ROBotic Open-architecture Technology for
Cognition, Understanding and Behavior

Project No. 004370

RobotCub
Development of a Cognitive Humanoid Cub

Instrument: Integrated Project
Thematic Priority: IST - Cognitive Systems

D6.4 Interaction Histories for Action and Gesture Communication: Implementation and Results
WP6 - Gesture Communication

Due date: **01/03/2008**
Submission Date: **15/04/2008**

Start date of project: **01/09/2004** Duration: **60 months**

Organisation name of lead contractor for this deliverable: **University of Hertfordshire**

Responsible Person: **Kerstin Dautenhahn**

Revision: **1.0**

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Service)	
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Contents

1 Executive Summary

WP6 continues to focus on interaction dynamics of social interaction during robot-human play and the prerequisites for gesture and non-verbal communication between robots and humans, as well as the realization of these capabilities in a robot. In this particular research work, techniques for achieving this capability in an autonomous robot through grounded sensorimotor experience and interaction histories, are investigated.

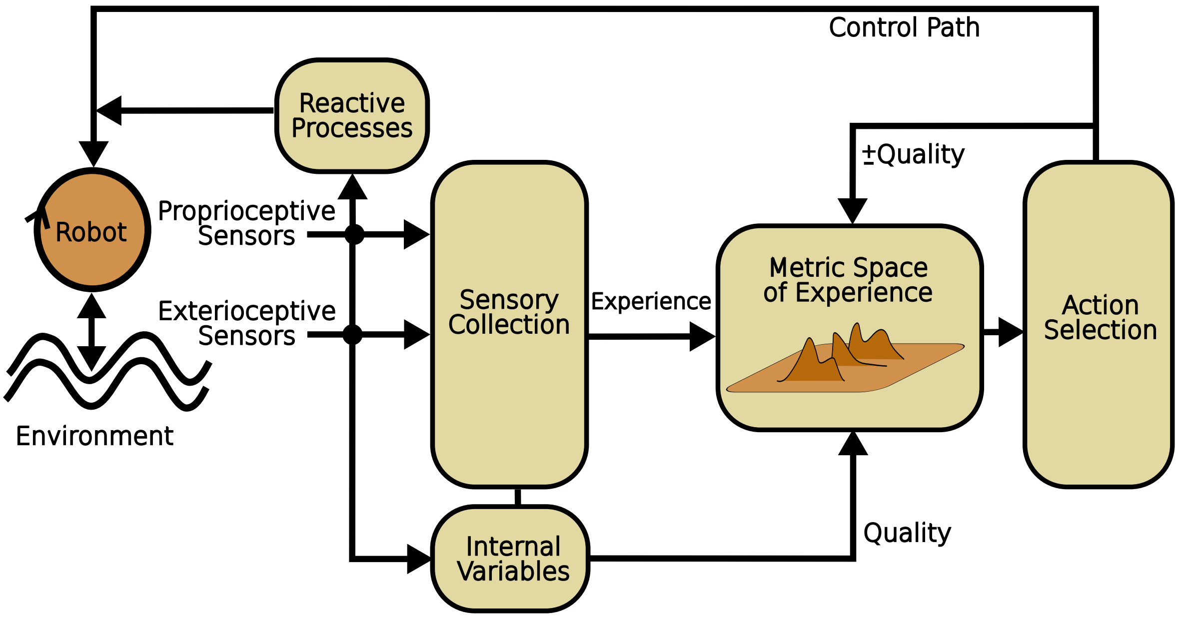
We present an architecture by which a robot can ontogenetically develop through social interaction and grounded sensorimotor experience detailing the architecture and recent experiments using the early interaction game, “peekaboo”, between an upper-body humanoid robot and human interaction partner. In earlier experiments (Mirza et al., 2007), the interaction history was shown to be capable of supporting development of a turn-taking interaction in a robot which took appropriate actions or gestures based on its own grounded sensorimotor experience.

Herein, we report the results from a recent experiment demonstrating the development of the capability of the robot to play the simple early-interaction game peekaboo.

1.1 Contents of the Report

Section 2 explains the operation of the Interaction History Architecture. Section 3 describes an experiment using a humanoid robot (Kaspar2) demonstrating the architecture. The research scenario is first discussed in Section 3.2 followed by a description of the robot, the experiment, the configuration of the Interaction History Architecture and finally the presentation and discussion of the results.

2 Interaction History Architecture



The Interaction History Architecture is shown schematically in Figure 2. The approach is as follows:

1. to continually gather sensorimotor data and find "suitable" episodes of sensorimotor experience in the history *near* (in terms of the experience metric) to the current episode;
2. depending on the course of subsequent experience, to choose from among actions that were executed when these episodes were previously encountered;
3. where no suitable experiences are found, to choose random actions.

There are two key aspects of this architecture. The first is the *metric space of experience* whereby new experiences appear as points in a growing and changing metric space. In this architecture the metric space is enhanced with *quality* information from the environment, internal drives or affective state. Each experience is also associated with actions executed during the experience. The second is the *action selection* system. This "closes the perception-action loop" and also closes an internal loop feeding back and modifying the experience space. The quality associated with each experience combined with proximity in the metric space is used to select experiences from the history and select actions associated with those experiences.

2.1 Interaction History Space

The metric space of experience in the Interaction History architecture is described in detail in previous deliverables (D6.2 and D6.1) as well as the published paper (Mirza et al., 2007) (also to be found as an appendix to D6.2). Briefly, the *Interaction History Space* consists of:

Sensorimotor "Experiences" These are time-series of sensor readings from all available sensors of a robot, from time t to another time $t+h$ where h is the *horizon length* of the experience.

A Metric A metric measure of distance between the experiences. This is the "Experience Metric", that utilizes the information distances between sensor time-series viewed as values of random variables.

Next Action information The next action executed after an experience is associated with that experience.

Quality information A value representing environmental reward received after the experience (for a particular time span).

The metric space is constructed continuously as the robot experiences its environment. A new experience is created every *Granularity* G timesteps, and consists of Horizon h timesteps counting back from the current timestep. Where $h > G$ the experiences will overlap. Each sensor reading is quantized into Q evenly-sized bins. Each new quantized experience is compared to other experiences in order to determine its neighbours. This process, if all experiences are compared, results in a distance matrix between experiences which defines the structure of the metric space as it is experienced by an individual robot. A *quality* value is assigned to the quantized experience, determined by factors such as environmental reward/punishment, internal drive and affective state. The actual formula for calculation of *quality* is specific to the application and goal and can be a determining factor in the eventual behaviour and course of development, although it can be fairly general and thus applicable to a wide range of situations. Finally, the last action executed during the experience is also noted and stored with the quantized experience.

Thus the metric space of experience in the Interaction History Architecture, the *interaction history space*, can be described by the tuple $(\boldsymbol{\epsilon}, \boldsymbol{D}, \boldsymbol{q}, \boldsymbol{a})$, where $\boldsymbol{\epsilon}$ is a collection of quantized "experiences", $\boldsymbol{D}$ is the a matrix of distances between elements of $\boldsymbol{\epsilon}$, $\boldsymbol{q}$ is a vector of quality values and $\boldsymbol{a}$ a vector of actions.

Issues concerning the efficient construction of the metric space as new experiences arrive were discussed in the RobotCub deliverable D6.2.

2.2 Action Selection

A simple mechanism is adopted for action selection whereby the robot can execute one of a number of "atomic" actions (or no action) at any timestep. This is seen as a tractable first-step, and a more sophisticated action or behaviour generation capability would allow for more open-ended development. The actual action selected will either be a random selection of one of the atomic actions, or will be an action that was previously executed after an experience in the history that is *near* to the current episode. An advantage of this approach is that behaviour can be bootstrapped from early random activity, and later behaviour built on previous experience.

The process of action selection is as follows:

1. up to K *candidate experiences* from the experience space within a given information distance *radius*[1] r_{max} of the current experience $E_{current}$ are initially selected;

2. these K experiences are ranked as $E_1, \ldots, E_K$ according to how close they are to $E_{current}$;

[1] The radius can be fixed, but, note that this could be adapted on-line.

3. then, *sequentially*, experience E_i is chosen with probability a linear function of the *quality* of E_i until either an experience is chosen or the ranked list is exhausted;

4. if an experience is chosen from the candidate list, then the particular action that was executed following the chosen experience is then chosen as the action to be executed next, otherwise a random action is chosen.

The linear mapping from quality to probability ensures that, with small probability, the robot may still choose a random action as this may potentially help to discover new, more salient experiences. This has the advantage of emulating body-babbling, i.e. apparently random body movements that have the (hypothesized) purpose of learning the capabilities of the body in an environment (Meltzoff and Moore, 1997). Early in development, there are fewer, more widely spread experiences in the space, so random actions would be chosen more often. Later in development, it is more likely that an the action selected will come from past experience.

2.2.1 Roulette-Wheel Action Selection

In later implementations (including the T-Maze implementation described in Deliverable D6.2), the process was improved to use a *roulette-wheel selection* from a probability list. The chance of random action selection is also represented in that list. The probabilities are calculated using a "gravitational model" where each experience is represented as a point mass a particular distance from the $E_{current}$. The probability of selecting an experience E_i from $E_1, \ldots, E_K$ is:

$$p_i = C_h \frac{m_i q_i}{D(E_{current}, E_i)^2} \tag{1}$$

where q_i is the *quality value* of E_i, m_i is the mass (*i.e.*how many experiences have been merged into this experience) and $D(E_{current}, E_i)$ is the experience distance. C_h is an optional quantity that is used to adjust for the "horizon effect"[2] and is given by

$$C_h = \frac{\sqrt{h}}{\sqrt{H_{max}}} \tag{2}$$

The chance of random is added to the list as:

$$p_0 = \frac{\sum_{i=1}^{K} p_i}{(r_{max}/\tau)^2} \tag{3}$$

where r_{max} is the radius of the ball that includes the ranked experiences and τ is a *temperature* factor, that controls the chance of random action selection.
Then the weighting on the "roulette wheel" is given by:

$$w_i = \frac{p_i}{\sum_{i=0}^{K} p_i} \tag{4}$$

[2]The "horizon effect" occurs when using multiple horizon length metric spaces for experience selection. If one of the horizons is too short for the task, the experience selection process tends to choose this in preference to potentially more relevant experience of longer horizon.

2.3 Update of Environmental Reward

Each experience in the interaction history space is associated with a quality value q, see Section 2.1. This value has bearing on the selection of the experience, and in turn on the action-selection process. The quality value is intended to reflect how useful the experience is in terms of positive or negative environmental feedback, and is derived directly from the internal reward function or an external reward measured by the robot's sensors.

In the simplest case, the immediate (instantaneous) reward received from the environment is associated with the current experience. An alternative scheme is for the quality associated with an experience to be dependent not only on the current reward, but also on the future reward. The *future reward* for an experience $E_{t,h}$ for some given horizon h_{future} is a function $\mathcal{F}()$on all reward values received for h_{future} timesteps after time t. Of course, this value cannot be known completely until at least h_{future} timesteps have passed, but it estimated until that point. Two functions have been used in the implementations in this thesis. The first, $\mathcal{F}_{min\_max}()$, returns the most proximal maximum or minimum reward. The second. $\mathcal{F}_{max}$ simply returns the maximum reward over the horizon.

2.4 Feedback Loop

Finally, a feedback process evaluates the result of any action taken in terms of whether there was an *increase in quality* after the action was executed, and then adjusts the quality of the candidate experience, from which the action was derived, up or down accordingly. By this mechanism, the metric space is effectively altered from the point of view of the action-selection system. Closing of the perception-action loop in this way with feedback together with growth of the experiential metric space, results in the construction of modified behaviour patterns over time. This can be viewed as a form of ontogenetic development and adaptation, that is, a process of change in structure and skills through embodied, structurally coupled interaction.

2.5 Merging and Deletion of Experiences in the Interaction History Space

As discussed in the RobotCub Deliverable D6.2, it is necessary to employ strategies such as *merging* and *forgetting*, if storage and computation requirements are to be controlled. The merging strategy in the Interaction History Architecture is to merge any two experiences closer than a threshold T_{merge} (see Algorithm 1). T_{merge} was fixed for the most part, however alternative strategies were trialled during development of the algorithm, including adapting the threshold such that the maximum number of experiences in the space remained constant.

Algorithm 1 shows how the meta-information associated with experiences that are merged are also assimilated. Actions from both merged experiences are accumulated, resulting in an action probability distribution; the quality values are averaged; and, a weight value, indicating the number of experiences that have been merged together, set to the sum of the weights of the merged experiences.

Experiences may also be deleted, that is, forgotten. This serves two particular purposes in the present architecture. The first is to provide a mechanism where the interaction history space can be continually modified and so be adaptive to changes in the environmental interaction. The second, more practically, is to reduce the number of experiences in the space and so reduce computational complexity in estimating distances to new experiences inserted into the space (see also Deliverable D6.2). There are a number of different strategies to decide which experiences should be forgotten, and the one used here is to forget those experiences which have lower quality values and thus will have little or no impact on future action selection. Specifically, experiences older than h_{future} with a quality less than or equal to T_{purge} will be deleted.

Algorithm 1: Algorithm `IHA_MERGET`: Choose and Merge 2 experiences using a threshold

for E^i ***in*** *all experiences* **do**
  for E^j ***in*** *neighbours of* E^i **do**
    if $d(E^i, E^j) \leq T_{merge}$ **then**
      $actions(E^i) = actions(E^i) + actions(E^j)$
      $quality(E^i) = \left(quality(E^i) + quality(E^j)\right)/2$
      $weight(E^i) = weight(E^i) + weight(E^j)$
      delete all distances to and from E^j in the metric space
      delete E^j
    end
  end
end

3 Development using Interaction Histories Through Playful Interaction

We describe an experiment that illustrates how a robot can develop action capabilities based on its history of interaction with the environment through the use of the architecture presented. The scenario is a simple communicative interaction game, "peekaboo", that uses simple non-verbal gestures. The peekaboo game as a research tool is described, followed by a description of an experiment using an upper-body humanoid robot that uses its interactionhistory to develop the capability to engage in a peekaboo interaction with a human partner.

3.1 Peekaboo as a Research Tool

The development of gestural communicative interaction skills is grounded in the early interaction games that infants play. In the study of the ontogeny of social interaction, gestural communication and turn-taking in artificial agents, it is instructive to look at the kinds of interactions that children are capable of in early development and how they learn to interact appropriately with adults and other children. A well known interaction game is "peekaboo" where classically, the caregiver having established mutual engagement through eye-contact, hides their face momentarily. On revealing their face again the care-giver cries "peek-a-boo!', "peep-bo!", or something similar. This usually results in pleasure for the infant which, in early development, may be a result of the relief[3] in the return of something considered lost (*i.e.* the emotionally satisfying mutual contact), but later in development also may be a result of the meeting of an expectation (*i.e.* the contact returning as expected along with the pleasurable and familiar sound), and the recognition of the pleasurable game ensuing (Montague and Walker-Andrews, 2001; Veatch, 1998).
Bruner and Sherwood Bruner and Sherwood (1975) studied peekaboo from the viewpoint of play and learning of the rules and structures of games. They also recognize that the game relies on (and is often contingent with) developing a mastery of object permanence as well as being able to predict the future location of the reappearing face. The individual parts of the game can be viewed as gestures in a non-verbal communicative interaction. The hiding of the face is one such gesture, and the vocalization, and the showing of pleasure (laughing) are others. In order for the interaction game to proceed successfully, the gestures must be made by either party at the times expected by the players, and that

[3] In the context of humour, peekaboo in its early stages is an example of relief laughter. That is relief that the caregiver that is thought to have disappeared, actually has not (Veatch, 1998).

absence or mis-timing can result in the game cycle being broken. Learning of the game is supported by further gestures such as a rising expectant intonation of the voice during hiding, as a reassurance or cue of the returning contact. Later in development the roles of the game can become reversed with the child initiating the hiding, while still obeying the established rules by, for instance, uttering the vocalization on renewed contact.

In all this, the rhythm and timing of the interaction are crucial and, Bruner and Sherwood suggest that the peekaboo game and other early interaction games act as scaffolding on which later forms of interaction, particularly language and the required intricate timing details, can be built (Pea, 2004, pp 424-5). Discussing scaffolding, Roy Pea notes that "*. . . there are regularly structured situations in which the range of meanings is actually quite limited and that these simple formats provide a highly constrained situation in which the child can bootstrap some of the conventions of turn taking and meaning making with words that are required of a language user.*" (Pea, 2004, pp434-425), emphasizing, therefore, the importance of early communication games such as peekaboo in the development of language.

