Applying Reinforcement Learning to the Determination of Crystal Structures with Neutron Diffraction

August 8th, 2018

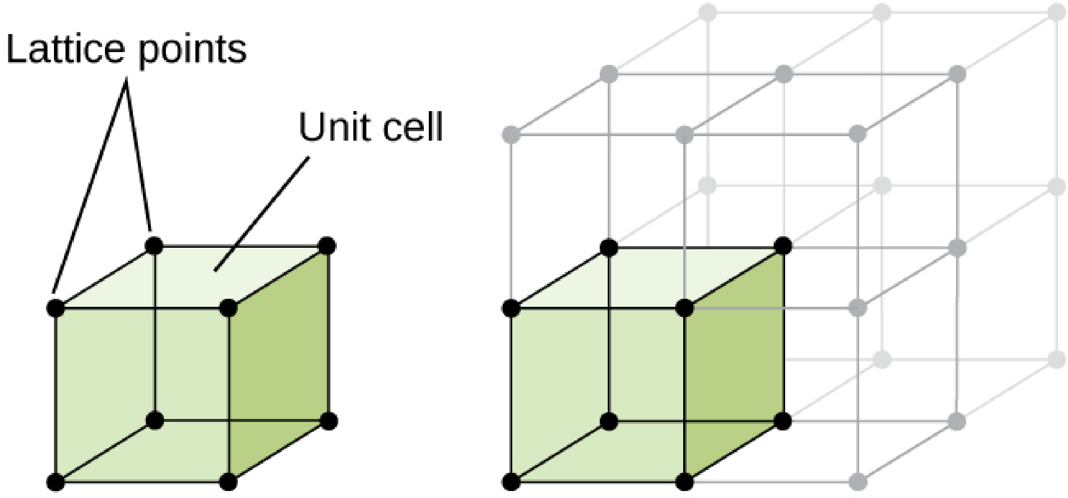
Abigail Wilson
NIST SURF Fellow, Tufts University Class of 2021

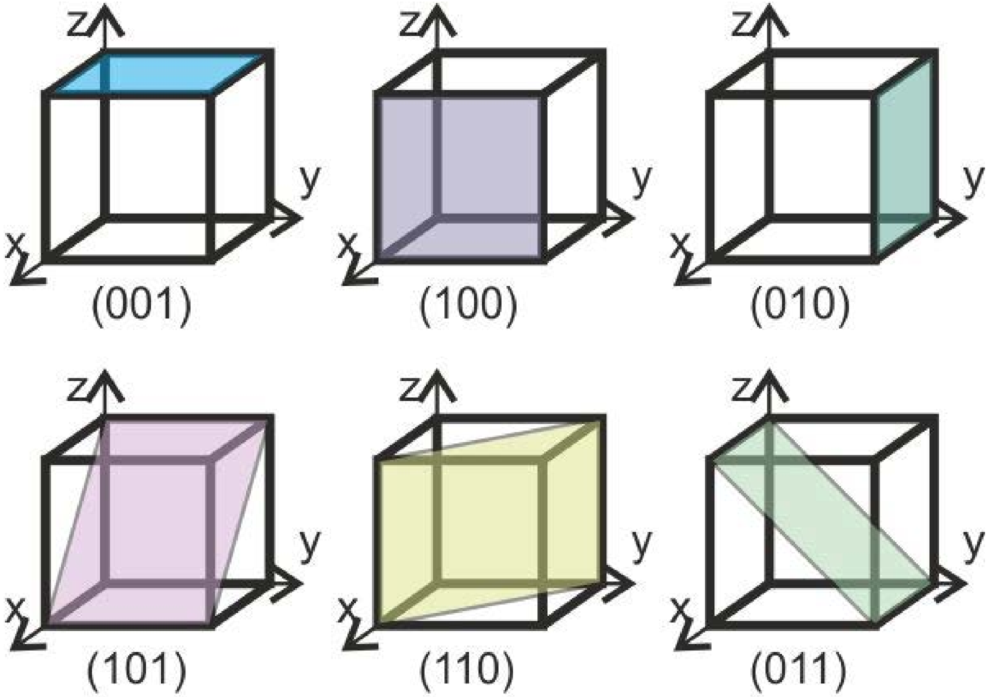
Dr. William Ratcliff
NIST Center For Neutron Research, Neutron Condensed Matter Science Group

Outline

- Background
  - Crystallography
  - Neutron Diffraction
- Problem
- Algorithms
- Performance
- Reinforcement Learning in this Problem Space
- Future Steps

