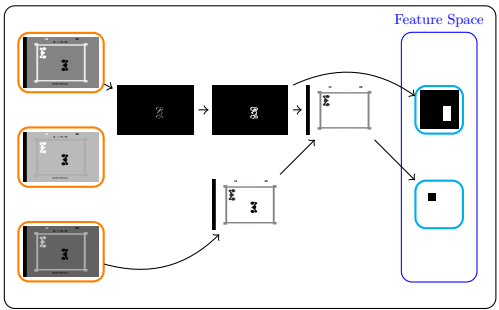
Playing Atari games with an Interpretable Agent

October 27, 2021

Interpretable CGP Architecture

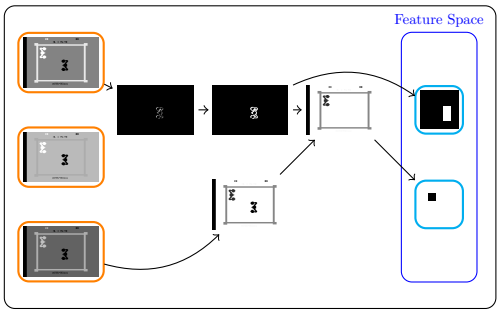
Interpretable CGP Architecture

CGP Encoder

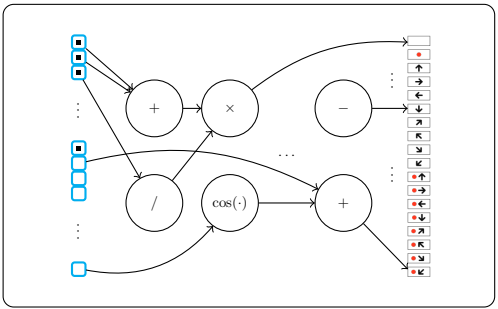


Interpretable CGP Architecture

CGP Encoder

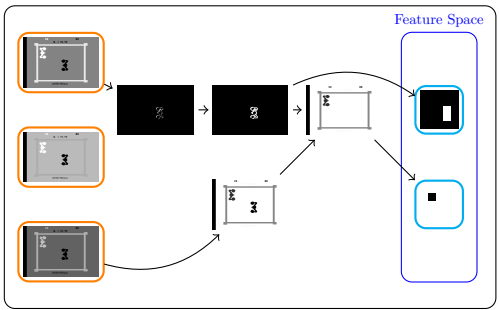


CGP Controller

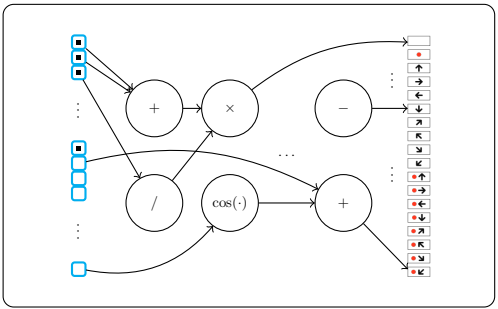


Interpretable CGP Architecture

CGP Encoder



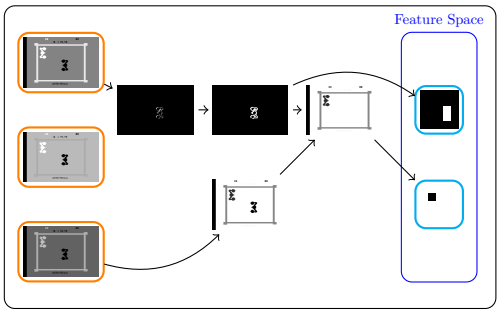
CGP Controller



Functions: dilate, erode, subtract, threshold, binary, NOT, AND, OR, XOR, motion-capture

