POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Steps toward the implementation of an improvisational robot

Laurea Magistrale in Computer Science Engineering - Ingegneria Informatica

Author: Aniello De Santis

Advisor: Prof. Andrea Bonarini

Academic year: 2022-2023

1. Introduction

Nowadays, programmable machines as robots are replacing humans in a variety of tasks. Besides being employed as industrial tools, a growing body of research and commercial applications is envisioning robotic social companions for everyday life. The goal of a robotic social companion able to interact with humans in *unconstrained* settings is, however, still far from being achieved, due to both technical and ontological problems. For this purpose the art of theatre is often employed as a testing ground, being it a representation and simulation of reality, and in particular of social interactions. The objective of this master thesis is to further the research on a robot named "Robocchio," which was initially developed in AIRLab by Lorenzo Bonetti [1] and improved with a recognizing system able to perceive actor's actions by Lorenzo Farinelli [4], and a behavioural model of reaction by Claudia Chiroli [2]. At the start of my thesis, Robocchio already had all the necessary modules for improvisation in a controlled context; they simply needed to be integrated with each other. Following a period of maintenance and integration, the system was prepared to progress to the next phase of research. Specifically, this study explores three main directions:

1. The redesign and implementation of reaction features performed by Robocchio.
2. Enhancements to the software and hardware systems to improve the robot's lifelikeness.
3. The integration of an innovative behavioural model using a *Machine Learning*(ML) approach.

2. The status quo

Before delving into the details of our research steps, it is crucial to provide a brief overview of the starting point of our work. As mentioned in the introduction Section 1, our initial focus involved integrating pre-existing modules and making minor modifications.

2.1. Hardware

The robot, Robocchio, is designed with a structure reminiscent of a main character from an animated film. Standing at a height of 1 meter, it features two protruding eyes that resemble those of snails. The robot is supported by a mobile base comprising three DC motors, which enable holonomic motion. This motion system allows Robocchio to navigate in space by combining both linear and angular movements. In terms of expressing emotions, Robocchio possesses the capability to move various limbs. Its eyes are connected to two servo-motors each, which, con-

Figure 1: Robocchio's front image.

trolled by a microcontroller, enable linear movements in four primary directions: right, left, up, and down. Additionally, there are four additional servo-motors responsible for controlling body movements. For a visual representation of Robocchio, please refer to Figure 1.

2.2. Software

Concerning the invisible parts of the robot, it was developed in ROS framework. At the outset of my research, two sub-modules were already in place: a perception system and a reaction system.

The perception module is designed to perceive features from the stage where interactions take place. Its main objective is to categorize the actions performed by the person interacting with the system.
The variables considered are:

- The emotion, expressed through the face and the body of the actor;
- The proximity, which is the distance between interlocutor and the robot, and characterize the familiarity between them [5];
- The proxemic movement of the actor with respect to the robot in terms of direction and speed.

The above mentioned extracted features are combined together with a Fuzzy system to generate a discrete classification of actor action. The defined output space is composed of 21 scenic actions, they are the following: Attack, Intimidation, Scolding, Holding Grudge, Refuse, Perplexity, Share Fear, Caution, Hesitancy, Shock, Escape, Share Joy, Greet, Happy Person, Satisfaction, Share Sadness, Sad Person, Disappointment, Share Surprise, Astonishment, and Disbelief.

The output system is composed of all the components responsible for processing the information about the scenic action happening on the stage and generating an output. A reaction refers to a sequence of atomic movements executed by Robocchio's limbs and base. At the initial stage of our project, we had an existing behavioural model. This model employs a *Non-Deterministic Finite Automaton* (NFA) to determine the appropriate reaction. A behaviour is associated with a specific personality and reactions have been qualitatively defined per each behaviour.
Our contribution consists of three main software improvements:

- One of the improvements we made to the originally designed reactions was the introduction of parallelism through multithreading. Initially, the atomic movements of Robocchio's limbs were performed sequentially, one after another. By implementing parallelism, we enabled simultaneous execution of multiple limb movements. This enhancement makes the robot's movements appear more fluid and lifelike, closely resembling human actions. Parallel execution allows for a more dynamic and responsive behavior, adding to the overall perception of the robot being "alive".
- The second main aspect in which we have enhanced Robocchio is in the robot-actor alignment algorithm. We discovered high inaccuracies in the original algorithm as it only considered the actor's position at the beginning of the movement, without considering any subsequent moves. To overcome this limitation, we integrated a closed loop algorithm that incorporates the webcam frame to achieve a more precise and continous alignment between the robot and

the actor.

