MOREHOUSE COLLEGE

HUME CENTER FOR NATIONAL SECURITY AND TECHNOLOGY
VIRGINIA TECH

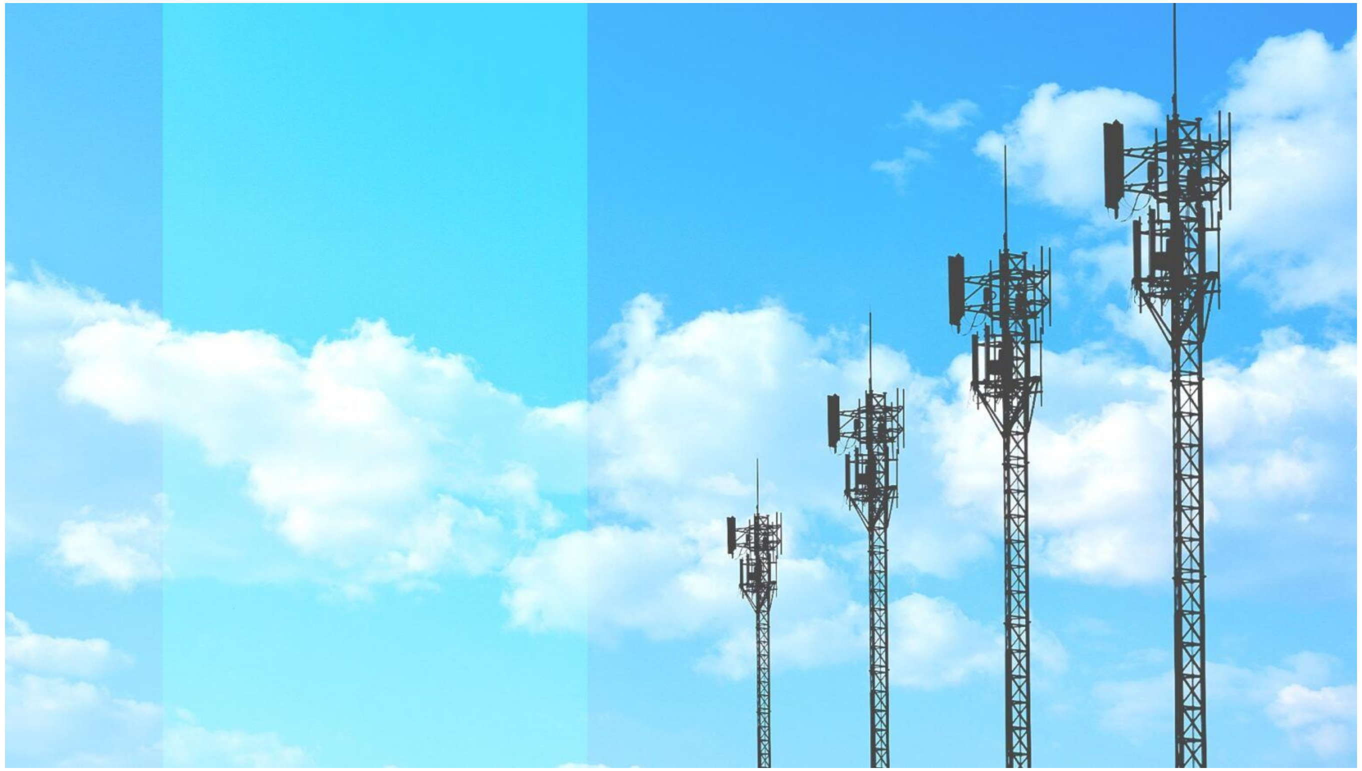
The RFRL Gym:

A testbed for training and evaluating reinforcement learning for cognitive radio systems

