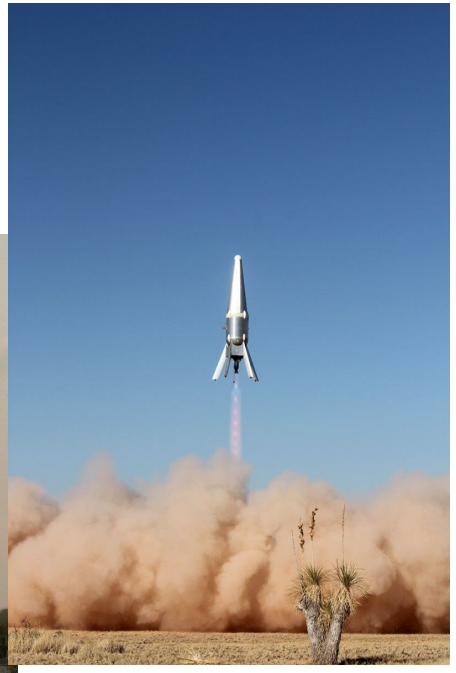
Keen Technologies

Research Directions

Quick Background: Id Software

Quick Background: Armadillo Aerospace



Quick Background: Oculus

Quick background: Keen Technologies

- OpenAI
- Doing the reading
- Research, not product
  - There are key, fundamental discoveries to be made
- Richard Sutton and Alberta
- Six researchers today
  - John Carmack
  - Gloria Kennickell
  - Lucas Nestler
  - Richard Sutton
  - Joseph Modayil
  - Khurram Javed

Where I thought I was going

- Not LLMs
  - LLMs can *know everything without learning anything*
  - Learn from experience, not an IID blender
- Virtual environments and games
  - Bot history
- Video understanding
- Infinite video wall

Missteps

- Too low level
- Avoided larger experiments too long
- Sega Master System
- Video can wait

Settling in with Atari

- Deep research history
  - Mostly replicable, but assumptions varied
  - RL framework rot over the years
- Unbiased and diverse
- Isn’t Atari solved?
  - MEME, Muesli, BBF, etc
  - Scores are there for most games, given enough time
  - Lots of critical questions are still open!
- Why not Pokemon or Minecraft?
  - Scrolling and 3D will definitely add more challenges
  - Tempts people to look at internal data structures

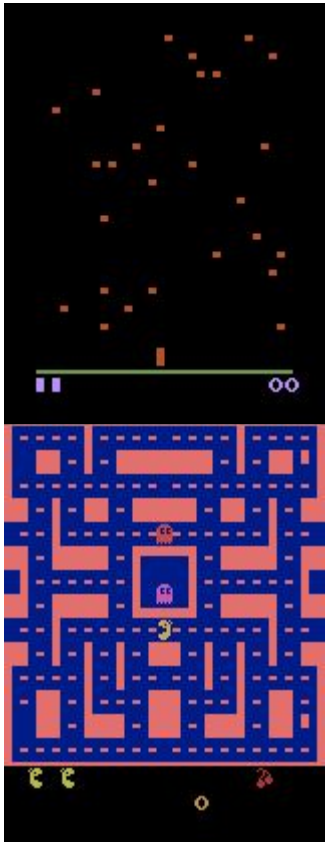
Reality is not a turn based game

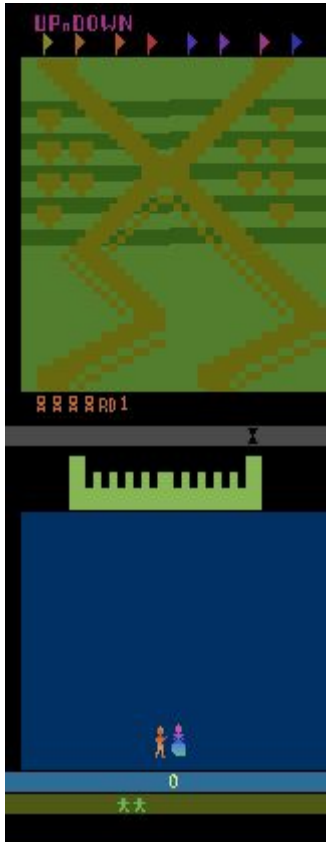
- The real world keeps going
  - regardless of if you are ready
- Invert the typical RL research environment
  - the agent gets called, not the environment
- Single stream of experience
  - Parallel environments are a crutch
- Speed
- Latency
  - Add a little action queue!
- How else might reality differ?

