Human-Swarm Interaction through Natural Language Commands

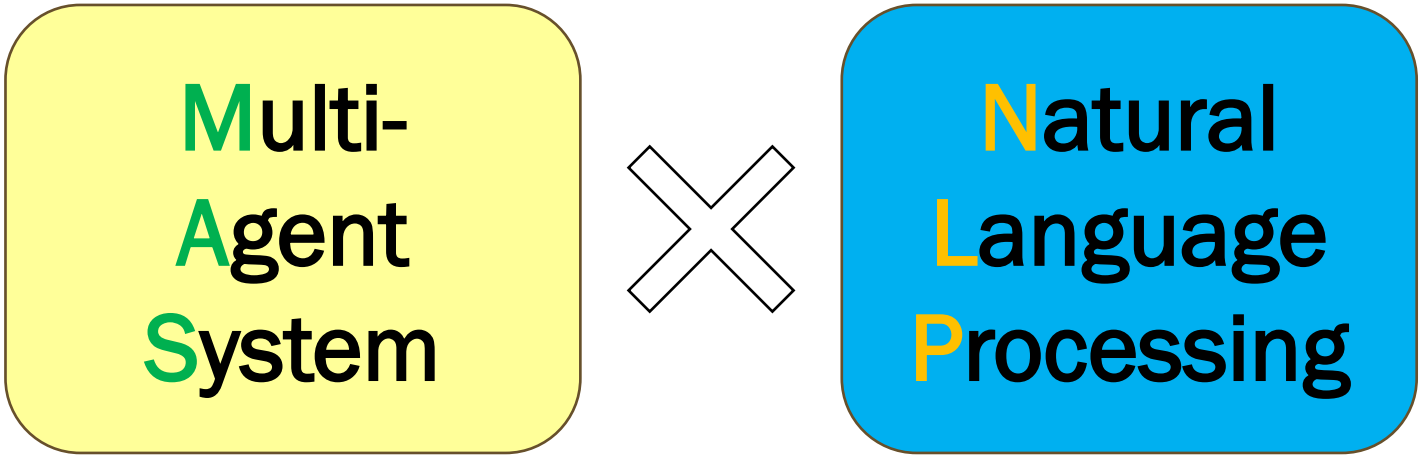
CNIT 581-SDR [Spring 2021]: Final Presentation

Roman Ibrahimov, Yifei Hu

PURDUE UNIVERSITY | Polytechnic Institute

Project Introduction

Motivation



Project Introduction

Motivation

- Why multi-agent systems are important?
  - *Widely exist in nature and engineering applications.*
  - *Offers better autonomy, robustness, flexibility, etc.*
  - *Can achieve sophisticated missions that are well beyond individual system's capability*
- Why using natural language to interact with the robots?
  - Natural language is the most “natural” way for human to communicate with each other
  - No specific training required

