

# Neuro-evolucioni flappy bird

Stevan Jankulov – Fakultet tehničkih nauka

## Uvod

Igra se sastoji od igrača (ptice) i nasumično postavljenih preprekama (cevi). Cilj igre je da što duže ptica izbegava prepreke i tako ostvaruje poene. Osim prepreka, neophodno je i ostati na ekranu. Ukoliko se ptica sudari sa preprekom ili izađe van ekrana igra se završava. Jedini unos koji igrač ima jeste mogućnost da zamahne krilima i tako dobije na visini. Korišćenjem neuralnih mreža i genetskog algoritma, moguće je istrenirati računar da sam igra igru bez eksplicitnog programiranja pravila ponašanja ptica.

## Učenje

Učenje se vrlo brzo može primetiti, ptice već posle samo nekoliko generacija nauče da održavaju visinu i time maksimizuju score. Sledeći bolji napredak je dostizanje visine prepreke zamahinjavem ili prestanak dok se ptica ne spusti. Uglavnom se razvije jedan napredak, potrebno je više generacija kako bismo dobili mrežu koja ima sve tri sposobnosti. Jedan od primera koji je trenirao 20 minuta na maksimalnoj brzini je postigao score od čak 40.000 što je otprilike 11 minuta neprekidnog života ptice bez sudara na normalnoj brzini.

## Algoritam

Svaka ptica ima neuronsku mrežu koja za unos ima svest o sledećim detaljima sveta:

- Y pozicija ptice
- udaljenost do prve naredne cevi
- Y pozicija centra te cevi
- vertikalna brzina ptice.

Koristeći ove informacije, neuralna mreža nam govori koliko je sigurna da je potrebno zamahnuti krilima. Ako je sigurnost veća od 50%, ptica će zamahnuti, u suprotnom neće.

Pri prvoj generaciji, sve ptice dobijaju neuronske mreže sa nasumičnim vrednostima za težine. Ova generacija ptica se ponaša nepredvidljivo i potpuno nasumično. Kada se sve ptice sudare sa preprekom ili izađu van ekrana, generiše se nova generacija ptica. Ova nova generacija uzima sve ptice iz prethodne generacije u obzir kada bira naslednike. Šansa da se izabere konkretna ptica je direktno proporcionalna njenim fitness score-om. Score predstavlja dužinu života ptice, ptice sa većim score-om imaju veću šansu da budu izabrane. Iste ptice se mogu više puta izabrati. Čak i najgore ptice imaju šansu da budu izabrane ali je ona veoma mala.

Kada se popuni nova generacija, potrebno je promeniti ponašanja mreža kako bismo dobili poboljšanje ponašanja ili čak nova otkrića. Svaka nova ptica ima relativno malu šansu da se njene težine u neuralnoj mreži promene za malu količinu, ovaj korak se naziva mutacija. Mutacija je jedini način da se stekne novo ponašanje, ne postoji ukrštavanje gena već se geni prenose takvi kakvi jesu i retko nasumično promene. Pored toga, najbolja ptica iz svih prethodnih generacija je uvek prisutna i odabrana bar jednom.

## Zaključak

Ovaj pristup učenja predstavlja prvo elegantan način za učenje agenata kako da igraju igre bez ikakvog programiranja logika.

Jedna od makra ovog postupka jeste što nam rezultat nije deterministički pa svako drugo pokretanje izaziva drugačija ponašanja i brže (ili sporije) razvijanje neuralnih mreža.

Posmatranjem više generacija možemo uočiti kako ptice dobijaju nova ponašanja jednostavnom selekcijom i vrlo malim "suptilnim" izmenama.

Generation	Avg Score
1	70.3
5	230.6
25	623.2