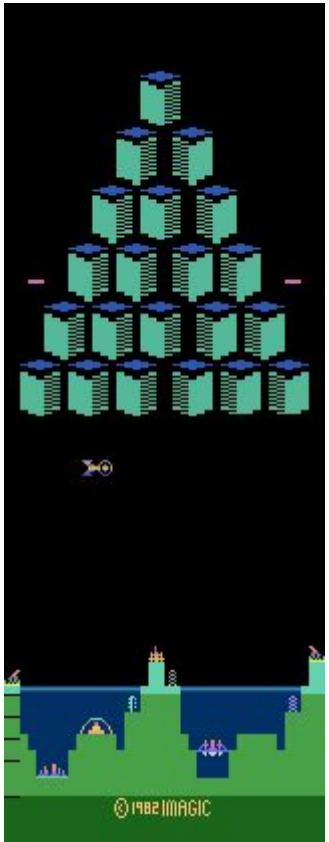
Physical Atari!

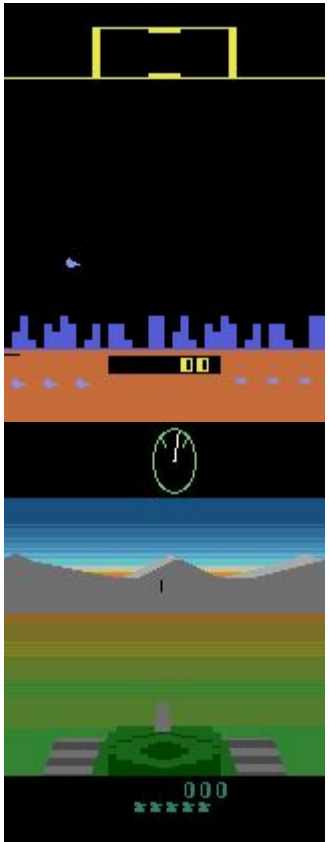
- Kind of a stunt, but novel AFAIK
- Could a robot play a video game?
- Contact with reality
- Felt like a product push
- We will Open Source

Physical Atari: Games









Physical Atari: Compute

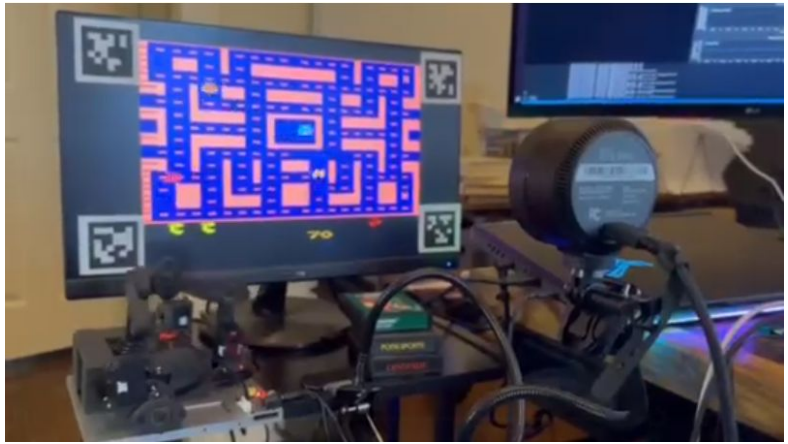
- ASUS ROG Strix SCAR 16
- 16GB 4090 Laptop GPU
- Realtime training at frame_skip 4

Physical Atari: Camera

- Resolution
  - Atari only 160x210
- Frame rate
- Uncompressed images
- USB interface
  - GPUDirect for Video
- Scanout concerns

Physical Atari: Rectification

- Learns OK with a fixed camera
  - No ability to transfer to another setup
- Manual corner picking
- April Tags
  - Can work with a moving camera
  - Lighting challenges
- General purpose screen detection
  - Future research

