Efficient Evolutionary Methods for Game Agent Optimisation: Model-Based is Best

Simon M Lucas, Jialin Liu, Ivan Bravi, **Raluca D. Gaina**, John Woodward, Vanessa Volz, Diego Perez-Liebana

AAAI-19 Game AI Group QMUL SUSTech

1 Main Contributions

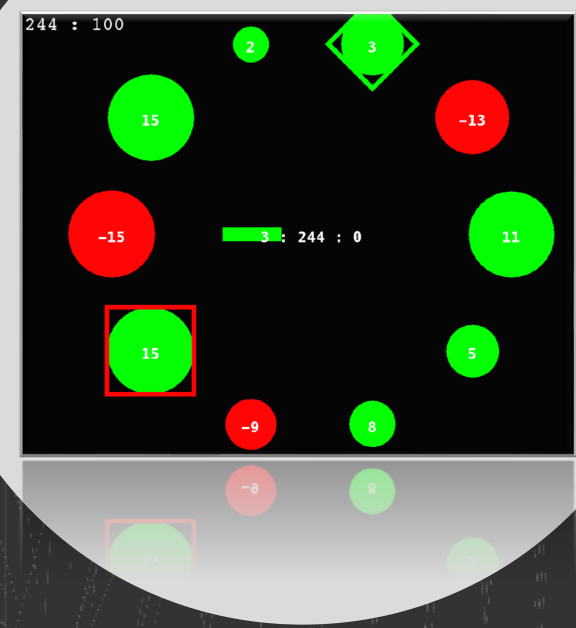
- Show effectiveness of model-based optimization approaches.
- Insight into N-Tuple Bandit EA (NTBEA).
- Emphasize importance of parameter tuning in game AI.

2 Motivation

Noisy optimisation problem:

- Stochastic games
- Stochastic agents

Develop efficient games for testing AI agents, rapid experimentation.



3 Problem

Optimisation of Rolling Horizon Evolutionary Algorithm (RHEA) parameters, playing Fast Planet Wars against opponent RHEA with **fixed** parameters.

Parameter	Type	Legal values
Nb. Mutated Genes	Integer	0, **1**, 2, 3
Flip Min. One Gene	Boolean	False, **True**
Use Shift Buffer	Boolean	False, **True**
Nb. Resamples	Integer	**1**, 2, 3
Sequence Length	Integer	**5**, 10, 15, 20, 25, 30

4 Fast Planet Wars

- Random initial state (ownership, planet sizes and number of ships).
- Planets grow ships at rate ∝ to size.
- Ships transferred via each player's buffer and their planet of focus.
- Aim: have most ships at time limit.
- Win/loss game.
- Constant speed regardless of number of planets.

Multi-armed bandit

$$\mathrm{UCB}_i = \hat{x}_i + k\sqrt{\frac{\ln n}{n_i + \epsilon}}$$

Combinatorial multi-armed bandit

$$v_{UCB}(x) = \frac{1}{m}\sum_{j=1}^{m} UCB_{N_j(x)}$$

(m = 16)

5 NTBEA

Efficient, effective and informative. Uses a standard EA but does most of the search in the space of a learned model.

N-Tuple Bandit Fitness Landscape Model ⟺ Evolutionary Algorithm ⟷ Noisy Fitness Evaluator

The model is similar to a combinatorial multi-armed bandit, offering explicit control of exploitation / exploration dilemma and natural escape from local optima.

NTBEA Insight

Detailed statistics for each parameter combination, including number of times selected, average fitness, standard error, min and max.

6 Results

Algorithm	Avg. ± St. Err.
RMHC (1)	-0.29 ± 0.01
CMA-ES (1x)	-0.12 ± 0.04
UH-CMA-ES (1x)	-0.09 ± 0.04
RMHC (5)	0.01 ± 0.01
SGA	0.03 ± 0.01
CMAE-ES (5x)	0.29 ± 0.02
UH-CMA-ES (5x)	0.32 ± 0.03
SWcGA	0.36 ± 0.01
SMEDA	0.38 ± 0.01
NTBEA (1,2,5)-	0.41 ± 0.01
UH-CMA-ES (10x)	0.44 ± 0.02
CMA-ES (10x)	**0.48 ± 0.02**
SMAC	**0.49 ± 0.01**
NTBEA (1)	**0.50 ± 0.01**
NTBEA (1,2)	**0.51 ± 0.01**
NTBEA (1,2,5)+	**0.51 ± 0.01**

7 Conclusions

- New game with significant skill depth.
- Importance of parameter tuning demonstrated, well tuned RHEA agents hugely outperformed poorly tuned ones.
- Restricted sample budget for tests (288 fitness evaluations per optimisation run).
- Model-based clearly performed best.
- NTBEA is especially simple and informative: high mutation rates work surprisingly well.

Future Work

- Add continuous parameters.
- Add tree-structured parameterization.