Research proposal

Enhancing Human-Robot Interaction through the Use of Verbal and Non-verbal Cues & Emotion Recognition: A study of the Effectiveness of Natural Language Understanding in Robot Learning

Francois Gonothi Toure
francoisgonothitoure@gmail.com

Research Master's candidate @ the University of Québec in Montréal - UQAM

Major: Master of Science in Computer Science
Minor: Electronic Systems / Artificial Intelligence
Canada

1 Keywords

Human-Robot Interaction, Verbal Cues, Non-verbal Cues, Emotion Recognition, Natural Language Understanding, Robot Learning, Effectiveness, Artificial Intelligence, Natural Language Processing, Computer Vision, Deep Learning, Reinforcement Learning, Machine Learning, Supervised Learning, Machine Learning Algorithms, Machine Learning Models, Algorithms, Robotics, Sensors, Input Devices, Social Robotics, Cognitive Robotics, Robotic agent, Humanoid Robots, Companion Robots, Autonomous Systems, Robot Learning Systems, Human-Machine Communication, Human-Computer Interaction, Human-Centered Design, Human-Centered Computing, Human-Robot Communication, Human-Robot Emotional Intelligence, Visuospatial Information-Processing, Language acquisition, Emotion Detection, Emotional Intelligence, Psychological Signals, Neuro-Cognitive Architectures, Problem-Solving.

2 Executive Summary

This research proposal explores the design and implementation of a machine learning algorithm to enhance human-robot interaction through the use of verbal and non-verbal cues and emotion recognition. The ability for robots to learn new skills from human tutors has the potential to greatly expand their capabilities and increase their utility in a variety of settings. However, traditional methods of robot learning rely heavily on pre-programmed rules and tasks, which can be time-consuming and inflexible. In this proposal, we propose a new approach to robot learning that utilizes human tutor recognition of verbal and non-verbal inputs and emotions to enhance the learning process. By using machine learning techniques to interpret a human tutor's verbal instructions, facial expressions, and body language, the robot will be able to learn new skills more efficiently and effectively. This suggested research work will be completed by conducting in-depth research in the field of Natural Language Processing (NLP) and Robotics, developing a final prototype of the real secure machine learning algorithm and testing the intelligent agent on different types of humanoid robots to get a desired outcome. Also, this research will be answering research questions like: How does the use of NLP and robotics in the proposed research compare to current approaches in these fields? What are the key challenges and opportunities in using machine learning algorithms for secure natural language processing in the context of humanoid robots? What are the potential practical applications and implications of the intelligent agent in the context of humanoid robots? How can the results of this research contribute to the broader field of NLP and robotics? can emotions detection and language ability improve the ability of a robot to learn from real world?

3 Survey of Background Literature

3.1 Overview

In recent years, there has been a significant increase in the development of robots that can perform tasks and assist humans in various settings. However, most of these robots are designed to perform pre-programmed tasks and do not have the ability to adapt and learn new skills on their own. This limits their capabilities and hinders their

potential to be useful in a wide range of settings. One way to improve the capabilities of robots is to enable them to learn from humans. This can be done through direct human-robot interaction, where the human serves as a tutor and the robot learns by observing and imitating the human's actions. However, for this to be effective, the robot must be able to recognize and interpret the verbal and non-verbal inputs and emotions of the human tutor. As a result, I believe it is noteworthy to have an optimal and secure machine learning algorithm with high accuracy that can be embodied into a humanoid robot that not only assists humans, but also learns new skills from their verbal and non-verbal cues and emotions.

3.2 Objectives

The primary objective of this research is to develop a robot learning system that is able to recognize and interpret a human tutor's verbal and non-verbal inputs and emotions in order to learn new skills more effectively. This will be accomplished through the following specific objectives:

1. Develop and train machine learning algorithms to recognize and interpret a human tutor's verbal instructions, facial expressions, and body language.
2. Integrate the machine learning algorithms into a robot learning system that can adapt and adjust its learning approach based on the tutor's inputs and emotions.
3. Test and evaluate the effectiveness of the robot learning system in a variety of scenarios and tasks.

