

### Lernziel 3: Evolutionäre Algorithmen

Mithilfe des gym Packets und der Cartpole Umgebung wurde ein evolutionärer Algorithmus implementiert.

Bei diesen Algorithmen in Zusammenspiel mit Neuronalen Netzen (wie hier angewendet) wird das Netzwerk nicht durch einen Gradient Descent trainiert. Stattdessen wird eine Anzahl an Individuen, jeweils mit einer eigenen Zufallsinitialisierung der Gewichte des Neuronalen Netzes, erzeugt. Für jedes dieser Individuen wird die Fitness berechnet. Diese ist in diesem Fall als der erzielte Score beim Cartpole spielen definiert. Die Individuen werden daraufhin anhand ihrer Fitness absteigend sortiert. Die besten 10 Individuen werden miteinander gekreuzt, was bedeutet, dass die Parameter (Gewichte) der Neuronalen Netze von jeweils 2 „Eltern“ gekreuzt werden. Die entstandenen Kinder werden noch durch einen Faktor sigma zufällig „mutiert“, um den Evolutionsprozess am Laufen zu halten.

Diese Abfolge wird in diesem Fall 100 mal ausgeführt, es werden also 100 Generationen durchlaufen. (Grafik siehe Statistic.png) Es steht zu erkennen, dass schon nach einer kurzen Zeit das beste Individuum den maximalen Score von 500 (in Cartpole-v1, 300 in Cartpole-v0) erreicht. Der Average Score aller 100 Individuen einer Generation steigt zu Beginn ebenfalls sprunghaft, um dann ab etwa 60 Epochen einen Seitwärts-Trend einzuschlagen.