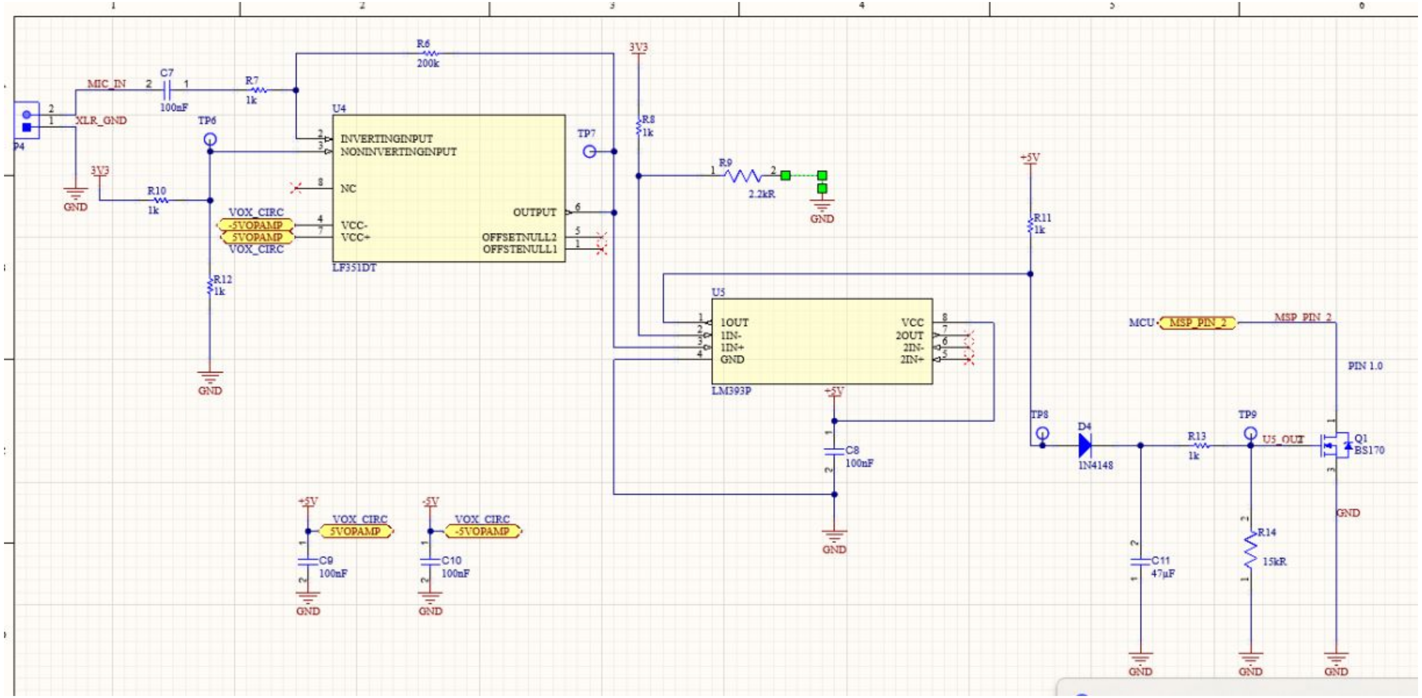
- Provides voltages for the MCU and IOCB peripherals
- Multiple circuits used to prevent overheating from current
- Powered directly from the 12V power supply



