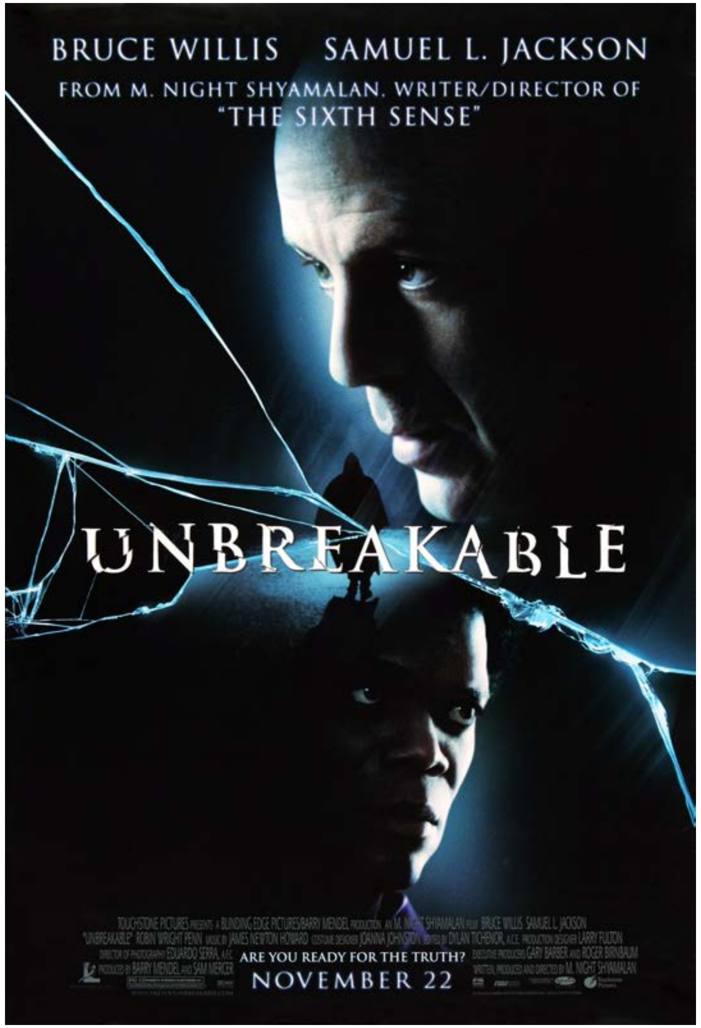
Weight agnostic neural networks

NeurIPS 2019, Inria / Google

Introduction (joke ㅎ)

- Movie “Unbreakable” (2000)
  - Mr. Glass vs David Dunn
  - Neural networks are sensitive to their parameters…



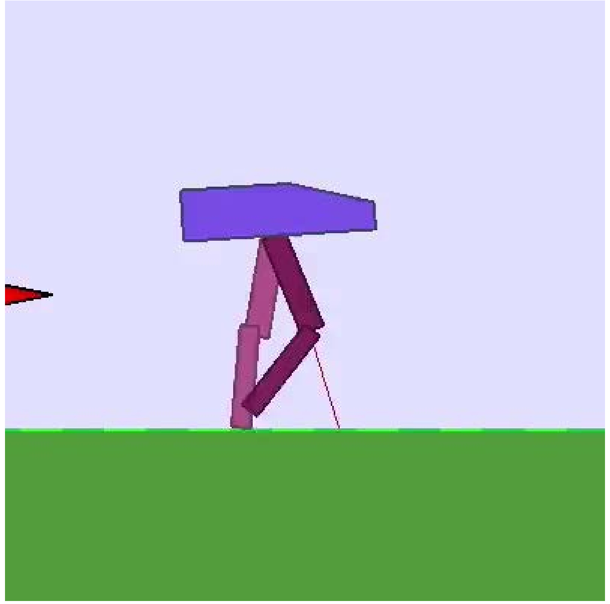
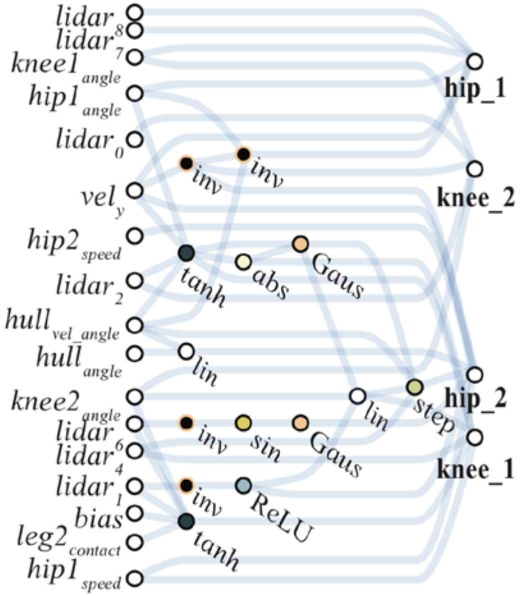
Question