3.3 Impact

The ability for robots to learn new skills from human tutors has the potential to greatly expand their capabilities and increase their utility in a variety of settings. This research will contribute to the development of more advanced and intelligent robots that can learn and adapt in a more natural and human-like manner. The results of this research will have significant implications for a wide range of industries, including manufacturing, healthcare, and education, where robots can play an increasingly important role in assisting and augmenting human tasks.

3.4 Expected Results

It is expected that the results of this research will provide a better understanding of the role of verbal, non-verbal, and emotional cues in enhancing human-robot interaction for the purpose of robot learning. This will include identifying the most effective combinations of cues and the ways in which they influence the learning process. Additionally, the results of the interviews with human tutors will provide valuable insights into the subjective experiences of the tutors and how they perceive the influence of these cues. We expect that the robot learning system developed in this research will be able to significantly enhance the learning process by accurately recognizing and interpreting a human tutor's verbal and non-verbal inputs and emotions. This will allow the robot to adapt its learning approach and better understand the intent of the tutor, resulting in more efficient and effective learning of new skills.

4 Proposed Methodology

The overall mechanism will be carried as problems and their solutions throughout the research process till I submit my work for review and approval. First, the robot needs to be able to recognize and understand the verbal input of the human tutor. This can be done through natural language processing techniques, such as speech recognition and language parsing. The robot should also be able to recognize and understand non-verbal inputs, such as gestures, facial expressions, and body language. This can be done through machine vision and image recognition techniques. The robot should also be able to recognize and understand the emotions of the human tutor. This can be done through techniques such as facial expression recognition and voice emotion recognition. Once the robot has recognized and understood the verbal, non-verbal, and emotional inputs of the human tutor, it should use this information to infer the intended meaning and action of the tutor. The robot should then attempt to perform the task or action that it has inferred from the tutor's inputs. The robot should observe the result of its actions and compare it to the expected outcome. If the result is successful, the robot should store this information and continue to the next task. If the result is unsuccessful, the robot should seek additional input or clarification from the tutor. The robot should continue this process of receiving input, inferring meaning, and attempting tasks until

it has learned all of the skills it was meant to learn from the human tutor. Most importantly, I will create a novel cognitive model to supervise the mentioned cognitive activities and this cognitive model will be designed to analyze and interpret the various cues and emotions present during human-robot interactions, and use this information to guide the robot's learning and decision-making processes. For example, the model could help the robot to recognize when a human is happy, sad, angry, or neutral, and use this information to adjust its own behavior accordingly.

Initial description of the research methodology:

- First, the robot should be equipped with sensors or other input devices that allow it to recognize and interpret verbal and non-verbal inputs from the human tutor, such as speech, gestures, facial expressions, and posture.
- The robot should also be equipped with a machine learning model that can classify the tutor's emotions based on their verbal and non-verbal inputs.
- As the human tutor demonstrates the skill to the robot, the robot should use its input devices to gather data on the tutor's verbal and non-verbal inputs.
- The robot should then use its machine learning model to classify the tutor's emotions based on this data.
- If the robot detects that the tutor is feeling positive (e.g. happy or proud), it should continue to observe and learn from the tutor.
- If the robot detects that the tutor is feeling negative (e.g. frustrated or annoyed), it should remember the experience as a negative example for the future.
- The robot should continue this process until it has successfully learned the skill from the tutor.
- Once the skill has been learned, the robot should be able to demonstrate it to the tutor for feedback. If the tutor is satisfied with the robot's performance, the learning process is complete. If not, the robot should continue to adjust its learning strategy and seek further feedback until the tutor is satisfied with its performance and each example should be store in a long term memory of the cognitive model for future tasks.

