

From Motor Babbling to Planning

Cornelius Weber

Learning and behaviour of mobile robots has many limitations. In reinforcement learning, for example, an agent learns a strategy to get to only one specific target point within a state space. However, we can grasp a visually seen object at any point in space, or navigate to any position in a room. I will present a model in which an agent learns a model of the state space that allows it to get from any position to an arbitrarily chosen goal via a short route. By random exploration (motor babbling), the agent learns associations between two adjoining states. Furthermore, it learns a forward- and an inverse model of the action that links any two states. Given a starting and a goal position, route-finding (planning) is done in two phases. In the first phase, an activation gradient spreads around the goal travelling along the associative connections that link any two adjoining states and reaches as far as to the agent. These activations mimic the value function in reinforcement learning that gets larger toward the goal. The second phase consists of two steps that are done in alternation. First, the agent uses its current state and the adjoining state with highest value on the activation gradient to determine the action that leads it to that state up the gradient. This involves the inverse model. The next step is either to perform or to imagine the action using the forward model. In the latter case, the agent uses its current state and the chosen action to predict and activate that next state, which then becomes the current state. By repeating the two steps, the agent ascends the gradient toward the goal. A maze example shows the generality of the method. All mechanisms are biologically justifiable, and I will ask whether the model is a biologically plausible account for planning.