In relation to the development of social cognition in infants, "peekaboo" and other social interaction games, that are characterized by a building and then releasing of tension in cyclic phases, are important as they are considered to contribute developmentally to infant understanding and practise of social interaction. Peekaboo provides the caregiver with the scaffolding upon which infants can co-regulate their emotional expressions with others, build social expectations and establish primary intersubjectivity (Rochat et al., 1999).

3.2 Peekaboo with the Humanoid Robot Kaspar2

This section describes an experiment that continues the ongoing investigation of the Interaction History Architecture as the basis for developing appropriate actions in response to the ongoing history of the robot-environment interactions. The implementation in this experiment includes an audio used both as an extra sensory modality as well as an additional environmental reward feedback for the peekaboo game that results directly from the human-robot interaction. This is also the first use of an upper-body humanoid robot and is a step towards the implementation in the iCub. The architecture is fully implemented and includes both merging of and deletion of experiences as the mechanism for modifying the metric space of experiences.

3.3 Experimental Setup

This section details any additions or variations to the general architecture described in Section 2, as well as the specific setup parameters of the metric space creation and control architecture used in these experiments. Reasonable values were chosen for the various parameters, such as horizon length and merging threshold, based on the results of previous experiments and the nature of the present experiment. Furthermore, this section describes the setup for conducting the experiments and retrieving results.

3.3.1 Motivational Dynamics

In this experiment, motivation feedback (reward) is provided through two mechanisms: observation of a face, and audio feedback.

Face: As before, a face can be detected in the robot's camera image and this provides direct positive reward. Habituation causes this reward to drop-off over time. The reward for face detection, R_f, constrained to be in the range $[0, 1]$, is a function of the number of consecutive timesteps a face is seen. First the reward rises linearly, then holds at 1 for a period before decaying towards 0. R_f is calculated incrementally as follows:

$$R_f^{t+1} = R_f^t + \begin{cases} 1/\mathcal{T}_{rise} & t < \mathcal{T}_{rise} \\ 0 & t < \mathcal{T}_{rise} + \mathcal{T}_{hold} \\ -R_f^t/\mathcal{T}_{fall} & t \geq \mathcal{T}_{hold} \end{cases} \tag{5}$$

where $\mathcal{T}_{rise}$, $\mathcal{T}_{hold}$ and $\mathcal{T}_{fall}$ are paremeters that control the length of the attack, hold and decay phases. At any time a face is not detected, $R_f^{t+1} = 0$.
In this experiment the parameters were set as follows: $\mathcal{T}_{rise} = 4$, $\mathcal{T}_{hold} = 2$ and $\mathcal{T}_{fall} = 20$.
Sound: New to this experiment, sound is captured from a microphone, and used both as an additional sensory signal as well as providing further environmental reward. The “energy” of the sound over the period of a timestep, ε_{sound}, provides a new sensory input to the robot. It is calculated as the sum of the amplitude of the sound signal for every sound sample in a period of a timestep, and is normalized to take values in the range [0,1]. In converting ε_{sound} to a reward signal R_s, low level background noise is attenuated by taking the square of the sound sensor variable for all values below a threshold T_{sound}, above which the reward value is set to 1.

$$R_s = \begin{cases} \varepsilon_{sound}^2 & \varepsilon_{sound} < T_{sound} \\ 1 & t\varepsilon_{sound} \geq T_{sound} \end{cases} \tag{6}$$

Resulting Reward Signal: The final reward signal generated by the robot in response to it’s environmental interaction is a combination of the sound and face reward signals, as follows:

$$R = \max(1, \alpha(R_f + R_s)) \tag{7}$$

where α, in the range [0,1] attenuates the reward signal. With $\alpha = 0.5$, R is the average of the reward signals, and with $\alpha = 1$, either of the reward signals can result in a maximum resulting reward. For these experiments, α is set between these two values at $\alpha = 0.75$, meaning that neither reward signal on its own can result in a maximum R, but requires support from the other reward signal.

3.3.2 Interaction History Architecture Components and Settings

Metric Space of Experiences:
Experiences older than h_{future} timesteps were deleted (forgotten) where they were associated with a quality value of less than or equal to $T_{purge} = 0.9$. Experiences were merged where both their distance in the metric space of experiences was less than $T_{merge} = 0.6 bits$ and they were associated with the same next action. A combination of the merging and forgetting processes resulted in a manageable sized metric space for real-time operation.
Action Selection: The closest $K = 4$ neighbours of the current experience within a radius of $r_{max} = 2.0$*bits* of $E_{current}$ were considered in the action-selection process (see Section 2.2).

3.3.3 Experimental Materials and Methods

Robot: The robot used was the upper-body humanoid Kaspar2 robot created at the University of Hertfordshire, see Figure 1. The robot has 17 individually controlled motors: three in the neck controlling head orientation, two controlling the eyes (the eyeballs are connected and move in unison -

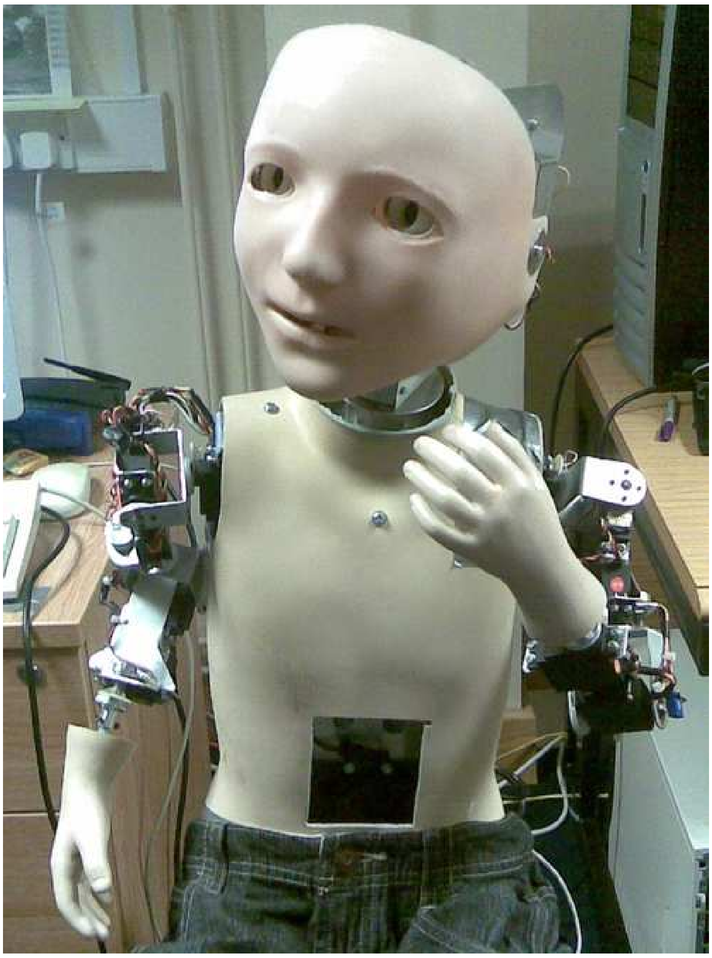


Figure 1: The Kaspar2 robot (University of Hertfordshire) used in the experiments.

there is no vergence control), two controlling the mouth for facial expression, and five controlling each arm. The motor control boards provide a serial link and the control software was written in C++. The interaction history architecture was written in C++ as multiple interacting modules, with the communication layer and abstraction of hardware control provided by the YARP framework (Metta et al., 2006).

Actions:

A total of 17 actions were available to the robot, and these can be considered in 3 groups: movement actions, facial expressions and resetting actions. These are listed in Table 1 and selected actions and expressions are shown in Figures 2 and 3. The types of action that the robot can execute at any time depends on which action was last executed. This is so that the robot does not attempt to execute actions that could possibly damage it. The configuration therefore defines the set of next actions possible after any given action and the action selection process is responsible for ensuring that these conditions are met. For reference, these action state dependencies are illustrated in Figure 30 in Appendix C.

Table 1: Kaspar2 Peekaboo: Actions

Group	Number	Action	Description
Movement Actions	3	HL	Head Left
	4	HR	Head Right
	6	HID	Hide Head with Hands
	8	RAU	Right Arm Up
	9	LAU	Left Arm Up
	12	RAW	Wave Right Arm
	13	LAW	Wave Left Arm
	14	TR	"Think" Right - raise right arm to chin and look right
	15	TL	"Think" Left - raise left arm to chin
Facial Expressions	1	Smi	Smile
	2	Neu	Neutral
	16	Frn	Frown
Resetting Actions	0	Rst	All motors to resting position
	7	NA	No Action
	5	HF	Head to forward position
	10	RAD	Right Arm Down
	11	LAD	Left Arm Down

3.3.4 Defining a Peekaboo Sequence

A "peekaboo" sequence is defined to be a sequence of actions beginning with the robot hiding its face (action 6 - HID), followed by any number of "no-action" actions (action 7 - NA) and ending with the robot back in the resting position (action 0 - Rst). Furthermore, for the purposes of evaluating the results of this experiment the actions should be selected from previous experience rather than executed randomly.
To measure the relative amounts of peekaboo in any given period of behaviour, $p_{sel}(A^{HID})$, the percentage of times the hiding action was *selected* as compared to other "movement" actions, was used as a measure and is calculated as follows. Given N possible actions $\{A^1, A^2, \ldots A^N\}$ and a period of behaviour consisting of K actions executed (selected or random), action A^n will be executed $F(A^n) = F_{rand}(A^n) + F_{sel}(A^n)$ times, where F_{rand} indicates the frequency of random executions and F_{sel} the frequency of the action being deliberately selected. Then the percentage of times the Hiding action A^{HID} was selected is given by

$$P_{sel}(A^{HID}) = 100\frac{F_{sel}(A^{HID})}{K} \tag{8}$$

Note that for the purpose of evaluating "peekaboo", only actions in the "movement actions" group were considered (see Table 1).

3.3.5 Method

The robot and human partner were positioned facing each other at a distance of a few feet, with their eye-level at approximately the same height. The robot control software was started with the interaction

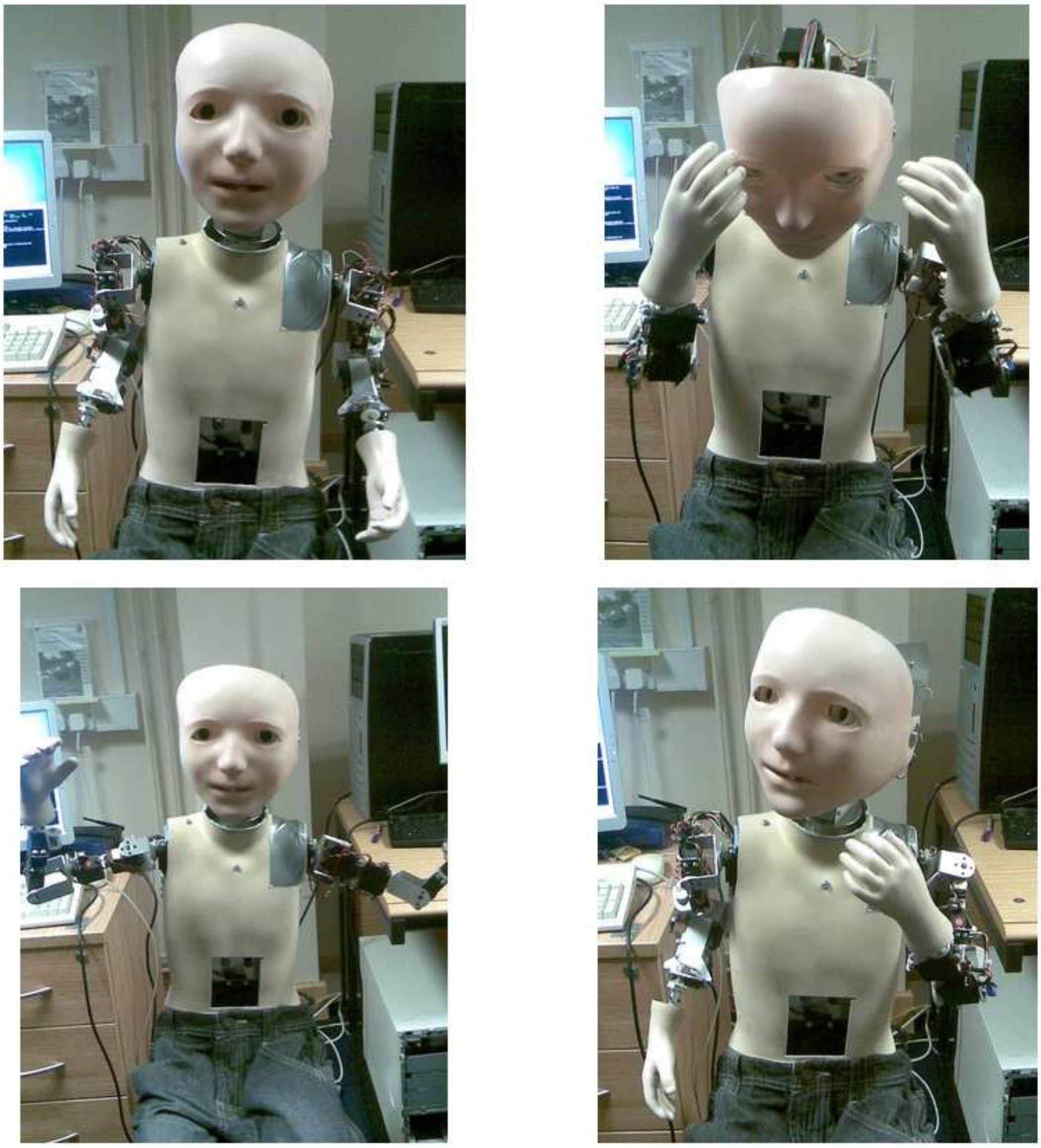


Figure 2: *Kaspar2 Sample Actions.* (top-left) Normal resting position, (top-right) Hiding action, (bottom-left) both arms are raised (a combination of two actions required), (bottom-right) The "think right" (TR) action.

history containing no previous experiences. Interaction then commenced with the robot executing various actions and the human offering vocal encouragement when it was thought appropriate. The interaction then continued for approximately two to three minutes.

Three different conditions were tried. Firstly, any hiding action was encouraged with a call of "peek-aboo" when the robot revealed its face again. The second condition encouraged an alternative action which also turned the robot's head away from the interaction partner. Both "head left" and "think right" were used for this purpose. The final condition was to offer no vocal encouragement at all during the interaction.

The experimental hypothesis was that encouraging the hiding action would result in a higher rate of

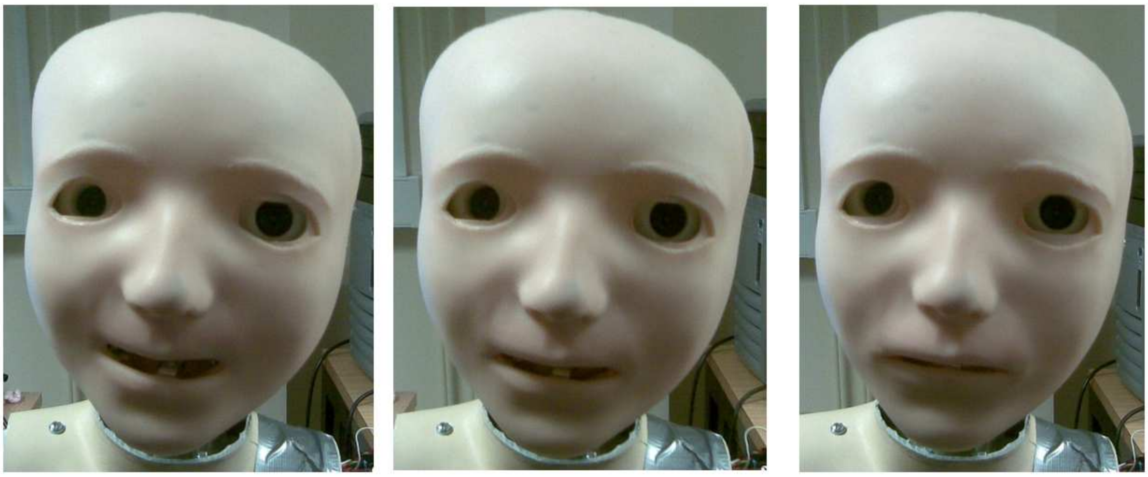


Figure 3: *Kaspar2 Expressions.* (left) Smile, (middle) Neutral, (right) Frown.

peekaboo sequences than would be expected from random action selection. Furthermore, this should also be the case when other actions are encouraged instead. Finally, this hypothesis was also tested by the no-encouragement condition with the expectation that no action would be selected in preference to any other.
Note that for all these experiments I personally took the role of the human partner and so was fully aware of the capabilities of the robot and of the software.

