Constructing a Reinforcement Learning Environment for Determining a UB Matrix

Cynthia Liu
Mentor: William Ratcliff

In Layman's Terms

- NCNR uses machines (gonoimeters) to analyze crystals
  - Neutron diffraction
- Crystals have UB matrix – describes crystal structure and orientation
- Reinforcement learning - learn by trying
- "Letting the gonoimeters learn how to best find a UB matrix by trying"

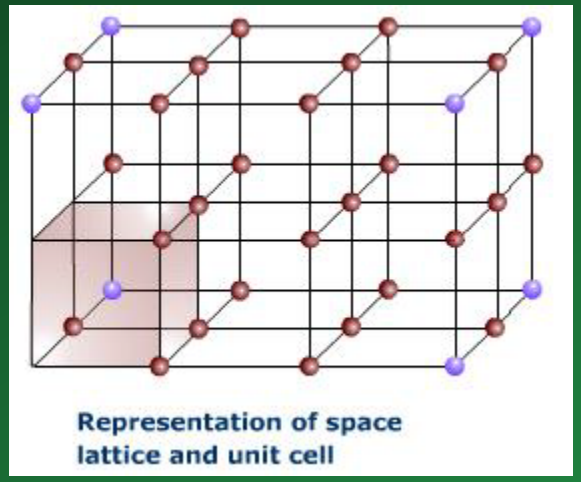
Importance

- Reinforcement learning applications:
  - General, self-teaching AI
  - Personalization
- UB Matrix problem:
  - Logic behind the “best” measurements?
  - Most efficient way of handling incorrect initial measurements?
- Neutrons are difficult to make
  - Want fewer measurements, same results

http://blogthinkbig.com/google-wants-to-create-the-ultimate-personal-assistant/

Background - Crystallography

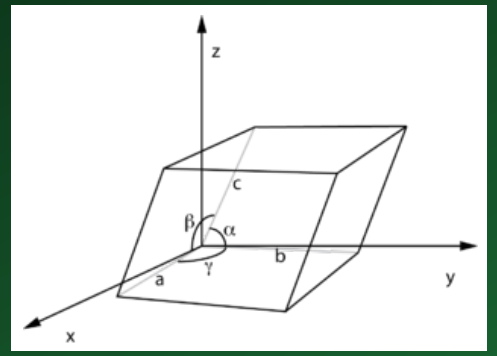
- Crystal is repeating, ordered arrangement of atoms
  - Smallest component is unit cell
- Crystallography – structure and properties of crystals



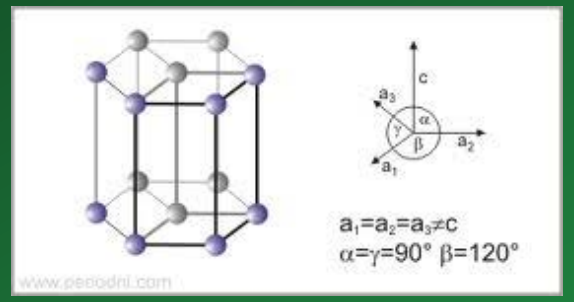
https://alienspacesciencenews.wordpress.com/2013/06/12/line-22f1fa19c3m1b7c8e13c9c6c2-tetrahedron-crystal-miller-matrix-core-ufo-5g-wow-seti/