Project Introduction

Background

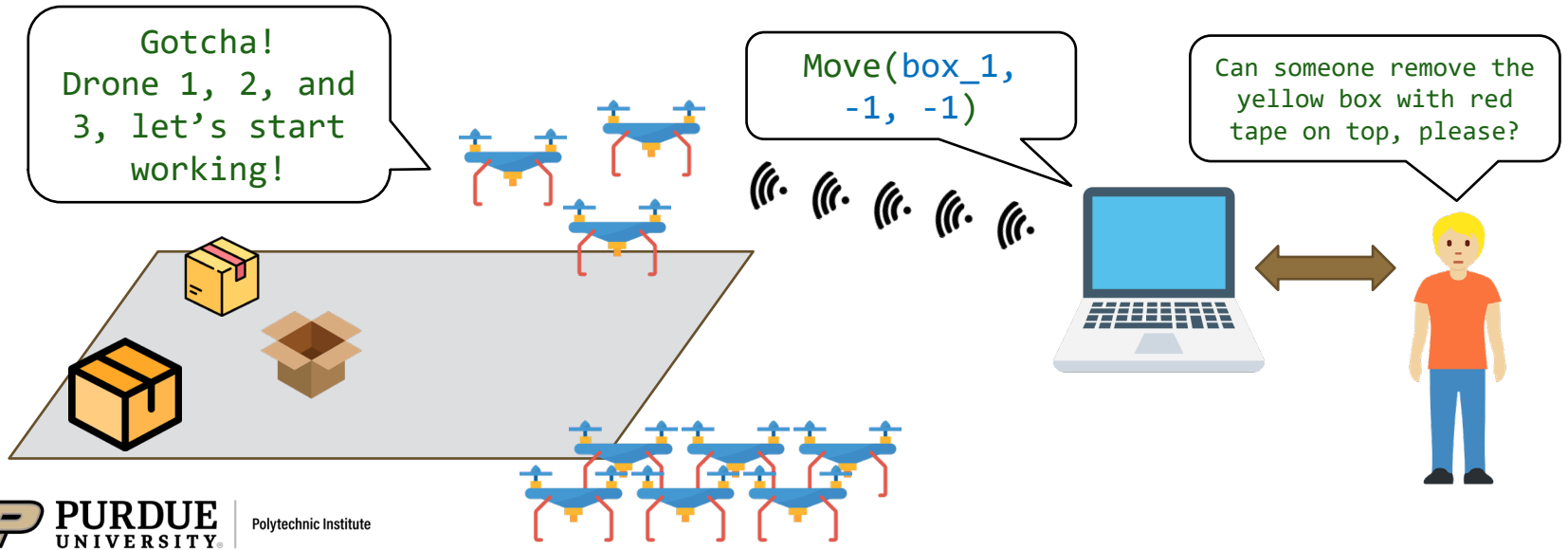
- NLP technology is used to extract explicit command(s) from natural language
  - Ontology-based NL control system [1][2]
  - Pre-defined Syntax [3]
  - Robots can also use NL to respond [4]
- Human-swarm interaction is a relatively new field
  - Human-swarm interaction based on augmented reality tablet application [5]
  - An inherently collaborative task: collective transport [5]
  - Human-swarm interaction based on tablet application [6]
  - Quick manipulation of the swarm [6]

[1] Hong, J. H., Min, B. C., Taylor, J. M., Raskin, V., & Matson, E. T. (2012, October). NL-based communication with firefighting robots. In 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1461-1466). IEEE.
[2] Matson, E. T., Taylor, J., Raskin, V., Min, B. C., & Wilson, E. C. (2011, December). A natural language exchange model for enabling human, agent, robot and machine interaction. In *The 5th International Conference on Automation, Robotics and Applications* (pp. 340-345). IEEE.
[3] Matuszek, C., Herbst, E., Zettlemoyer, L., & Fox, D. (2013). Learning to parse natural language commands to a robot control system. In *Experimental robotics* (pp. 403-415). Springer, Heidelberg.
[4] Raman, V., Lignos, C., Finucane, C., Lee, K. C., Marcus, M. P., & Kress-Gazit, H. (2013, June). Sorry Dave, I'm Afraid I Can't Do That: Explaining Unachievable Robot Tasks Using Natural Language. In Robotics: Science and Systems (Vol. 2, No. 1, pp. 2-1).
[5] Diaz-Mercado, Yancy, Sung G. Lee, and Magnus Egerstedt. "Human–swarm interactions via coverage of time-varying densities." *Trends in Control and Decision-Making for Human–Robot Collaboration Systems* (2017): 357-385.
[6] Patel, Jayam, Yicong Xu, and Carlo Pinciroli. "Mixed-granularity human-swarm interaction." *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.