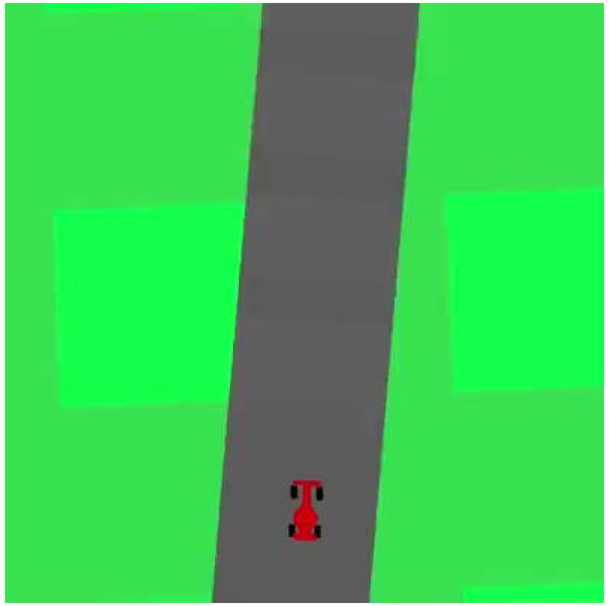
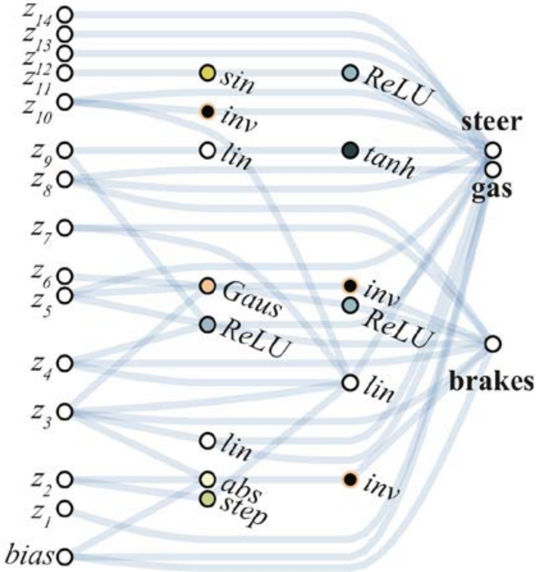
- Not all neural network architectures are created equal, some perform much better than others for certain tasks
  - Is there a network that can already perform a task without any explicit weight training?
  - Is there a network architecture born to accomplish a certain task without training?
  - ...

  - The answer is “yes”, but how can we find it?