Danny Rosen, Alyse Jones, Dr. Chris Headley

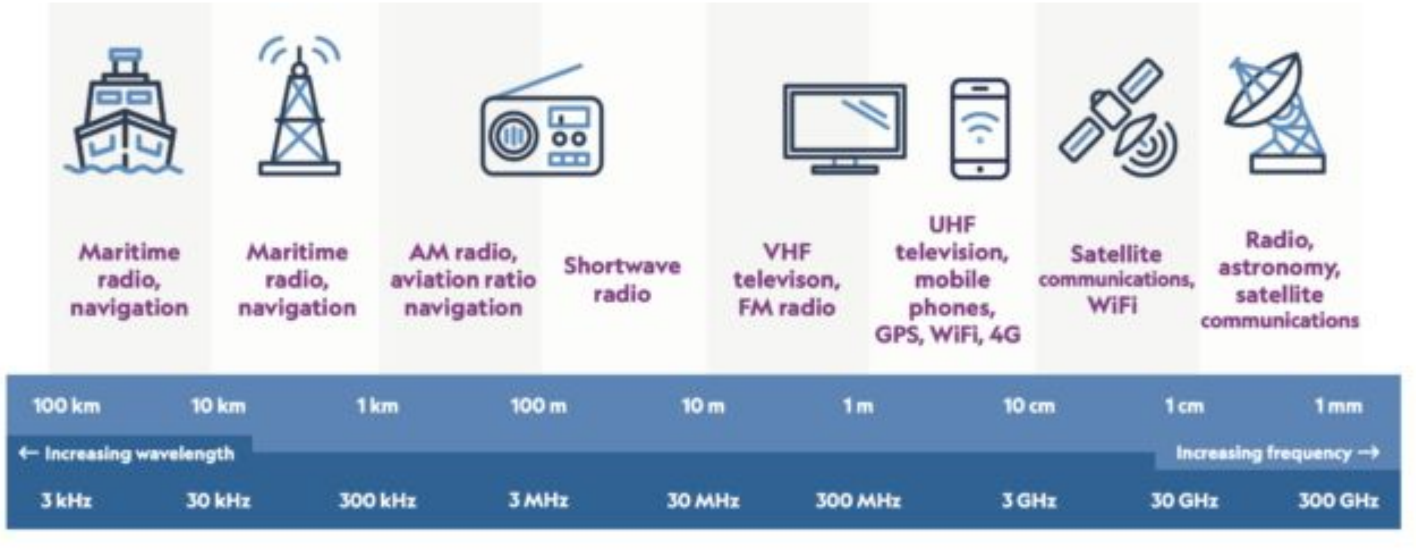
Overview

- Introduction to problem space
  - Wireless Spectrum Problems: Interference and congestion
  - DSA and Jamming
- Motivation
- Contributions
- RFRL Gym Framework
- Validation and Analysis
- Future Work and Conclusions



Wireless Spectrum

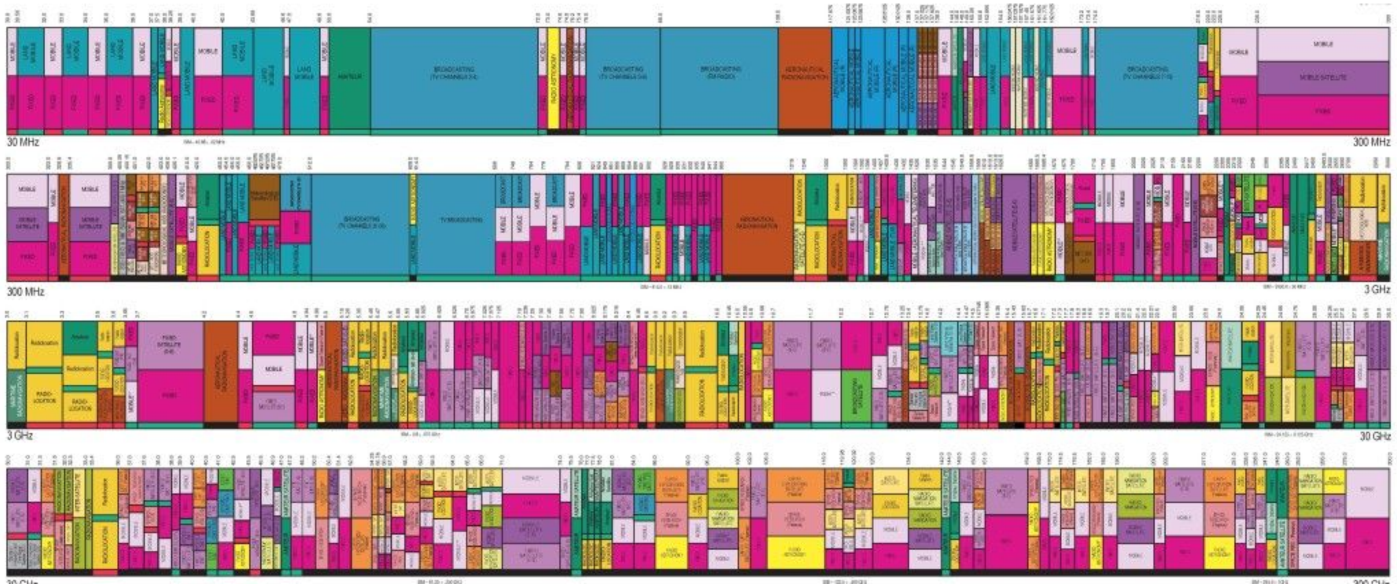
- Set of electromagnetic frequencies that are used for wireless communication
- Continuous Frequencies grouped into discrete “bands” or “channels”
- Spectrum in our daily lives - Cellular, WIFI, Bluetooth
- Problem: Interference
    - Multiple transmissions at the same frequency
    - Receiver can not correctly interpret correct signal



Acquired from [1]

Causes of Interference

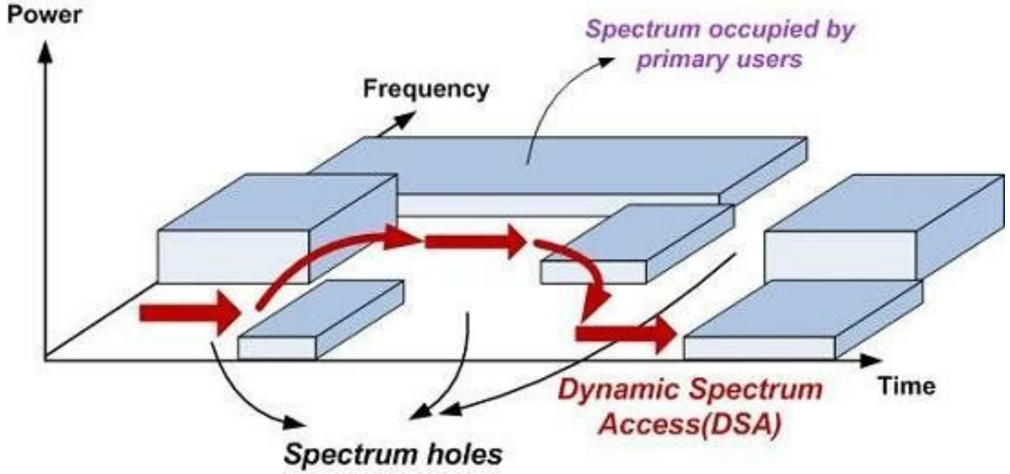
- Military: Electronic Warfare
- Commercial: Interactions between consumer devices
    - 2010-2021: 183x increase in wireless spectrum use [2]
- FCC Spectrum Allocation