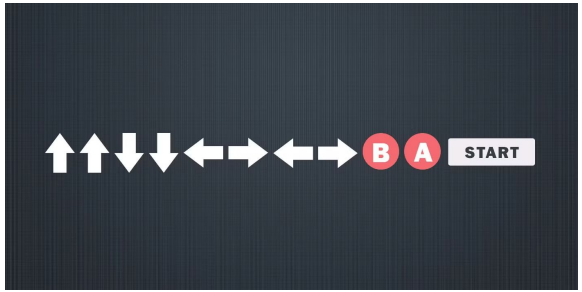


Physical Atari: Virtual Joystick

- Digital IO board wired to joystick connector

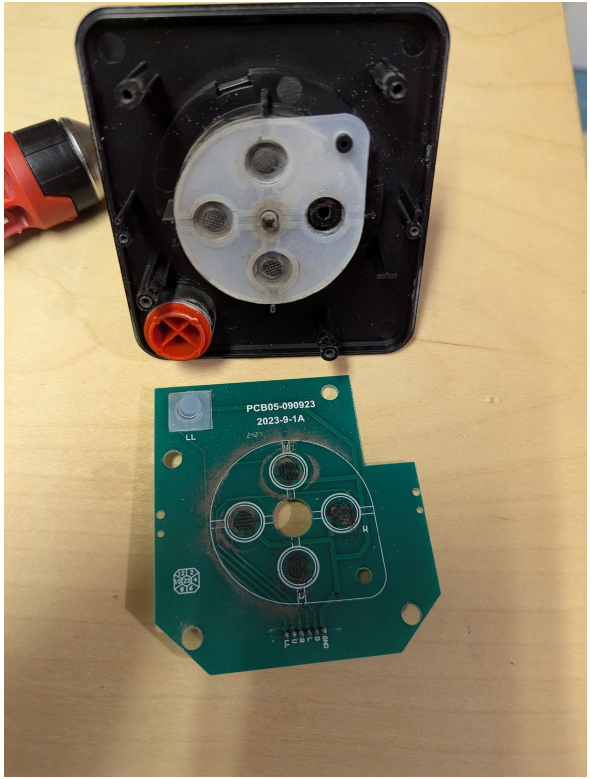
Physical Atari: Robotroller

- Three servos
- Additional latency
- Spurious action issue
  - Atlantis side firing
  - Actions in observations