3.4 Results

A total of 22 runs were completed. 16 of these for the first condition (encouraging the Hiding action), 3 for the second condition and 3 for the no-encouragement condition. The results are summarized in Table 2 and more details of the results from the individual experiments are given in Appendix B. In most of the experimental runs it was fairly straightforward to estimate whether the experiment successfully supported, or clearly failed, the hypothesis that the interaction history would result in increases in frequency of the encouraged action. However, in 2 of the runs, this was not possible ("borderline" in Table 2). In run d0039, the hiding action was the only one to be selected (rather than chosen randomly) however the run was too short for successful evaluation. In run d0052, the figures for the whole run do not indicate success, however, the results are borderline as the peekaboo behaviour was clearly beginning to occur towards the end of the run.
Where a result could be determined, 14 out of 20 runs (70%) were successful. In the following sections representative results from each condition are discussed.

3.4.1 Peekaboo Encouragement Condition

Figure 4 shows for the first run (d0032), how the motivational variables (face, sound and resultant reward) vary with time, along with the actions being executed . The interaction partner encourages the first "peekaboo" sequence ("hide-face" on the diagram). Note that a "peekaboo" action is actually a combination of the action to hide the face (action 6), any number of "no-action" actions (action 7) and an action to return to the forward resting position (action 0) (for clarity only the primary action is shown on the trace). This results in a maximal reward shortly after the hide-face action, and as the

Table 2: IHA on KasparII: Experimental Runs Summary

Run	Encouragement Type	Horizon	Comment	HID Chosen %	Result
d0032	Peekaboo	16	HID action executed early and repeated many times	55.17%	Success
d0033	Peekaboo	16	HID action executed early and repeated many times	41.18%	Success
d0034	None	16	HID action only twice randomly	0.00%	Success
d0035	Encourage HL	16	HL action chosen often. HID also chosen. HL=36.59%	14.63%	Success
d0036	Peekaboo	16	HID chosen often.	42.11%	Success
d0037	Peekaboo	16	3 HID actions selected, but RAW selected more often	13.64%	Fail
d0038	Peekaboo	16	No random HID to encourage.	0.0%	Fail
d0039	Peekaboo	16	Hid was only action chosen (once) but run too short	12.50%	Borderline
d0041	Peekaboo	16	Mixed actions - some peekaboo	5.49%	Fail
d0042	Peekaboo	16	Mixed actions	9.68%	Fail
d0043	Peekaboo	16	HID only twice	1.09%	Fail
d0044	Peekaboo	16	Peekaboo throughout	18.87%	Success
d0045	None	16	Few random HID actions	0.00%	Success
d0046	Encourage HL	16	HL action chosen many times, HID a few times. HL=11.84%	2.63%	Success
d0049	Peekaboo	20	Only a few random HID actions	3.26%	Fail
d0050	Peekaboo	20	HID chosen often	26.32%	Success
d0051	Peekaboo	20	HID chosen often	19.32%	Success
d0052	Peekaboo	20	HID not chosen enough for success over run. However, regular peekaboo was begining to occur at the end.	4.96%	Borderline
d0053	Peekaboo	20	HID chosen often	17.46%	Success
d0054	Peekaboo	20	HID chosen very much. HID was 1st action	61.76%	Success
d0055	Encourage TR	20	TR (Think-Right) encouraged. TR=26.00%	0.00%	Success
d0056	None	20	Some HID actions chosen	2.53%	Success

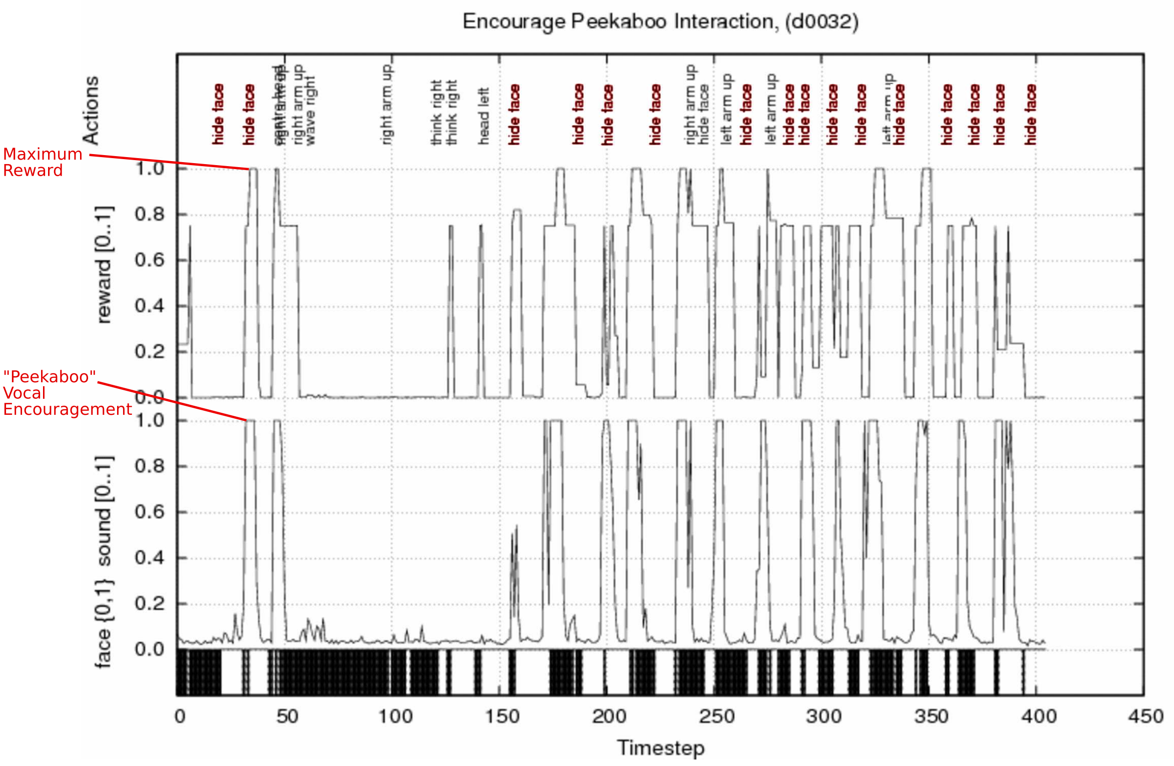


Figure 4: *Kaspar2 Results d0032. Example of Peekaboo Encouragement Condition.* The trace shows, against time, the detection of the face and audio encouragement as well as the resulting reward. Along the top are shown the actions executed.

interaction partner continues to reinforce the peekaboo behaviour with vocal reward, this pattern can be seen repeated throughout the trace.

As the chance of choosing a random action rather than selecting one using the history gradually declines the early part of the run will be more exploratory (have more randomly selected actions) whereas towards the end of the run, actions will be more likely to be deliberately selected using past experience. It can be seen that during the first half of the run various different actions are tried, but during the second half of the run, the "hide-face" action is chosen regularly.

The timing of the motivational feedback given by the interaction partner to the robot is important in determining what actions are executed. In Figure 5 from run d0050, the encouragement for the hiding action (and subsequent actions to return the robot to the resting position) is only received *after* the robot additionally turns its head to the side. The result is that when the robot decides to repeat the hiding action, it generates experiences which are likely to generate the actions that were executed following the original hiding action, *i.e.* the robot hides its face, returns to face the front and immediately turns its head to the side.

This behaviour (of the architecture) is an important part of how not just single actions are repeated, but instead how sequences of actions and robot behaviour are replayed, and it is this that encourages a fuller development of capabilities of the robot. It is important to note also that a specific sequence of actions are not learnt, instead it is the continuing generation of experience through the structural coupling of the embodied agent and its environment that drives this observed repeated behaviour. This can be clearly seen from Figure 5 in that the timing of the subsequent head-turn following a hiding

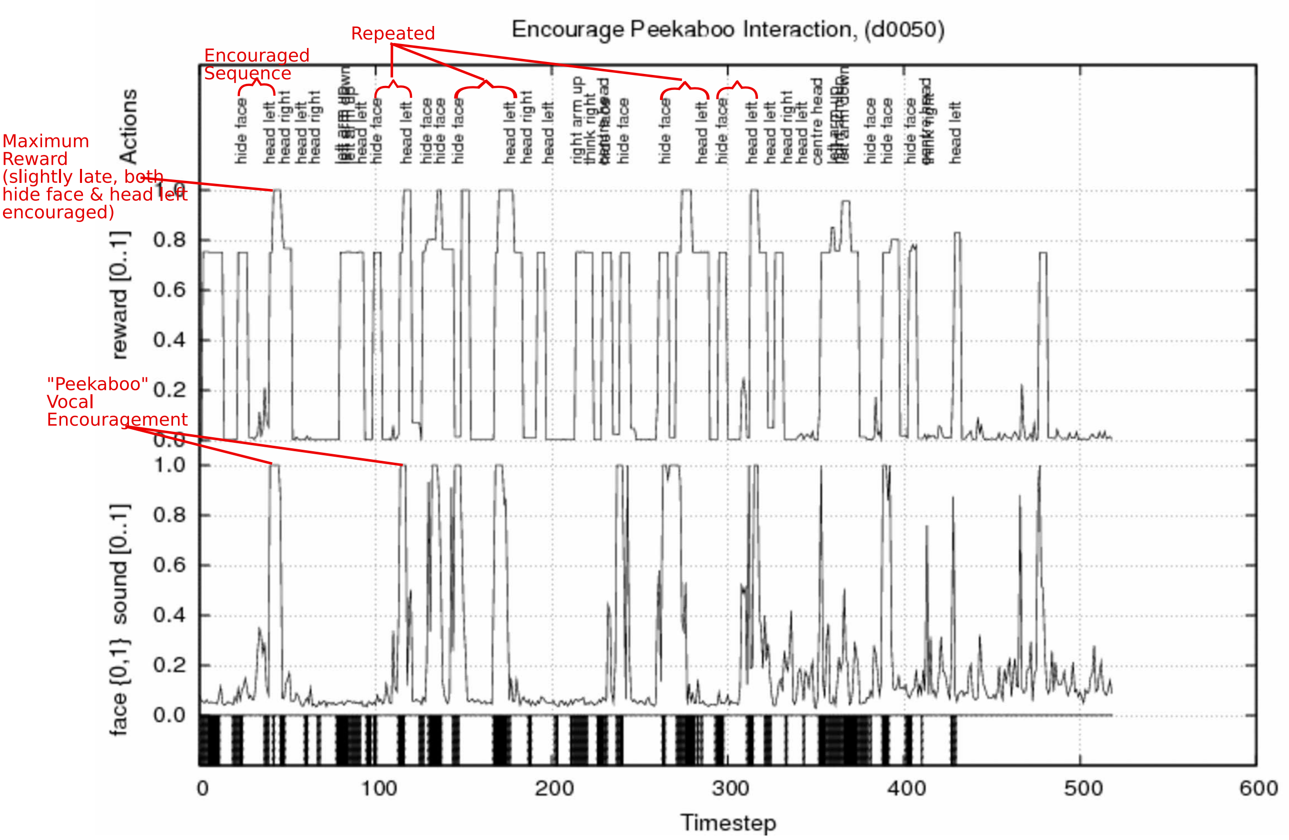


Figure 5: *Kaspar2 Results d0050. Showing a repeated action sequence.* A multiple action sequence is encouraged and repeated here.

action is not always the same, and indeed does not always occur.

3.4.2 Alternative Action Encouragement Condition

To illustrate that the operation of the interaction history is not limited to the peekaboo behaviour, the interaction partner also encouraged certain alternative actions rather than hiding. In two cases the "head left" (HL) action was encouraged (once also with a different call of "hello!" instead of "peekaboo!") and in one case the "think right" (TR) action was encouraged instead. In each of these cases the predominant action after some time was the encouraged one. Figure 6 from run d0035 shows a situation where the head-left action was encouraged, and it can be seen that the HL action was chosen in 36.9% of the "movement" actions whereas the Hiding action, for reference was chosen in 14.63%.

3.4.3 No Encouragement Condition

The final condition where the interaction partner offered no or very little encouragement resulted in various kinds of behaviour, none of which reinforced any particular action over any other, other than "doing nothing". An example is shown in Figure 7, where no encouragement at all is offered. In this case, some random actions are chosen but as time goes on, movement actions are not chosen and the robot executed actions that keep it stationary (the resetting actions in Table 1). In this case 152 actions

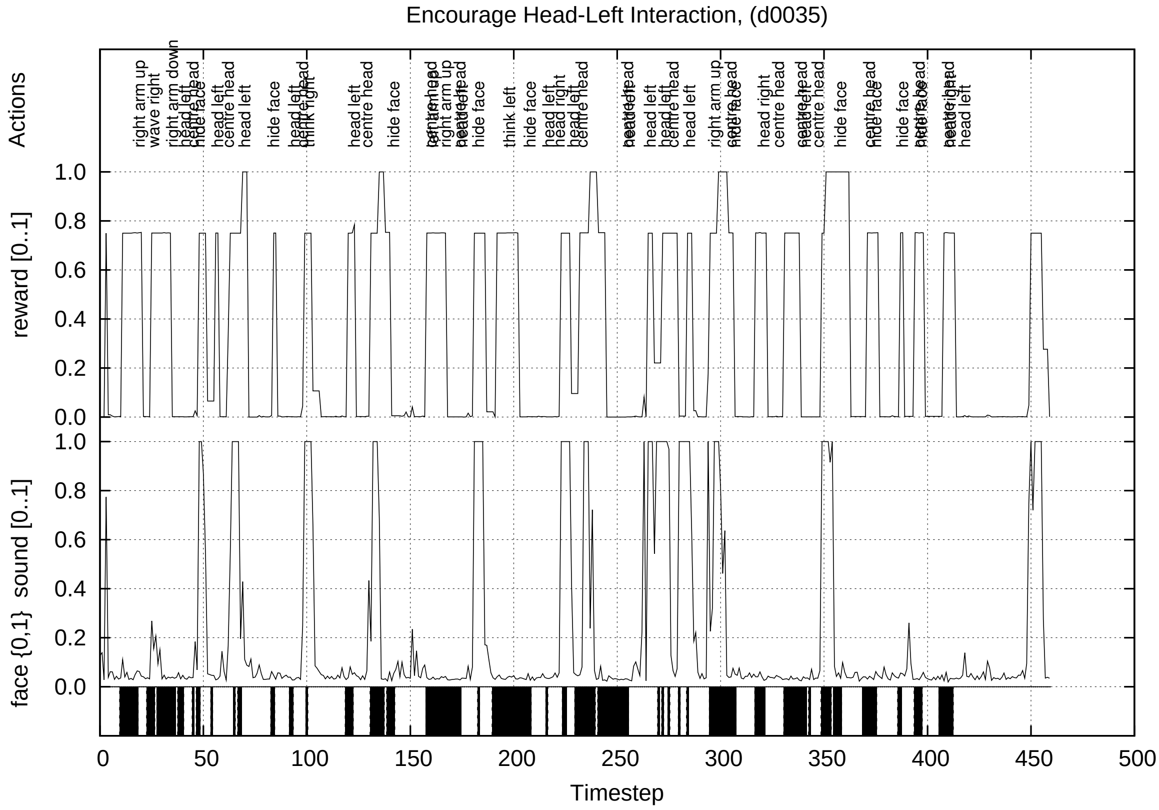


Figure 6: *Kaspar2 Results d0035. Encouraging and alternative action.* The “head left” (HL) action is encouraged and repeated.

are executed with only 32 actins of the “movement” type, evenly spread among these actions. The remaining 120 being mainly “Rst” and “NA”.

In the other cases where no encouragement was offered (runs d0034 and d0056 - see Appendix B) the robot did receive some reward albeit not a maximal reward. In these cases the robot did have actions from recent behaviour to choose from, however, the behaviour did not become repeated over the long term as continual merging and purging of experiences that do not result in near maximal reward resulted in only transitory behaviour. Thus the modification of the space through merging and deletion plays an important role.

3.5 Future Work

While short term behaviour acquisition is illustrated here, future research work should look at how behaviour can be altered over the long term in response to changing encouragement and reward by the interaction partner. Furthermore, showing how different behavioural responses can be developed for different experiences would be important next step.

Further experiments should also utilize interaction partners that do not have prior knowledge regarding the operation of the robot and software.