Acquired from [3]

Spectrum Management

Cognitive Radio (CR)	Dynamic Spectrum Access (DSA)
Radio with the ability to dynamically change transmission parameters (center frequency, bandwidth) based on spectrum sensing results.	Algorithmically switch to vacant frequencies to avoid other entities. Spectrum holes found by Spectrum Sensing.



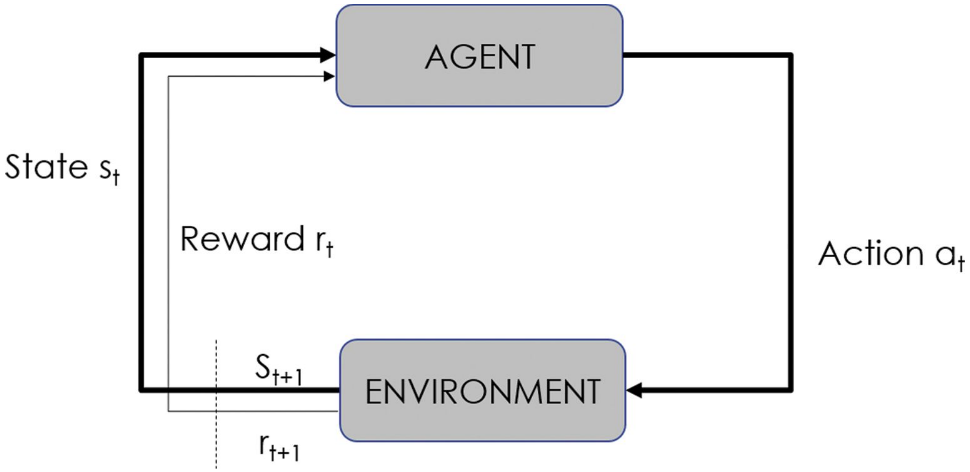
Acquired from [4]

CRs with DSA alone cannot understand where spectrum users will go next, resulting in interference

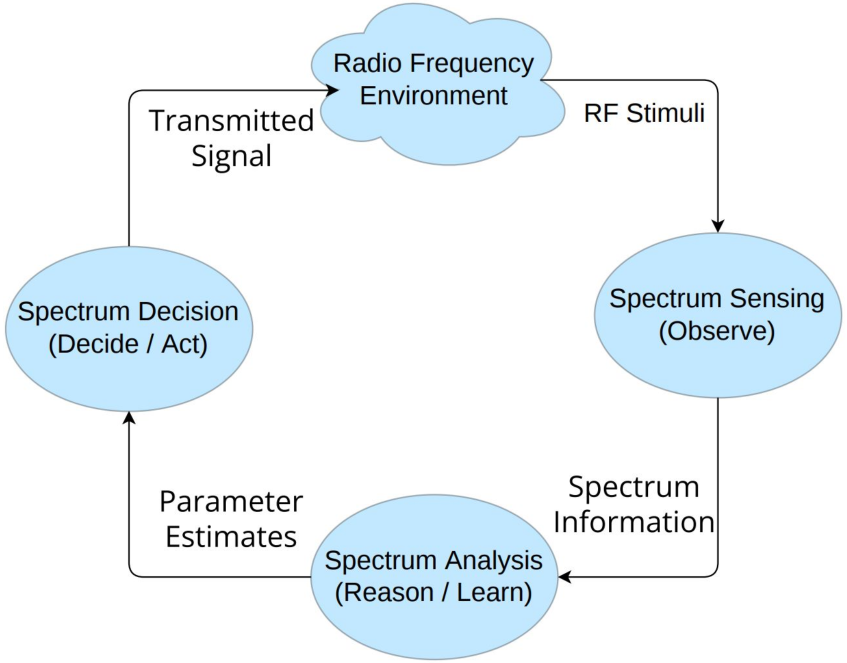
Reinforcement Learning in Spectrum Management

OODA: Observe Orient Decide Act

Understanding the expected behavior of spectrum activity through **prediction** with Reinforcement Learning



Acquired from [5]



Can use Reinforcement Learning (RL) with CRs to enable **prediction** in its cognition to understand the dynamics and behavior of external factors affecting the RF environment.