Crystal Structure

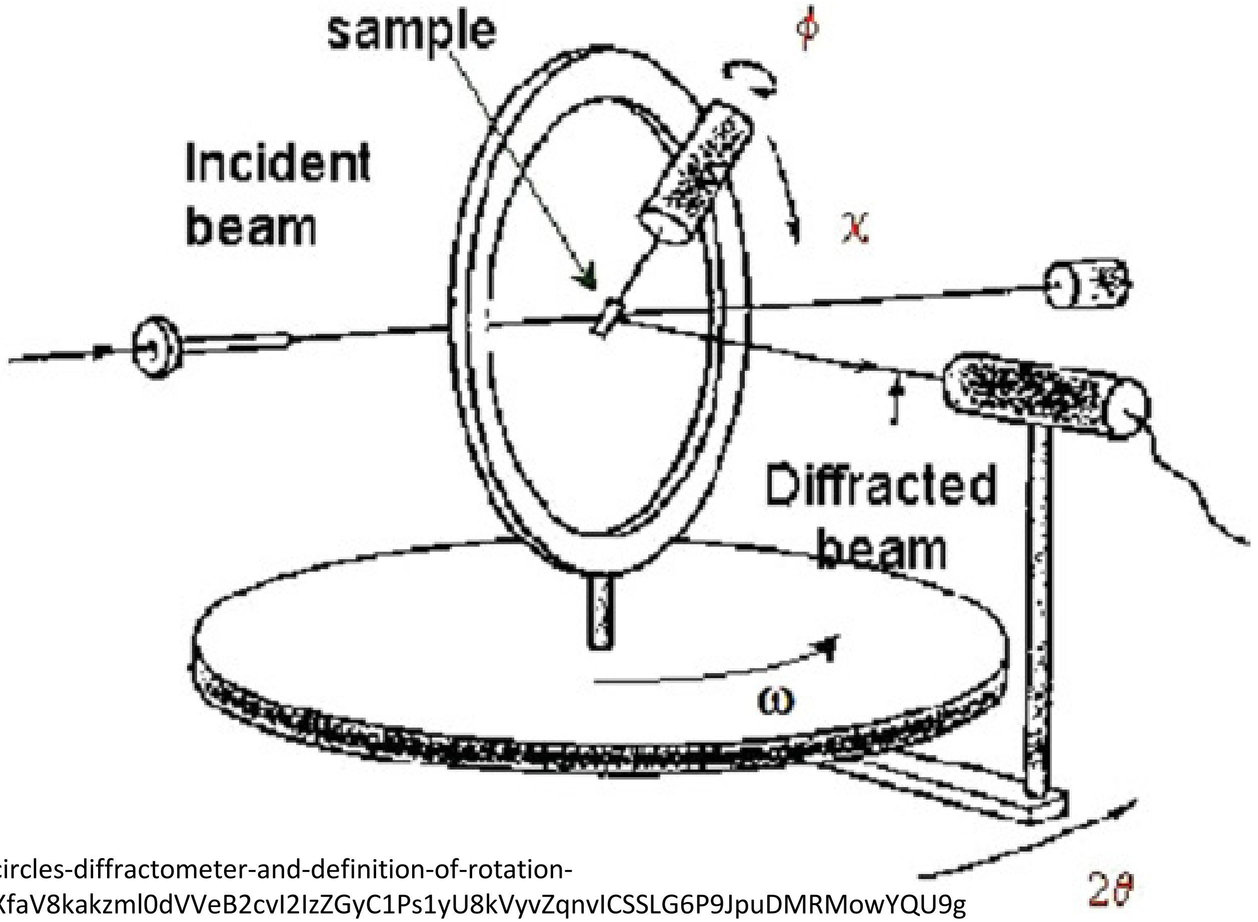




https://opentextbc.ca/chemistry/chapter/10-6-lattice-structures-in-crystalline-solids/

https://www.researchgate.net/figure/Miller-indices-indicating-the-plane-perpendicular-to-the-vector-given-for-the-cubic_fig7_302838100

Neutron Diffraction



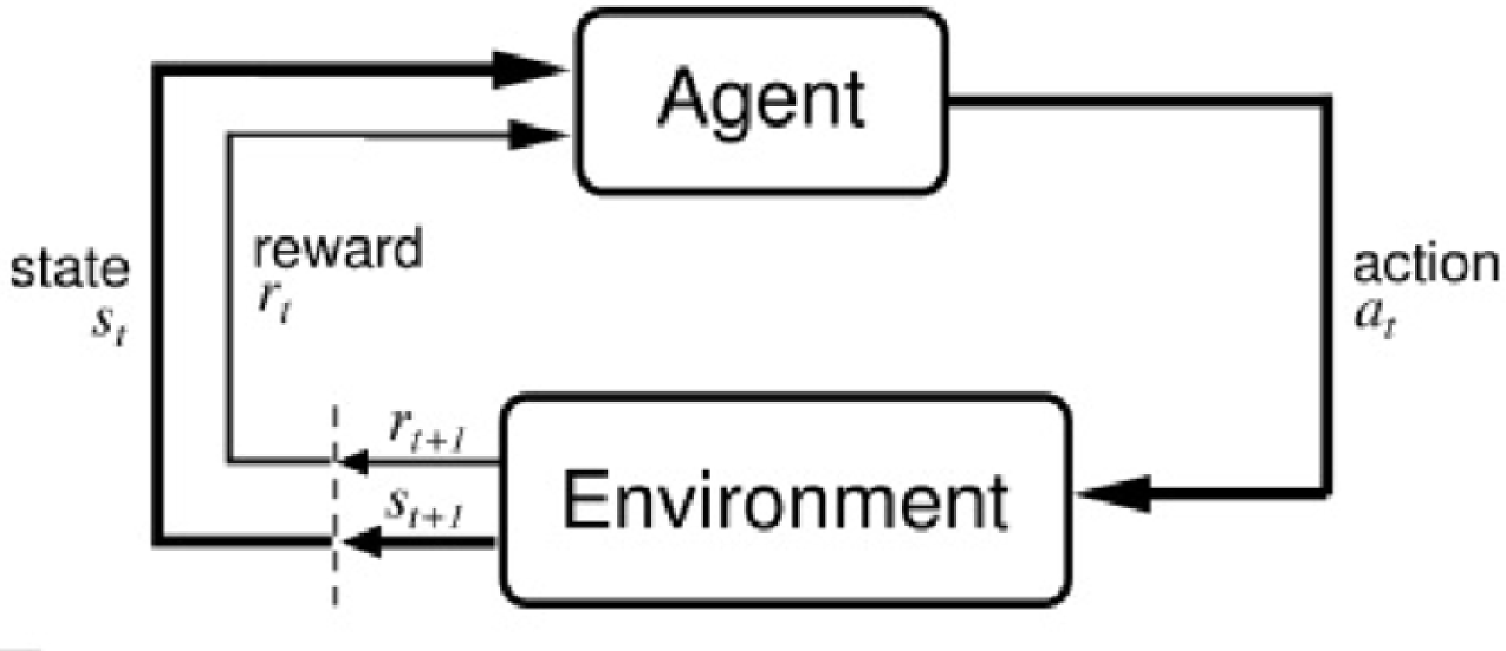
https://www.researchgate.net/figure/Principle-of-a-4-circles-diffractometer-and-definition-of-rotation-angles_fig5_320672016?_sg=2SkZLSPV7hSnuAUpNs3TXfaV8kakzml0dVVeB2cvI2IzZGyC1Ps1yU8kVyvZqnvICSSLG6P9JpuDMRMowYQU9g

Problem

- Determining atomic structure is key to understanding novel materials
- Beam time is highly limited
  - Few neutron sources
- Need to optimize experiments to minimize required beam time for experiments

Can Reinforcement Learning be applied to optimize Neutron Diffraction experiments?

