Developmental Robotics
Final report

Djamari Oetringer - s4464559 Tjalling Haije - s1011759
Paul Koch - s1015256

December 5, 2018

Abstract

The aim of this project is to implement joint attention using a Nao robot. Joint attention is achieved by following the gaze of a given person. The final model learned to follow the gaze using Q-learning. Furthermore, multiple principles of developmental robotics have been implemented. All code can be found on Github.[1]

1 Introduction

For the course Developmental Robotics our task was to create a developmental model to learn joint attention and run it using a NAO robot. For this purpose, we used an already existing model to detect the gaze [1]. Here, joint attention is defined as the robot looking at the same object as a human is looking at.

Figure 1 shows the final pipeline created for this project. This pipeline visualises the behaviour of the robot during one trial. The following list describes the individual components of the pipeline.

- Robot looks around: the robot follows a predefined path of head movements in order to search for a face.
- Robot turns toward face: given that a face has been found, the robot will centre the face, such that the face is at the centre of the image.
- Establishing eye contact: the robot waits until the gaze of the human is pointing straight at the robot. Thereby, eye contact is established.
- Image capture: an image is captured from the robot camera and stored such that gaze and object detection can be performed.
- Face detection: the face closest to the centre of the robot camera image is returned and saved in the memory.

[1] `https://github.com/KochPJ/follow_gaze_developmental_robotics`

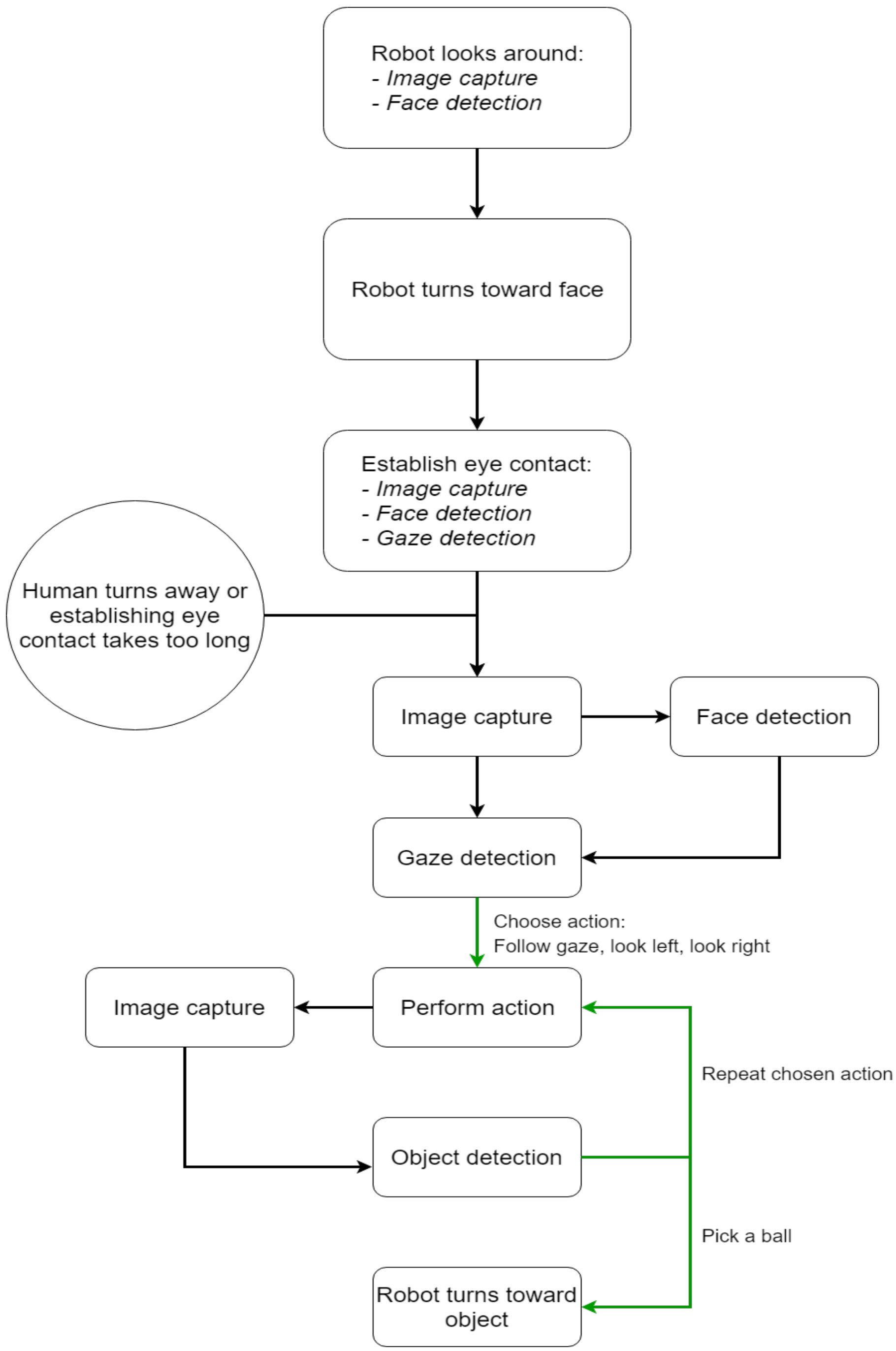


Figure 1: The final pipeline. The italic words are the subcomponents that the robot performs continuously while within the main component. The green arrows and box indicate that at those points, the robot has to make a decision.

- Gaze detection: the gaze of the given face is returned and saved in the memory.
- Perform action: given the chosen action, the robot looks left, right, or follows the gaze.
- Object detection: given the captured image, all objects are returned.
- Robot turns toward object: given a selected object, the robot centres that objects and points with an arm toward it.

2 Inspiration for the implementation

Studying the development of joint attention in infants has lead us to add a motivation system to the architecture, which mimics the preference of infants to look at faces [2], to perform an action if the other person breaks mutual eye contact [3], and to look at interesting objects [4]. [5] successfully used such a motivation system which generated motivations for the system to do certain actions, based on valences and external stimuli. The manner in which the motivation system is created and works is purely pragmatic; it is solely the motivations which are generated for certain situations which are developmentally inspired.

Furthermore, our implementation is in general based on multiple developmental principles that we tried to achieve. More about this can be found in section 3.2.