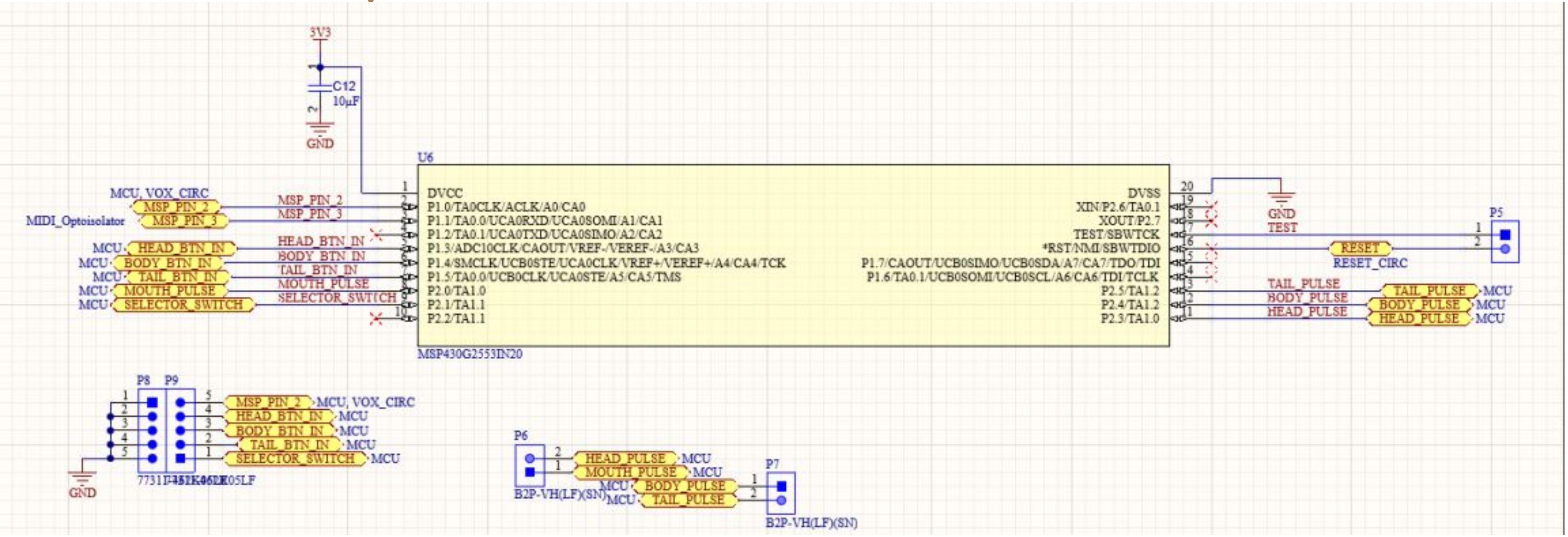
Ham Radio VOX Circuit

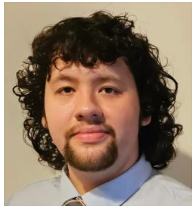
Voice activated switch for hands free mouth actuator movement and voice response board voice detection.

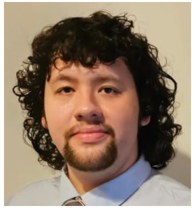




MCU chip connections







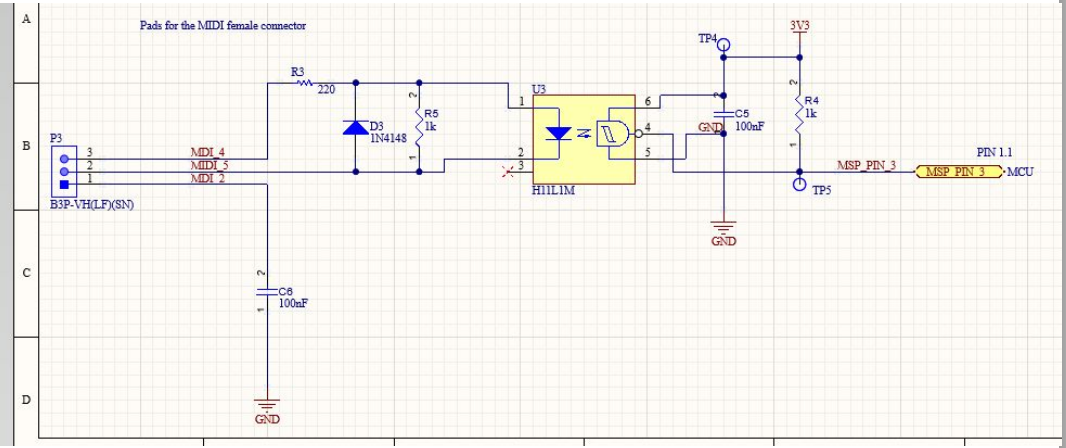
MIDI Optocoupler

UART communication

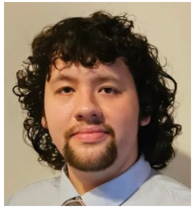
Controls movements with familiar standardized language