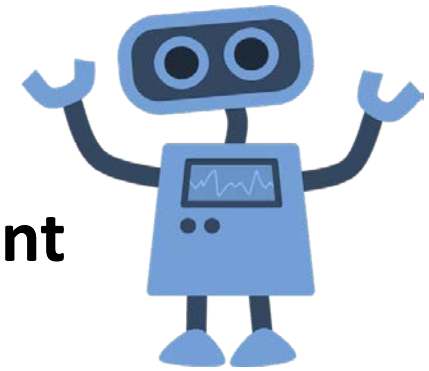
Reinforcement learning



Whiteson, Shimon. (2010). Adaptive Representations for Reinforcement Learning. 10.1007/978-3-642-13932-1.

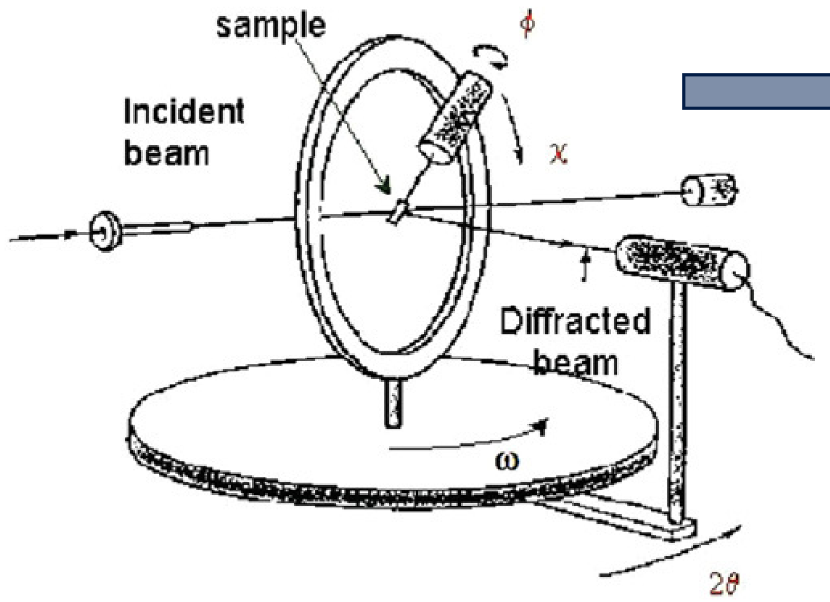
Action: Take Measurement

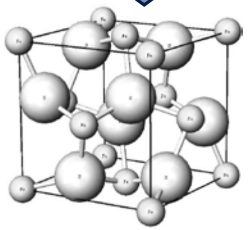
Agent



https://publicdomainvectors.org/en/free-clipart/Green-robot/38871.html

Environment





http://www.geocities.jp/ohba_lab_ob_page/structure6.html

Reward: +/-

State: Measurements Taken

https://www.researchgate.net/figure/Principle-of-a-4-circles-diffractometer-and-definition-of-rotation-angles_fig5_320672016?_sg=2SkZLSPV7hSnuAUpNs3TXfaV8kakzml0dVVeB2cvI2IzZGyC1Ps1yU8kVyvZqnvICSSLG6P9JpuDMRMowYQU9g

Our Algorithms

ALGORITHMS

- Q Learning
- Deep Q Learning
- Epsilon Greedy
  - Implemented by Ryan Cho and Telon Yan
- Actor Critic
  - Implemented in collaboration with Ryan Cho and Telon Yan

RESOURCES

- BLAND [1]
  - Crystallographic library
- Training Data
  - Pr_2NiO_4 – single crystal data from FullProf

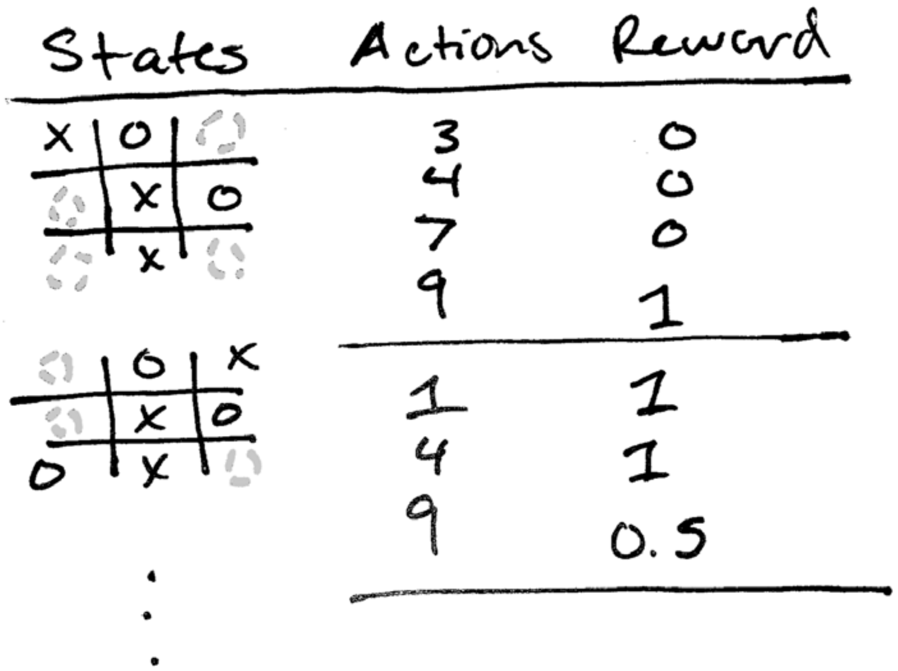
[1] J. E. Lesniewski, S. M.T. Disseler, D. J. Quintanta, P. Kienzle, W. D. Ratcliff. “Bayesian Method for the Analysis of Diffraction Patterns using BLAND,” *Journal of Applied Crystallography*, vol. 49, December 2016.