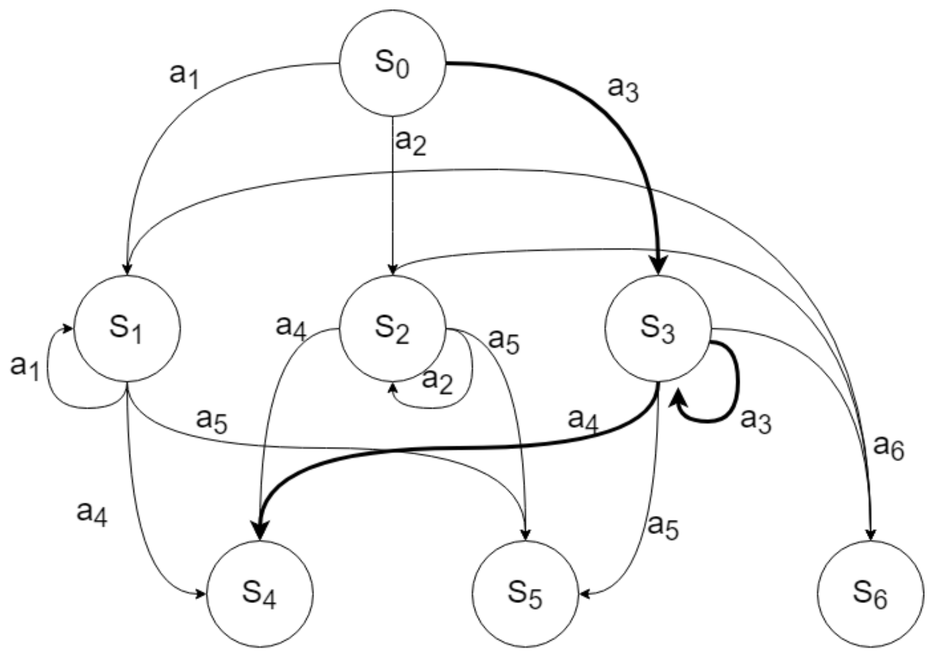
3 Aspects of developmental psychology

In short, we implemented reinforcement learning to try to achieve multiple developmental principles. In this section, we will explain our reinforcement learning model and the achieved developmental aspects in more detail.

3.1 Reinforcement learning

Our robot has learned to make the right decisions to achieve joint attention and locate an interesting object, and thus the robot's decision making is the developmental aspect of our implementation. We implemented reinforcement learning for this purpose. We decided to let the robot make decisions at two different stages in the pipeline:

1. When the human looks away after having established eye contact, or after the robot has failed to make eye contact after a certain duration, the robot has to choose between:
    - Follow a horizontal line to the right.
    - Follow a horizontal line to the left.



s0: starting state
s1: looked left
s2: looked right
s3: followed gaze
s4: chose ball close by
s5: chose ball far away
s6: terminated

a1: look left
a2: look right
a3: follow gaze
a4: choose ball close by
a5: choose ball far away
a6: terminate

Figure 2: Visualisation of the states and actions. The bold arrows are the actions that the robot should learn to perform.

    - Follow the gaze line.

2. When at least one ball is within the visual field of the robot, choose between:
    - Select a ball that is far away from the line that the robot followed, based on threshold.
    - Select a ball that is close to the line that the robot followed, based on a threshold.
    - Repeat the previous action (look left, look right, or follow the gaze).
    - Terminate.

The above decisions that have to be made result into a set of actions, states, and relationships between them that are visualised in Figure 2.

We have decided to implement Q-learning, which is a form of reinforcement learning. To determine the exploration versus exploitation rates, we used Value-Difference Based Exploration (VDBE), as proposed by [6]. The feedback is based on the colour of the ball: a green ball gives a positive reward, while a pink ball gives a negative reward. Furthermore, we set the parameters of the algorithm based on a simulation.

3.2 Developmental principles

According to our project proposal, we wanted the following principles to be involved in our implementation, taken from [7]:

- Online, open-ended, cumulative learning.
- Embodied and situated development.
- Intrinsic motivation and social learning.
- Nonlinear, stage-like development (emergent from the previous principles).

We achieved *online, open-ended, cumulative learning.* Our robot starts without any knowledge, and then updates its Q-values over time. The performance increases with more successions, and thus the learning is online and cumulative. The learning process is also open-ended, since it is still updating the Q-values after any given number of trials.

Since we ran the model on a robot during training and because our robot uses the environment to make decisions, we also achieved *embodied and situated development.*

Whether we achieved *intrinsic motivation and social learning* is less straightforward. Intrinsic motivation is something we tried to achieve by looking at the intrinsic motivations of infants according to literature (see Section 2), and then implementing those motivations in our model. That is why our code contains a motivation module [2]. The robot then makes a decision based on the current motivation. Furthermore, to achieve social learning, the setup of the training sessions should be such that the robot uses social cues to learn. However, this is not something our robot does. It does look for a face and for the gaze of that face, but social cues like speech or gestures are not picked up by the robot.

In our project proposal, we argued that *non-linear, stage-like development* could emerge from the other principles, especially because [8] stated that a *basic set* of structures and mechanisms is sufficient to account for typical development, which includes non-linear, stage-like development. We especially hoped for the stages as proposed by [4], of which the first stage is called sensitivity to field. In this stage, the infant looks to the left or right. Our robot also does that at the start, although it differs in two ways:

1. At the start of the learning process, thus without any knowledge, there is already a probability of 1/3 that the robot will follow the gaze.
2. If the robot decides to either look left or right on a horizontal line, the direction does not depend on the gaze. Thus, it could be that the robot looks in the opposite direction, while the infant would look at the side towards which the gaze of the human is pointing.