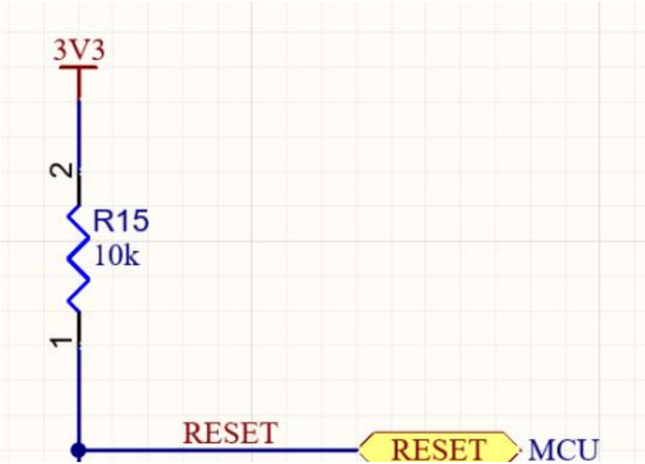
Does not require complete synchronization between devices

Easily implemented

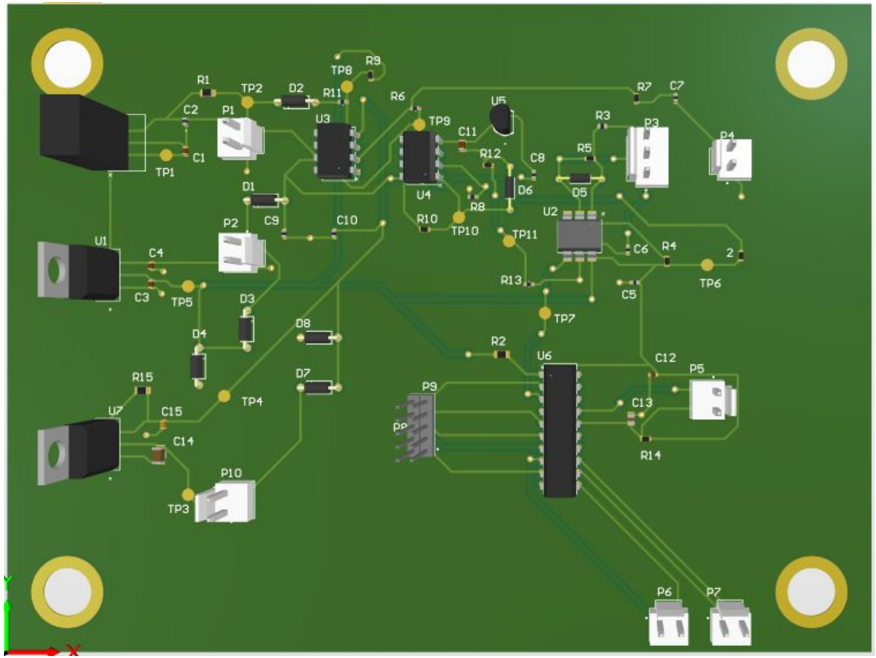