Epsilon Greedy

- Implemented by Ryan Cho and Telon Yan
- Agent chooses the best action at each state, with a small likelihood of choosing a random action
  - The “best” action is determined by recording the average reward earned from taking that action previously
- Preliminary results:
  - The agent learned a preference for certain measurements
  - Have not yet identified a pattern

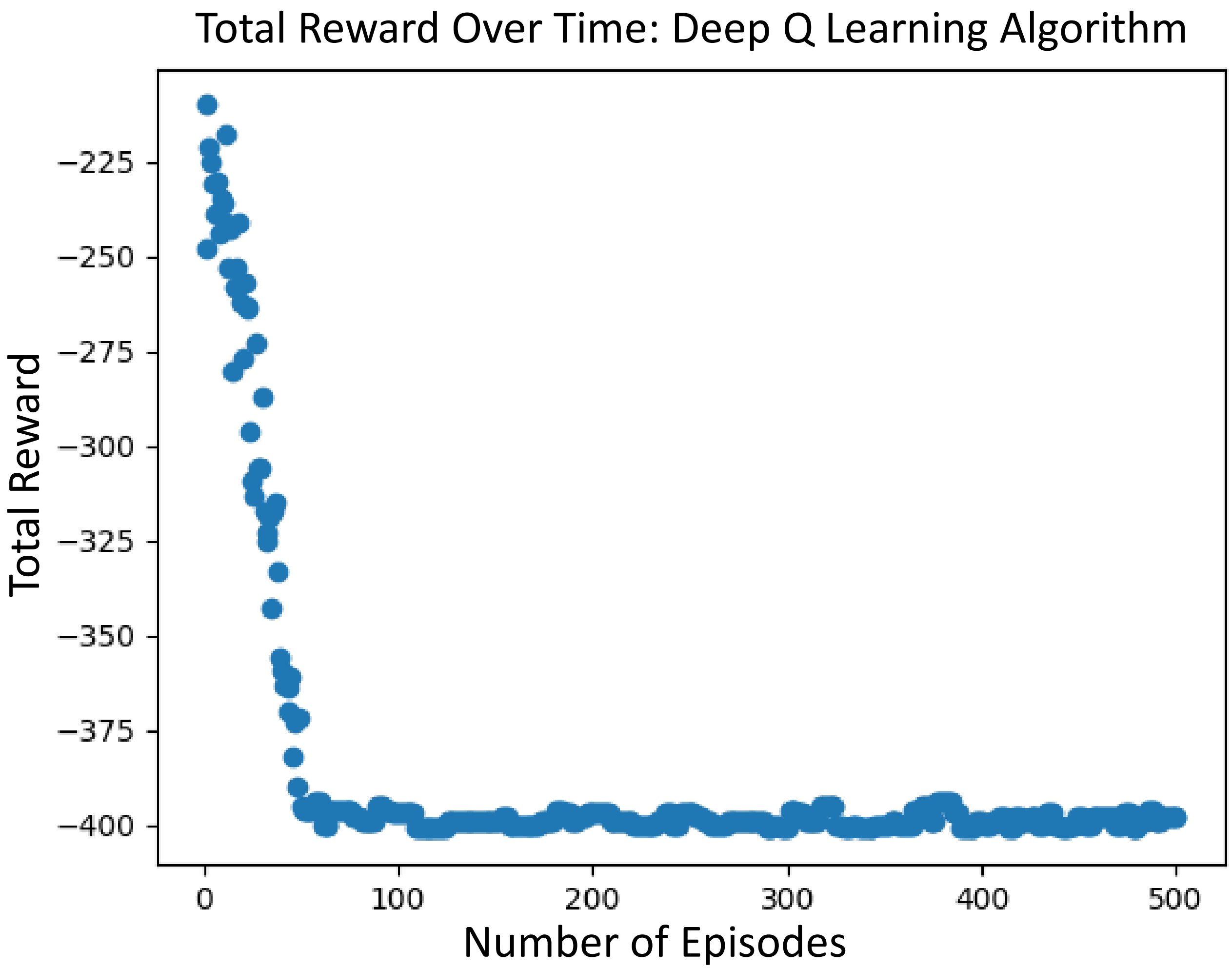
Q Learning

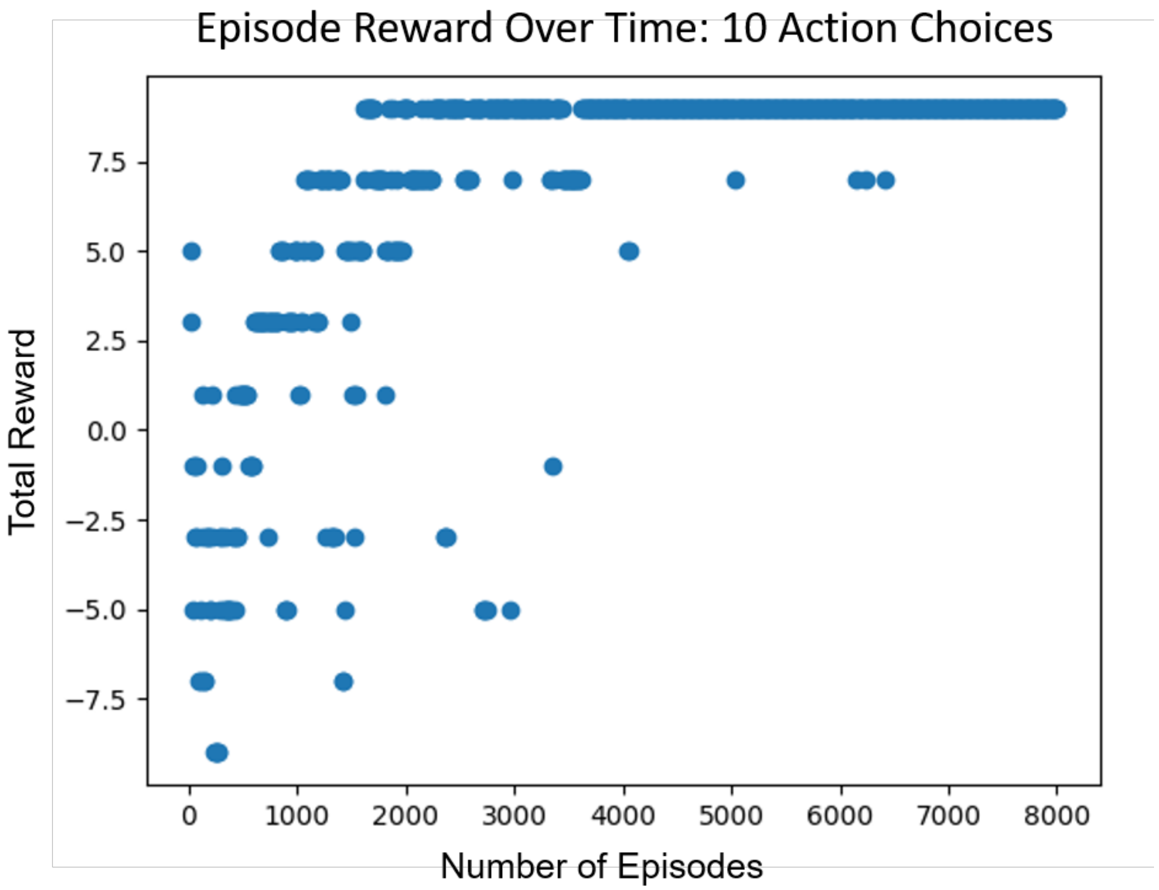
A sample Q table for Tic-Tac-Toe



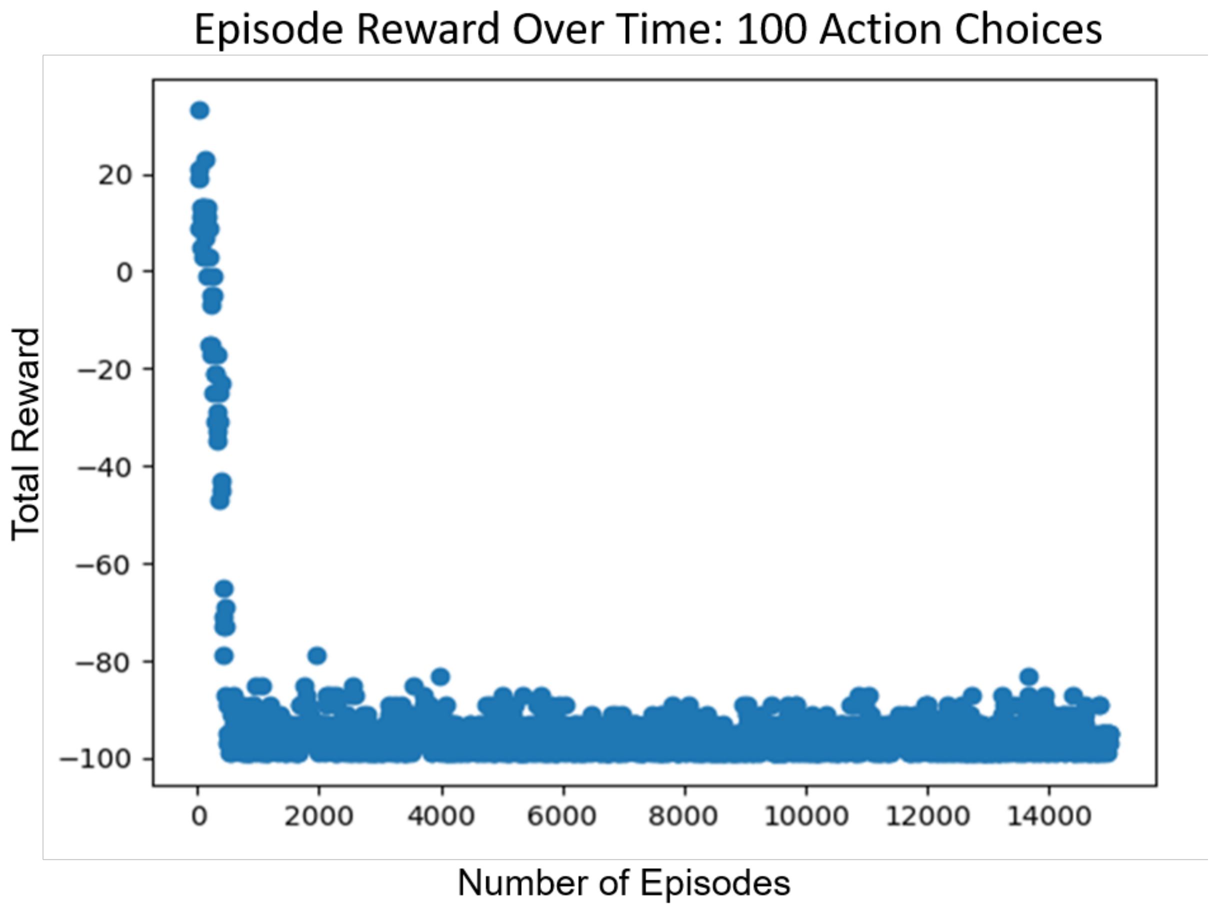
Deep Q Learning

- Adds a Deep Neural Network to standard Q Learning
- Implemented an Environment in TensorForce to represent this problem space