Project Introduction

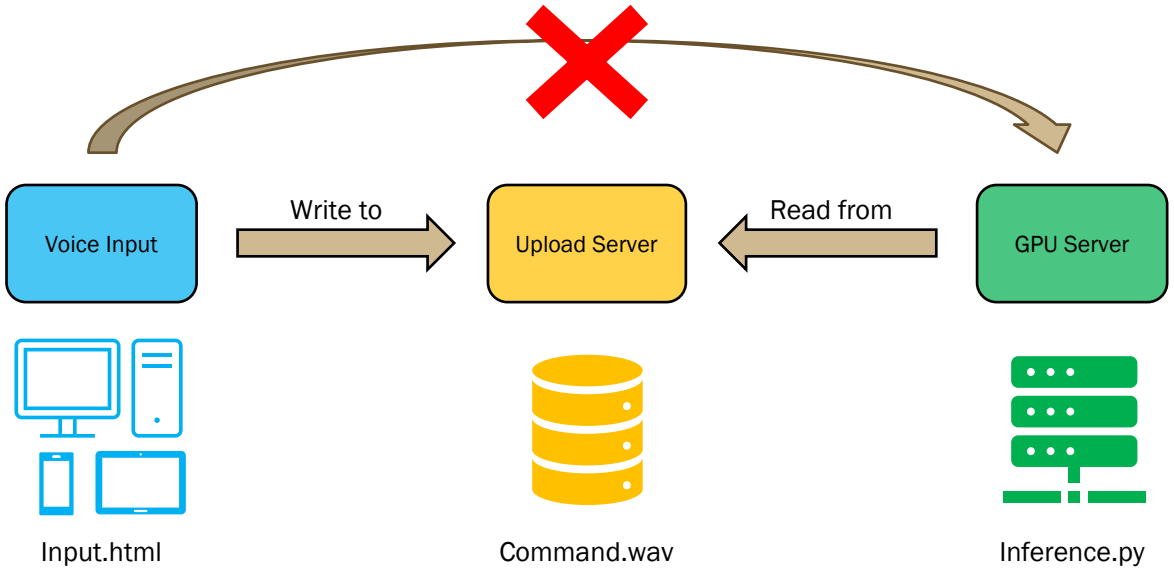
Objectives

- Build a swarm capable of executing various tasks
- Allow any users to control and interact with the swarm using natural language



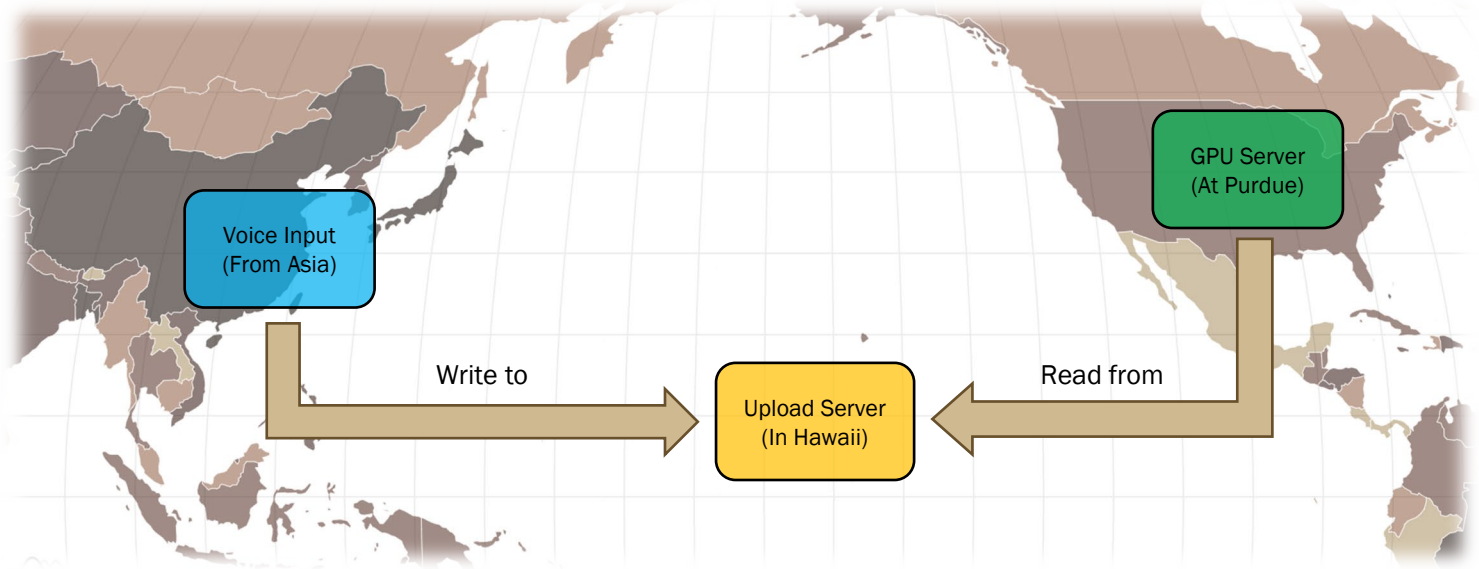
Approaches: Communication

Across different programming languages, operating systems, hardware, and locations