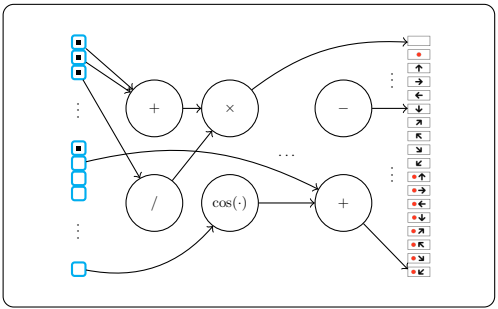
Interpretable CGP Architecture

CGP Encoder



Functions: dilate, erode, subtract, threshold, binary, NOT, AND, OR, XOR, motion-capture

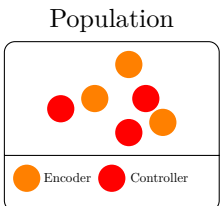
CGP Controller



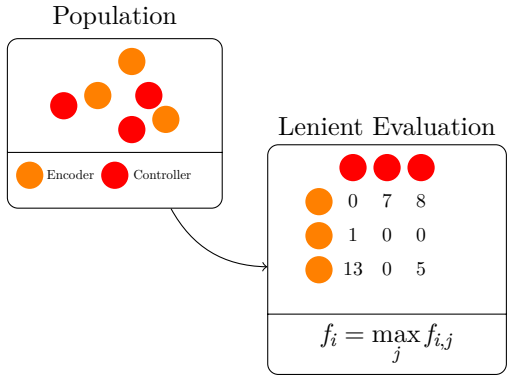
Functions: $+$, $-$, $\times$, $\div$, $|\cdot|$, $\sqrt{\cdot}$, $\leq$, $\geq$, NOT, AND, OR, XOR

Experiment 1: $1 + \lambda$ Co-evolution

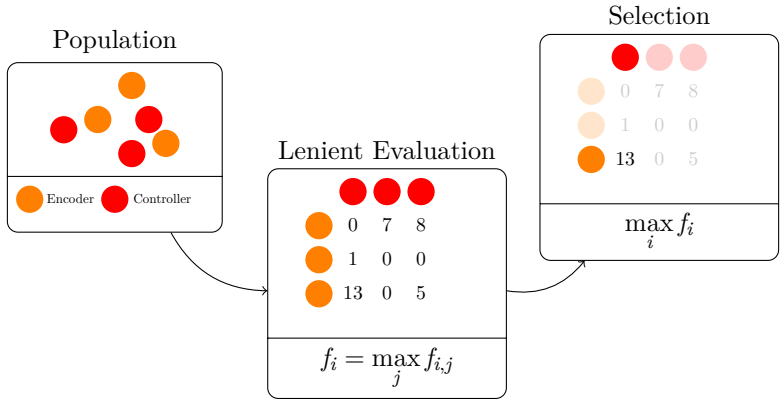
Experiment 1: $1 + \lambda$ Co-evolution



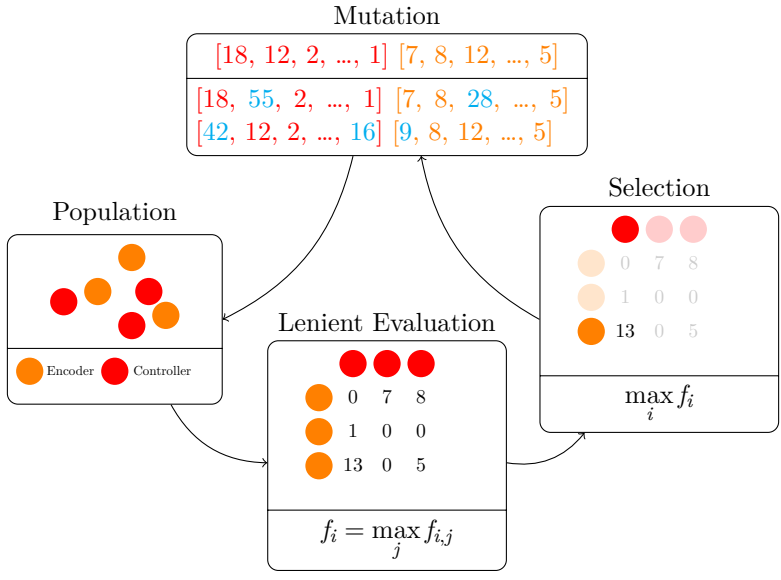
Experiment 1: $1 + \lambda$ Co-evolution