Background - Crystallography

- Crystal coordinate systems
- Axes a, b, c
- Angles α, β, γ

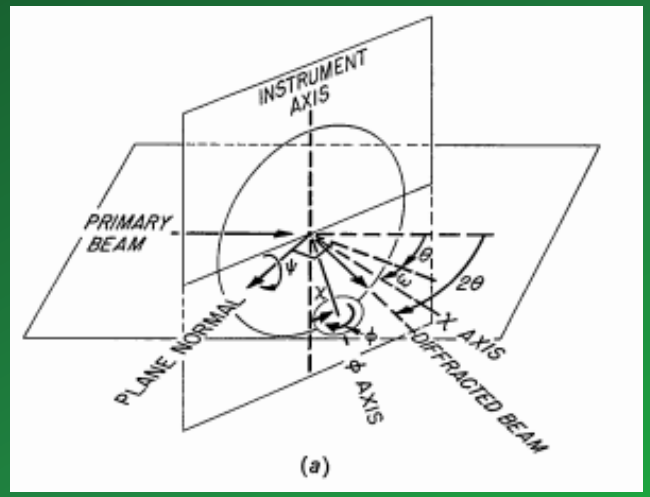


https://en.wikipedia.org/wiki/Lattice_constant



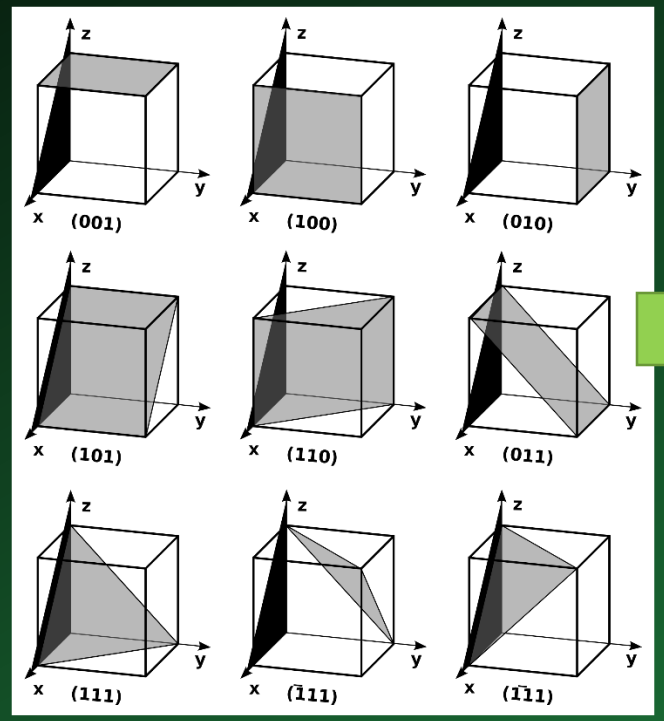
https://www.quia.com/jg/2506550list.html

- Instrument coordinate systems
- Fixed on instrument axis
- Angles 2θ, χ, φ



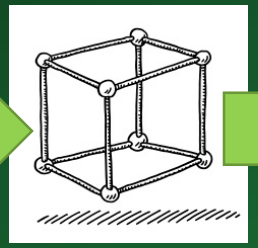
http://journals.iucr.org/q/issues/1967/04/00/a05492/a05492.pdf

Background - Crystallography



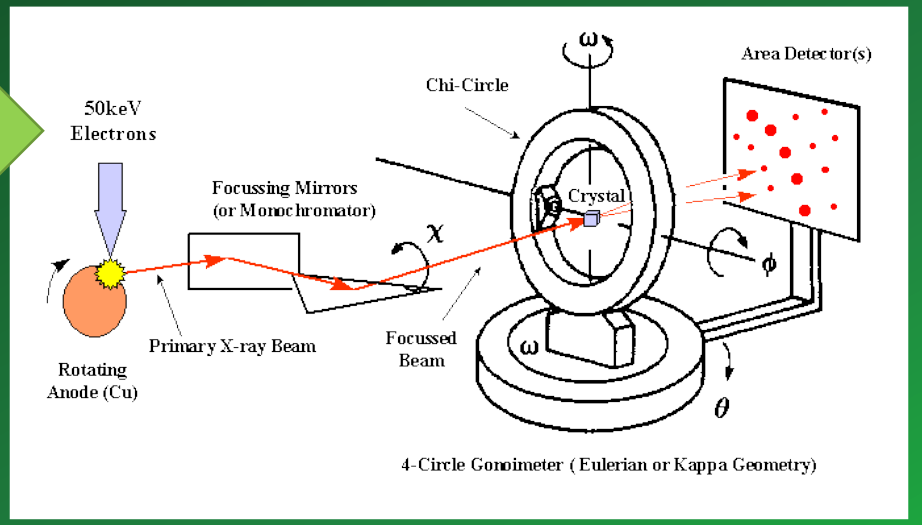
https://commons.wikimedia.org/wiki/File:Miller_Indices_Cubes2.svg