Then in the following stage, an infant is more inclined to choose a ball close by and it does not look at a ball that is outside of the initial visual field. Our robot also shows this behaviour, but only when there are other objects within the visual field on the way from the face of the human to the object. If an object would be placed such that the robot has to turn its head maximally to detect it without any other object on the way, it would still turn and end up at the object, regardless of how much it has or has not learned.

Although the behaviour of our robot does not correspond completely to the stages that are found in infants, we do believe that our robot showed stage-like development. Namely:

- Stage 1: random decisions.
- Stage 2: learned to follow the gaze, but not yet about which ball to pick.
- Stage 3: learned to follow the gaze and to choose the correct ball.

For more information about the developmental behaviour we observed, see Section 4.1.

4 Robot behaviour

Testing of the designed developmental joint attention model in the robot will be done with a number of scenarios, shown in Table 3. This variety in the number of balls, people, and environment will test if the correct object can be located using joint attention, how robust it is, and how accurate.

[2] See https://github.com/KochPJ/follow_gaze_developmental_robotics

Levels	1	2	3	4	5	6	7	8
What kind of object?	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored	Ball, brightly colored
How many objects?	One	One	Two	One	Two	One	Three	Five
Clean or cluttered environment?	Clean	Clean	Clean	Cluttered	Clean	Clean	Clean	Clean
How many humans?	One: human looks at ball constantly	One: human looks at the ceiling	One: human looks at same ball constantly	One: human looks at ball constantly	One: human switches gaze randomly	Two: Both humans look at same ball	Two: Both humans look at different balls	Two: Both humans look randomly at various balls, randomly at ceiling

Figure 3: The scenarios as designed by P. Haselager, G. Ras and S Hunnius to evaluate the accuracy and robustness of the developed joint attention model.

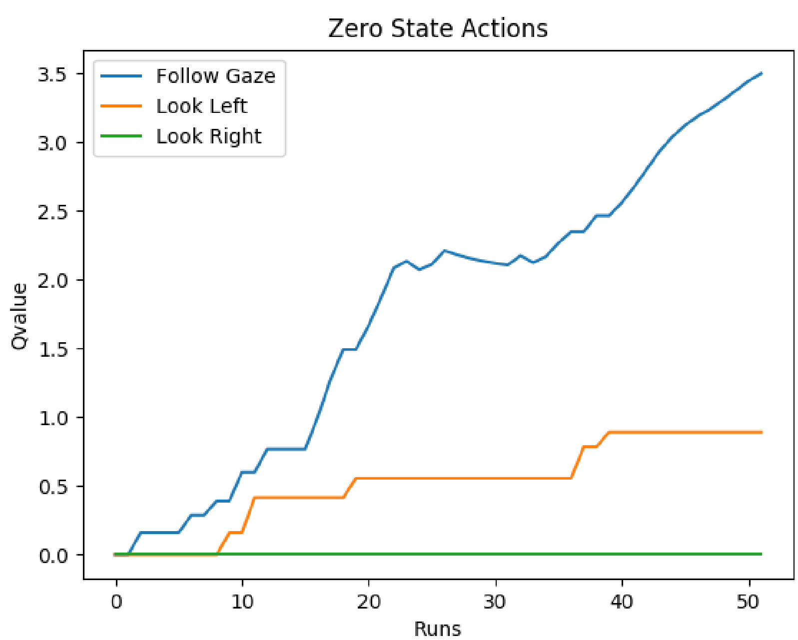


Figure 4: Visualization of the Q-values of the zero state, which is the starting state.