Example

- Demo: https://weightagnostic.github.io/

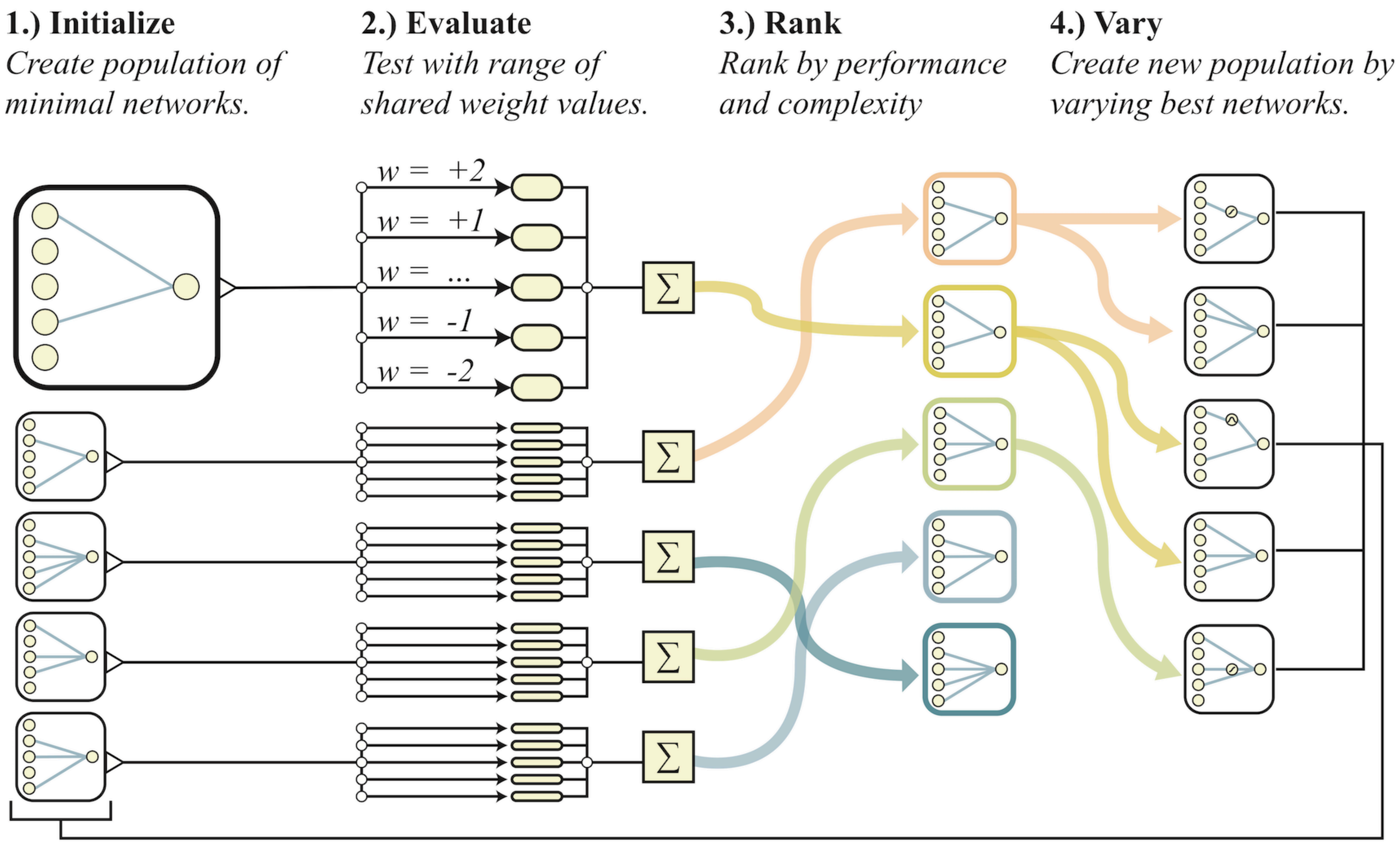



- Bipedal Walker (left) and Car Racing (right)

Problem setup

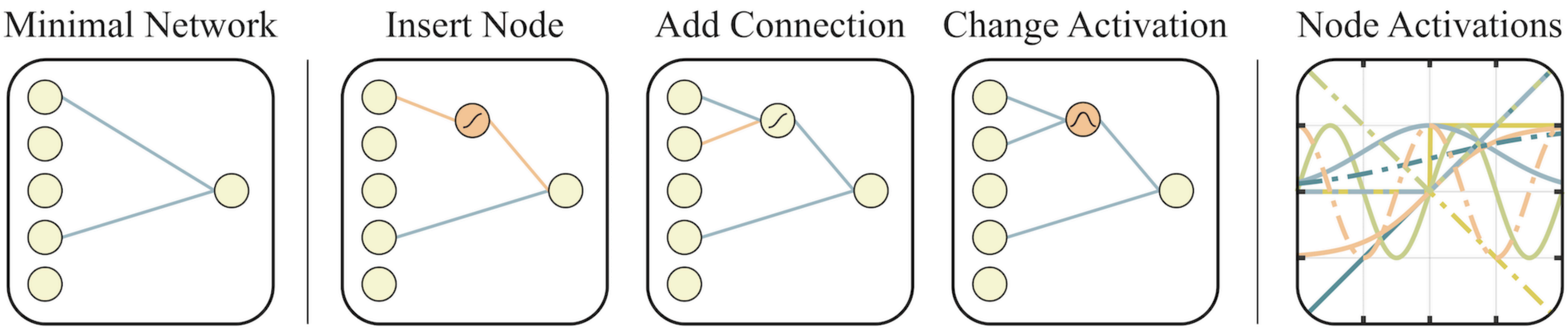
- Weight agnostic neural networks (WANNs) for a given task
  - An initial population of minimal neural network topologies is created
  - Each network is evaluated over multiple rollouts, with a different weight value
    - Weight-sharing on all weights; the number of parameter is one
  - Networks are ranked according to their average performance and complexity
  - A new population is created by varying the highest ranked network topologies
  - Repeat procedures above

