Connectionist Learning of Expert Preferences by Comparison Training

Authors: Gerald Tesauro

Abstract: A new training paradigm, caned the "eomparison pa.radigm," is introduced for tasks in which a. network must learn to choose a prdcrred pattern from a set of n alternatives, based on examplcs of Imma.n expert prderences. In this pa.radigm, the inpu t to the network consists of t.wo uf the n alterna tives, and the trained output is the expert's judgement of which pa.ttern is better. This para.digm is applied to the lea,rning of hackgammon, a difficult board ga.me in wllieh the expert selects a move from a. set, of legal mm·es. \Vith compa.rison training, much higher levels of performance can hc a.chiew~d, with networks that are much smaller, and with coding sehemes t.hat are much simpler and easier to understand. Furthermorf', it is possible to set up the network so tha.t it always produces consisten t rank-orderings .1.