Motivation of the RFRL Gym

1. **Enhance** research capabilities in wireless spectrum applications
   - Easy-to-use cognitive radio-centric research tool
2. **Address** pressing problems of wireless spectrum management
   - Reducing interference in the wireless spectrum
3. **Expand** interest and knowledge in RL-based wireless communications

Contributions

- **The RFRL Gym** - a customizable tool built in Open AI gymnasium to simulate and test the performance of reinforcement learning (RL) for cognitive radio applications (such as DSA and jamming)
- The RFRL Gym contains the following **novel features**:
  - Integration with Third Party RL packages
  - Customizable Scenarios
  - Out of the box entities
  - Multiple rendering options
  - GUI

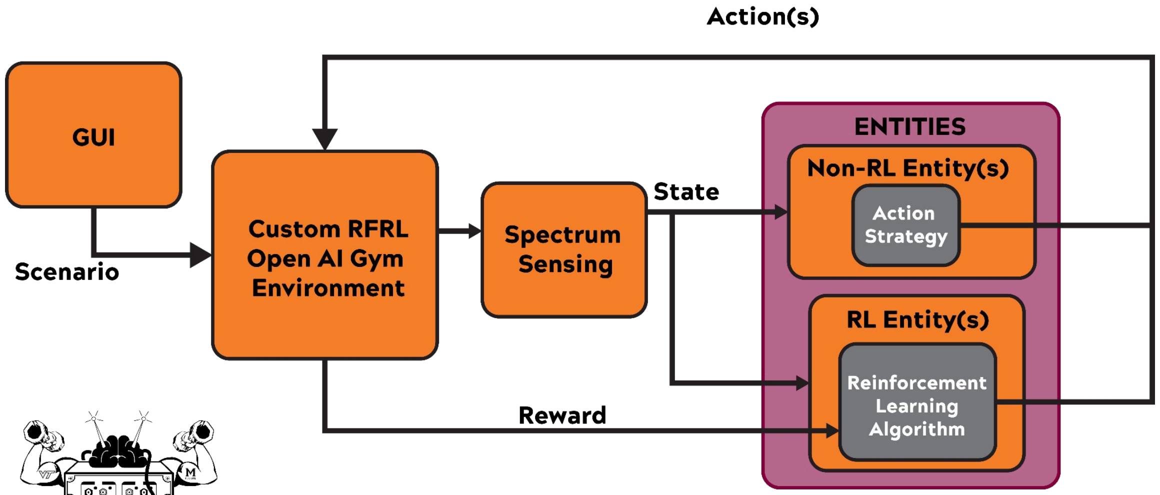
Features	RFRL Gym	GrGym [14], [15]	ns3-Gym [17]
Flexible Scenario Design	✓	✓	
RL Package Compatibility	✓	✓	✓
Spectrum Sensing Capabilities	✓		✓
Multi-Agent Capabilities			✓
Ease of Use	✓		
Graphical User Interface	✓		
Hardware Compatible		✓	

The RFRL Gym Framework

The Development of a Custom OpenAI Gym Environment for Cognitive Radio Applications

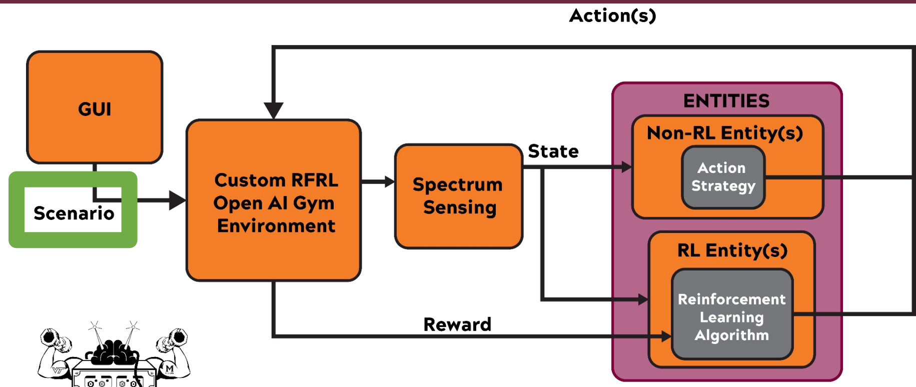
Why OpenAI Gymnasium?

- OpenAI Gymnasium Integration
  - Standardized API
  - State space
  - Action space
- Easy integration with external RL packages
  - Stable Baselines, Hugging Face, Ray RLLib, etc.
  - Pytorch, Tensorflow, etc.