Problem setup



Topology search

- NeuroEvolution of Augmenting Topologies (NEAT)
  - The initial population is composed of sparsely connected networks
    - Networks with no hidden nodes and only a fraction of the possible connection between input and output
  - New networks are created by modifying existing networks using one of three operators
    - Insert node / Add connection / Change activation



Performance and complexity

- Performance
  - Network topologies are evaluated using several shared weight values
    - [−2, −1, −0.5, +0.5, +1, +2]
  - Mean performance over all weight values
  - Max performance on the single best weight value
- Complexity
  - Minimal description length (MDL) → Occam's razor
  - Number of connections in the network
- Ranking
  - 80%: mean performance / number of connections
  - 20%: mean performance / max performance

Experimental results

- Continuous control
  - CartPoleSwingUp / BipedalWalker-v2 / CarRacing-v0

- Mean performance over 100 trials under 4 conditions:
  - Random weight: individual weights drawn from $U(-2,2)$
  - Random shared weight: a single shared weight drawn from $U(-2,2)$
  - Tuned shared weight: the highest performing shared weight in $[-2,2]$
  - Tuned weight: individual weights tuned using population-based REINFORCE

Experimental results

Swing Up	Random Weights	Random Shared Weight	Tuned Shared Weight	Tuned Weights
WANN	**57 ± 121**	**515 ± 58**	**723 ± 16**	**932 ± 6**
Fixed Topology	21 ± 43	7 ± 2	8 ± 1	918 ± 7

Biped	Random Weights	Random Shared Weight	Tuned Shared Weight	Tuned Weights
WANN	**-46 ± 54**	**51 ± 108**	**261 ± 58**	332 ± 1
Fixed Topology	-129 ± 28	-107 ± 12	-35 ± 23	**347 ± 1**

CarRacing	Random Weights	Random Shared Weight	Tuned Shared Weight	Tuned Weights
WANN	**-69 ± 31**	**375 ± 177**	**608 ± 161**	893 ± 74
Fixed Topology	-82 ± 13	-85 ± 27	-37 ± 36	**906 ± 21**