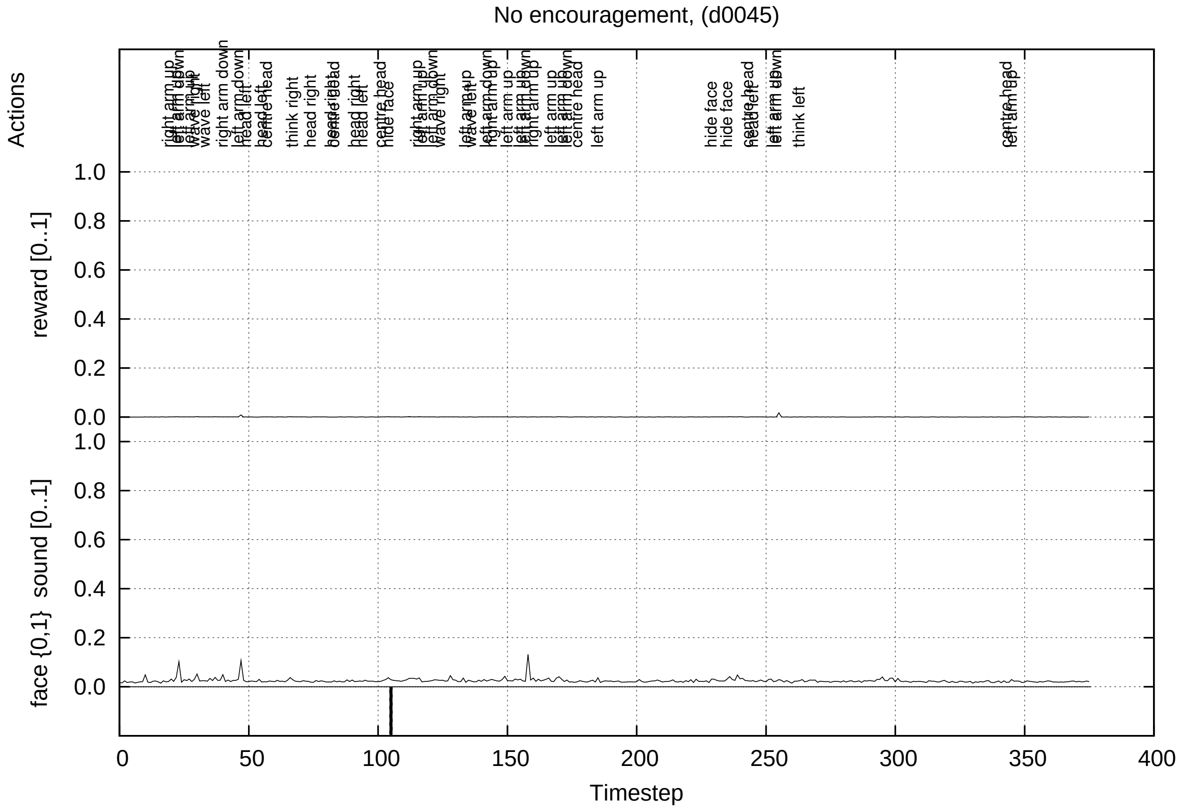


Figure 7: *Kaspar2 Results d0045. No Encouragement condition.* No encouragement is offered and the robot develops no action pattern.

3.6 Summary

The Interaction History Architecture was implemented for the upper-body humanoid robot Kaspar2. The peekaboo interaction game was used to evaluate the architecture in terms of how the robot could use its own personal interaction history to develop the capability to engage in the game. Results show that giving appropriate encouragement to the robot as it executes certain series and groups of behaviours can result in those behaviours being selected in preference to others in equivalent conditions. This result supports the hypothesis that encouraging the hiding action would result in a higher rate of peekaboo sequences than would be expected from random selection. Furthermore, encouraging alternative action sequences resulted in those actions being repeated, inviting the conclusion that this behaviour of the architecture is general and not limited to the peekaboo game. Additional support for the hypothesis was found in the conditions that offered no encouragement. In these cases no single action or sequence was selected in preference to any other, emphasizing the importance of the interaction of the environment with the robot in producing a history of interaction that can be used to develop action capabilities.

References

Jerome S. Bruner and V. Sherwood. Peekaboo and the learning of rule structures. In J.S. Bruner, A. Jolly, and K. Sylva, editors, *Play: Its Role in Development and Evolution*, pages 277–285. New York: Penguin, 1975.

A. Meltzoff and M. Moore. Explaining facial imitation: a theoretical model. *Early Development and Parenting*, 6:179–192, 1997.

Giorgio Metta, Paul Fitzpatrick, and Lorenzo Natale. YARP: Yet Another Robot Platform. *International Journal of Advanced Robotic Systems*, 3(1):43–48, 2006.

Naeem Assif Mirza, Chrystopher L. Nehaniv, Kerstin Dautenhahn, and René te Boekhorst. Grounded sensorimotor interaction histories in an information theoretic metric space for robot ontogeny. *Adaptive Behaviour*, 15(2):167–187, 2007.

Diane P. F. Montague and Arlene S. Walker-Andrews. Peekaboo: A new look at infant's perception of emotion expression. *Developmental Psychology*, 37(6):826–838, 2001.

Roy D. Pea. The social and technological dimensions of scaffolding and related theoretical concepts for learning, education and human activity. *The Journal of the Learning Sciences*, 13(3):423–451, 2004.

Philippe Rochat, Jane G. Querido, and Tricia Striano. Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology*, 35(4):950–957, 1999.

Thomas C. Veatch. A theory of humour. *International Journal of Humour Research*, 11(2):161–175, 1998.

A Kaspar2: Sensors

Table 3: Kaspar2 Sensors

Sensor	Min	Max	Description
HEAD_LR	740	2100	Head Pan Left-Right
HEAD_UD_L	640	2200	Left Neck Elevation Motor
HEAD_UD_R	820	2200	Right Neck Elevation Motor
EYES_LR	930	2060	Eyes Pan Left-Right
EYES_UD	980	1920	Eyes Up-Down
EYELIDS	1150	1700	Eyelids Open-Close
MOUTH_OPEN	600	1730	Mouth Open
MOUTH_SMILE	600	2200	Mouth Corner Elevation
ARM_R_1	650	2200	Right Shoulder Rotate
ARM_R_2	1090	2200	Right Shoulder Elevate
ARM_R_3	910	2200	Right Arm Rotate
ARM_R_4	600	2200	Right Elbow Bend
ARM_R_5	780	2200	Right Forearm Rotate
ARM_L_1	600	2200	Left Shoulder Rotate
ARM_L_2	780	2000	Left Shoulder Elevate
ARM_L_3	600	2140	Left Arm Rotate
ARM_L_4	600	2200	Left Elbow Bend
ARM_L_5	600	2200	Left Forearm Rotate
FACE	0	1	Face detection signal
SOUNDS	0	1	Sum of Sound Amplitudes for Timestep
reward	0	1	Resulting Reward sensor
action	0	20	Action

B Kaspar2 Peekaboo: Full Results

This appendix contains the results from 15 experimental runs of the history architecture running on the KasparII robot. The human interaction partner either encourages Peekaboo, another action or gives no encouragement at all.

Table 4: Actions executed (consolidated): Run d0032

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	16	2	0	1	1	16	4	1	0	0	1	0	1	0	43
chosen	33	0	0	0	16	21	0	2	0	0	0	0	1	0	73
both	49	2	0	1	17	37	4	3	0	0	1	0	2	0	116

Table 5: Actions executed (primary): Run d0032

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	20.00	0.00	10.00	40.00	10.00	10.00	0.00	10.00	0.00	100.00
chosen :	0.00	0.00	84.21	0.00	10.53	0.00	0.00	5.26	0.00	100.00
both :	6.90	0.00	58.62	13.79	10.34	3.45	0.00	6.90	0.00	100.00
Percentage Random v Chosen Actions										
random	100.00	0.0	5.88	100.00	33.33	100.00	0.0	50.00	0.0	
chosen	0.00	0.0	94.12	0.00	66.67	0.00	0.0	50.00	0.0	
Overall Chosen %:	0.00	0.0	55.17	0.00	6.90	0.00	0.0	3.45	0.0	65.52

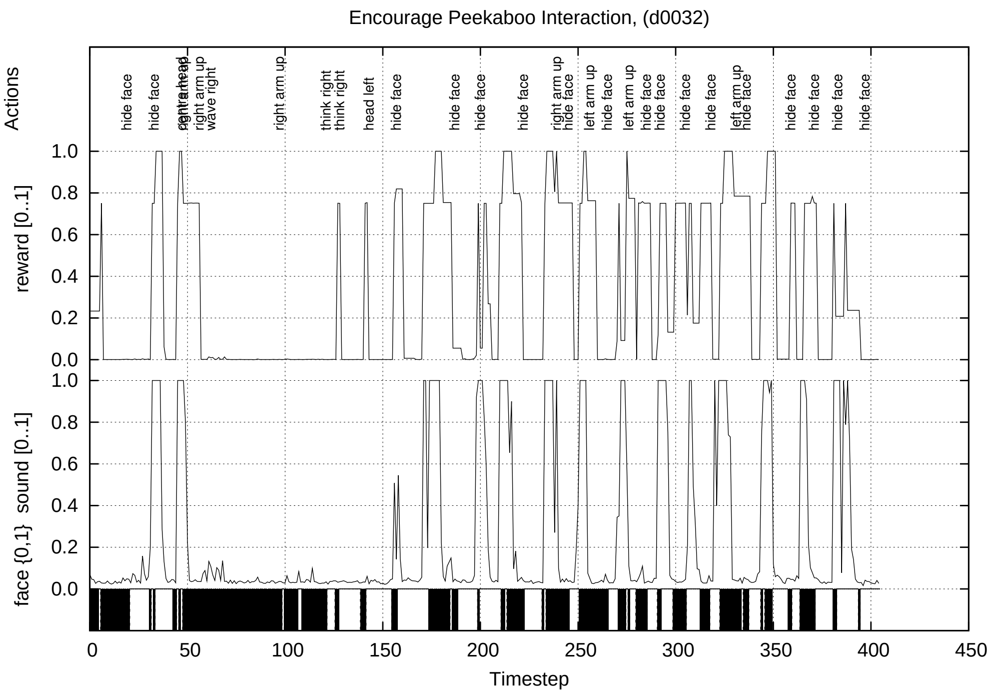


Figure 8:

Table 6: Actions executed (consolidated): Run d0033

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	27	0	0	1	3	17	5	3	1	2	1	0	4	2	66
chosen	9	0	0	0	14	3	0	0	0	0	0	0	1	1	28
both	36	0	0	1	17	20	5	3	1	2	1	0	5	3	94

Table 7: Actions executed (primary): Run d0033

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	0.00	0.00	16.67	27.78	16.67	5.56	0.00	22.22	11.11	100.00
chosen :	0.00	0.00	87.50	0.00	0.00	0.00	0.00	6.25	6.25	100.00
both :	0.00	0.00	50.00	14.71	8.82	2.94	0.00	14.71	8.82	100.00
Percentage Random v Chosen Actions										
random	0.0	0.0	17.65	100.00	100.00	100.00	0.0	80.00	66.67	
chosen	0.0	0.0	82.35	0.00	0.00	0.00	0.0	20.00	33.33	
Overall Chosen %:	0.0	0.0	41.18	0.00	0.00	0.00	0.0	2.94	2.94	47.06

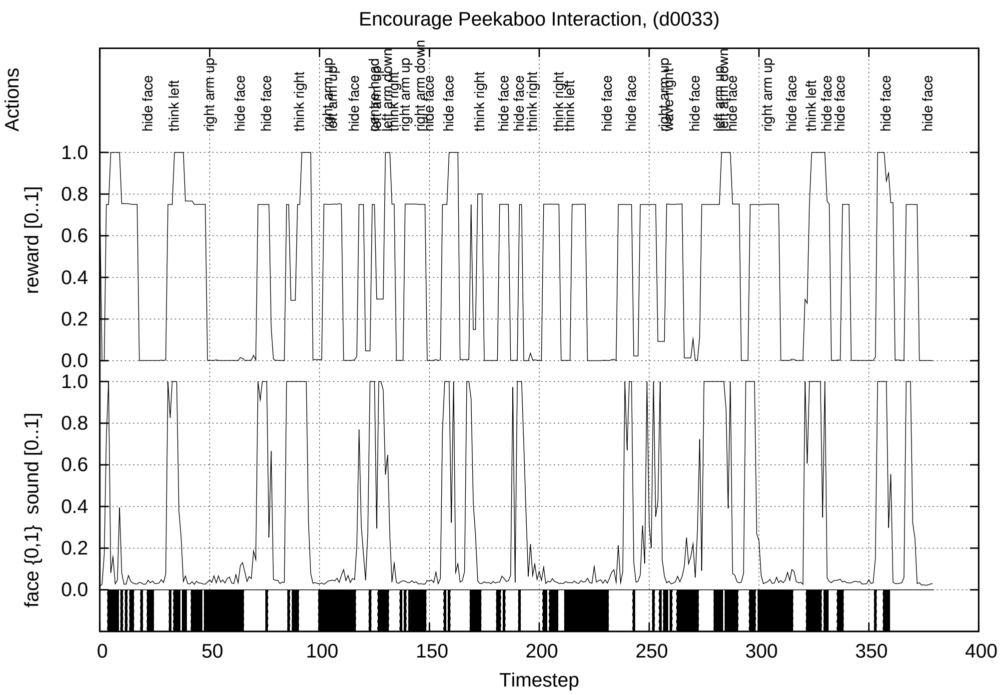


Figure 9:

Table 8: Actions executed (consolidated): Run d0034

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	7	1	0	5	2	9	4	2	2	1	2	0	0	2	37
chosen	60	1	0	0	0	0	1	0	1	0	0	0	0	0	63
both	67	2	0	5	2	9	5	2	3	1	2	0	0	2	100

Table 9: Actions executed (primary): Run d0034

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	7.69	0.00	15.38	30.77	15.38	15.38	0.00	0.00	15.38	100.00
chosen :	50.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	13.33	0.00	13.33	33.33	13.33	13.33	0.00	0.00	13.33	100.00
Percentage Random v Chosen Actions										
random	50.00	0.0	100.00	80.00	100.00	100.00	0.0	0.0	100.00	
chosen	50.00	0.0	0.00	20.00	0.00	0.00	0.0	0.0	0.00	
Overall Chosen %:	6.67	0.0	0.00	6.67	0.00	0.00	0.0	0.0	0.00	13.33

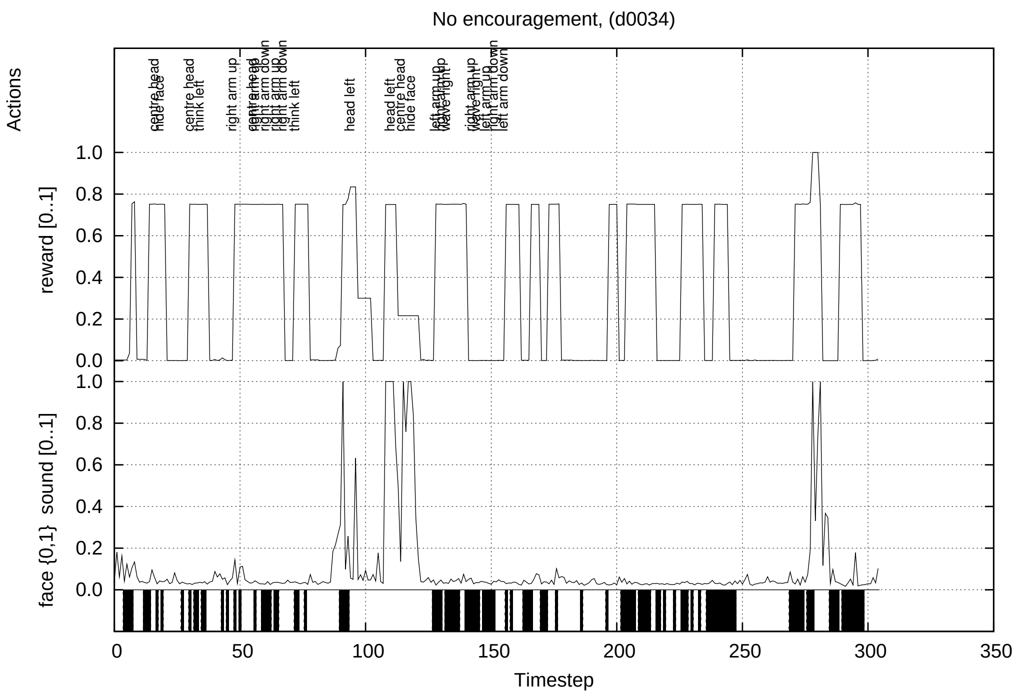


Figure 10:

Table 10: Actions executed (consolidated): Run d0035

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	17	5	2	9	4	15	3	1	1	0	1	0	1	1	60
chosen	13	15	2	11	6	2	0	0	0	0	0	0	0	0	49
both	30	20	4	20	10	17	3	1	1	0	1	0	1	1	109

Table 11: Actions executed (primary): Run d0035

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	27.78	11.11	22.22	16.67	5.56	5.56	0.00	5.56	5.56	100.00
chosen :	65.22	8.70	26.09	0.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	48.78	9.76	24.39	7.32	2.44	2.44	0.00	2.44	2.44	100.00
Percentage Random v Chosen Actions										
random	25.00	50.00	40.00	100.00	100.00	100.00	0.0	100.00	100.00	
chosen	75.00	50.00	60.00	0.00	0.00	0.00	0.0	0.00	0.00	
Overall Chosen %:	36.59	4.88	14.63	0.00	0.00	0.00	0.0	0.00	0.00	56.10

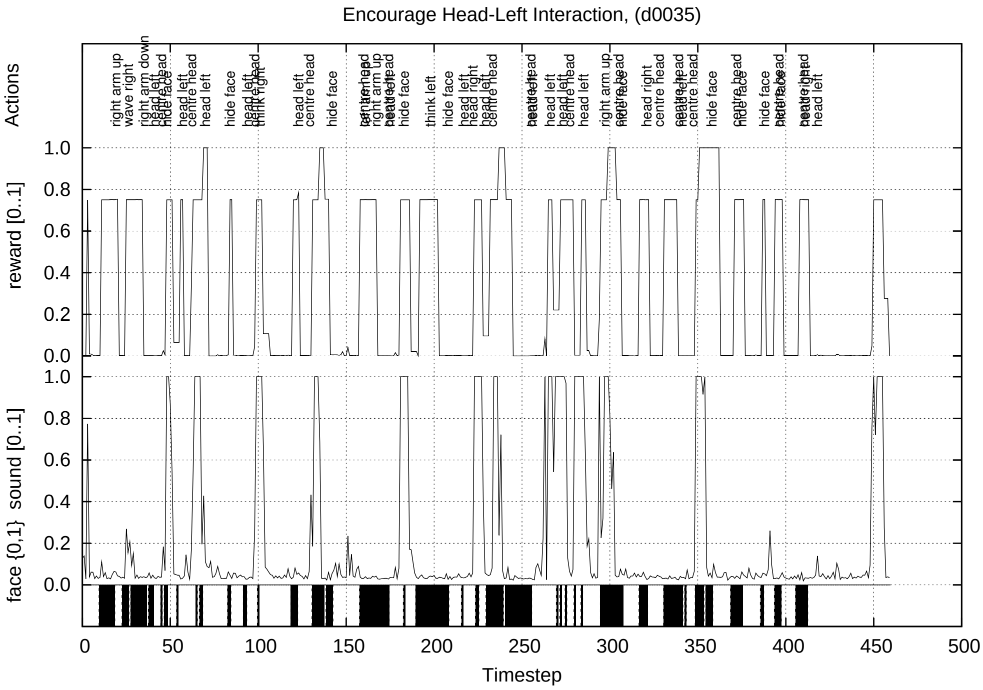


Figure 11:

Table 12: Actions executed (consolidated): Run d0036

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	12	3	1	4	3	14	1	2	1	2	0	0	1	0	44
chosen	4	0	0	0	8	7	0	0	0	0	0	0	0	0	19
both	16	3	1	4	11	21	1	2	1	2	0	0	1	0	63

Table 13: Actions executed (primary): Run d0036

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	27.27	9.09	27.27	9.09	18.18	0.00	0.00	9.09	0.00	100.00
chosen :	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	15.79	5.26	57.89	5.26	10.53	0.00	0.00	5.26	0.00	100.00
Percentage Random v Chosen Actions										
random	100.00	100.00	27.27	100.00	100.00	0.0	0.0	100.00	0.0	
chosen	0.00	0.00	72.73	0.00	0.00	0.0	0.0	0.00	0.0	
Overall Chosen %:	0.00	0.00	42.11	0.00	0.00	0.0	0.0	0.00	0.0	42.11

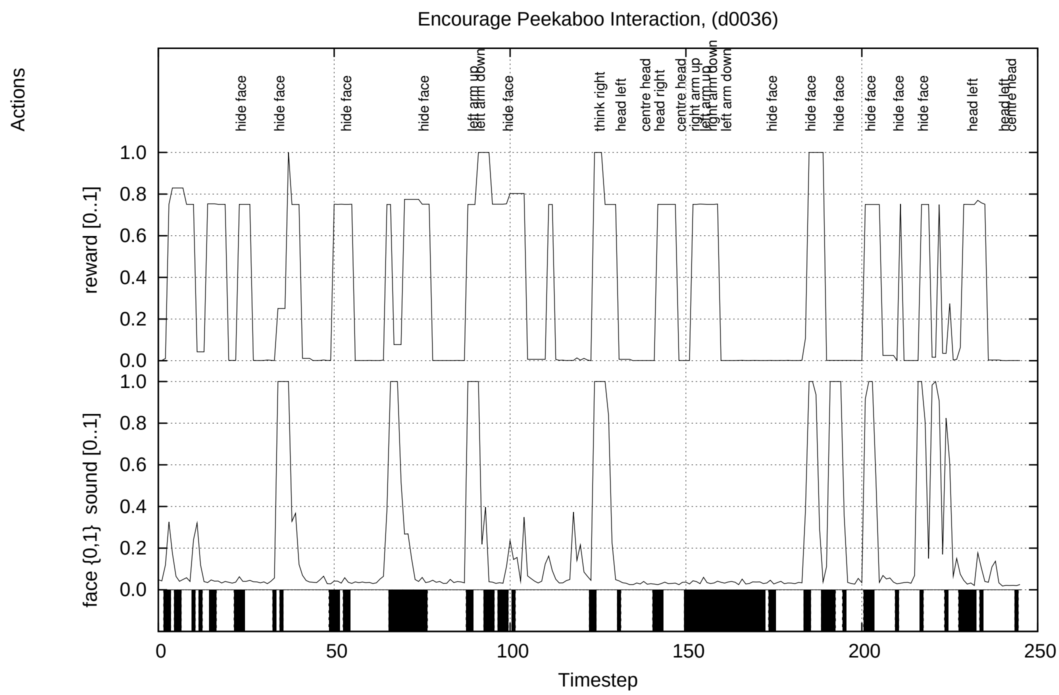


Figure 12:

Table 14: Actions executed (consolidated): Run d0037

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	5	1	2	1	1	13	3	1	1	0	1	0	2	1	32
chosen	57	0	0	6	3	0	0	0	0	0	4	0	3	0	73
both	62	1	2	7	4	13	3	1	1	0	5	0	5	1	105

Table 15: Actions executed (primary): Run d0037

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	8.33	16.67	8.33	25.00	8.33	8.33	0.00	16.67	8.33	100.00
chosen :	0.00	0.00	30.00	0.00	0.00	40.00	0.00	30.00	0.00	100.00
both :	4.55	9.09	18.18	13.64	4.55	22.73	0.00	22.73	4.55	100.00
Percentage Random v Chosen Actions										
random	100.00	100.00	25.00	100.00	100.00	20.00	0.0	40.00	100.00	
chosen	0.00	0.00	75.00	0.00	0.00	80.00	0.0	60.00	0.00	
Overall Chosen %:	0.00	0.00	13.64	0.00	0.00	18.18	0.0	13.64	0.00	45.45

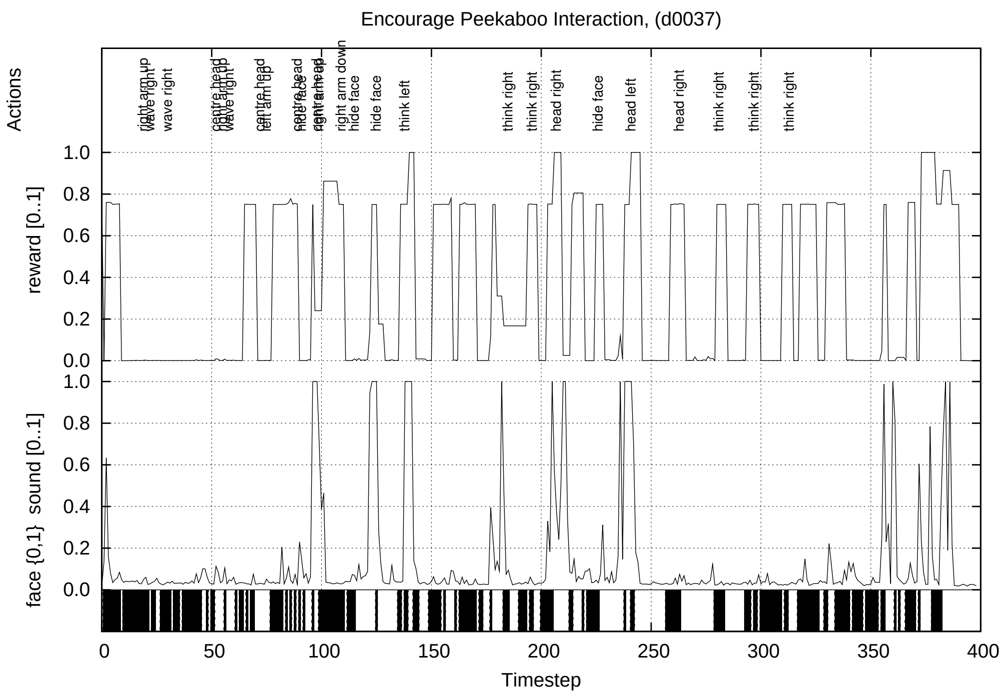


Figure 13:

Table 16: Actions executed (consolidated): Run d0038

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	3	0	0	0	0	2	0	1	0	1	0	0	0	2	9
chosen	140	0	0	0	0	22	0	0	0	0	0	0	0	5	167
both	143	0	0	0	0	24	0	1	0	1	0	0	0	7	176

Table 17: Actions executed (primary): Run d0038

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	0.00	0.00	0.00	0.00	33.33	0.00	0.00	0.00	66.67	100.00
chosen :	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	100.00
both :	0.00	0.00	0.00	0.00	12.50	0.00	0.00	0.00	87.50	100.00
Percentage Random v Chosen Actions										
random	0.0	0.0	0.0	0.0	100.00	0.0	0.0	0.0	28.57	
chosen	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	71.43	
Overall Chosen %:	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	62.50	62.50

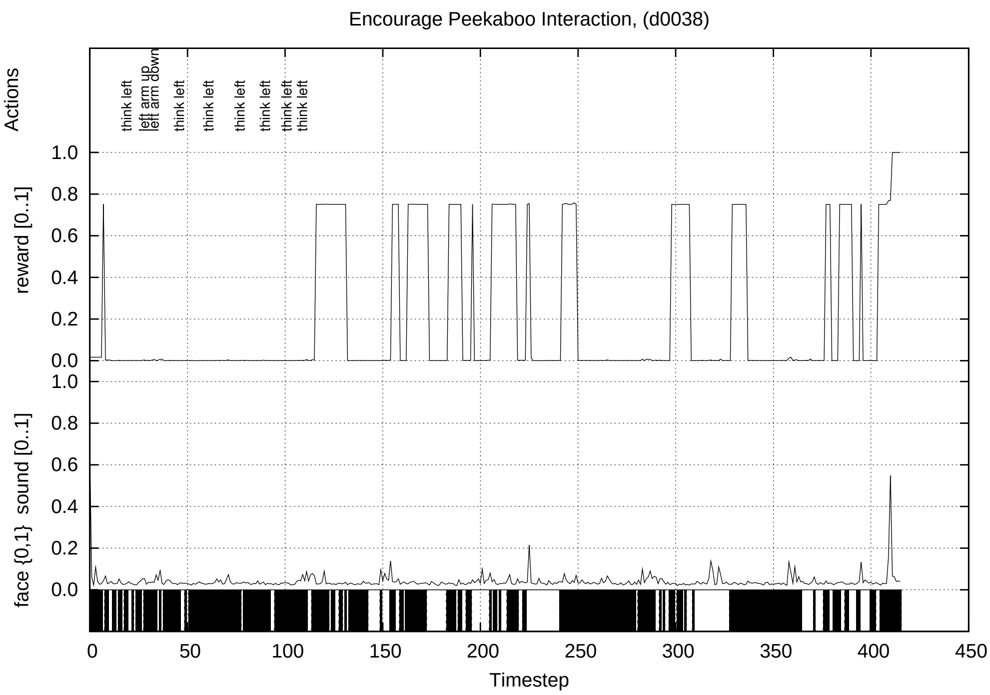


Figure 14:

Table 18: Actions executed (consolidated): Run d0039

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	5	2	0	1	2	3	1	1	0	1	0	0	0	1	17
chosen	4	0	0	0	1	14	0	0	0	0	0	0	0	0	19
both	9	2	0	1	3	17	1	1	0	1	0	0	0	1	36

Table 19: Actions executed (primary): Run d0039

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	28.57	0.00	28.57	14.29	14.29	0.00	0.00	0.00	14.29	100.00
chosen :	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	25.00	0.00	37.50	12.50	12.50	0.00	0.00	0.00	12.50	100.00
Percentage Random v Chosen Actions										
random	100.00	0.0	66.67	100.00	100.00	0.0	0.0	0.0	100.00	
chosen	0.00	0.0	33.33	0.00	0.00	0.0	0.0	0.0	0.00	
Overall Chosen %:	0.00	0.0	12.50	0.00	0.00	0.0	0.0	0.0	0.00	12.50

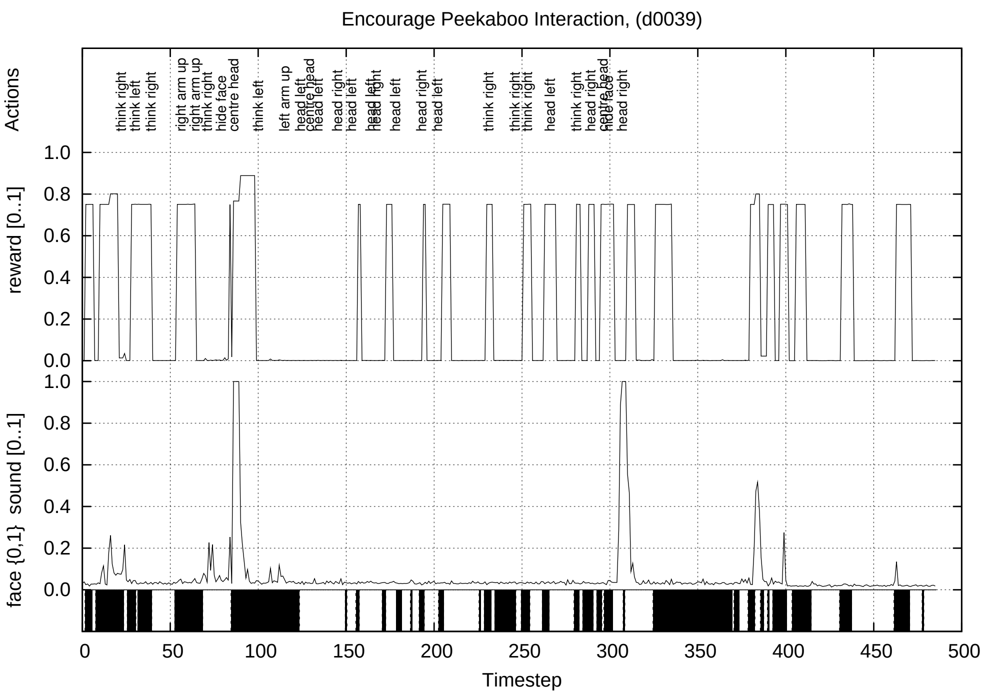


Figure 15:

Table 20: Actions executed (consolidated): Run d0041

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	25	14	6	5	4	25	5	6	1	2	0	3	1	4	101
chosen	96	16	6	0	5	10	13	5	0	0	0	0	1	2	154
both	121	30	12	5	9	35	18	11	1	2	0	3	2	6	255