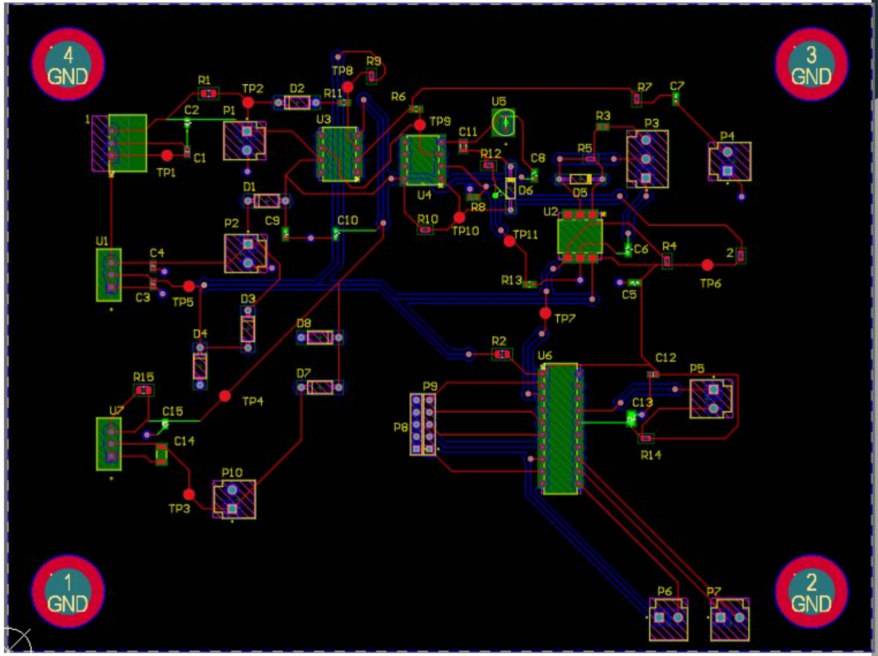
Reset circuit



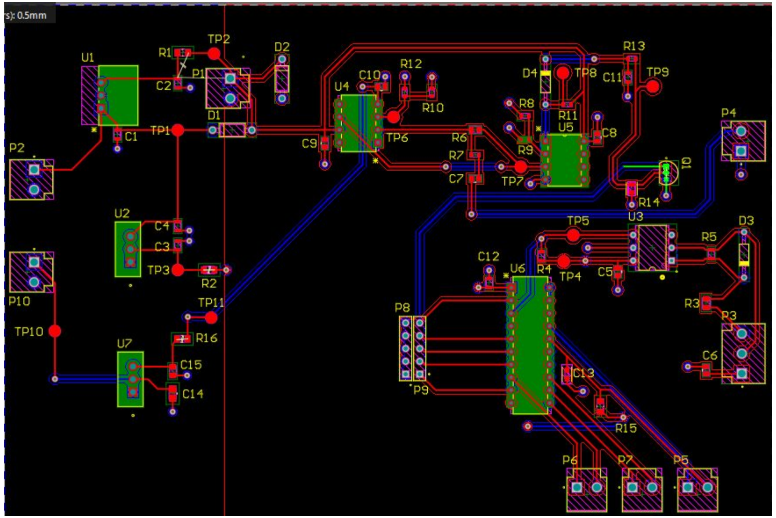


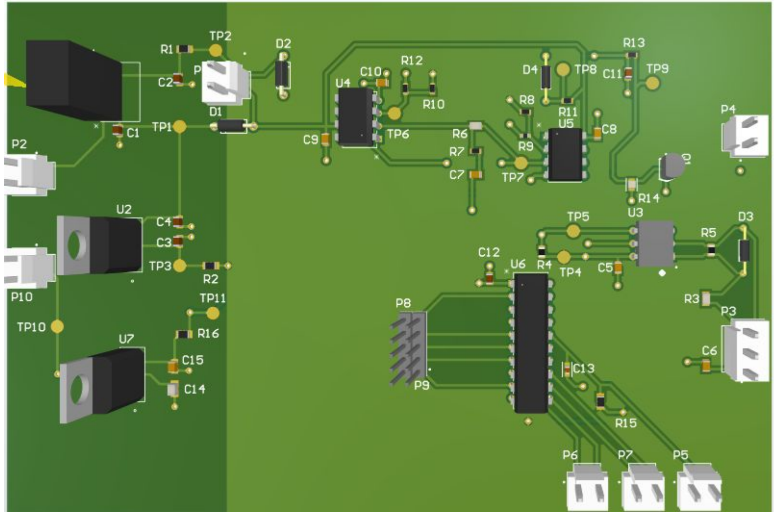
Second PCB layout





Final PCB Layout (third layout)