- In conclusion, modifications to the map navigation system have been applied. The original system implemented a virtual map navigation approach. Virtual boundaries were defined at the beginning of the sketch, confining the robot's actions within those limits. Due to restrictive security boundaries and localization inaccuracies, the system caused several false positives, bringing interruptions in the robot's reaction. The navigation system was revolutionalized removing virtual boundaries, and implementing a simple obstacle detection system exploiting laser data. This new system ensures that the robot's reactions are never interrupted unless it is moving towards an obstacle. In such cases, only the base movement is temporarily halted, without affecting the other atomic moves of the robot. This modification allows for smoother and uninterrupted reactions, improving the overall performance and user experience of Robocchio.

2.3. Integration

Once the two sub systems were available, we proceeded with making minor changes to the code. Additionally, we discovered that the expected set of actions from the reaction system was slightly different from the initially designed ones. To ensure coherency and consistency we remapped actor actions and robot reactions within a unified discrete space of cardinality 14.
The final action space is:

1. Attack
2. Scolding
3. Intimidate
4. Grudge
5. Sharing Happiness
6. Happy Person
7. Satisfaction
8. Sharing Fear
9. Running Away
10. Sharing Sadness
11. Disappointment
12. Surprise
13. Disbelief
14. Astonishment

3. Strategy development

When improvisation system was ready, our ideas were put into practice. In this section, we further explore the details of our two main advances.

3.1. Reactions redesign

The first step is the redefinition of robot reactions. Since initially defined reactions were deemed overly simplistic, our objective was to introduce greater complexity by adhering to a set of criteria outlined in the relevant literature on expressive behaviours. Main theories considered in the research are:

- The Ekman Model of emotion [3] which discretize emotions into anger, disgust, fear, happiness, sadness, and surprise;
- The circumplex model of affect [7] which modelize emotions depending on arousal (a sleepiness-alertness continuum) and valence (a pleasure-displeasure continuum);
- Laban movement analysis [6] which divides emotion expressions into four categories: body, shape, space and effort.

The research's literature has been enriched with the twelve Disney principles of animation [8], with the objective of adding liveliness to our robot.
In addition, with the aim of imbuing the robot with a more lifelike character, we have devised two simple routines for the robot to execute during moments of environmental perception. Specifically, we have incorporated straightforward body movement routines into the robot's behavior.

3.2. Reinforcement Learning behaviour

Simultaneously, a new behavioural model for Robocchio has been engineered. The main concept behind this model is enabling the robot to learn, through Machine Learning, from the environment a reaction policy coherent with the context. The improvisational environment can be represented as a Markov Decision Process (MDP) in which:

- A state S_t is defined as a pair of robot reaction rr_t and actor action aa_t performed at that specific time t.
- An action A_t, represents one of the fourteen reactions the robot can perform at time t.

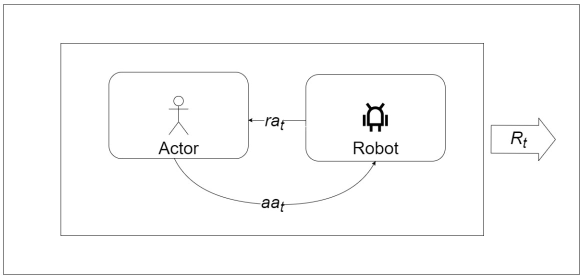

Figure 2: Modelized improvisational framework.