Experiment 1: $1 + \lambda$ Co-evolution



Experiment 1: $1 + \lambda$ Co-evolution

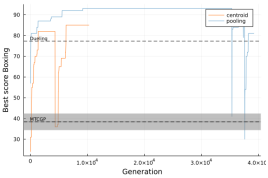


Experiment 1: $1 + \lambda$ Results

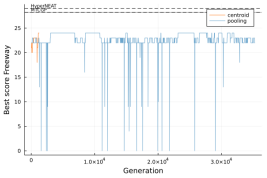
https://github.com/erwanlecarpentier/ICGP-results

Experiment 1: $1 + \lambda$ Results

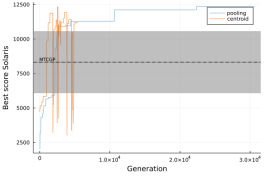
Boxing



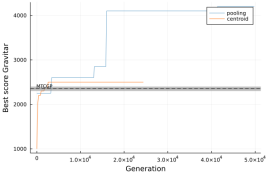
Freeway



Solaris



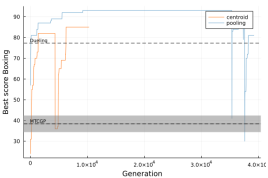
Gravitar



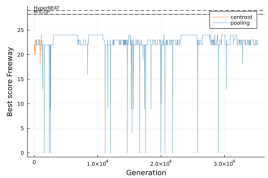
https://github.com/erwanlecarpentier/ICGP-results

Experiment 1: $1 + \lambda$ Results

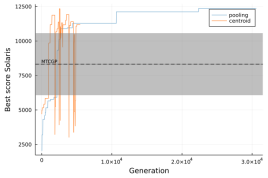
Boxing



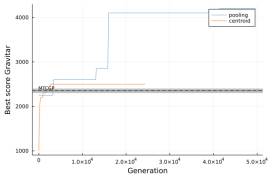
Freeway



Solaris



Gravitar



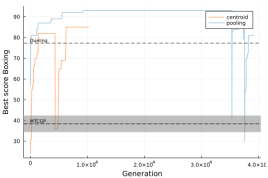
Conclusions:

- "Rote Learning" of deterministic game is efficient in this setting
- Interpretability (Boxing video) (Solaris video)
- Introducing stochasticity is catastrophic

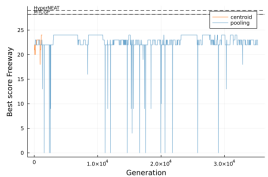
https://github.com/erwanlecarpentier/ICGP-results

Experiment 1: $1 + \lambda$ Results

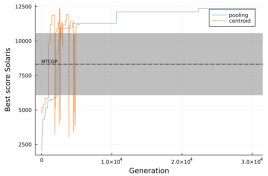
Boxing



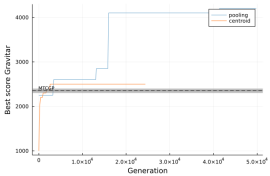
Freeway



Solaris



Gravitar



Conclusions:

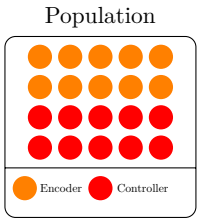
- "Rote Learning" of deterministic game is efficient in this setting
- Interpretability (Boxing video) (Solaris video)
- Introducing stochasticity is catastrophic

→ Tackle the stochastic problem
→ What optimizer in this setting?