Scenario: Configuration of the simulation

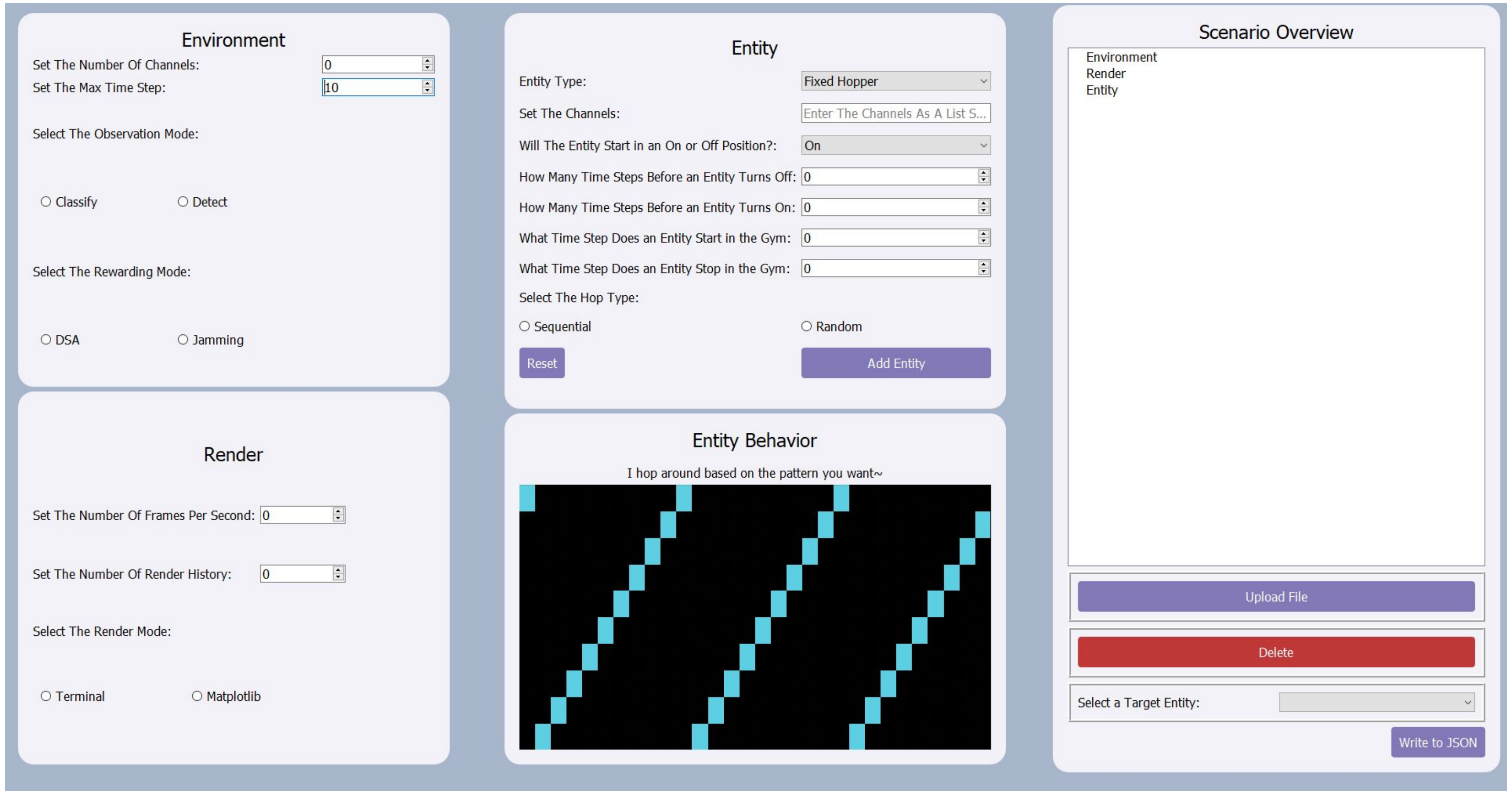
- Entities: Signals moving in spectrum
- Observation modes: Classification and Detection
- Reward modes: Jamming or DSA
- Render Modes: Terminal or PyQt



$$DSA\ Reward = \begin{cases} 1 & no\ collision \\ -1 & collision \end{cases}$$

$$Jamming\ Reward = \begin{cases} 1 & collision \\ -1 & no\ collision \end{cases}$$

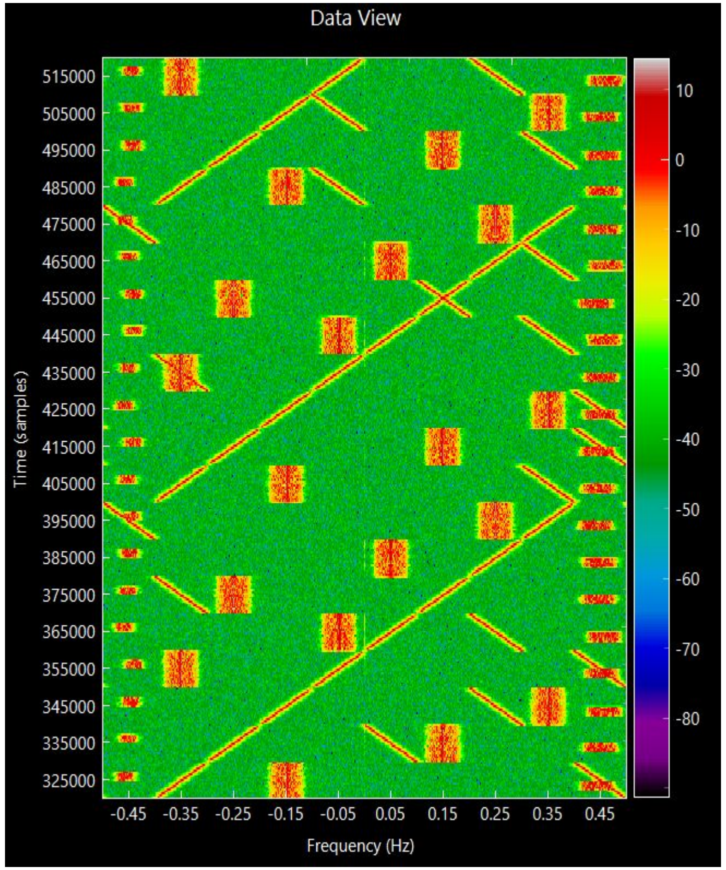
Scenario File Generation




```json
{
    "environment":
    {
        "num_channels": 10,
        "max_steps": 100,
        "observation_mode": "detect",
        "reward_mode": "jam",
        "target_entity": "constant_freq_1"
    },
    "entities":
    {
        "constant_freq_1":
        {
            "type": "ConstantFreq",
            "channels": [5],
            "onoff": [1,1,0],
            "modem_params":
            {
                "type": "qam",
                "order": 16,
                "filter": "RRC",
                "center_frequency": [-0.1,0.1],
                "bandwidth": 0.25,
                "start": 0.25,
                "duration": 0.25
            }
        }
    },
    "render":
    {
        "render_mode": "pyqt",
        "render_fps": 2,
        "render_history": 20,
        "render_background": "black"
    }
}
```

