When solving complex machine learning tasks, it is often more practical to let the agent find an adequate solution by itself using e.g. reinforcement learning rather than trying to specify a solution in detail. The only information required for reinforcement learning is a reward that gives the agent reinforcement about the desirability of his actions. Our experiments suggest that good results can be achieved by reinforcement learning with online learning neural networks. The functionality of such neural network may be further extended by allowing it to model the environment and/or by providing it with recurrent connections. In this thesis, we show that for a given network predicting the reward, it is NP-complete to find the agent action that maximizes this reward. We describe three neural network models, one of them being an original modification of Sutton's TD(¸) algorithm that extends its domain to non-Markovian environments. All three models were thoroughly tested with our predator-prey simulator. The most powerful of them, the modified TD(¸) was then applied to control of a real mobile robot. Simultaneously, we have discussed the principles of rewarding the agents, the biological plausibility of the algorithms, the importance of the exploration capabilities and general bounds of reinforcement learning. As a signicant part of the thesis, a C++ neural networks library was developed and the described models were implemented.