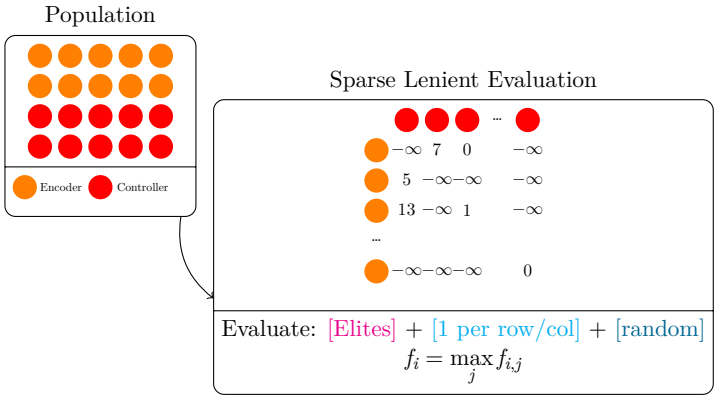
https://github.com/erwanlecarpentier/ICGP-results

Experiment 2: GA Co-evolution

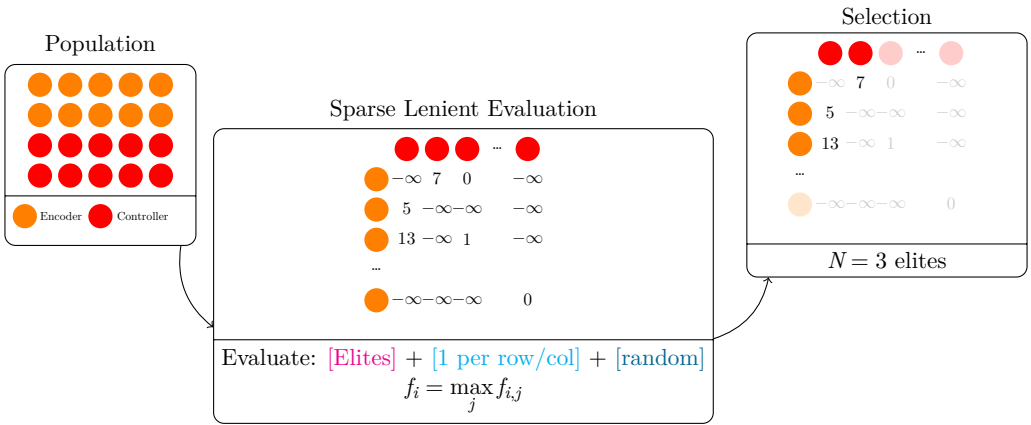
Experiment 2: GA Co-evolution



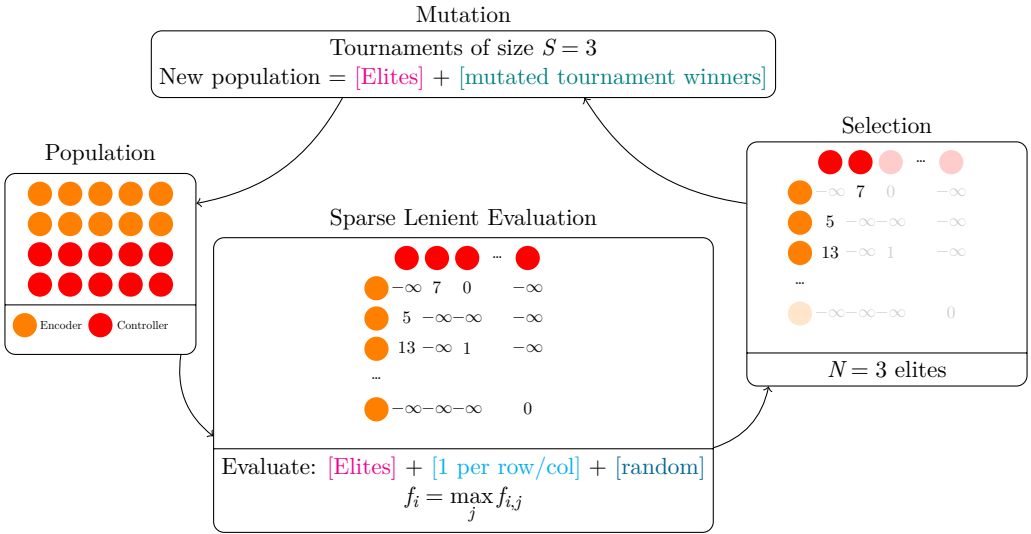
Experiment 2: GA Co-evolution



Experiment 2: GA Co-evolution



Experiment 2: GA Co-evolution

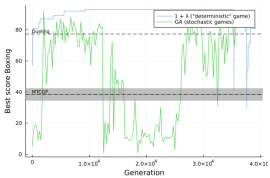


Experiment 2: GA Results

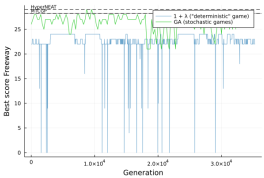
https://github.com/erwanlecarpentier/ICGP-results

Experiment 2: GA Results

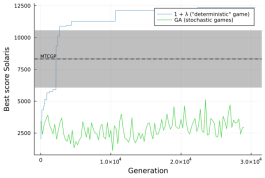
Boxing



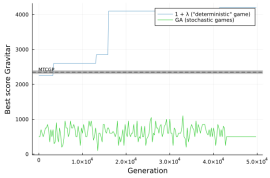
Freeway



Solaris



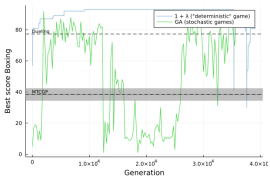
Gravitar



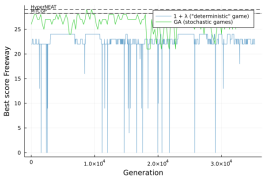
https://github.com/erwanlecarpentier/ICGP-results

Experiment 2: GA Results

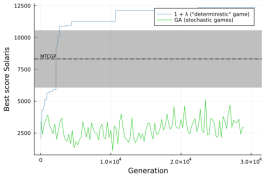
Boxing



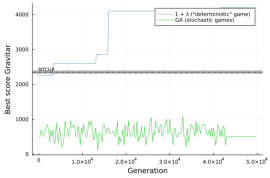
Freeway