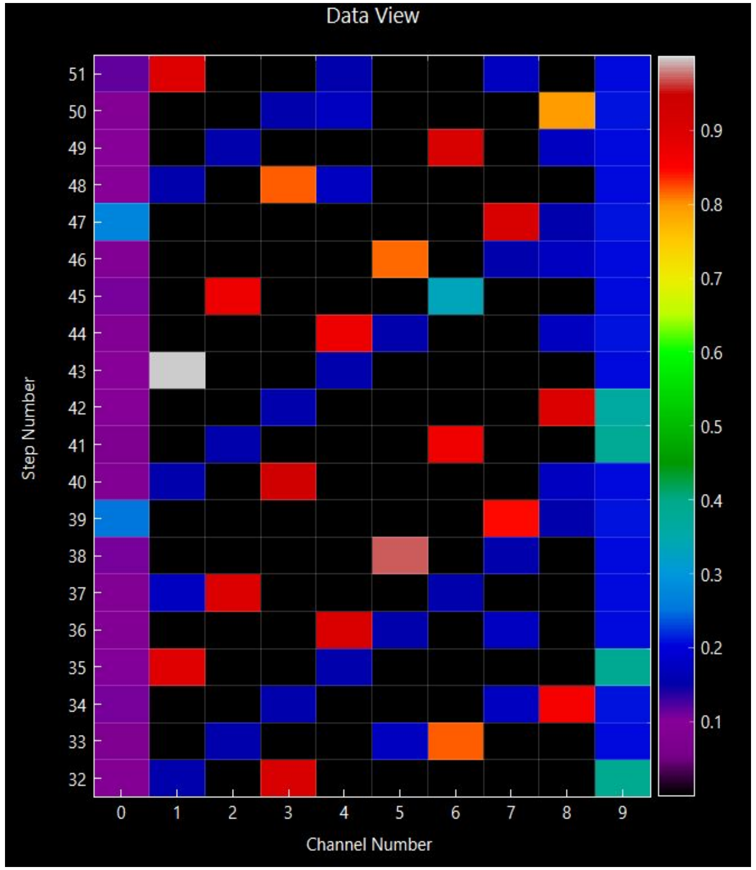
Translating the Wireless Spectrum into the RFRL Gym

Spectrum View



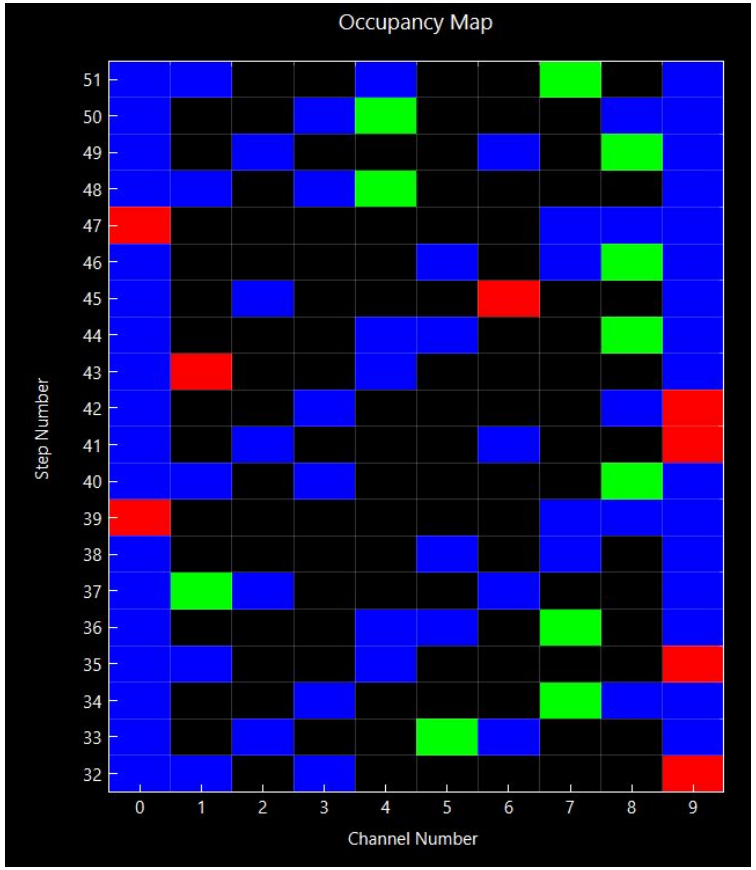
$$X(f,\tau) = \frac{1}{N}\sum_{t=1}^{N} x(t)w(t-\tau)e^{-2\pi jft/N}$$