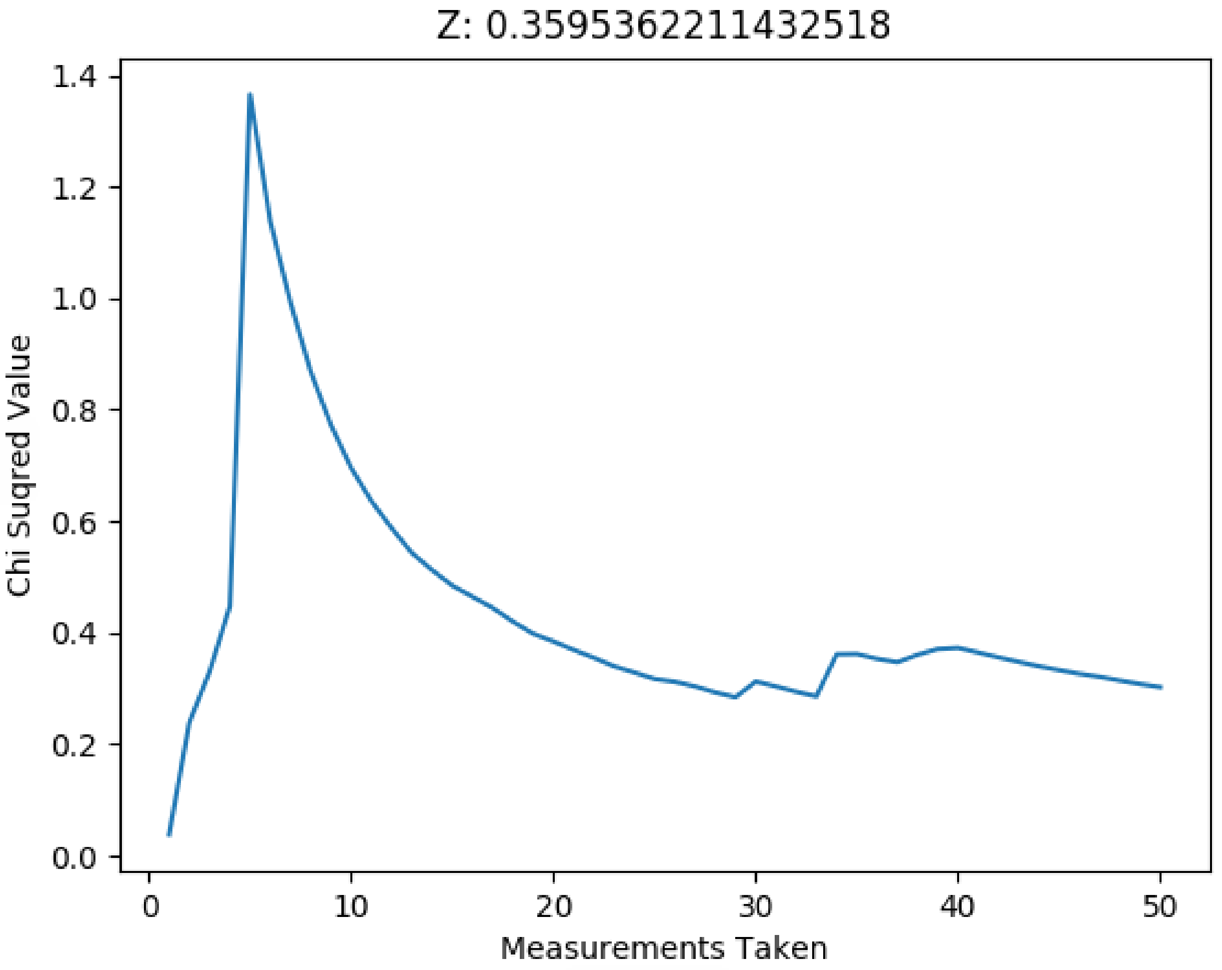

Total Reward Over Time: Deep Q Learning Algorithm
Total Reward
−225
−250
−275
−300
−325
−350
−375
−400
0
100
200
300
400
500
Number of Episodes



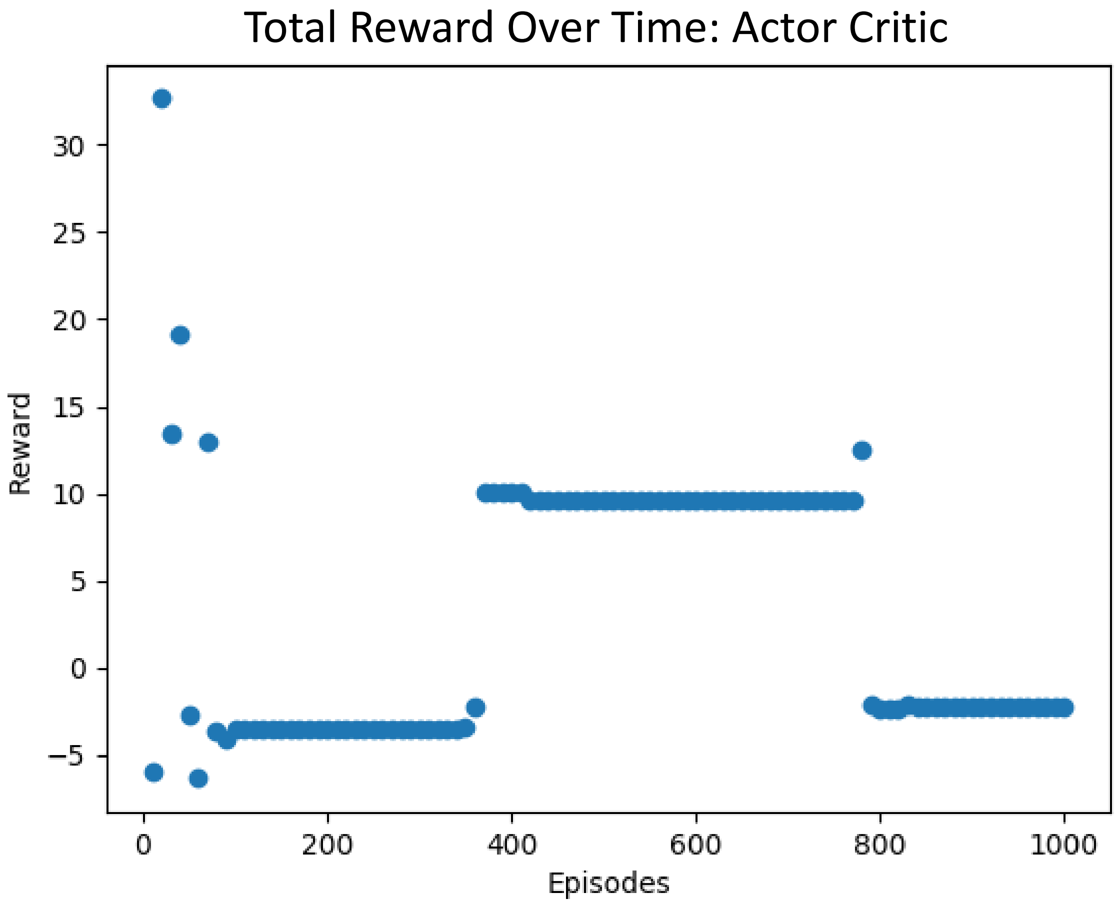
Episode Reward Over Time: 10 Action Choices
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
Total Reward
0
1000
2000
3000
4000
5000
6000
7000
8000
Number of Episodes