- A probability matrix P summarizes the actor behaviour probability, given a particular state and robot reaction.
- A reward r, associated with the coherency degree of certain robot reactions.
- A final state is defined as the fifteenth scene that occurs in the interaction. This choice is based on the assumption that significant human-robot interactions are typically brief and unlikely to exceed fifteen interactions.

Figure 2 summarizes the designed improvisational model.
Before training RL algorithms the first step was the data acquisition process. The decision was made to design an interactive survey where the audience could express their opinions on a video featuring Robocchio engaging in random interactions with an actor. This approach allowed to collect positive and negative opinions on each scene. The coherence votes provided by the audience were subsequently translated into an algebraic sum, representing the reward associated with a particular state within the improvisational MDP.

4. Experiments

After that the guidelines of our project were established, we proceeded with the execution of our experiments. The initial phase involved conducting a validation experiment to assess the effectiveness of the reactions' redesign. Following this validation, we embarked on the construction of a comprehensive dataset to serve as input for our RL algorithms. All video resources used in the surveys can be found at [1].

[1] https://github.com/Aniello98/ImprobotProject/blob/main/Video_Links.md

```
{
    "attack": "intimidate",
    "scolding": "scolding",
    "intimidate": "astonishment",
    "grudge": "scolding",
    "sharing_happiness": "sharing_happiness",
    "happy_person": "scolding",
    "satisfaction": "satisfaction",
    "sharing_fear": "sharing_fear",
    "running_away": "sharing_sadness",
    "sharing_sadness": "disappointment",
    "disappointment": "grudge",
    "surprise": "surprise",
    "disbelief": "disbelief",
    "astonishment": "sharing_sadness"
}
```

Figure 3: Saved answer from the reactions survey, representing a dictionary where keys are the actual reaction performed and values are the guessed ones.

4.1. Reactions validation

To assess the extent to which Robocchio fulfils action expectations, we employed an empirical approach. An online survey was conducted, involving individuals randomly selected from the general population, to gather their opinions. In this survey, a series of videos showcasing different robot reactions were recorded. Participants were presented with randomized sequences of reactions and were asked to identify the corresponding action performed by the robot for each sequence. At the end of the experiment, 45 participants' responses were collected and saved as JSON files, as shown in Figure 3.

4.2. RL preparation and training

After the analysis done on the reactions, we proceeded with the next activity: data collection. To gain information about the defined RL improvisational environment another online survey has been done along the lines of the previous one, with the difference that this time the robot interacts with an actor. The audience in this case is urged to interactively react to what they are seeing. Particularly, the user sees one of the four improvisational videos recorded and successively edited at AIRLab. Each video is labelled with a series of improvisational states happening at a defined time range. The scope of the survey is to make the audience vote the scenes through two buttons: "yes" and "no", expressing the coherency in Robocchio's reactions. Each interaction has been recorded and transformed in a reward point added or subtracted to the associated MDP state action pair.

actor action / robot reaction	Attack	Scolding	Intimidate	Grudge	Sharing Happiness	Happy Person	Satisfaction	Sharing Fear	Running Away	Sharing Sadness	Disappointment	Surprise	Disbelief	Astonishment
Attack	0	0	12.5%	0	12.5%	12.5%	12.5%	12.5%	0	25%	12.5%	0	0	0
Scolding	14.3%	14.3%	0	0	0	0	0	14.3%	0	28.6%	14.3%	0	0	14.3%
Intimidate	0	0	25%	8.3%	8.3%	16.7%	8.3%	8.3%	0	25%	0	0	0	0
Grudge	0	0	12.5%	0	0	0	0	12.5%	0	50%	25%	0	0	0
Sharing Happiness	0	10%	0	0	10%	20%	0	20,00%	0	20%	10%	0	10%	0
Happy Person	0	8.3%	8.3%	0	8.3%	33.3%	0	8.3%	0	0	16.7%	8.3%	0	8.3%
Satisfaction	0	0	25%	0	0	0	12.5%	0	0	37.5%	12.5%	12.5%	0	0
Sharing Fear	0	0	20%	0	10%	0	10%	40,00%	0	20%	0	0	0	0
Running Away	0	0	0	0	0	33.3%	0	0	0	66.6%	0	0	0	0
Sharing Sadness	0	8.3%	16.7%	8.3%	0	8.3%	0	16.7%	0	16.7%	0	16.7%	8.3%	0
Disappointment	0	0	0	8.3%	8.3%	25%	0	8.3%	0	33.3%	8.3%	0	8.3%	0
Surprise	0	0	0	5.3%	5.3%	26.3%	5.3%	10.5%	0	26.3%	10.5%	5.3%	0	5.3%
Disbelief	11.1%	11.1%	0	0	11.1%	11.1%	0	0	0	33.3%	22.2%	0	0	0
Astonishment	0	12.5%	16.6%	4.17%	8.3%	12.5%	12.5%	0	0	16.7%	8.3%	0	0	8.3%

