

# Embodied Reinforcement Learning

René Ahn, Loe Feijs, Saskia Bakker, Sibrecht Bouwstra, Jos Verbeek, Bram van der Vlist, Arne Wessels en Rick van de Westelaken. Eindhoven University of Technology, Department of Industrial Design.

## Technology Masterclass

As part of the curriculum of the Master Industrial Design, 6 students participate in the "Masterclass Technology". This masterclass is a six week course in which several subjects are taught. One of these six weeks is devoted to the subject 'adaptive learning', taught by dr ir René Ahn and dr ir Emilia Barakova of the Designed Intelligence Group. During this week, different learning mechanisms were discussed, such as unsupervised and supervised learning and reinforcement learning or Q-learning. To put these theories into practice, two robots were built that implement the Q-learning algorithm; Johnny Q and The Crawler. The robots are built using Lego Mindstroms NXT and the algorithms are written in Java.

**TU/e** Technische Universiteit Eindhoven University of Technology department of industrial design



reward button



light sensor



ultrasonic sensor

## Johnny Q

Johnny Q has wheels and left-right motor drives to move forward, backward, rotate left and right. Johnny Q measures the brightness of the floor and "sees" the distance from a wall or reference object. Inside is an NXT control brick, an embedded processor programmed in Java to execute the reinforcement learning algorithm (Q-learning). The reward is being hit on the shoulder; a simple button serves to count touches. Johnny Q learns by being trained. Depending on what the human user does or does not reward, Johnny Q learns behaviours, such as turning away from a dark spot, or running backwards near an obstacle. But it can also learn the opposite behaviour, bumping against the wall. It explores its possibilities and learns how to accumulate maximal rewards. The observer engages in a training session, teaching tricks and little games, much like training a dog. Usually this algorithm is demonstrated through screen demos but here the potential of embodied learning is visible in a truly embodied model. From a semantic point of view, it is interesting to sculpt the behaviour which (of course) requires some patience. Johnny Q will gradually forget although desired behaviour can be maintained through continued training.