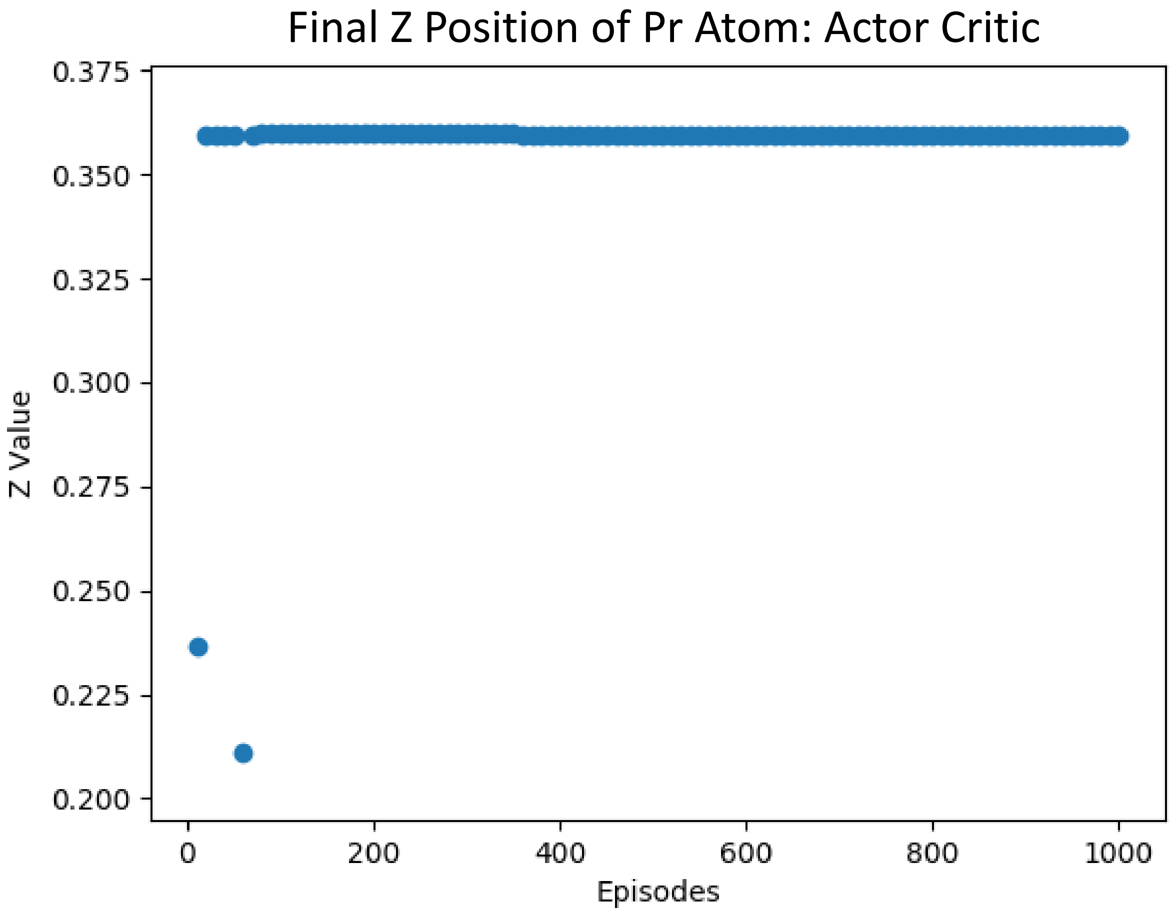


Episode Reward Over Time: 100 Action Choices
20
0
−20
−40
−60
−80
−100
Total Reward
0
2000
4000
6000
8000
10000
12000
14000
Number of Episodes



Z: 0.3595362211432518
Chi Suqred Value
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
Measurements Taken


Total Reward Over Time: Actor Critic
Reward
30
25
20
15
10
5
0
−5
0
200
400
600
800
1000
Episodes


Final Z Position of Pr Atom: Actor Critic
Z Value
0.375
0.350
0.325
0.300
0.275
0.250
0.225
0.200
0
200
400
600
800
1000
Episodes

Reinforcement Learning in Crystallographic Problem Space

STRENGTHS

- Problem can be framed as a game
- Has the potential to identify patterns researchers wouldn't find
- If effective, could optimize experiments

WEAKNESSES

- Has a large discrete action space
- Challenging to develop an effective reward

Future Steps

- Complete Actor Critic
- Continue exploring rewards functions
- Implement the Wolpertinger architecture
  - Designed for large discrete action spaces
  - Proposed by Dulac-Arnold et al in “Deep Reinforcement Learning in Large Discrete Action Spaces” (arXiv:1512.07679 [cs.AI])

Acknowledgements

Dr. William Ratcliff, Paul Kienzle

NIST Center For Neutron Research, Neutron Condensed Matter Science Group

Ryan Cho and Telon Yan (SHIP)