Figure 4: Table representing the estimated matrix probability P belonging to the environment. Each entry corresponds to $\hat{P}(aa_i|ra_i)$.

To model the environment we needed to estimate the P matrix which represents the actor's behaviour after a robot reaction. The formalization of this element has been done by analyzing the actor's behaviour during the experiments. Furthermore, through improvisational samples, we have computed the frequency of instances in which, given a robot reaction rr_i, the actor performed an action aa_i.
Estimated results are in Figure 4.

$$\hat{P}(aa^i|ra^i) = \frac{\#(ra^i, aa^i)}{\#(ra^i, aa)}$$

After the datatset preparation, we started to apply ML techniques. The focus is on a Reinforcement Learning algorithms from which the robot can maximize rewards associated with a coherency measure of reaction. Main state-of-the-art solutions have been used, such as Q-Learning, SARSA, Expected-SARSA and Double Q-Learning. Several hyperparameter combinations of these have been tested, and the chosen solution was on a Double Q-Learning algorithm trained with an $\epsilon - greedy$ policy and a linear annealing exploring rate. The discriminative factors have been the computed average reward of the model and the policy distribution of actions. Two different trained models have been taken for the analysis, they are the same but the bootstrapping parameter n that determines the length of a sampled episode considered for the update of a Q-value. The final choice and results are further discussed in Section 4.2.

5. Discussion of results

5.1. Expressivity measures

To understand how humans perceive Robocchio's actions, an analysis on the guessed actions and emotions have been led. To have an idea of the distribution of guesses, we have exploited the confusion matrix. The resulting matrix in

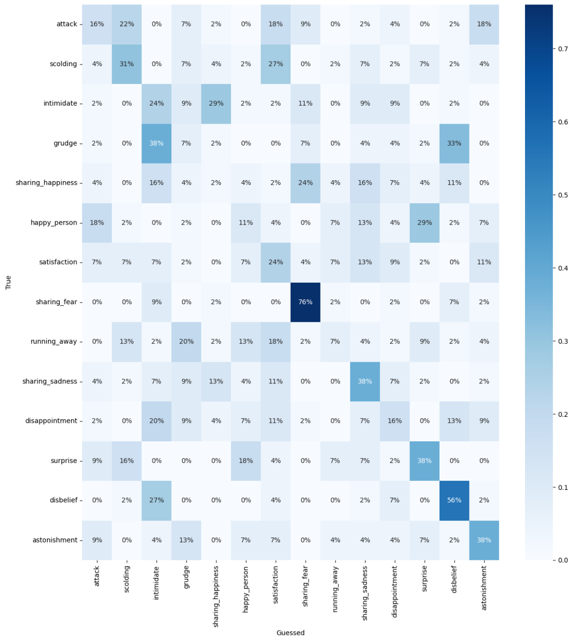

Figure 5: Confusion matrix of action per action analysis representing on raws the actual action, and on the columns the guessed one.

Figure 5 highlights that certain actions such as "scolding", "sharing fear", "sharing sadness", "surprise", "disbelief", and "astonishment" were generally easier to recognize correctly, indicating that they possessed distinguishing features. However, some actions were highly confused such as "attack" and "scolding", "intimidate" and "sharing happiness", "grudge" and "disbelief", and "disbelief" and "intimidate." Possible explanations for the confusion include the physical appearance of the robot, lack of clear distinguishing features, and low arousal expressed in certain actions. Surprisingly, the "scolding" action was often confused with "satisfaction", despite lacking shared movements or characteristics. Furthermore, participants associated the "sharing happiness" action with fear, possibly due to the quick and asynchronous eye movements of the action.