Experimental results

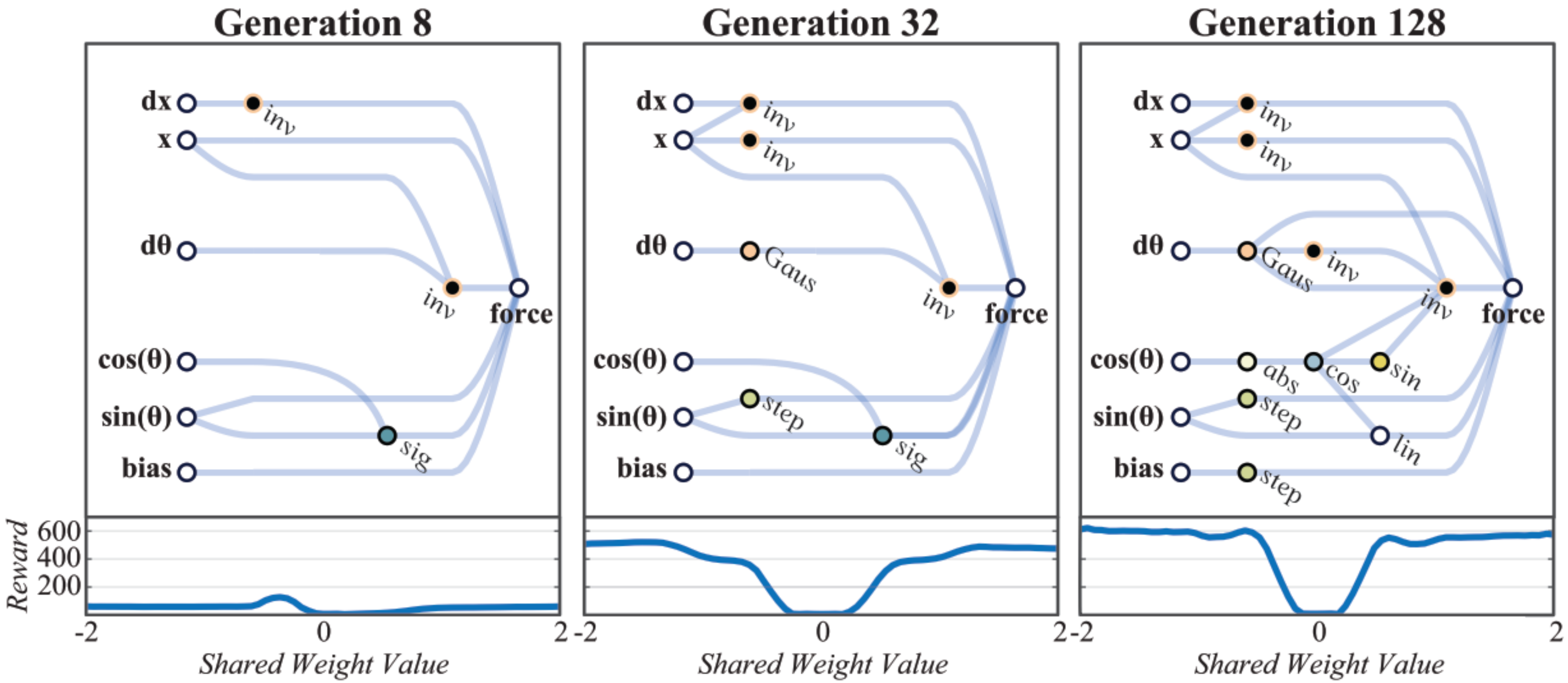


Figure 4: *Development of Weight Agnostic Neural Network Topologies Over Time*
Generation 8: An early network which performs poorly with nearly all weights.
Generation 32: Relationships between the position of the cart and velocity of the pole are established. The tension between these relationships produces both centering and swing-up behavior.
Generation 128: Complexity is added to refine the balancing behavior of the elevated pole.