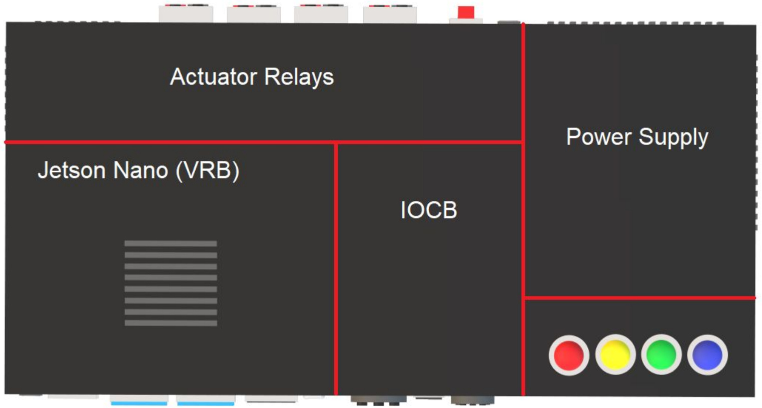




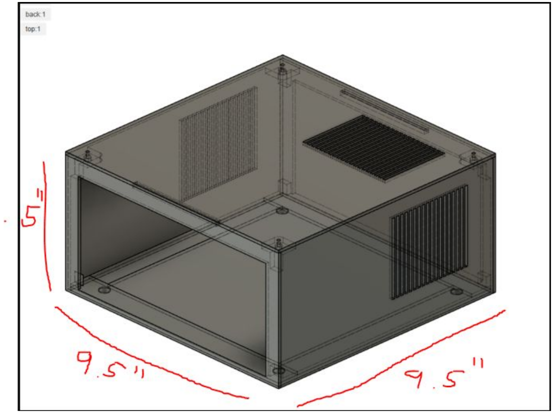
Chassis Layout

- Stores the IOCB, VRB, actuator relays, and DC power supply in one location
- Protects circuit boards
- Support structure

