

Autonomna vožnja pomoću distribuiranog algoritma učenja uslovljavanjem

SW-21-2018 LUKA BJELICA

Algoritam

Korišćen je algoritam distribuiranog **duplog dubokog Q učenja** (eng. DDDQ Learning) koji predstavlja algoritam učenja uslovljavanjem bez modela

Eksperimentisanje sa algoritmom feedback funkcije najviše sa njenom kompleksnošću, kao i opsegu vrednosti koje vraća

Prelazak na distribuirani sistem sa više instanci simulacija i generatora gradijenata

Poboljšanje algoritma u sinhronizaciji target i Q mreža

Treniranje

Implementacija podsistema za praćenje procesa treniranja slanjem podataka na cloud

Podrška za praćenje u realnom vremenu, kao i pregled agregiranih podataka istorije treniranja

Potrebno je dobiti uvid u proces treniranja kako bi se mogao doneti zaključak o valjanosti određenih parametara

Cilj je postići stabilno treniranje, tako da agent inkrementalno napreduje, a ne da naglo uspe da nauči (pređe poligon) i podjednako brzo zaboravi naučeno

Parametri

Implementacija podsistema za dinamičko podešavanje parametara radi brže i efikasnije optimizacije

Stavljen fokus na sledećim parametrima:

- Epsilon (epsilon greedy policy)
- Gamma (MDP discount factor)
- Batch size
- Replay Buffer size
- Learning Rate
- Target and Q model sync frequency

Ostali parametri:

- Number of simulations
- Agent call frequency
- Max speed
- Number of lanes
- Car creation batch size
- Car wave order
- Network Architecture
- Optimizer
- Activation functions
- Loss function