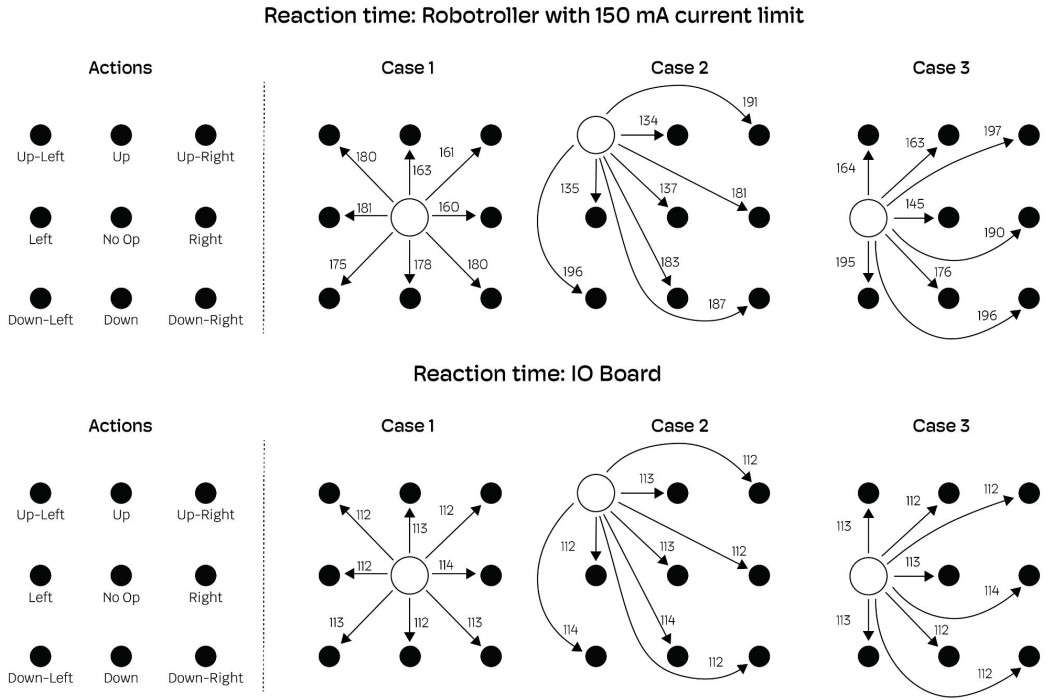
Physical Atari: Robustness

- Servos and joysticks wear out!
  - Reduce max current



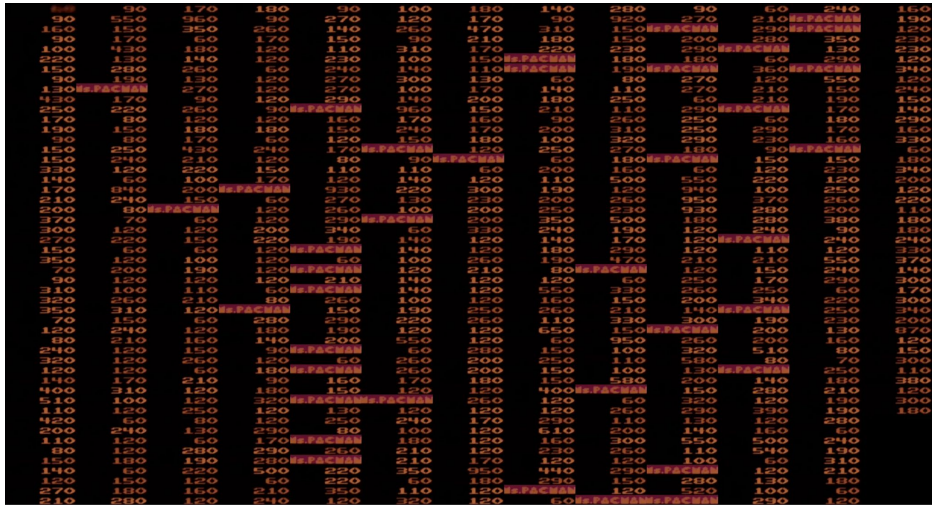
Physical Atari: control latencies

- Total latency is substantial
- Not far off from human
  - Try a reaction tester



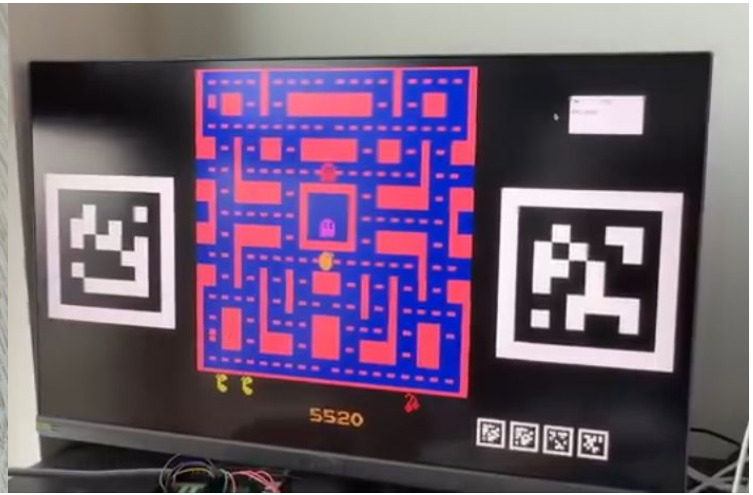
Physical Atari: Score and life detection

- Surprisingly, the trickiest part!
  - Hasn’t this been solved since MNIST?
  - Ask an LLM to tell you the score…
  - Game dependent
  - Heuristics



Physical Atari: Custom Dev Box

- Raspberry Pi running ALE
- Atari joystick port
- Emissive tags for rectification and score



Physical Atari: Lessons

- ConvNets tolerate a wide range of image distortions
  - Not yet clear how well the models transfer across lighting conditions
- TD learning tolerates system latency
- Top simulation agents perform poorly with added latency
  - SPR and world models that condition on action taken
  - Need a robust fix for this, not just a matching delay
- Action-to-action paths can be a problem
- Decent learning in a few hours of experience
- Achieving human level performance looks feasible
  - Especially if talking to a remote big GPU

Sequential Multi-Task Learning

- Sequential is much harder than parallel
- No TaskID or weight switching
- No hyperparameter ramp alignment
  - Exploration, optimizer, model reset
- The global weight impact of online learning
- Is a giant replay buffer enough?
  - Using TB of storage isn’t crazy
- Offline RL can bootstrap into a coherent fantasy untested by reality

Transfer Learning

- Don’t play like an idiot at the start
  - After years of subjective time playing game, should do better
- OpenAI’s Gotta Learn Fast
  - 1M steps is plenty to learn without any transfer learning
- GATO: More effective to learn a game from scratch than to fine tune
  - Negative transfer learning
- Distinct from continuous forgetting

New Benchmark

- Eight games?
- Three cycles through all the games?
- 400k frames per cycle?
  - Handle truncation, or only switch at end of episode?
- Close to the compute for 26 games of Atari100k
- No dedicated evaluation phase, sum over last cycle
- Full action set
- Sticky actions
- Add a control latency?
- Real time?
- BBF from scratch as baseline performance
- Standard benchmark harness

Sparse rewards

- Individual scores isn’t the driver for humans like it is for RL agents
  - Often don’t even look at score until the end!
- The “hard exploration games” – Pitfall, Montezuma’s Revenge, etc
  - Make any game into hard exploration by only giving reward at loss of life or game over
- Dense rewards for human tasks are the exception, not the rule
- Intrinsic rewards
- Curiosity
- Meta-curiosity across games
  - How would a human play a library of arcade games?