4.1 During training

During the training phase, scenario 3 from Table 3 was used, where the robot is sitting in a crouched position opposite to a human, surrounded by multiple balls and a clean environment. In total a number between two and five balls were used at various heights and directions from the robot. The balls were laid out in a 180 degree area in front of the robot, such that every ball could be centred using solely head movement of the robot, and no repositioning of the torso was required. The decisions made were evaluated by using two types of balls, where the colour indicates the correct or wrong ball. During the ball detection and decision making process, pink and green balls are dealt with in the same manner; the ball colour was only used after having chosen a ball to determine the reward for the reinforcement learning. The advantage of this approach over oral feedback, or other forms of manual feedback, is that it is much faster.

A full training session was run by running the experiments as above, until the model had fully explored all options according to the VDBE algorithm (all epsilons < 0.1). This took a total of 51 trials. In Figure 4 the Q-values associated with the Starting state actions are visualised. Clearly visible is that after 20 trials, the model already has a high preference for using gaze following. This was also shown by the behaviour of the robot: after 20 trials, the robot almost always chose gaze following above the other two possible actions. Furthermore, Figure 5 shows the Q-values associated with the various actions that are possible from within the Followed gaze state. Here too it is clearly visible how the model learns a preference for choosing a ball close to the line being followed, although this learning occurs later than learning to follow the gaze. This was also shown by the behaviour of the robot: after having followed the gaze, it explored choosing a ball far away, but the rate of this exploration decreased over time. In addition, the robot has even learned to ignore balls within the visual field but far away from the gaze line, as they have a low probability of being the correct ball.

4.2 At demo

At the demo the trained Q-values will be loaded, and the model will be tested on the scenarios listed in table 3. The scenarios have been tested beforehand, with positive results. As such, given correct lighting for the face and ball detection, the robot should do well on all demo scenarios.

As scenario 3 was used for training, we considered scenario 1, 2 and 3 as a pass. Furthermore, the robot was tested in a cluttered environment, which also worked well if adequate lighting was present. Scenario 5, including a human switching its gaze constantly, also worked well. In this scenario the robot ran out of patience as it was not able to establish eye contact, and detected the gaze which was then followed.

Slight difficulties were encountered when multiple faces were encountered in the same frame, where the robot would switch its attention between the two as the face detection might not recognize the face every frame. This has

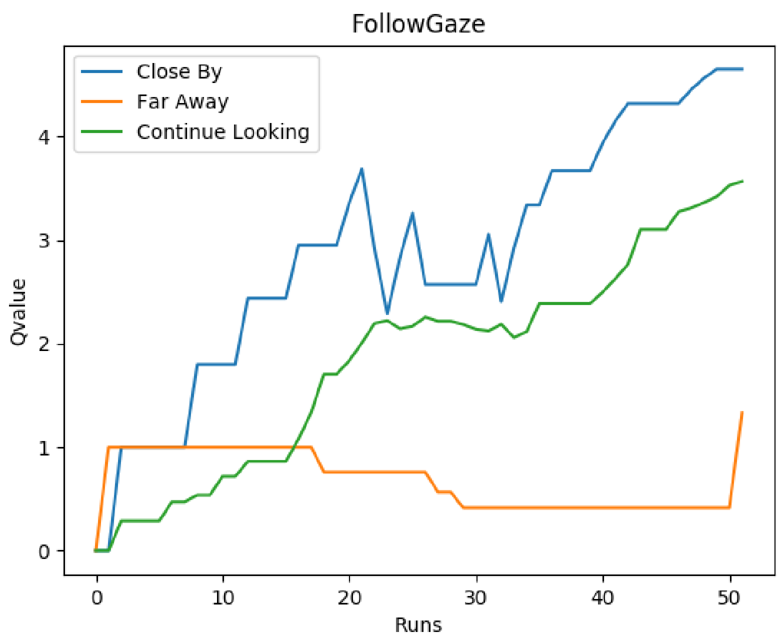


Figure 5: Visualisation of Q-values for the three possible actions from the follow gaze state.

been solved by making the robot check three frames before switching to another person when the face is lost.