B



http://i.istockimg.com/file_thumbview_approve/81543089/5/stock-illustration-81543089-cube-ball-and-stick-model-drawing.jpg

U

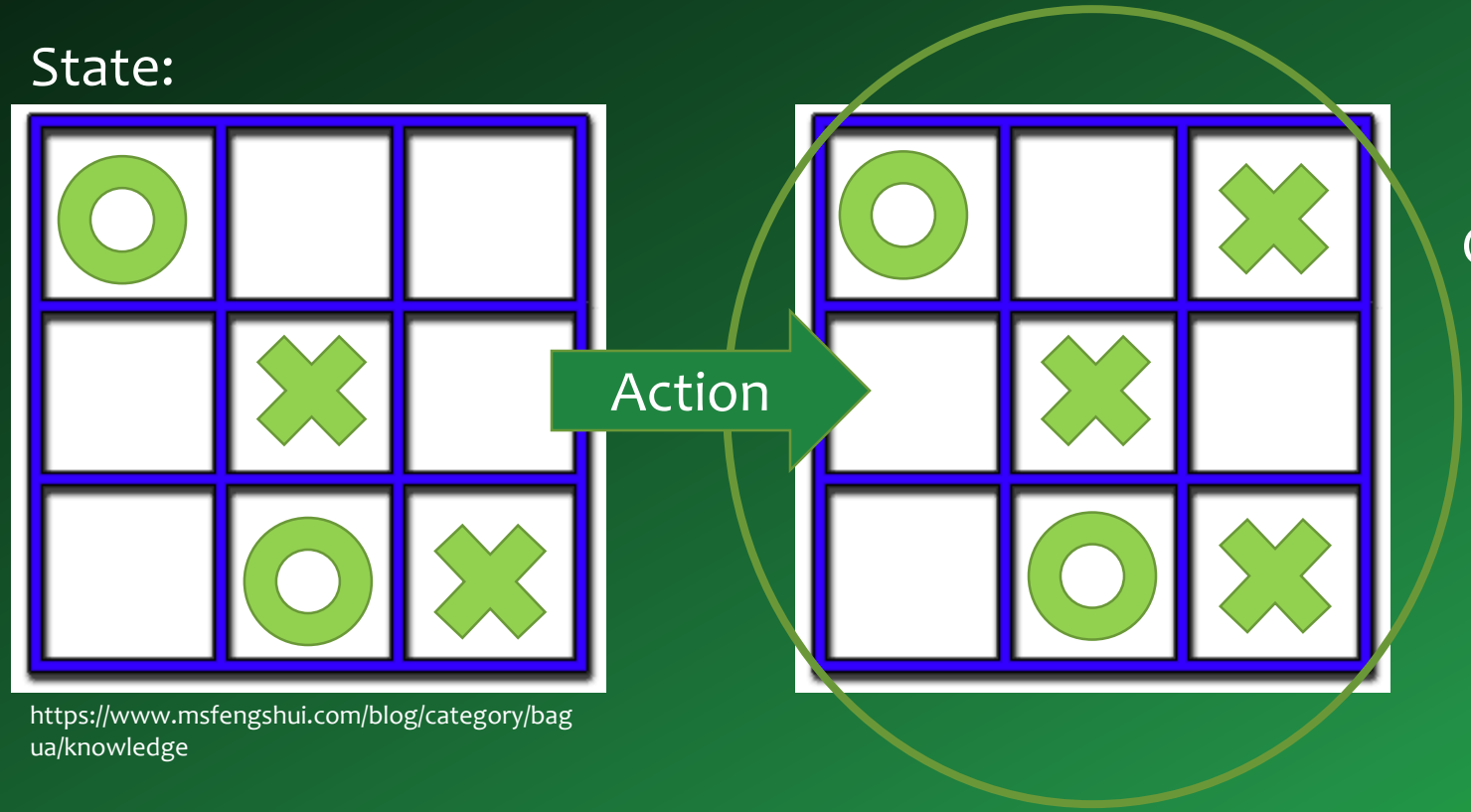


http://chem.suwon.ac.kr/~jwchoi/biophy2.htm

Background – Reinforcement Learning

Environment: Tic-tac-toe

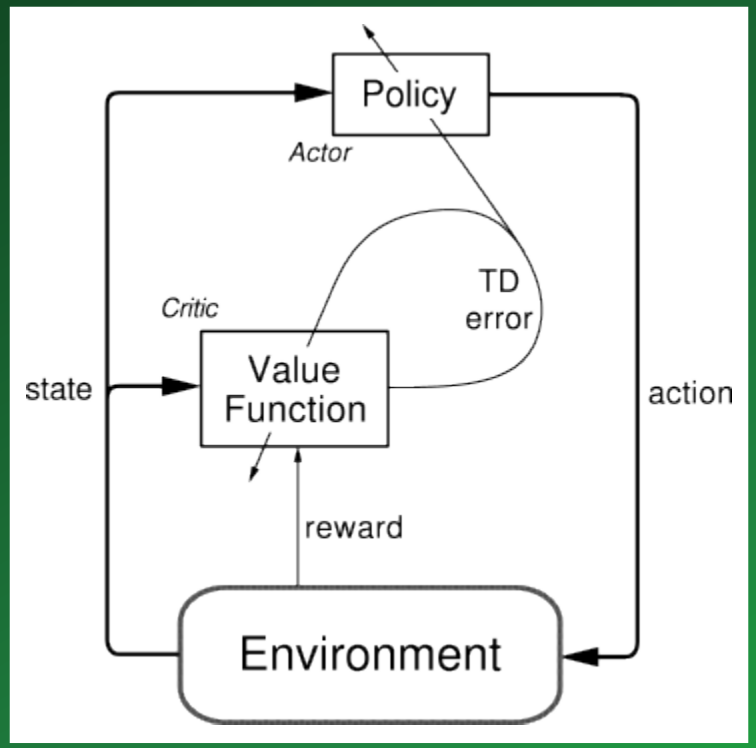
State:



https://www.msfengshui.com/blog/category/baguа/knowledge

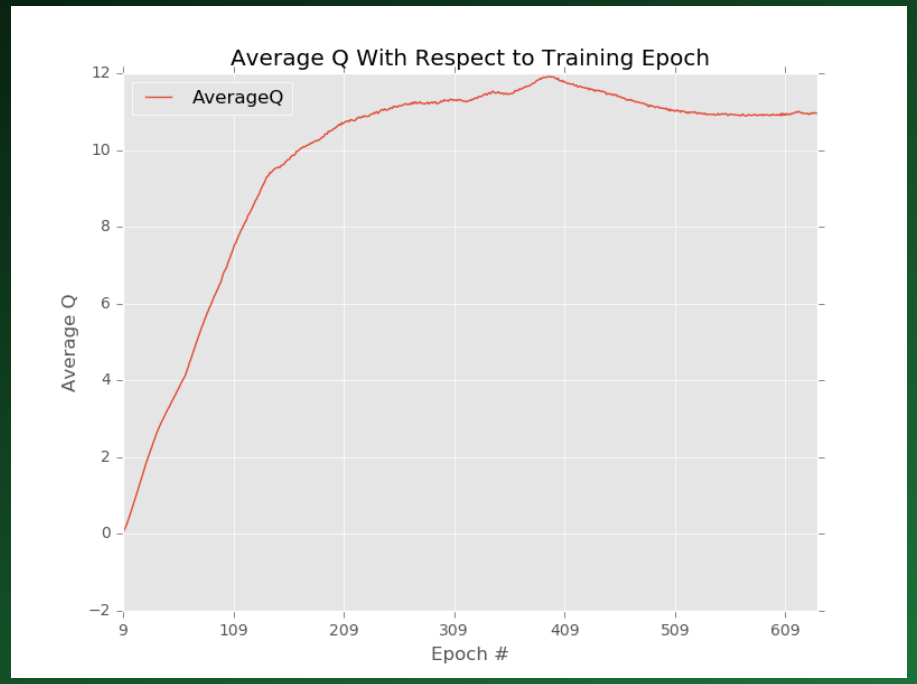
Background – Reinforcement Learning

- Actor-critic model
  - Actor: “Policy” for choosing actions; strategy
  - Critic: Evaluates quality of choice
    - Value function Q
- Actor and critic improve by minimizing loss and maximizing reward