Experimental results

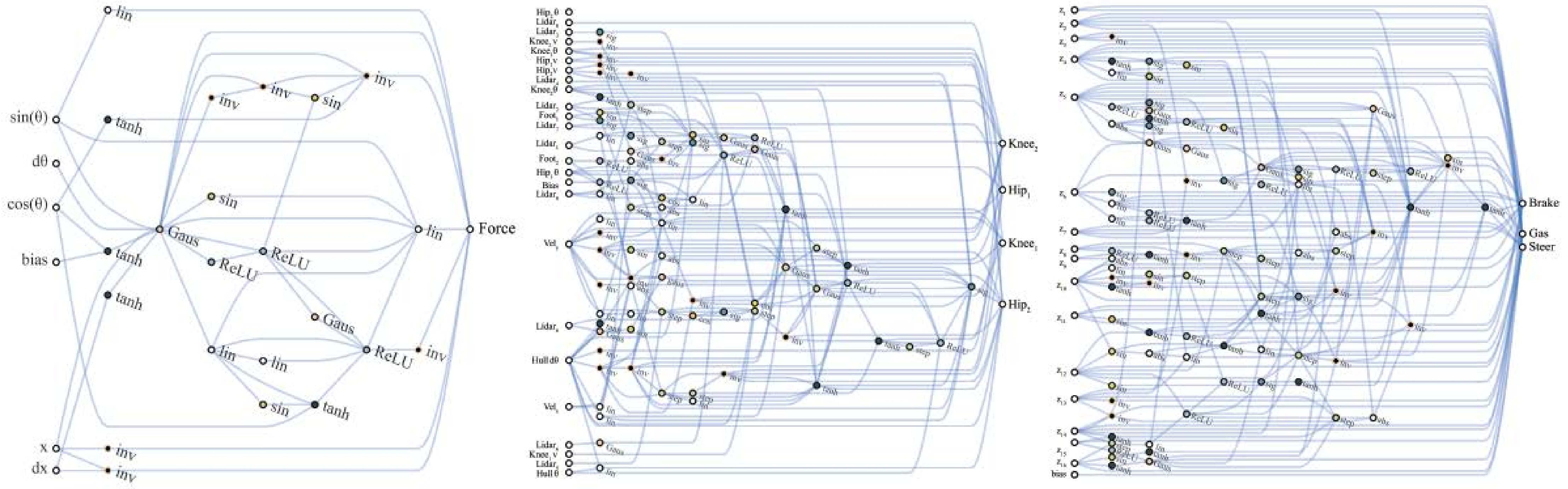
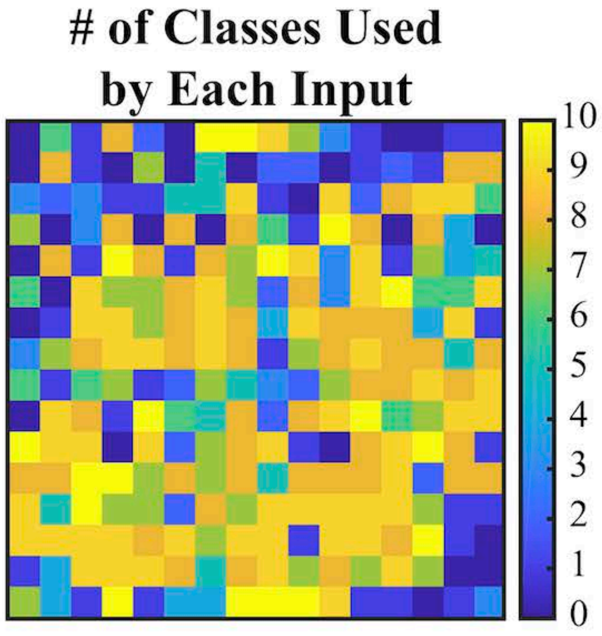
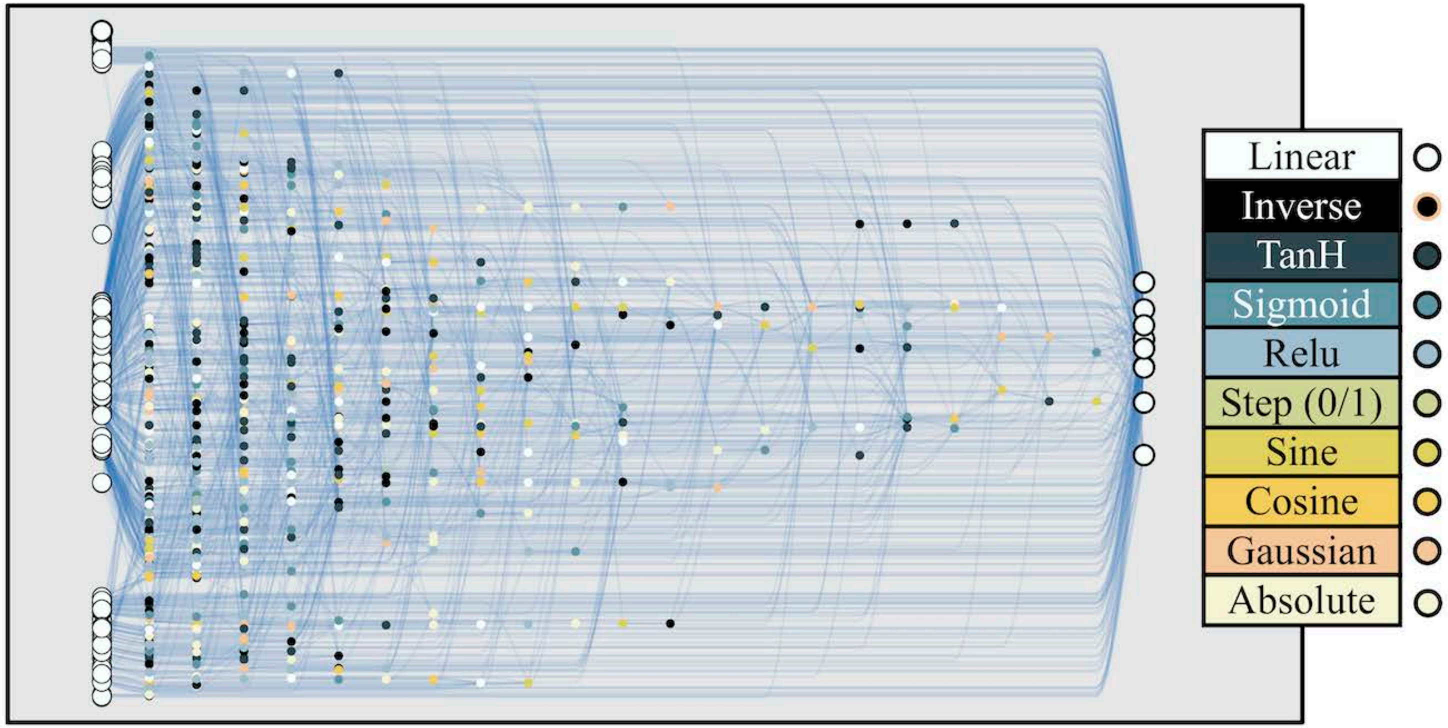