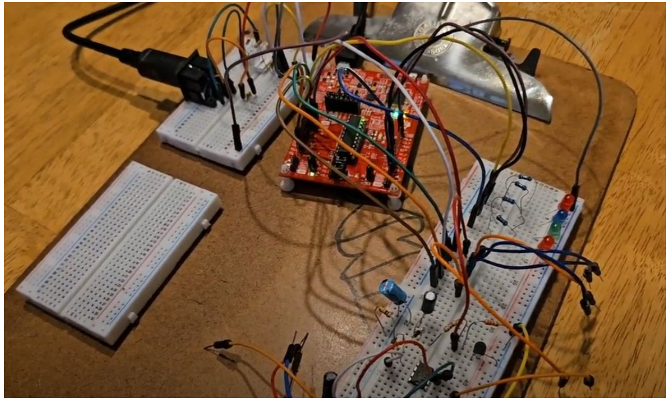
Potential Layout

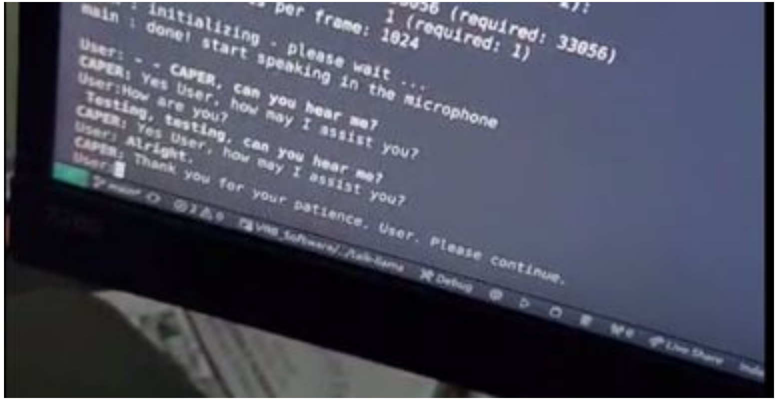


Final Design

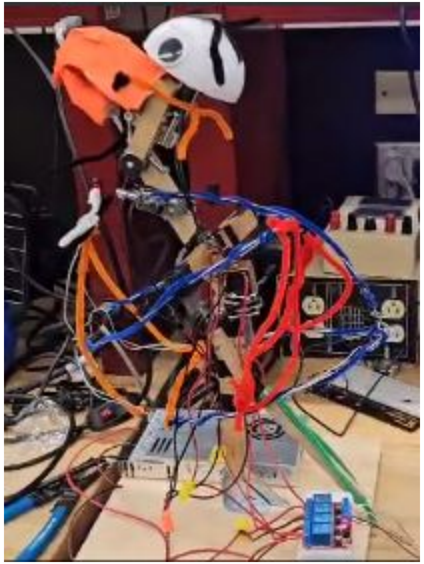


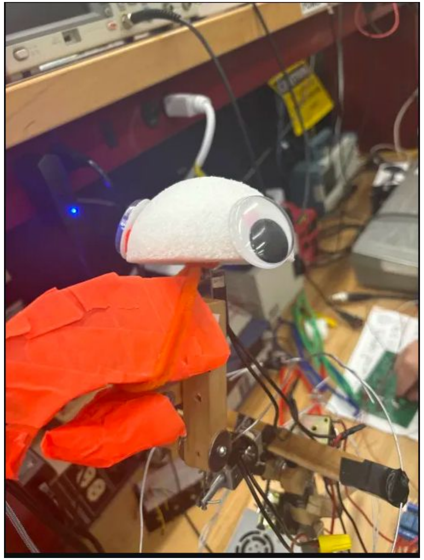
Prototyping
(Pre-Midterm Demo)





Final Prototype

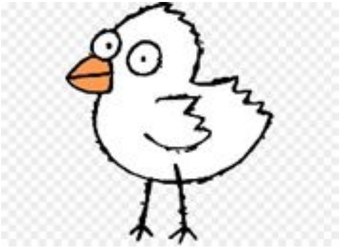




Work Administration table

Work	Primary Involvement
PCB design	Everyone
Embedded	William
Assembly	Sarah
Software	Paco
Circuitry	Kellen
Documentation	Everyone
Group Lead	William

Everyone was involved in every portion throughout the process of creating CAPER.



Budget (pg 1 of 2) – mechanical materials not included

Component	Model	Quantity	MSRP	Our Costs
+5 volt regulator	VXO7805-1000	6	$2.95 each	$17.70
-5 volt regulator	LM7905CT	6	$1.61 each	$9.66
+3.3 volt regulator	LM1117T-3.3	5	$1.69 each	$8.45
microcontroller	MSP430G2553IN20	5	$2.81	$14.05
remaining Board components	assorted components: resistors, capacitors, op-amps, etc.	5 boards worth of materials	N/A	~$70
solenoids	f190412ae059347	4	$12.25 each	$12 (for all four)
parrot Figure Structural Materials	N/A	1 figure's worth	N/A	~$60

Budget (pg 2 of 2) – mechanical materials not included

Nvidia Jetson Nano	900-13448-0020-000	1	$156.88	$0 (already owned)
relay modules	SRD-03VDC-SL-C	2 modules of 4 relays	$9.99	$17.98
microphone	MXL LSM-3	1	$79.99	$0 (already owned)
powered speakers	Numark	1	$99.99	$10 (used)
12V Power Bricks	JOVO JVN12V1ABK	2	$6.19	$12.38
power supply	MEAN WELL LRS-350-12	1	$31	$30
SD memory card	EVO select 256GB	1	$22.99	$0 (already owned)
2nd gen PCB fully assembled	JLCPCB	5	$135	$135
3rd gen blank PCB	JLCPCB	5	$17	$17

Goal costs:
<$1300

Estimated total:
$447.22
(excluding taxes)

What would we do differently?

Assemble in house

Use more up to date solenoids

Change physical construction of the bird to better support weight

Further Enhancements, Next Steps

- Fine-tune CAPER’s voice so she sounds more realistic
- Add sensors to create more real-time, life-like interactions, such as CAPER moving or speaking when she is touched or when she sees a person
- A finished body with skin, feathers, and all the fixins

Thank You!