https://webdocs.cs.ualberta.ca/~sutton/book/ebook/node66.html

Rllab on Cartpole

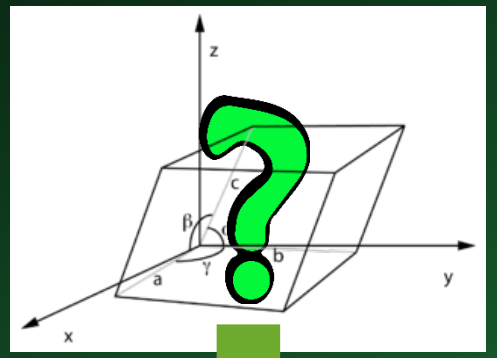


My Work

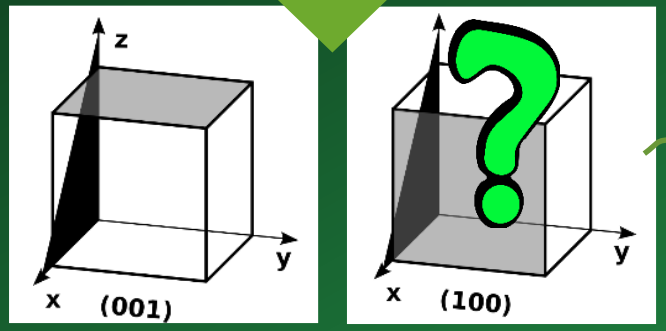
The problem, progress, and more

The Problem

https://en.wikipedia.org/wiki/Lattice_constant
http://www.freestockphotos.biz/stockphoto/16032



Take 2 Measurements



https://commons.wikimedia.org/wiki/File:Miller_Indices_Cubes2.svg http://www.freestockphotos.biz/stockphoto/16032

Calculate UB Matrix

```
[[ 0.03010972  0.07575927  0.14031317]
 [-0.17076072 -0.01488778  0.05298718]
 [ 0.04862107 -0.09920263  0.09920263]]
```

Measure more
Refine UB

The UB Matrix Problem

The Problem

- 3D action space (X, φ, 2θ)
  - X, φ are continuous, 2θ is discrete
- State: current UB and “guessed” initial indices
  - Program changes guesses if loss too high
  - Terminates when results match prediction for 3 measurements
- Reward:
  - Peak exists most important
  - Distance from predicted angles is loss
- Observation: Structure factor

UBEnv and Box2d_env_ub

- Box2d_env_ub
  - Original version for training in general 2D physics environment
  - Modified parameters, stepping, movement caused by action
- UBEnv
  - More environment-specific
  - Computes M, N, UB's, more

Environment Function	Completed
Added	Init_UB, calc_M, calc_N, calc_expected, calc_loss, add_ub, setup_spaces
Overrides	__init__, reset, compute_reward, is_current_done forward_dynamics, get_raw_obs, step

UBSpace

- New action type: discrete AND continuous
  - Discrete is mapped continuous
- Extra dimension for choice
- Most important, least code
  - Methods used in NormalizedEnv, which updates estimates
  - UBEnv and Box2d_env_ub depend on space

Ongoing and Future Work

- Current:
  - Incorporate first two measurements as part of training
  - What is realistic training data? - Automation
- Future:
  - Extracting data from machine
  - Training, implementation!

Acknowledgements

- Many, many thanks to:
  - Dr. William Ratcliff
  - Mr. Paul Kienzle
  - Dr. Julie Borchers
  - Ms. Yamali Hernandez