Table 21: Actions executed (primary): Run d0041

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	32.56	13.95	9.30	11.63	13.95	0.00	6.98	2.33	9.30	100.00
chosen :	33.33	12.50	10.42	27.08	10.42	0.00	0.00	2.08	4.17	100.00
both :	32.97	13.19	9.89	19.78	12.09	0.00	3.30	2.20	6.59	100.00
Percentage Random v Chosen Actions										
random	46.67	50.00	44.44	27.78	54.55	0.0	100.00	50.00	66.67	
chosen	53.33	50.00	55.56	72.22	45.45	0.0	0.00	50.00	33.33	
Overall Chosen %:	17.58	6.59	5.49	14.29	5.49	0.0	0.00	1.10	2.20	52.75

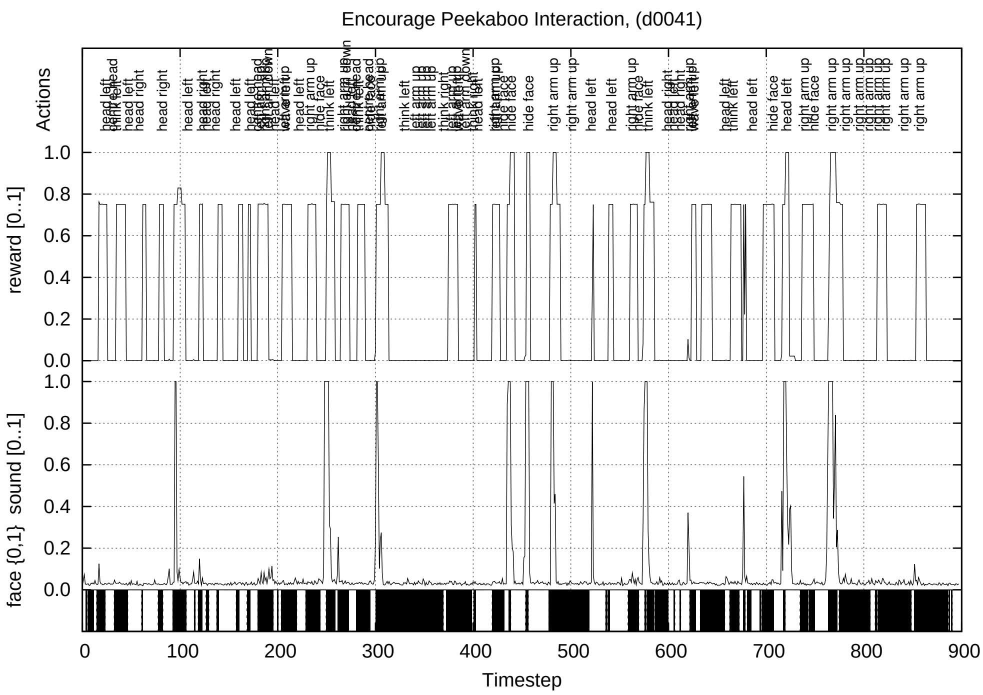


Figure 16:

Table 22: Actions executed (consolidated): Run d0042

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	28	6	8	11	4	22	5	3	1	0	0	2	3	1	94
chosen	8	1	5	8	6	13	0	1	0	0	0	16	0	1	59
both	36	7	13	19	10	35	5	4	1	0	0	18	3	2	153

Table 23: Actions executed (primary): Run d0042

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	18.75	25.00	12.50	15.62	9.38	0.00	6.25	9.38	3.12	100.00
chosen :	3.33	16.67	20.00	0.00	3.33	0.00	53.33	0.00	3.33	100.00
both :	11.29	20.97	16.13	8.06	6.45	0.00	29.03	4.84	3.23	100.00
Percentage Random v Chosen Actions										
random	85.71	61.54	40.00	100.00	75.00	0.0	11.11	100.00	50.00	
chosen	14.29	38.46	60.00	0.00	25.00	0.0	88.89	0.00	50.00	
Overall Chosen %:	1.61	8.06	9.68	0.00	1.61	0.0	25.81	0.00	1.61	48.39

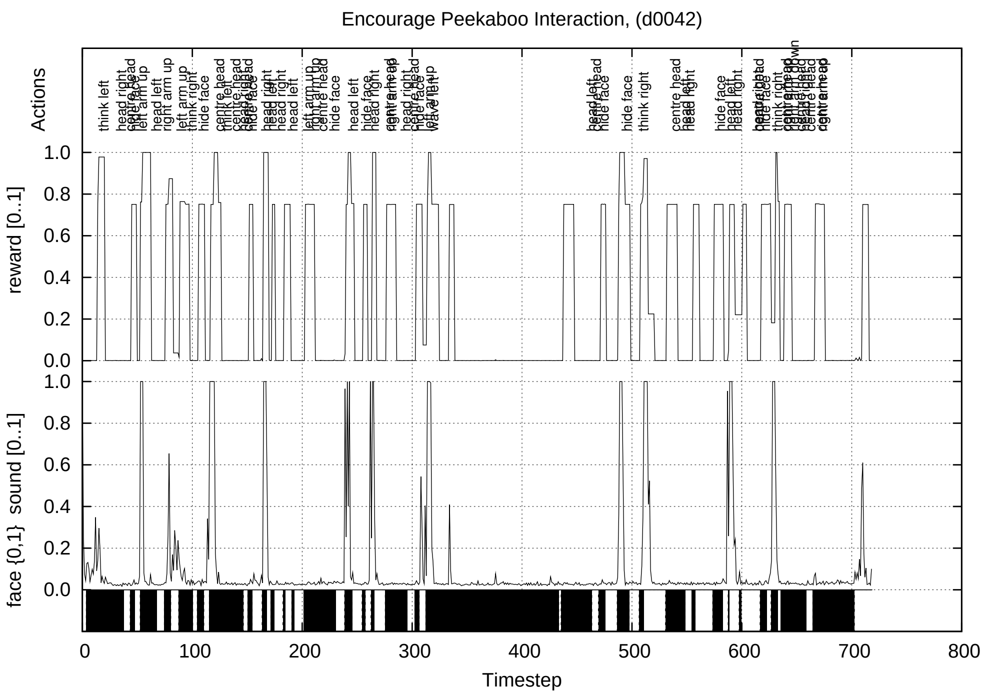


Figure 17:

Table 24: Actions executed (consolidated): Run d0043

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	23	11	9	13	1	22	5	4	2	1	3	2	4	5	105
chosen	32	19	26	4	1	11	0	0	0	0	0	0	0	2	95
both	55	30	35	17	2	33	5	4	2	1	3	2	4	7	200

Table 25: Actions executed (primary): Run d0043

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	25.00	20.45	2.27	11.36	9.09	6.82	4.55	9.09	11.36	100.00
chosen :	39.58	54.17	2.08	0.00	0.00	0.00	0.00	0.00	4.17	100.00
both :	32.61	38.04	2.17	5.43	4.35	3.26	2.17	4.35	7.61	100.00
Percentage Random v Chosen Actions										
random	36.67	25.71	50.00	100.00	100.00	100.00	100.00	100.00	71.43	
chosen	63.33	74.29	50.00	0.00	0.00	0.00	0.00	0.00	28.57	
Overall Chosen %:	20.65	28.26	1.09	0.00	0.00	0.00	0.00	0.00	2.17	52.17

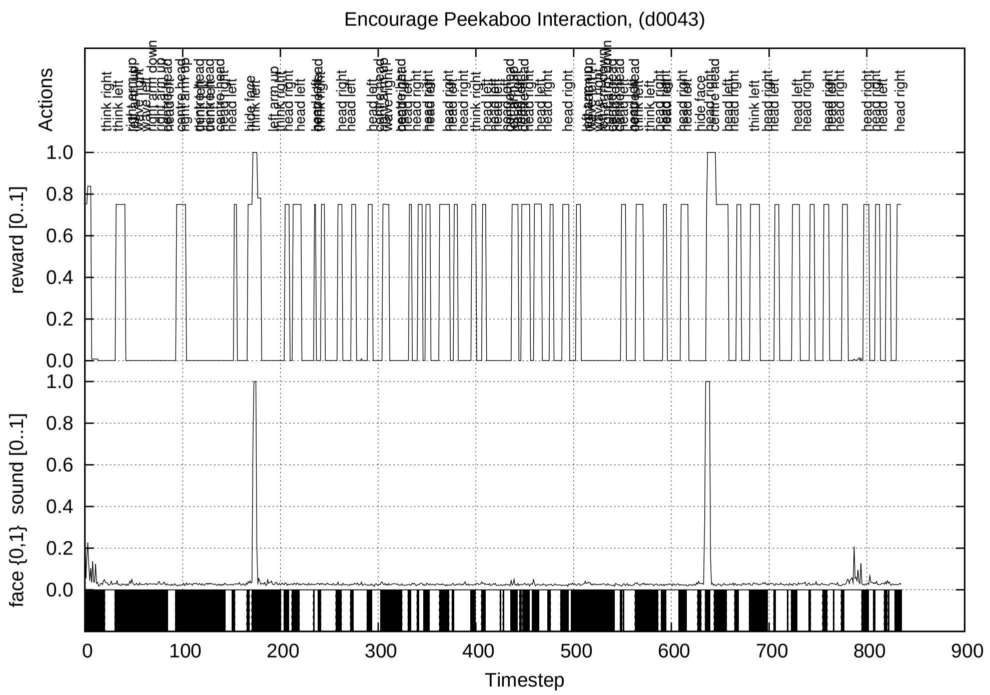


Figure 18:

Table 26: Actions executed (consolidated): Run d0044

	0	3	4	5	6	7	8	9	10	11	12	13	14	15	Tot
	Rst	HL	HR	HF	Hid	NA	RAU	LAU	RAD	LAD	RAW	LAW	TR	TL	
random	32	9	3	7	3	13	6	4	1	4	0	1	5	2	90
chosen	15	3	1	1	10	4	1	1	0	1	0	0	3	1	41
both	47	12	4	8	13	17	7	5	1	5	0	1	8	3	131

Table 27: Actions executed (primary): Run d0044

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	27.27	9.09	9.09	18.18	12.12	0.00	3.03	15.15	6.06	100.00
chosen :	15.00	5.00	50.00	5.00	5.00	0.00	0.00	15.00	5.00	100.00
both :	22.64	7.55	24.53	13.21	9.43	0.00	1.89	15.09	5.66	100.00
Percentage Random v Chosen Actions										
random	75.00	75.00	23.08	85.71	80.00	0.0	100.00	62.50	66.67	
chosen	25.00	25.00	76.92	14.29	20.00	0.0	0.00	37.50	33.33	
Overall Chosen %:	5.66	1.89	18.87	1.89	1.89	0.0	0.00	5.66	1.89	37.74

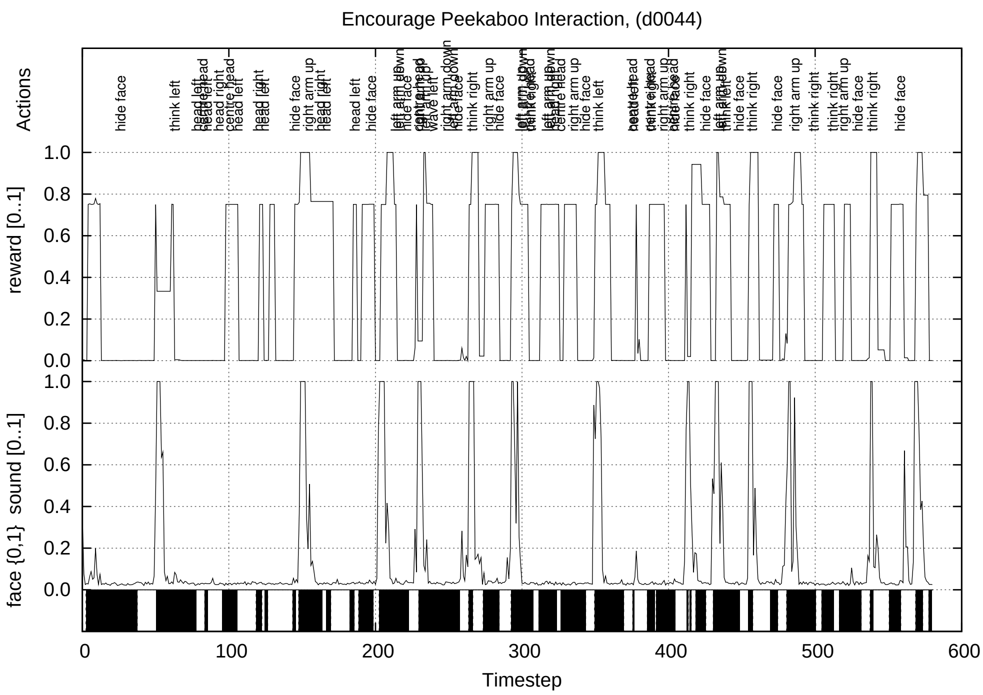


Figure 19:

Table 28: Actions executed (consolidated): Run d0045

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	7	4	3	6	3	10	3	8	1	5	2	2	1	1	56
chosen	57	1	0	0	0	32	1	3	0	2	0	0	0	0	96
both	64	5	3	6	3	42	4	11	1	7	2	2	1	1	152

Table 29: Actions executed (primary): Run d0045

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	14.81	11.11	11.11	11.11	29.63	7.41	7.41	3.70	3.70	100.00
chosen :	20.00	0.00	0.00	20.00	60.00	0.00	0.00	0.00	0.00	100.00
both :	15.62	9.38	9.38	12.50	34.38	6.25	6.25	3.12	3.12	100.00
Percentage Random v Chosen Actions										
random	80.00	100.00	100.00	75.00	72.73	100.00	100.00	100.00	100.00	
chosen	20.00	0.00	0.00	25.00	27.27	0.00	0.00	0.00	0.00	
Overall Chosen %:	3.12	0.00	0.00	3.12	9.38	0.00	0.00	0.00	0.00	15.62

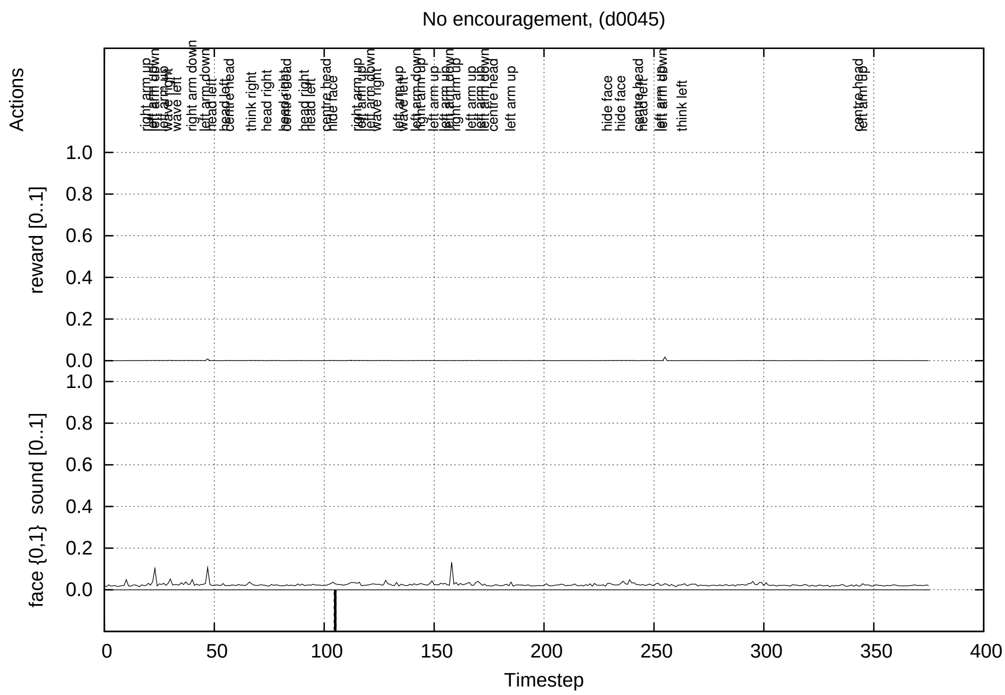


Figure 20:

Table 30: Actions executed (consolidated): Run d0046

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	21	6	2	5	3	20	4	6	0	2	1	1	4	3	78
chosen	32	9	0	3	2	19	1	1	0	0	0	30	3	0	100
both	53	15	2	8	5	39	5	7	0	2	1	31	7	3	178