Montgomery Blair High School

Tufts University

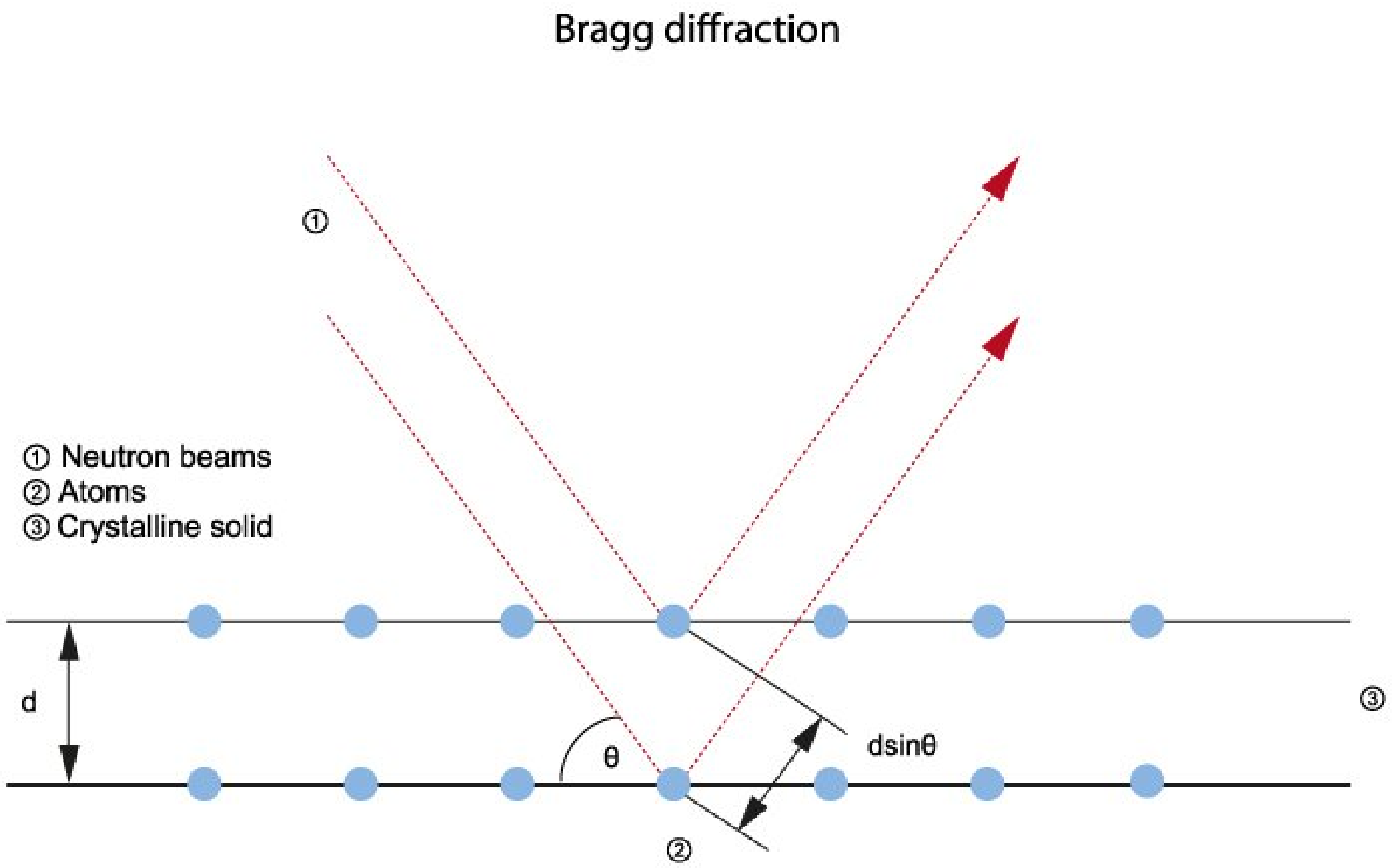
NIST SURF Program and Directors

Our Reinforcement Learning Game

- The environment contains an initial model of the crystal
- Agent chooses to make a certain measurement
  - Model is updated
  - Fit of the model is calculated
  - Agent gets reward
- Repeat until the uncertainty drops below threshold

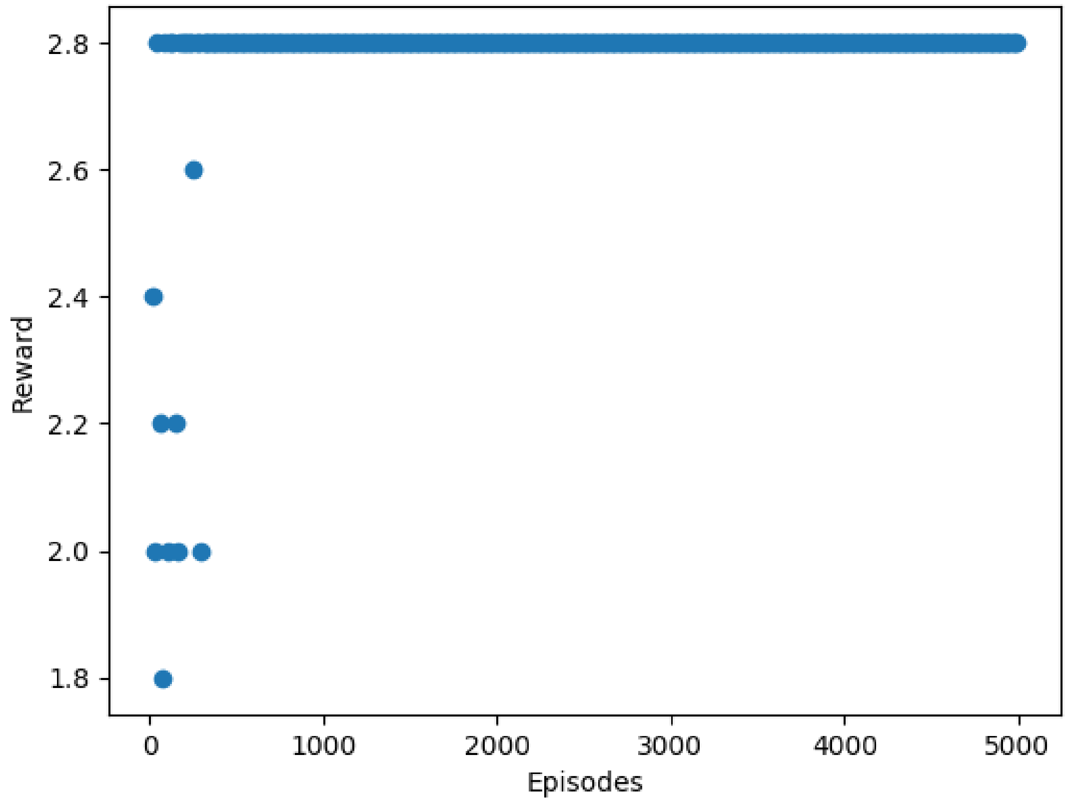
- The reward function
  - A penalty each step, to encourage efficiency
  - A measure of fit: χ^2
  - A reward for improving the fit of the model
    - Constant
    - Proportional to how significantly the fit improved
    - Inversely proportional to χ^2

Neutron Diffraction



https://nmi3.eu/neutron-research/techniques-for-/structural-research.html

Episode Reward Over Time: Actor Critic



Final Z Value: Actor Critic