Sensing View

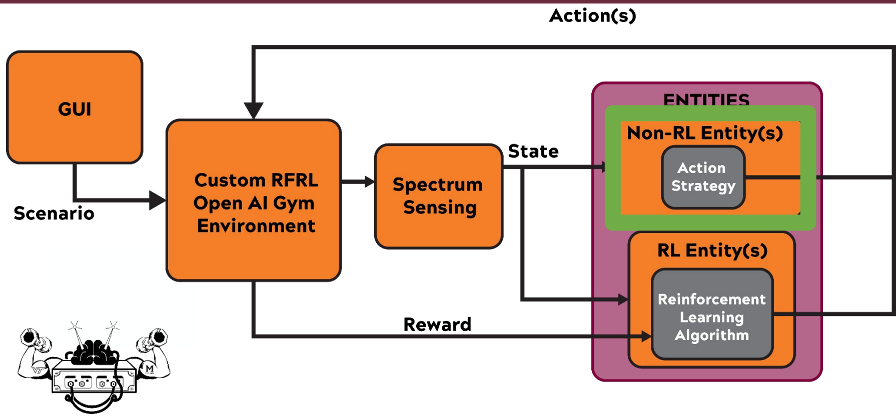


$$E[x_1] = \frac{1}{K}\sum_{i=1}^{K} |x_1[i]|^2$$

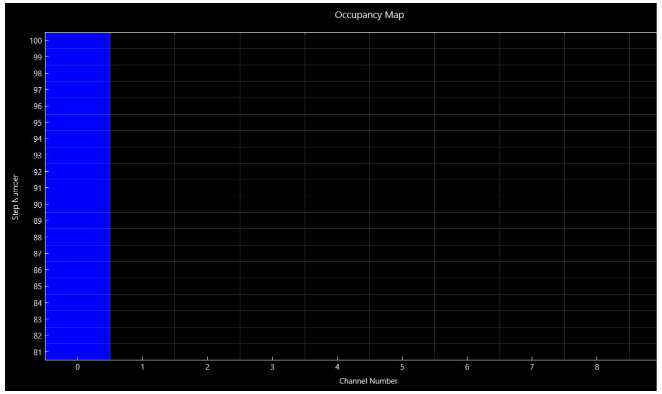
Occupancy Map

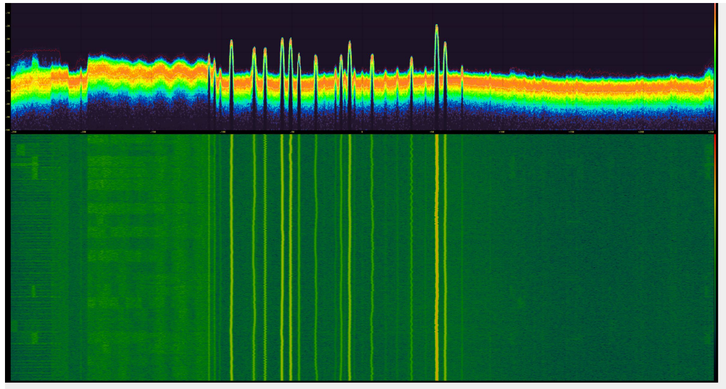


Entities: Non RL Control

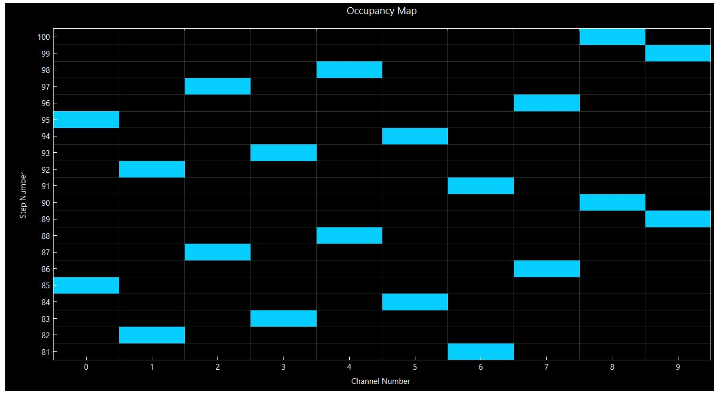


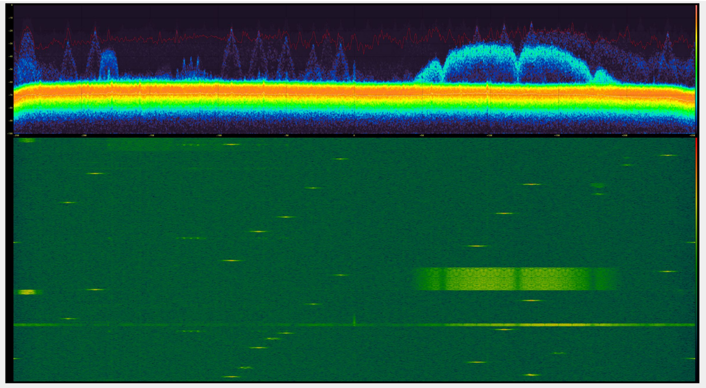
Constant Frequency Entity: Stays in one channel





