

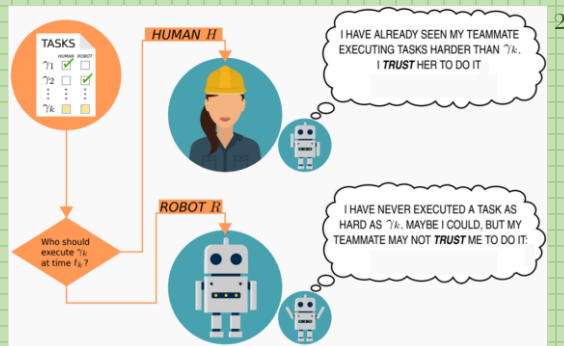
Trust in Human Robot Collaboration

Neelie Shah

Advised by Dr. Wollowski

Defining Trust

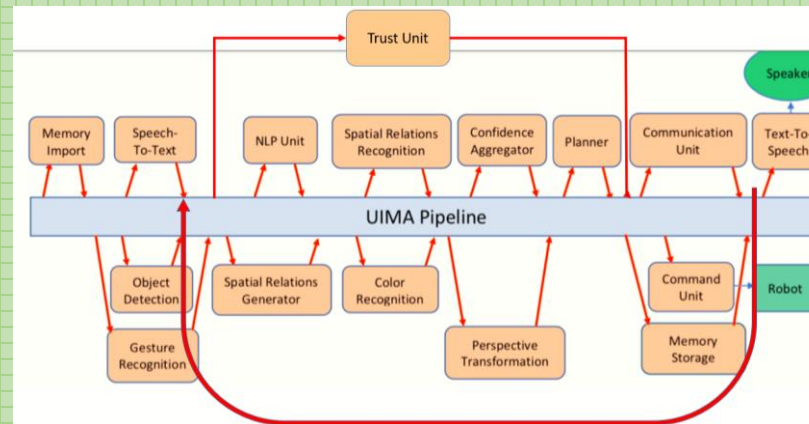
- Here, trust is defined as, “the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability”¹
- Use trust to build a better working balance with less errors



Refining Previous Code

- Unified project structure and repositories
- Transformation from Maven to Gradle
- Improvement in debugging logging
- Fixed JSON output generation of TextToSpeech Annotator
- Prototype speech output fixed
- Developer mode is now fully operational

Trust Unit



Integration of the Trust Unit into the current UIMA pipeline

- Trust unit keeps track of trust calibration
 - uses this to converse with the Human counterpart and make better judgements about delegation
- Mitigates the possibility of over or under reliance of the robot component**
- Simulates how artificial trust behaves with the robot
 - Successes of task accomplishments are tracked
 - Communicates with human counterpart to delegate tasks
 - Creates a trusting bond with human

Example Dialogue

What can I do for you today? (type a request, or type Q to quit)

pick up blue block and red block

Are you sure you want me to perform that task?

Given the importance of the task and my success rate in doing previous tasks, I trust you to it instead. Type Y if you would still like me to continue to do the task and N if you would like to do it yourself.

Y

- Given the importance of a task, the difficulty of a task, and the previous successes/failures with such a task, the system will be able to have a dialogue with the Human to let them know their opinion given their capabilities
- Metric to determine the type of notification to give the Human:

$$\text{Confidence} = \frac{[(\text{task 1 difficulty} \cdot \text{task 1 success}) + (\text{task 2 difficulty} \cdot \text{task 2 success}) + \dots]}{(\text{task 1 difficulty} + \text{task 2 difficulty} + \dots)}$$

<i>Proceed with task if:</i>	<i>Importance 1 confidence > 0.5</i>
	<i>Importance 2 confidence > 0.6</i>
	<i>Importance 3 confidence > 0.7</i>
	<i>Importance 4 confidence > 0.8</i>
	<i>Importance 5 confidence > 0.9</i>

References

¹Lee, J.D., and See, K.A. (2004). Trust in automation: Designing for Human Factors, 50-80.

²Azevedo-Sa, Herbert, Jessie Yang, Lionel P Robert, and Dawn M Tilbury. “A Unified Bi-Directional Model for Natural and Artificial Trust in Human–Robot Collaboration.” IEEE Robotics and Automation Letters, June 2021. <https://doi.org/10.1109/lra.2017.2691146>.