

Cílem práce je prozkoumat schopnost evolučních strategií trénovat architektury transformerů v prostředí zpětnovazebního učení. Provedeme experimenty s využitím vysoce paralelizovatelného algoritmu OpenAI-ES a dvou jeho variant využívajících konceptů novelty a quality-diversity prohledávání k trénování architektury Decision Transformeru v prostředí MuJoCo Humanoida a otestujeme tak schopnost těchto black-box optimalizačních technik trénovat i takto relativně velké (ve srovnání s dříve testovanými) a komplikované modely (využívajících self-attention vedle klasických plně propojených vrstev). Testované algoritmy se v našich experimentech ukázaly obecně jako schopné dosahovat silných výsledků a dokázaly vyvinout vysoce výkonné agenty — a to jak z náhodně inicializovaného modelu, tak z předtrénovaného modelu.