Table 31: Actions executed (primary): Run d0046

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	20.00	6.67	10.00	13.33	20.00	3.33	3.33	13.33	10.00	100.00
chosen :	19.57	0.00	4.35	2.17	2.17	0.00	65.22	6.52	0.00	100.00
both :	19.74	2.63	6.58	6.58	9.21	1.32	40.79	9.21	3.95	100.00
Percentage Random v Chosen Actions										
random	40.00	100.00	60.00	80.00	85.71	100.00	3.23	57.14	100.00	
chosen	60.00	0.00	40.00	20.00	14.29	0.00	96.77	42.86	0.00	
Overall Chosen %:	11.84	0.00	2.63	1.32	1.32	0.00	39.47	3.95	0.00	60.53

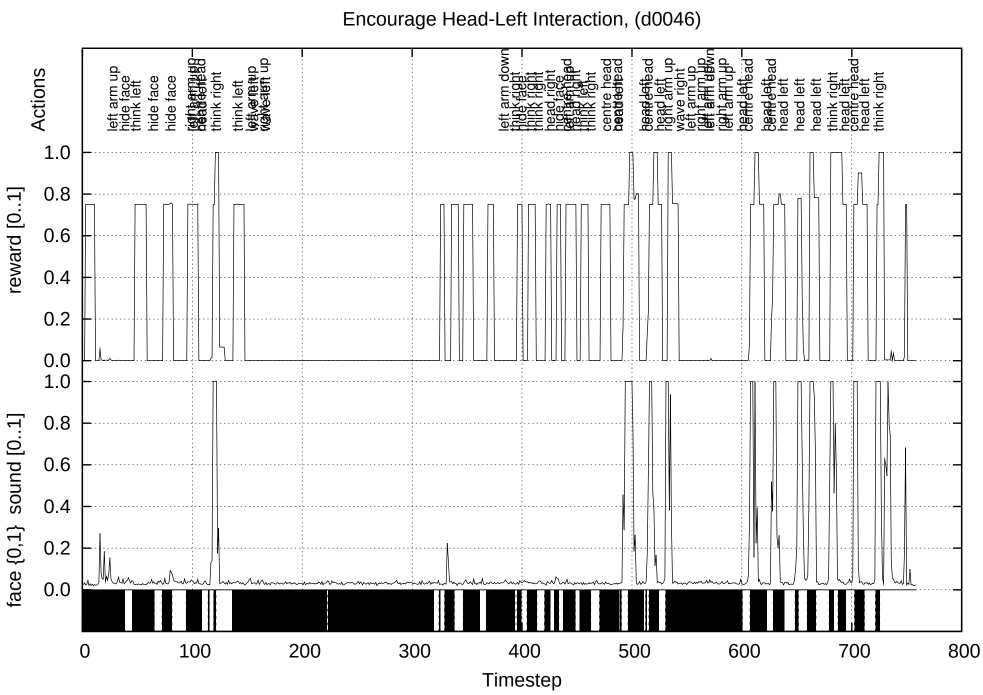


Figure 21:

Table 32: Actions executed (consolidated): Run d0049

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	33	14	6	10	4	29	7	6	4	4	1	2	7	6	133
chosen	87	13	9	3	3	25	0	0	0	0	0	0	12	2	154
both	120	27	15	13	7	54	7	6	4	4	1	2	19	8	287

Table 33: Actions executed (primary): Run d0049

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	26.42	11.32	7.55	13.21	11.32	1.89	3.77	13.21	11.32	100.00
chosen :	33.33	23.08	7.69	0.00	0.00	0.00	0.00	30.77	5.13	100.00
both :	29.35	16.30	7.61	7.61	6.52	1.09	2.17	20.65	8.70	100.00
Percentage Random v Chosen Actions										
random	51.85	40.00	57.14	100.00	100.00	100.00	100.00	36.84	75.00	
chosen	48.15	60.00	42.86	0.00	0.00	0.00	0.00	63.16	25.00	
Overall Chosen %:	14.13	9.78	3.26	0.00	0.00	0.00	0.00	13.04	2.17	42.39

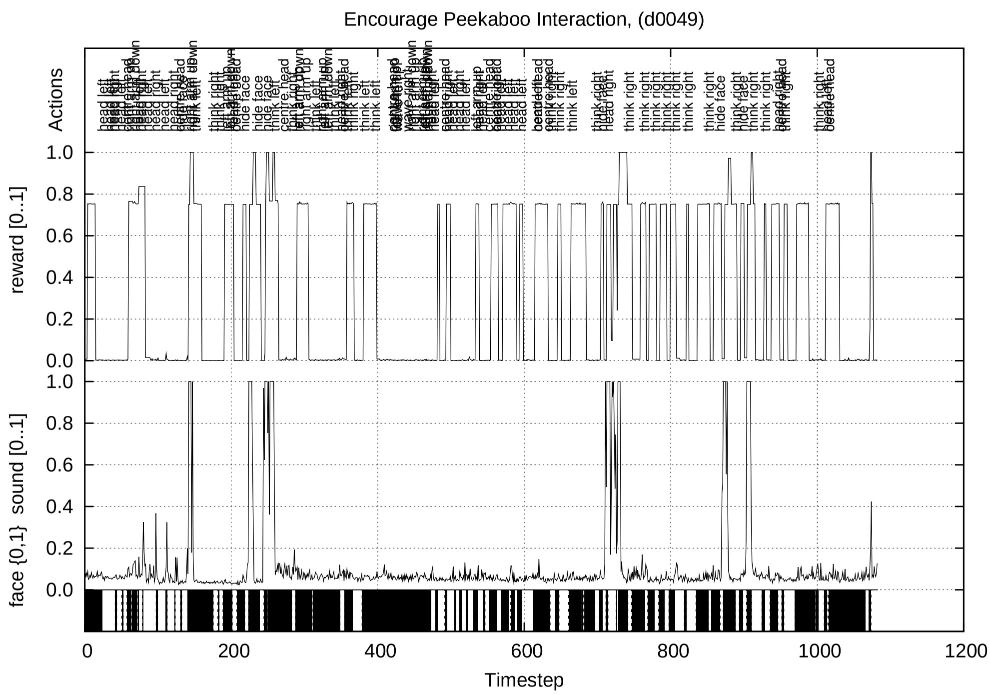


Figure 22:

Table 34: Actions executed (consolidated): Run d0050

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	19	7	7	3	2	12	2	3	0	2	0	0	2	0	59
chosen	7	5	0	0	10	2	0	0	0	0	0	0	0	0	24
both	26	12	7	3	12	14	2	3	0	2	0	0	2	0	83

Table 35: Actions executed (primary): Run d0050

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	30.43	30.43	8.70	8.70	13.04	0.00	0.00	8.70	0.00	100.00
chosen :	33.33	0.00	66.67	0.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	31.58	18.42	31.58	5.26	7.89	0.00	0.00	5.26	0.00	100.00
Percentage Random v Chosen Actions										
random	58.33	100.00	16.67	100.00	100.00	0.0	0.0	100.00	0.0	
chosen	41.67	0.00	83.33	0.00	0.00	0.0	0.0	0.00	0.0	
Overall Chosen %:	13.16	0.00	26.32	0.00	0.00	0.0	0.0	0.00	0.0	39.47

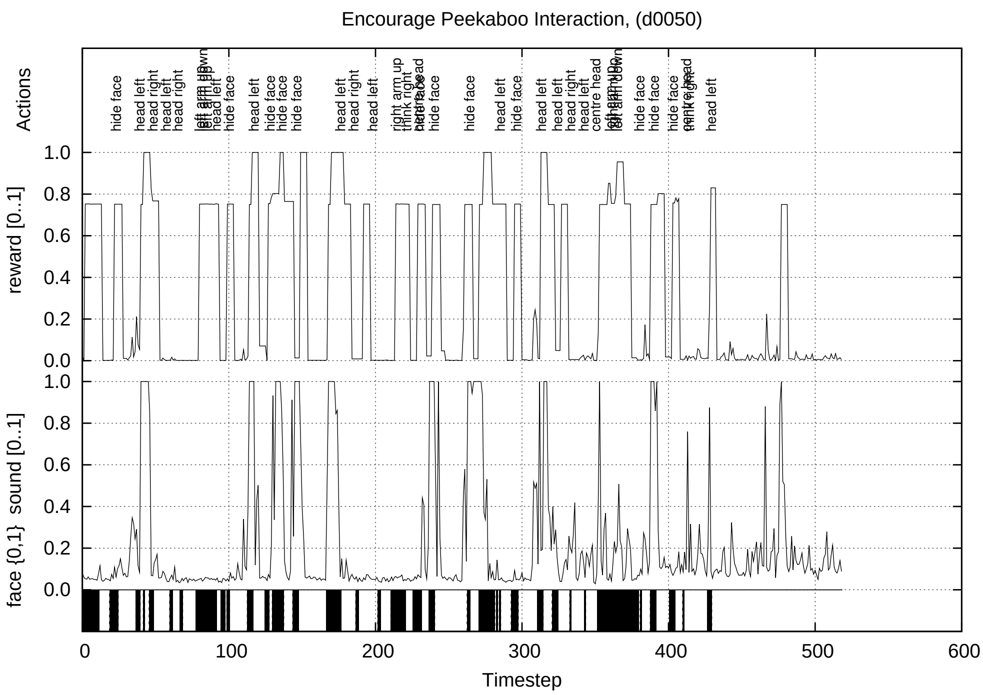


Figure 23:

Table 36: Actions executed (consolidated): Run d0051

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	32	6	3	7	4	27	10	12	5	2	2	6	1	6	123
chosen	53	3	0	0	17	5	3	2	0	0	0	0	0	13	96
both	85	9	3	7	21	32	13	14	5	2	2	6	1	19	219

Table 37: Actions executed (primary): Run d0051

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	12.00	6.00	8.00	20.00	24.00	4.00	12.00	2.00	12.00	100.00
chosen :	7.89	0.00	44.74	7.89	5.26	0.00	0.00	0.00	34.21	100.00
both :	10.23	3.41	23.86	14.77	15.91	2.27	6.82	1.14	21.59	100.00
Percentage Random v Chosen Actions										
random	66.67	100.00	19.05	76.92	85.71	100.00	100.00	100.00	31.58	
chosen	33.33	0.00	80.95	23.08	14.29	0.00	0.00	0.00	68.42	
Overall Chosen %:	3.41	0.00	19.32	3.41	2.27	0.00	0.00	0.00	14.77	43.18

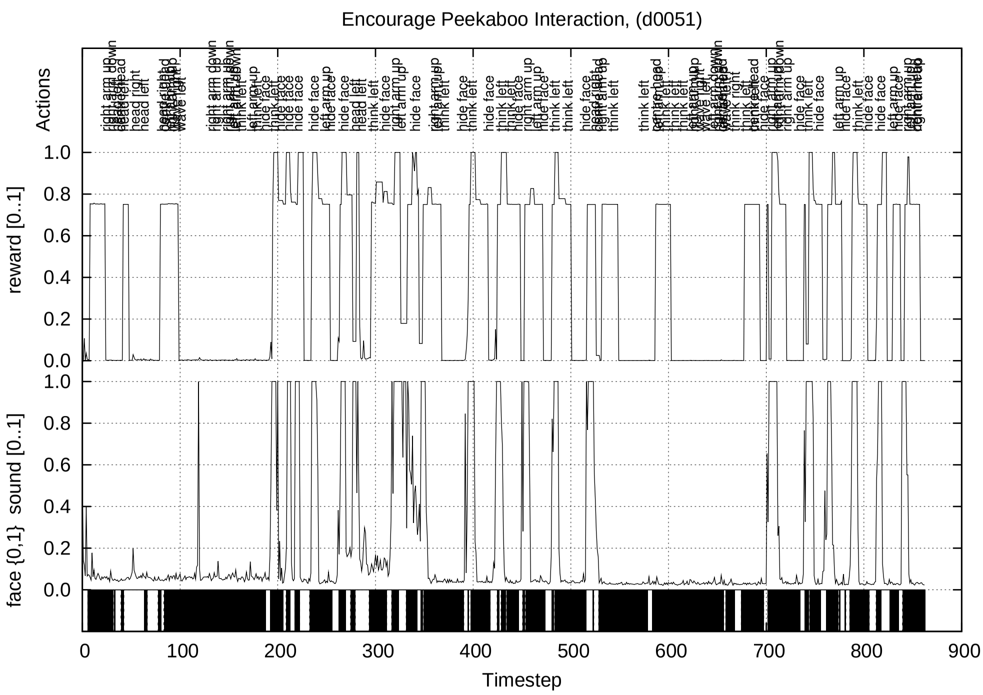


Figure 24:

Table 38: Actions executed (consolidated): Run d0052

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	22	23	8	13	2	24	7	3	4	0	6	1	7	5	125
chosen	74	31	9	16	6	11	3	0	0	0	2	0	6	2	160
both	96	54	17	29	8	35	10	3	4	0	8	1	13	7	285

Table 39: Actions executed (primary): Run d0052

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	37.10	12.90	3.23	11.29	4.84	9.68	1.61	11.29	8.06	100.00
chosen :	52.54	15.25	10.17	5.08	0.00	3.39	0.00	10.17	3.39	100.00
both :	44.63	14.05	6.61	8.26	2.48	6.61	0.83	10.74	5.79	100.00
Percentage Random v Chosen Actions										
random	42.59	47.06	25.00	70.00	100.00	75.00	100.00	53.85	71.43	
chosen	57.41	52.94	75.00	30.00	0.00	25.00	0.00	46.15	28.57	
Overall Chosen %:	25.62	7.44	4.96	2.48	0.00	1.65	0.00	4.96	1.65	48.76

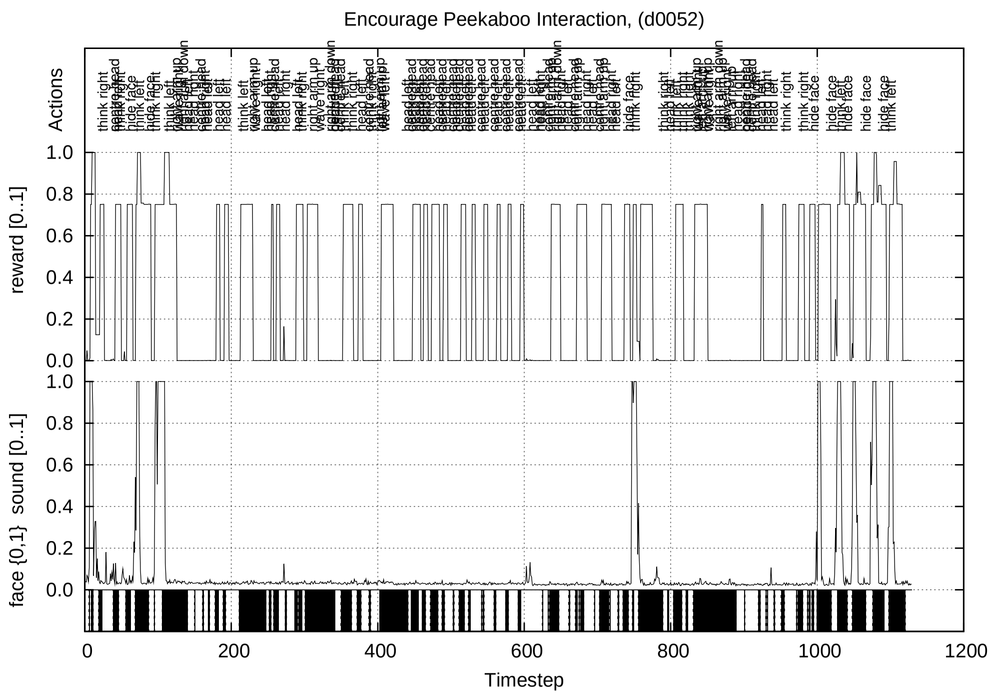


Figure 25:

Table 40: Actions executed (consolidated): Run d0053

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	31	3	5	6	6	21	5	7	5	4	5	2	3	3	106
chosen	26	1	1	21	11	4	7	0	1	0	0	0	3	1	76
both	57	4	6	27	17	25	12	7	6	4	5	2	6	4	182

Table 41: Actions executed (primary): Run d0053

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	7.69	12.82	15.38	12.82	17.95	12.82	5.13	7.69	7.69	100.00
chosen :	4.17	4.17	45.83	29.17	0.00	0.00	0.00	12.50	4.17	100.00
both :	6.35	9.52	26.98	19.05	11.11	7.94	3.17	9.52	6.35	100.00
Percentage Random v Chosen Actions										
random	75.00	83.33	35.29	41.67	100.00	100.00	100.00	50.00	75.00	
chosen	25.00	16.67	64.71	58.33	0.00	0.00	0.00	50.00	25.00	
Overall Chosen %:	1.59	1.59	17.46	11.11	0.00	0.00	0.00	4.76	1.59	38.10