5 Research Plan

The best way to break the work into smaller chunks is to spread them out throughout the six (4) semesters of my MS program as shown below. The first and second semesters will be spent conducting in-depth research on my research topic and taking classes to add new skills and improve existing ones for better and improved research, while the third and fourth semester will be spent implementing first and then final prototypes of the machine learning algorithm to enhance human-robot interaction through the use of verbal and non-verbal cues and emotion recognition, and finally submitting my final work to my supervisor(s) for review and approval before moving on to publishing. To create a robot that can learn skills from a human tutor by recognizing their verbal and non-verbal inputs as well as their emotions, I will likely need a strong background in a number of technical areas, including: Artificial intelligence (AI) and machine learning, Natural language processing (NLP), Computer vision, Robotics, Human-robot interaction, Emotional intelligence.

Semester 1:
Month 1: Conduct Literature review of related research papers.
Month 2: Develop and test sensors or other input devices that allow the robot to recognize and interpret verbal and non-verbal inputs from the human tutor, such as speech, gestures, facial expressions, and posture.
Month 3: Develop and test a machine learning model that can classify the tutor's emotions based on their verbal and non-verbal inputs.
Month 4: Begin demonstrating skills to the robot and gathering data on the tutor's verbal and non-verbal inputs.
Month 5: Continue demonstrating skills to the robot and using the machine learning model to classify the tutor's emotions based on this data.
Month 6: Continue demonstrating skills to the robot and using the machine learning model to classify the tutor's emotions based on this data & Submission of a journal publication.

Semester 2:
Month 1: Conduct Literature review of related research papers.

Month 2: If the robot detects that the tutor is feeling positive, continue to observe and learn from the tutor. If the robot detects that the tutor is feeling negative, remember the experience as a negative example for the future.
Month 3: Continue this process until the robot has successfully learned the skill from the tutor.
Month 4: Begin demonstrating the skill to the tutor for feedback. If the tutor is satisfied with the robot's performance, continue to the next skill. If not, continue to adjust learning strategy and seek further feedback until the tutor is satisfied with performance.
Month 5: Continue demonstrating skills to the tutor and seeking feedback until all skills have been learned.
Month 6: Begin designing the novel cognitive model to supervise the cognitive activities and analyze and interpret the various cues and emotions present during human-robot interactions & Submission of a journal publication.

Semester 3 & 4:
Month 1: Conduct Literature review of related research papers.
Month 2-3: Test the cognitive model and refine as needed. Also, Integrate the cognitive model into the robot's learning and decision-making processes.
Month 4-5: Test the completed robot's ability to recognize and understand the verbal, non-verbal, and emotional inputs of the human tutor and perform tasks based on these inputs.
Month 6-7: Gather data on the robot's performance and use it to make any necessary improvements. Also, Conduct final testing and gather data for the research proposal. And Submission of a journal publication.
Month 8-12: Thesis writing.

6 Conclusion

Overall, this research has the potential to significantly enhance our understanding of the role of verbal, non-verbal, and emotional cues in human-robot interaction and how they can be used to improve the learning capabilities of robots. By developing robots that are better able to recognize and interpret these cues, we can expand their potential applications and make them more effective assistants to humans.

7 References

- Han, J., Campbell, N., Jokinen, K. and Wilcock, G., 2012, December. Investigating the use of non-verbal cues in human-robot interaction with a Nao robot. In 2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom) (pp. 679-683). IEEE.
- Lee, J., 2017. A survey of robot learning from demonstrations for human-robot collaboration. arXiv preprint arXiv:1710.08789.
- Green, S.A., Billinghurst, M., Chen, X. and Chase, J.G., 2008. Human-robot collaboration: A literature review and augmented reality approach in design. International journal of advanced robotic systems, 5(1), p.1.
- Saran, A., Short, E.S., Thomaz, A. and Niekum, S., 2019, March. Enhancing robot learning with human social cues. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 745-747). IEEE.
- Mohammed, S.N. and Hassan, A.K.A., 2020. A survey on emotion recognition for human robot interaction. Journal of computing and information technology, 28(2), pp.125-146.