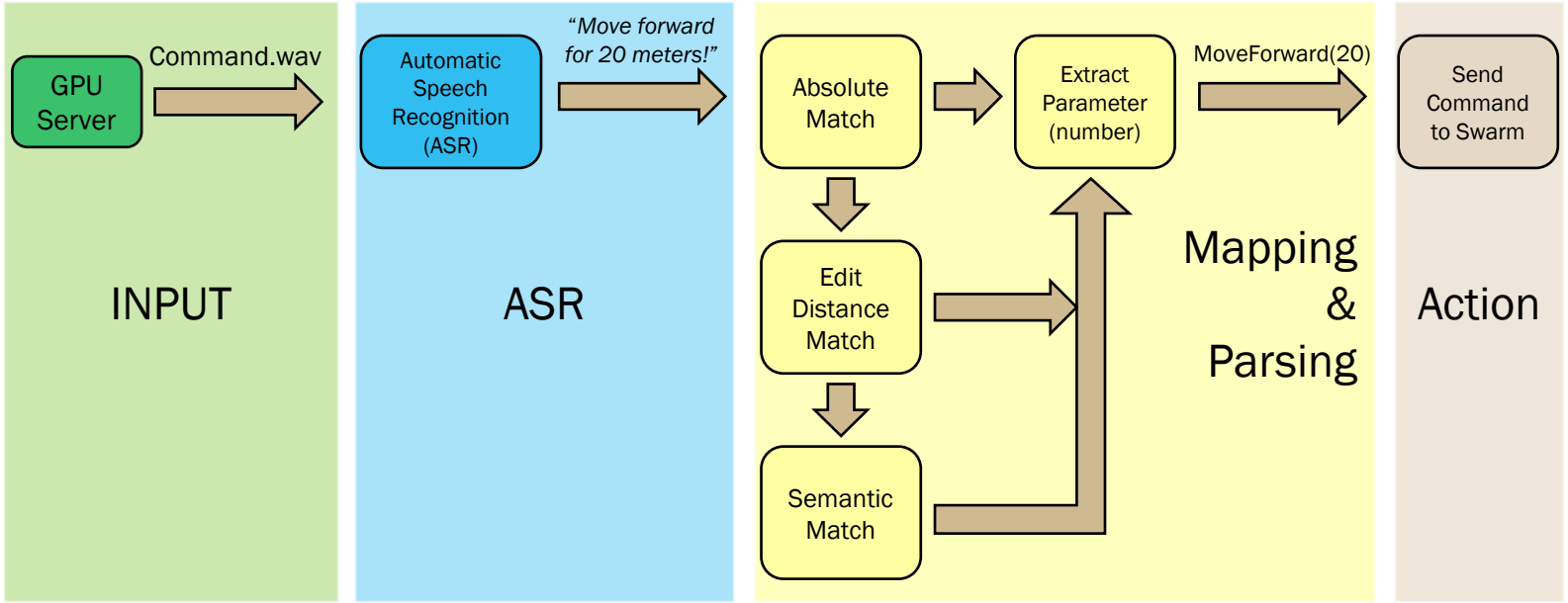
Approaches: Communication CONT.

Across different programming languages, operating systems, hardware, and locations