Solaris



Gravitar



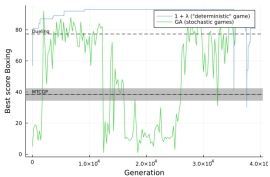
Remarks:

- Boxing: collapse
- Freeway Solaris Gravitar: no learning progress

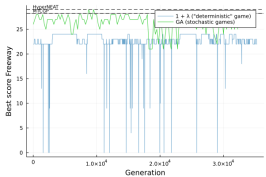
https://github.com/erwanlecarpentier/ICGP-results

Experiment 2: GA Results

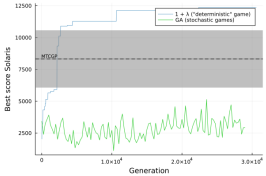
Boxing



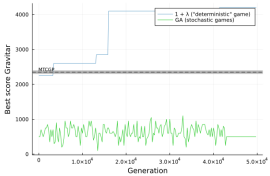
Freeway



Solaris



Gravitar



Remarks:

- Boxing: collapse
- Freeway Solaris Gravitar: no learning progress

→ Remove “cross-over” to avoid elites loss and focus on E/C pairs

→ Encourage diversity to avoid local minima

https://github.com/erwanlecarpentier/ICGP-results

Overall conclusion

Experimental results:

	Performance	Interpretability
Atari deterministic	OK	OK
Atari stochastic	NOT OK	OK

→ Need for a better optimizer in stochastic Atari games:

- Experiment 3: E/C pairs only
- We are unsure about ways to measure behavior diversity in Atari