Exploration

- Epsilon-greedy problems
- Soft-Q with temperature
- Action space factorization
  - Million+ actions on a modern controller
- Confidence
- Timescales
  - Frame_skip 4
  - Action gap
- Options

Recurrence vs Frame Stacks

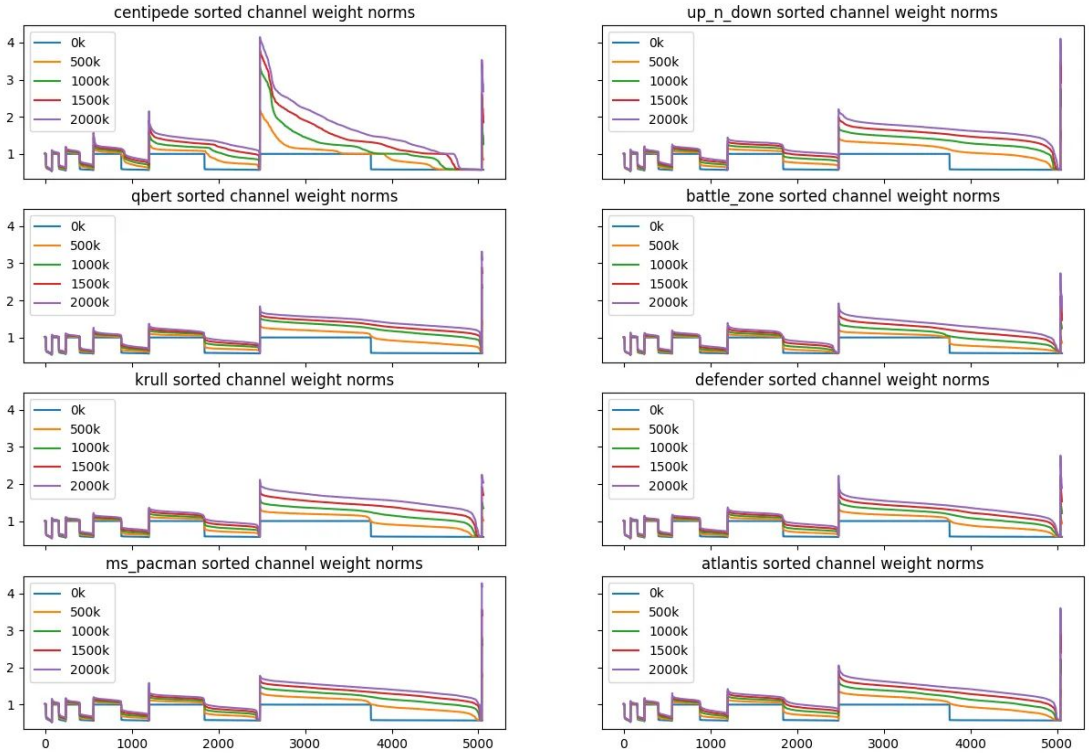
- Frame stacks are unfortunately effective for Atari
- Brains are recurrent neural networks

Function Approximation Dominates Performance

- Just a black box in classic RL
- Many duties, surprising it works as well as it does!
  - Learning new results for novel inputs
  - Generalizing across similar inputs
  - Averaging stochastic processes
  - Updating non-stationary processes
- Supervised learning practices don’t always transfer
- Adam is still very hard to beat
- Auxiliary losses just tweak the value / policy approximation
- Are neural nets and backprop even the right thing?

Value Representation

- DQN clamping
- Categorical values
- Quadratic value compression
- MSE was all we needed?
- Brains and values



Plasticity vs Generalization

- Generalization is ignoring details
- Plasticity involves noticing new details
- Every new online sample is a held-out validation sample
  - But how do we use it to improve?

Conv Nets

- Poor transfer from ImageNet to RL
- Kernel subsets
- Parameter sharing increases performance
- Factored 1D CNNs
- Isotropic CNNs
- Dilated Isotropic CNNs
- Isotropic DenseNet CNNs
- Recurrent Isotropic Semi-Dense CNNs