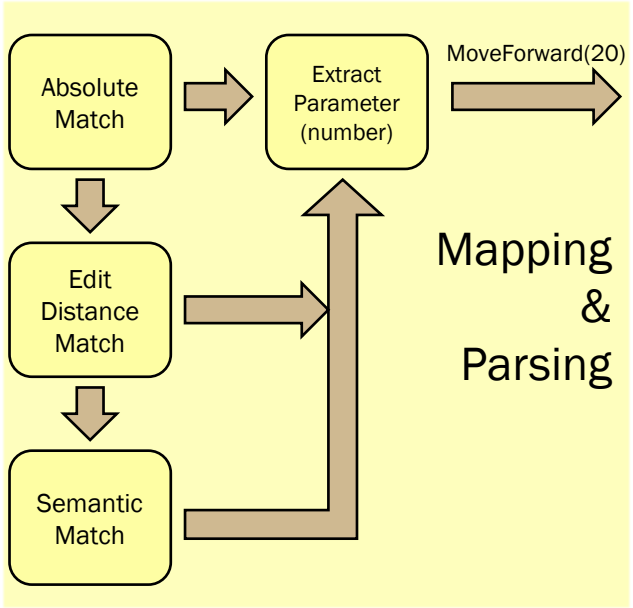
Approaches: Process the voice input

A hybrid model with multiple layers



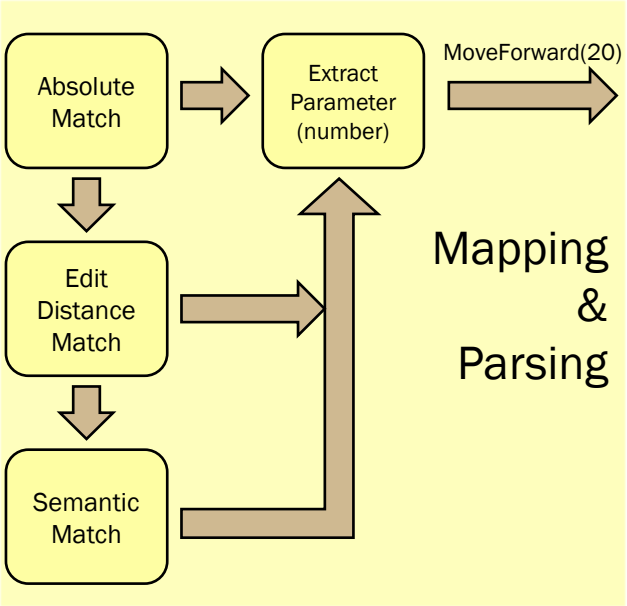
Approaches: Process the voice input

A hybrid model with multiple layers



Approaches: Process the voice input

The parsing and mapping module



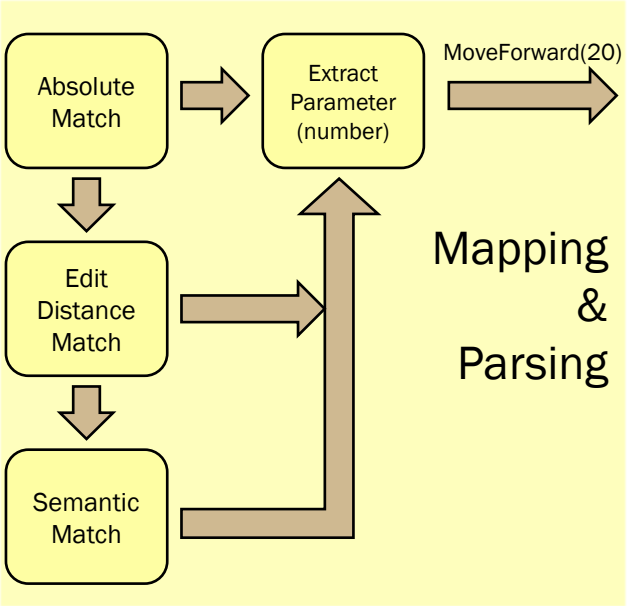
- Edit Distance [1] [2]
  - Compare the character-level differences of 2 given strings
  - “what” → “whatttt” = 3
- BERT-MRPC [3]
  - Compare the semantic similarity of 2 given sentences
  - “Speed up” = “Go faster” = “Move faster”
  - SOTA, but imperfect
- Stanza [4]
  - Recognize numbers in a given sentence

PURDUE UNIVERSITY | Polytechnic Institute

[1] Damerau, Fred J. "A technique for computer detection and correction of spelling errors." Communications of the ACM 7, no. 3 (1964): 171-176
[2] Levenshtein, Vladimir I. "Binary codes capable of correcting deletions, insertions, and reversals." In Soviet physics doklady, vol. 10, no. 8, pp. 707-710. 1966.
[3] https://huggingface.co/textattack/bert-base-uncased-MRPC
[4] Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. 2020. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. In Association for Computational Linguistics (ACL) System Demonstrations. 2020.