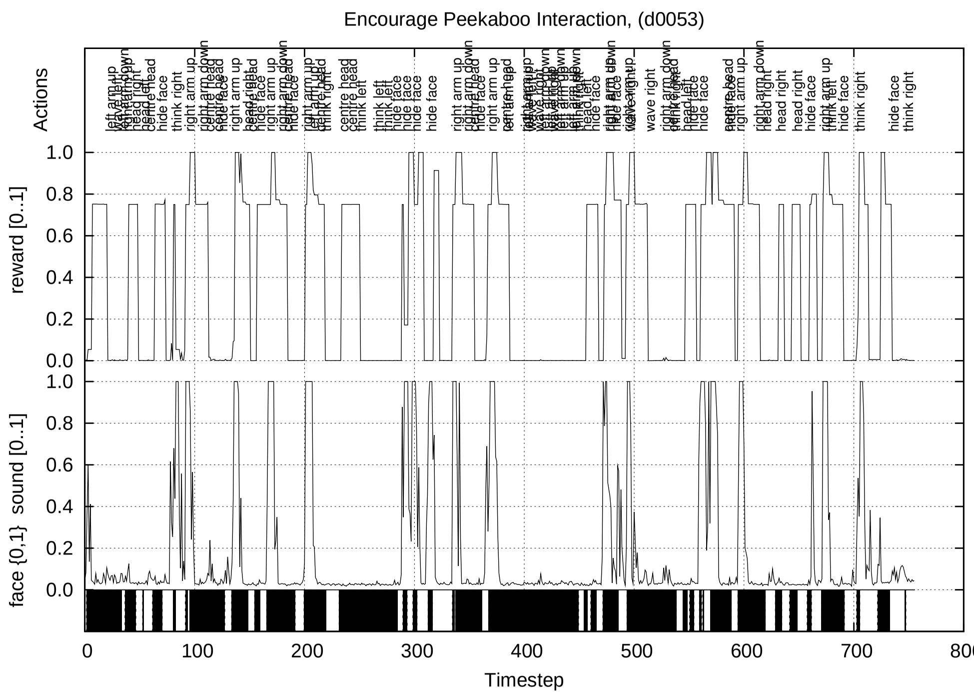


Figure 26:

Table 42: Actions executed (consolidated): Run d0054

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	47	3	6	8	3	57	3	2	2	2	1	0	3	1	138
chosen	19	3	1	49	42	8	0	0	0	0	0	0	0	0	122
both	66	6	7	57	45	65	3	2	2	2	1	0	3	1	260

Table 43: Actions executed (primary): Run d0054

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	13.64	27.27	13.64	13.64	9.09	4.55	0.00	13.64	4.55	100.00
chosen :	6.52	2.17	91.30	0.00	0.00	0.00	0.00	0.00	0.00	100.00
both :	8.82	10.29	66.18	4.41	2.94	1.47	0.00	4.41	1.47	100.00
Percentage Random v Chosen Actions										
random	50.00	85.71	6.67	100.00	100.00	100.00	0.0	100.00	100.00	
chosen	50.00	14.29	93.33	0.00	0.00	0.00	0.0	0.00	0.00	
Overall Chosen %:	4.41	1.47	61.76	0.00	0.00	0.00	0.0	0.00	0.00	67.65

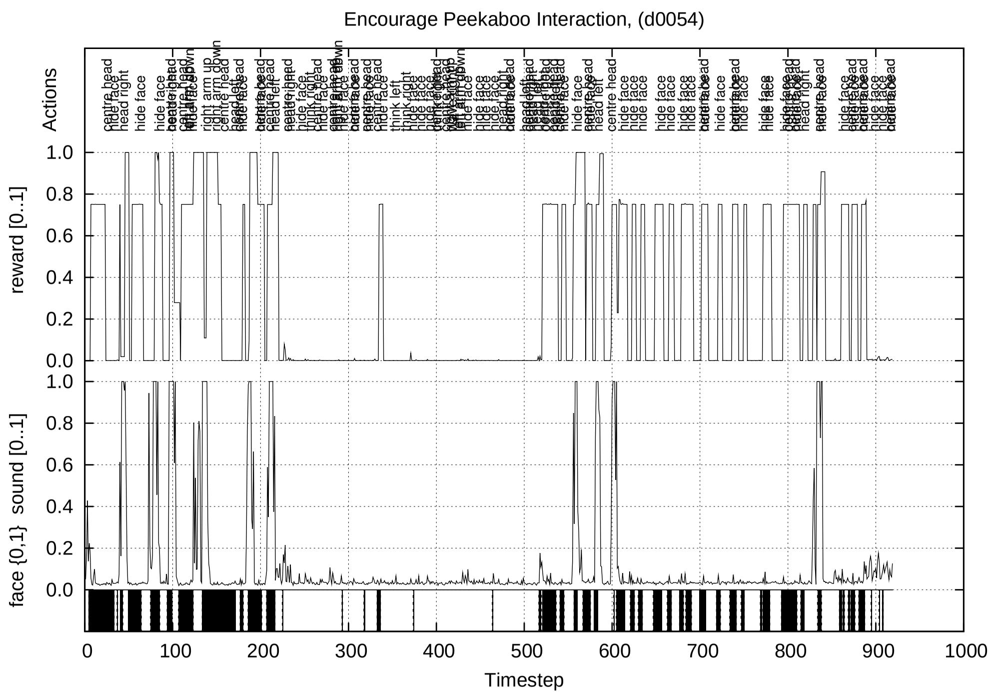


Figure 27:

Table 44: Actions executed (consolidated): Run d0055

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	23	7	5	10	4	19	3	7	2	3	0	2	4	0	89
chosen	20	5	0	1	0	39	0	0	0	0	0	0	13	0	78
both	43	12	5	11	4	58	3	7	2	3	0	2	17	0	167

Table 45: Actions executed (primary): Run d0055

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	21.88	15.62	12.50	9.38	21.88	0.00	6.25	12.50	0.00	100.00
chosen :	27.78	0.00	0.00	0.00	0.00	0.00	0.00	72.22	0.00	100.00
both :	24.00	10.00	8.00	6.00	14.00	0.00	4.00	34.00	0.00	100.00
Percentage Random v Chosen Actions										
random	58.33	100.00	100.00	100.00	100.00	0.0	100.00	23.53	0.0	
chosen	41.67	0.00	0.00	0.00	0.00	0.0	0.00	76.47	0.0	
Overall Chosen %:	10.00	0.00	0.00	0.00	0.00	0.0	0.00	26.00	0.0	36.00

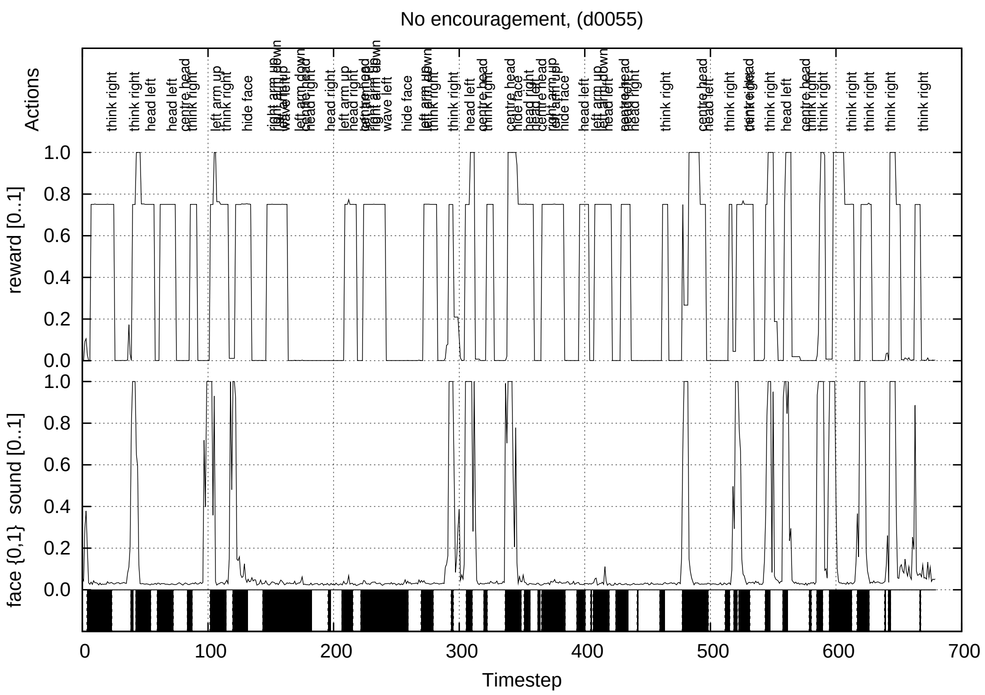


Figure 28:

Table 46: Actions executed (consolidated): Run d0056

	0 Rst	3 HL	4 HR	5 HF	6 Hid	7 NA	8 RAU	9 LAU	10 RAD	11 LAD	12 RAW	13 LAW	14 TR	15 TL	Tot
random	21	10	6	9	7	22	11	8	6	2	5	0	0	2	109
chosen	18	10	8	0	2	6	2	0	1	0	8	0	0	0	55
both	39	20	14	9	9	28	13	8	7	2	13	0	0	2	164

Table 47: Actions executed (primary): Run d0056

	HL	HR	Hid	RAU	LAU	RAW	LAW	TR	TL	total
Frequency As Percentage of Primary Actions										
random :	20.41	12.24	14.29	22.45	16.33	10.20	0.00	0.00	4.08	100.00
chosen :	33.33	26.67	6.67	6.67	0.00	26.67	0.00	0.00	0.00	100.00
both :	25.32	17.72	11.39	16.46	10.13	16.46	0.00	0.00	2.53	100.00
Percentage Random v Chosen Actions										
random	50.00	42.86	77.78	84.62	100.00	38.46	0.0	0.0	100.00	
chosen	50.00	57.14	22.22	15.38	0.00	61.54	0.0	0.0	0.00	
Overall Chosen %:	12.66	10.13	2.53	2.53	0.00	10.13	0.0	0.0	0.00	37.97

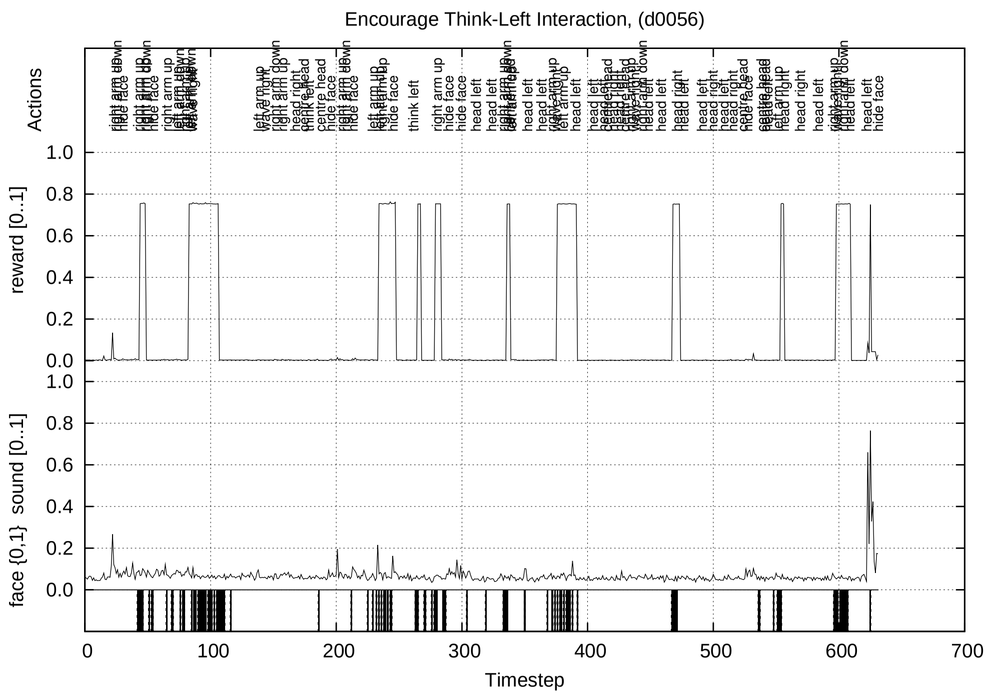


Figure 29:

C Kaspar2 Peekaboo: Action State Diagram

Actions					
0	Rest	8	RArm-Up	17	RArm-Up2
1	Smile	9	LArm-Up	18	LArm-Up2
2	Neutral	10	RArm-Down	19	RArm-Down2
3	head-left	11	LArm-Down	20	LArm-Down2
4	head-right	12	RArm-Wave		
5	head-forward	13	LArm-Wave		
6	Hide-face	14	ThinkR		
7	NA	15	ThinkL		
		16	Frown		

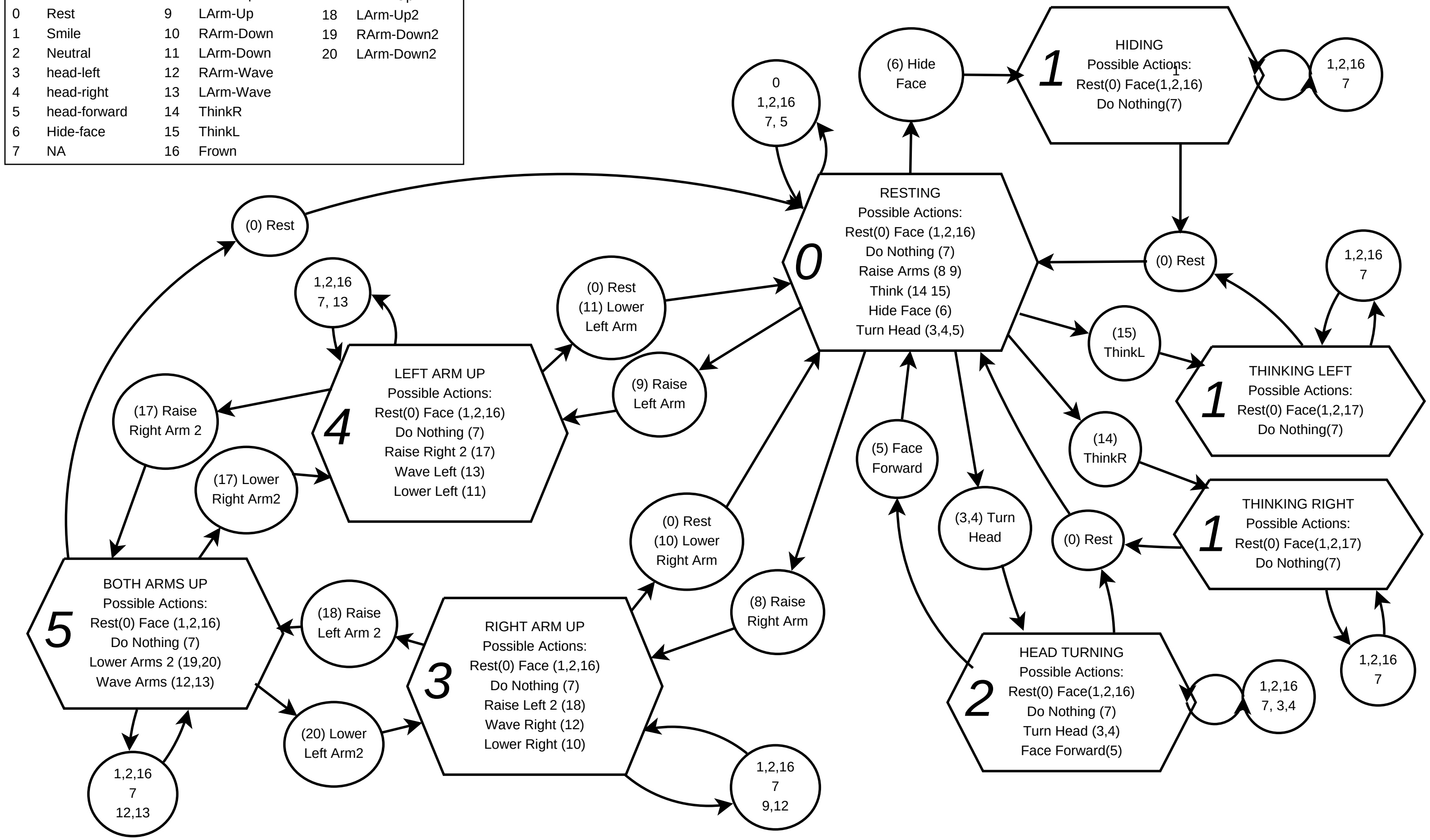


Figure 30: Kaspar2 Action States