5.2. Learned behaviour

The choice of the algorithm has been influenced by the average reward per episode and the distribution of the actions taken in the learned policy. The chosen one is summarized in Figure 6.
Some actions were rarely taken, possibly due to recognition difficulties or inadequate coverage in the dataset. Specifically, the Running

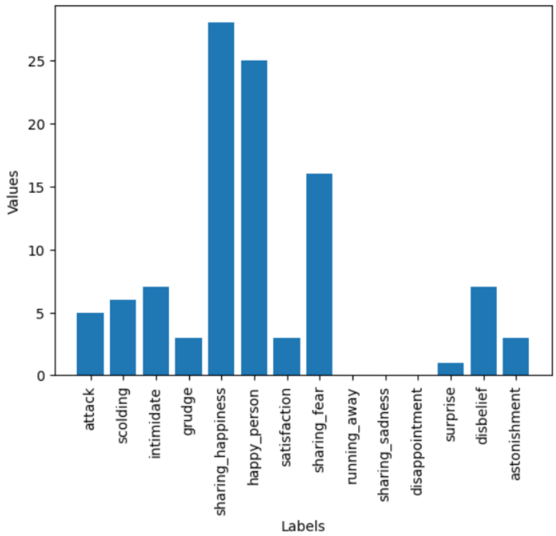

Figure 6: Histogram of the action distribution of the learned policy. Values are associated to count of states in which the action is taken.

Away and Disappointment actions received low guess scores during reaction validation, making it challenging for users to understand the agent's behavior. Additionally, the Sharing Sadness action had low rewards in the dataset, suggesting it may not be relevant in the improvisation context. On the other hand, Sharing Happiness, Happy Person, and Sharing Fear are found to be the most preferred ones. This finding implies that these actions are well-suited for a wide range of improvisational states. This analysis can be further considered to evaluate a possible redefinition of the improvisational reaction space. Overall, this analysis provides valuable insights that can guide further evaluation and potential adjustments in the design of the robot's improvisational reactions.

6. Conclusions

In conclusion, our research endeavors have led to significant advancements in the improvisational system of Robocchio. Throughout our study, we successfully addressed various weaknesses, including navigation limitations, obstacle detection, and parallel movements. These improvements have significantly enhanced the overall performance and fluidity of the robot's actions.

Furthermore, we have introduced a more sophisticated and adaptable framework for complex reactions, allowing for metric-wise modifications. This flexibility enables fine-tuning and customization of the robot's behavior according to specific requirements and contextual nuances.

One of the key achievements of our work is the development of a behavioral model for improvisation. This model has been derived from human votes, providing valuable insights into human perception and preferences. It opens up exciting avenues for future research opportunities, allowing for deeper investigations and refinements in understanding the dynamics of human-robot interactions.

In summary, our research efforts have culminated in a significantly improved improvisational system for Robocchio, addressing various limitations and introducing novel features. The availability of a sophisticated behavioral model and the successful mitigation of previous weaknesses pave the way for further advancements and exploration in the field of human-robot interaction.

References

[1] Lorenzo Bonetti. Design and implementation of an actor robot for a theatrical play. 2021.

[2] Claudia Chiroli. Simulation of emotional behaviour in a robot. 2021.

[3] Paul Ekman. An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200, 1992.

[4] Lorenzo Farinelli. Design and implementation of a multi-modal framework for scenic actions classification in autonomous actor-robot theatre improvisations. 2021.

[5] E. T. Hall and E. T. Hall. *The Hidden Dimension*, volume 609. Anchor, 1966.

[6] Rudolf Laban. Modern educational dance. 1963.

[7] James Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 12 1980.

[8] F. Thomas and O. Johnston. *The Illusion of Life: Disney Animation*. Disney Editions, 1995.