Approaches: Process the voice input

Add new commands



- The system is very extensive. New commands can always be added without programming.
- For one action, there can be multiple Text forms (to improve the matching accuracy)

#	Sample Text	Parameter
1	Turn left	Nullable
1	Make a left turn	Nullable
2	Move forward for # meters	Not Null
…	……	…
15	Hit the wall	Nullable

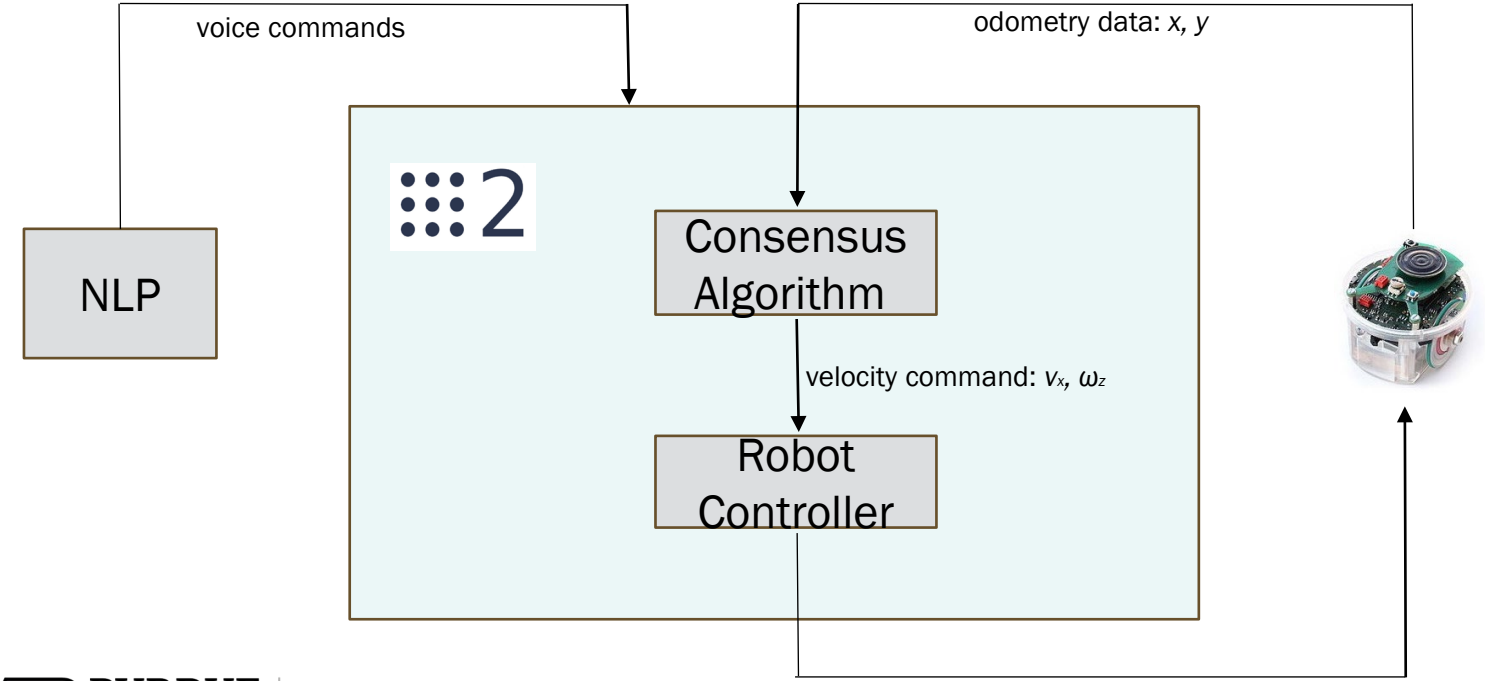
Approaches: ROS2 implementation

ROS1 vs ROS2

ROS1	ROS2
Supports single robot per ROS network	Supports multiple robots per ROS network
Architecture includes a central system controller "Master"	"Master"-less system with nodes capable of self-discovery
Concepts of QoS, Security absent	New concepts of QoS, Security introduced
Only CMake projects are supported	Supports multiple build systems with build isolation

Approaches: Swarm Control

What's happening after the swarm receiving command from the NLP pipeline



Result and Analysis

Demo video: simulation, real world experiments, failures, more failures, and success

https://www.youtube.com/watch?v=qRUkHuRrFbM

Conclusion

What have we achieved?