5 Evaluation

The final product of this project is a pipeline establishing joint attention using gaze following (see Section 1). We implemented multiple principles of developmental robotics (see Section 3). We ended up with the final product that we proposed. All proposed components are present and the proposed developmental principles have been achieved, although some more loosely than others. To wrap it up, we achieved what we planned to do, and are quite happy about the outcome. Overall we consider the project as a success.

5.1 Issues encountered

During the development, two major problems were encountered. The first issue was related to the gaze following, which gave difficulties during installation. These problems were solved by using a Linux computer. Although we had only one Linux computer, we could still work efficiently enough.

The second major issue was that the face, gaze and ball detection were all quite sensitive to lighting conditions, which resulted into mistakes in the perception of these. This was solved by finding a setup with good lighting, which solved most problems.

Some other minor issues were encountered:

- Following the gaze resulted in a proportional offset the further the robot moves its head, due to the two dimensional gaze being translated to three dimensional space of the robot. We minimised the offset by putting the robot in a posture with a vertical back such that head movement can be perfectly horizontal, and changing the threshold for the far away and close by balls the further the gaze is followed.
- We noticed that, when the human participant is looking downwards, the gaze model often gave a gaze vector that was too horizontal. We worked around this by not placing the balls too low as compared to the human.
- Due to lighting, especially the green ball is not always detected. During training, we tried to solved this by adding an extra light source. This did increase the performance, but not to the same level as the pink ball detection performance.

5.2 Main successes

We consider the project to be an overall success. We implemented each proposed pipeline component and eventually the finished project was able to successfully learn joint attention. Moreover, we managed to finish the training a week before the actual deadline.

5.3 Main failures

We consider our selection of the ball to be developmental, but it can be more realistic. We select either a ball far away or close to the line the robot is following. Hereby, a close by or far away ball is discriminated by a predefined threshold, which is chosen by us. It would be more developmental if this threshold could also be dynamic, or learned by the robot, instead of predefined.

Furthermore, the face and ball detection could be improved such that it is more robust to various lighting conditions.

5.4 Team work

We had group meetings twice a week during which we did the main work for our project. We worked mostly together, but split the work in components such that individual team members could work on them, possibly in pairs if desired. We always helped each other and were up-to-date on developments, and there was a very good team atmosphere and communication. When one of us had to cancel a meeting, we knew it well in advance. The project work was well distributed within the team and everybody did their part, all presentations and reports we did for the project were worked on together.

References

[1] A. Recasens, A. Khosla, C. Vondrick, and A. Torralba, "Where are they looking?," in *Advances in Neural Information Processing Systems*, pp. 199–207, 2015.

[2] F. Simion, V. Macchi Cassia, C. Turati, and E. Valenza, "The origins of face perception: specific versus non-specific mechanisms," *Infant and Child Development*, vol. 10, no. 1-2, pp. 59–65, 2001.

[3] C. Moore and V. Corkum, "Social understanding at the end of the first year of life," *Developmental Review*, vol. 14, no. 4, pp. 349–372, 1994.

[4] G. Butterworth and N. Jarrett, "What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy," *British journal of developmental psychology*, vol. 9, no. 1, pp. 55–72, 1991.

[5] F. Förster, J. Saunders, and C. L. Nehaniv, "Robots that say no. affective symbol grounding and the case of intent interpretations," *IEEE Transactions on Cognitive and Developmental Systems*, 2017.

[6] M. Tokic and G. Palm, "Value-difference based exploration: adaptive control between epsilon-greedy and softmax," in *Annual Conference on Artificial Intelligence*, pp. 335–346, Springer, 2011.

[7] M. S. Cangelosi, Angelo and L. B. Smith, *Developmental robotics: From babies to robots*. MIT Press, 2015.

[8] J. Triesch, C. Teuscher, G. O. Deák, and E. Carlson, “Gaze following: why (not) learn it?,” *Developmental science*, vol. 9, no. 2, pp. 125–147, 2006.