Figure 5: *Champion Networks for Continuous Control Tasks*
Left to Right (Number of Connections): Swing up (52), Biped (210), Car Racing (245)
Shown in Figure 1 (Page 1) are high performing, but simpler networks, chosen for clarity. The three network architectures in this figure describe the champion networks whose results are reported.

Experimental results

- Classification accuracy on MNIST



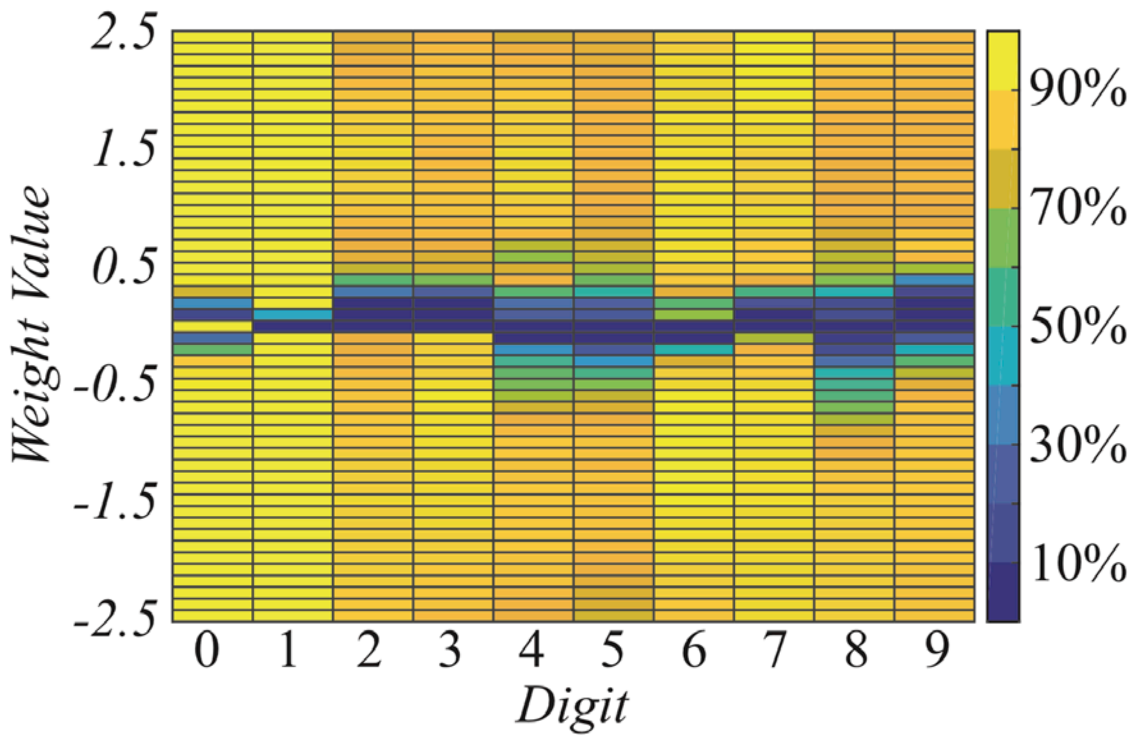


Experimental results

- Classification accuracy on MNIST

WANN	Test Accuracy
Random Weight	*82.0% ± 18.7%*
Ensemble Weights	*91.6%*
Tuned Weight	*91.9%*
Trained Weights	*94.2%*

ANN	Test Accuracy
Linear Regression	*91.6%* [62]
Two-Layer CNN	*99.3%* [15]



Discussion

- Weight agnostic neural network and natural / eigen parameters

- Neuroscience
  - *Animal behavior is innate, and does not arise from learning*
  - Born to ~~~ / born to be ~~~