- A Human-Swarm Interaction System with natural language user interface
  - Successfully used voice commands to control the robots
- A modular, extensive framework with good compatibility
  - Compare to end-to-end solutions, a modular framework is easier to optimize and debug
  - The framework can work with various programming language and operating system
  - Adding new commands does not require any programming
- A chance to discover the engineering challenges in a research project
  - Implementing an idea could be more difficult than producing new ideas.
  - Making things work in the physical world requires more efforts than simulations

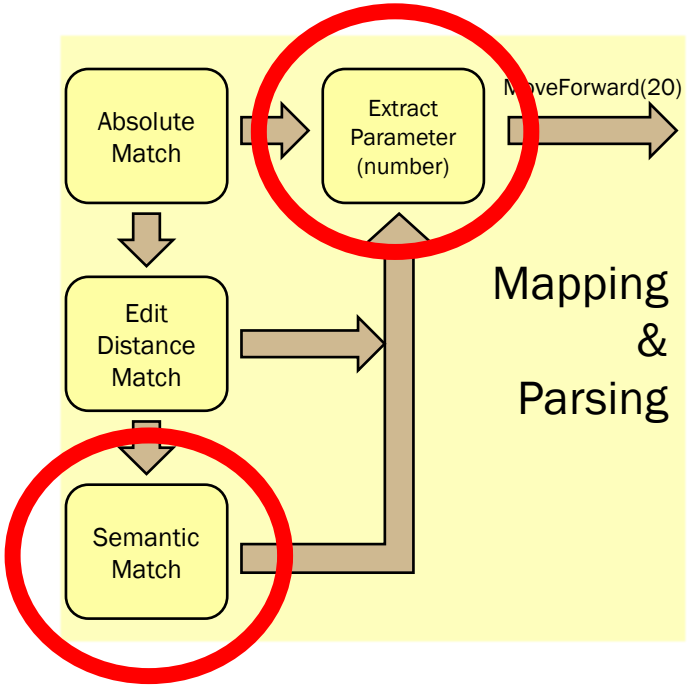
Conclusion CONT.

What did we fail to achieve?

- Bi-directional natural language communication
  - The swarm cannot respond in natural language
- Parsing more complicated tasks and fuzzy information
  - Tasks consist of multiple actions: “Bring Object A to XXX location without getting too close to Object B”
  - “Move forward a little bit.” How far is “a little bit”?

Future works

On the NLP side



- Fine-tune the BERT model with more data
  - The pre-trained language models can always be fine-tuned with new training data
  - Use generative models to create paraphrases of existing commands
- Parse the input unit (ex: from meter to feet)
  - Currently, our system only support 1 input unit

Future works

On the Robotics side

- Building multi-agent system in real life
  - Currently, we tested only a single agent
- Bigger drones with quad-decks that can lift heavy ogjects
  - We used only a Crazyflie 2.1 that can lift max 10 grams
- Controlled magnetic grabber
  - We tested uncontrolled magnetic grabber

Individual Contributions

Who did what

Tasks	Yifei	Roman
Project Planning	50%	50%
Literature Review	50%	50%
System Design	50%	50%
Swarm algorithm implementation	10%	90%
NLP algorithm implementation	90%	10%
Software Integration	50%	50%
Hardware implementation	1%	99%
Blog Maintenance	80%	20%

Thank you!

Roman Ibrahimov
Ph.D. student
Computer and Information Technology
ibrahir@purdue.edu

Yifei Hu
Ph.D. student
Computer and Information Technology
hu381@purdue.edu

polytechnic.purdue.edu

/ TechPurdue

PURDUE UNIVERSITY | Polytechnic Institute