Adaptive DQN for Breakout

https://github.com/manasjuneja/rlagentataribreakout

COMPREHENSIVE PERFORMANCE ANALYSIS

BASIC PERFORMANCE METRICS

Total Episodes: 1000
Average Score: 640.58
Best Score: 2050.00
Final 100 Episodes Average: 610.50
Score Standard Deviation: 148.96

CURRICULUM STAGE ANALYSIS

Normal Breakout:
  Episodes: 0-198 (199 episodes)
  Average Score: 760.20
  Best Score: 2050.00
  Improvement: [600, 600, 600, 600, 750, 600, 600, 600, 600, 600]
Dynamic Ball Speed:
  Episodes: 199-398 (200 episodes)
  Average Score: 615.00
  Best Score: 750.00
  Improvement: [600, 600, 600, 600, 600, 600, 600, 600, 600, 750]
Adaptive Paddle Size:
  Episodes: 399-598 (200 episodes)
  Average Score: 613.50
  Best Score: 750.00
  Improvement: [600, 600, 600, 600, 600, 600, 600, 600, 600, 600]
Full Adaptive Mode:
  Episodes: 599-798 (200 episodes)
  Average Score: 607.50

Best Score: 750.00
Improvement: [600, 600, 600, 600, 600, 600, 600, 600, 600, 600]
Stage 4:
Episodes: 799-999 (201 episodes)
Average Score: 607.46
Best Score: 750.00
Improvement: [600, 600, 600, 600, 600, 750, 600, 600, 600, 600]

ADAPTATION ANALYSIS

Total Adaptations Triggered: 24
Adaptation Types:
moderate_exploration: 15 times
increase_exploration: 9 times
Average Recovery Time: 2.2 episodes
Recovery Time Range: 0-17 episodes

DIFFICULTY CHANGE IMPACT

Total Difficulty Changes: 86439
Change Types:
ball_speed: 17967 changes
paddle_speed: 57060 changes
paddle_width: 11412 changes

LEARNED STRATEGY ANALYSIS

Action Distribution:
No-op: 100.0% (31441 times)
Left: 0.0% (0 times)
Right: 0.0% (0 times)

Most Successful Action: No-op
Success Rates by Action:
No-op: 98.93%
Left: 0.53%
Right: 0.53%

INTERNAL REPRESENTATION ANALYSIS

Analyzing 10 representation snapshots...

Q-VALUE EVOLUTION

Q-Value Trends:
  Episode 100:
    Action Preferences: [nan nan nan]
    Q-Value Ranges: [nan nan nan]
  Episode 400:
    Action Preferences: [nan nan nan]
    Q-Value Ranges: [nan nan nan]
  Episode 700:
    Action Preferences: [nan nan nan]
    Q-Value Ranges: [nan nan nan]

HIDDEN LAYER ACTIVATION ANALYSIS

HIDDEN_1 Analysis:
  Sparsity Evolution: 0.000 → 0.000
  Mean Activation: nan → nan
  Activation Variance: nan → nan

HIDDEN_2 Analysis:
  Sparsity Evolution: 0.000 → 0.000
  Mean Activation: nan → nan
  Activation Variance: nan → nan

HIDDEN_3 Analysis:
  Sparsity Evolution: 0.000 → 0.000
  Mean Activation: nan → nan
  Activation Variance: nan → nan

REPRESENTATION SHIFTS DURING DIFFICULTY CHANGES

Shift at Episode 199 (Variable Paddle Speed):
  Q-Value Shift: nan
  Hidden Layer Shifts: {'hidden_1': np.float64(nan), 'hidden_2': np.float64(nan), 'hidden_3': np.float64(nan)}

Shift at Episode 399 (Dynamic Ball Speed):
  Q-Value Shift: nan
  Hidden Layer Shifts: {'hidden_1': np.float64(nan), 'hidden_2': np.float64(nan), 'hidden_3': np.float64(nan)}

Shift at Episode 599 (Adaptive Paddle Size):
  Q-Value Shift: nan
  Hidden Layer Shifts: {'hidden_1': np.float64(nan), 'hidden_2': np.float64(nan), 'hidden_3': np.float64(nan)}

Shift at Episode 799 (Full Adaptive Mode):
  Q-Value Shift: nan
  Hidden Layer Shifts: {'hidden_1': np.float64(nan), 'hidden_2': np.float64(nan), 'hidden_3': np.float64(nan)}

SIMULATING GAMEPLAY VIDEOS

Normal Difficulty
Summary:
  Duration: 63 steps
  Final Score: 1350
  Paddle Hits: 0
  Bricks Destroyed: 29
  Action Distribution: No-op: 63, Left: 0, Right: 0
  Key Events: 30 significant events

Medium Difficulty

Summary:
Duration: 63 steps
Final Score: 1350
Paddle Hits: 0
Bricks Destroyed: 29
Action Distribution: No-op: 63, Left: 0, Right: 0
Key Events: 30 significant events

High Difficulty

Summary:
Duration: 32 steps
Final Score: 950
Paddle Hits: 0
Bricks Destroyed: 21
Action Distribution: No-op: 32, Left: 0, Right: 0
Key Events: 22 significant events

STRATEGY SUCCESS/FAILURE ANALYSIS

SUCCESSFUL STRATEGIES LEARNED

Difficulty Adaptation Strategy:
- Successfully triggered 24 adaptations
- Learned to increase exploration when performance drops
- Adaptation success rate: 91.7%

Curriculum Learning Strategy:
- Successfully progressed through 4 curriculum stages
- Advanced to 'Variable Paddle Speed' with performance: 615.0
- Advanced to 'Dynamic Ball Speed' with performance: 624.0
- Advanced to 'Adaptive Paddle Size' with performance: 606.0
- Advanced to 'Full Adaptive Mode' with performance: 609.0

Reward Optimization Strategy:

- Most successful action: No-op
- No-op success rate: 98.9%
- Left success rate: 0.5%
- Right success rate: 0.5%

STRATEGIES THAT FAILED TO DEVELOP

Predictive Difficulty Adaptation:
- Agent reacts to difficulty changes but doesn't predict them
- No evidence of proactive strategy adjustment
- Relies on reactive adaptation rather than anticipatory behavior

Advanced Ball Trajectory Prediction:
- Limited evidence of sophisticated ball trajectory prediction
- Paddle movements appear more reactive than predictive
- Could benefit from longer-term state sequence modeling

RECOMMENDATIONS FOR IMPROVEMENT

1. Implement LSTM-based sequence modeling (Priority: High)
   Add recurrent layers to detect patterns in difficulty changes
2. Physics-informed state representation (Priority: Medium)
   Include ball velocity and acceleration in state representation