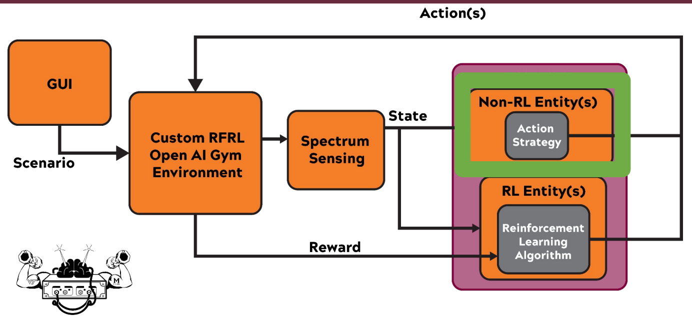
Hopper Entity: Move around with user-defined pattern



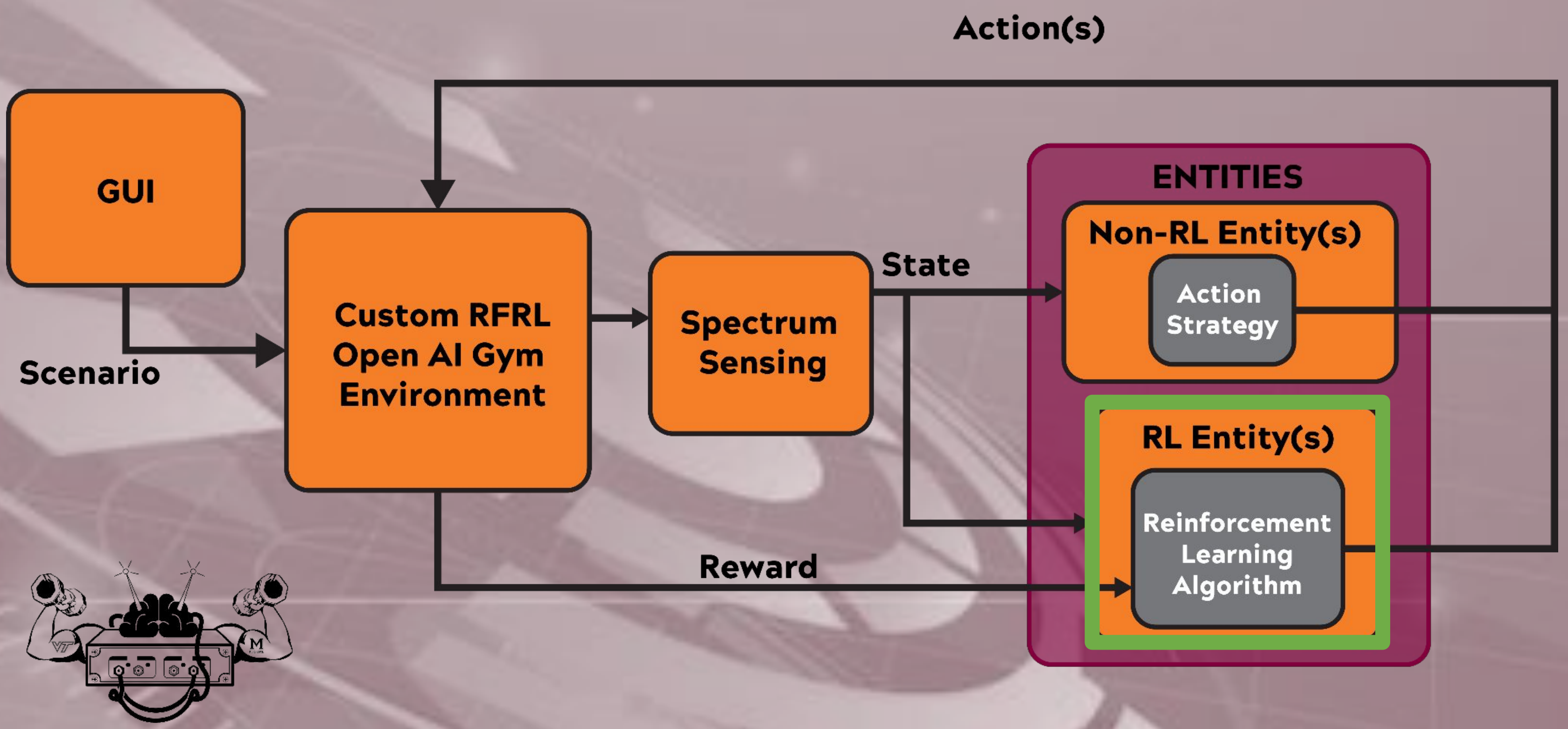


All Entities

Entity Type	Description	Real World Example
Constant	Remains in one channel	T.V. or Radio
Fixed Hopper	Moves in user-defined pattern	Bluetooth
Stochastic Hopper	Selects channel based on user-defined probabilities for each channel	Multi-Armed Bandit
Agile Hopper	Moves to empty channel if available. Has perfect knowledge of other entities in environment	CR with DSA
Simple Jammer	Moves to an occupied channel. Has perfect knowledge of other entities in environment	Intelligent Adversary
Custom	**Up to the user!**	



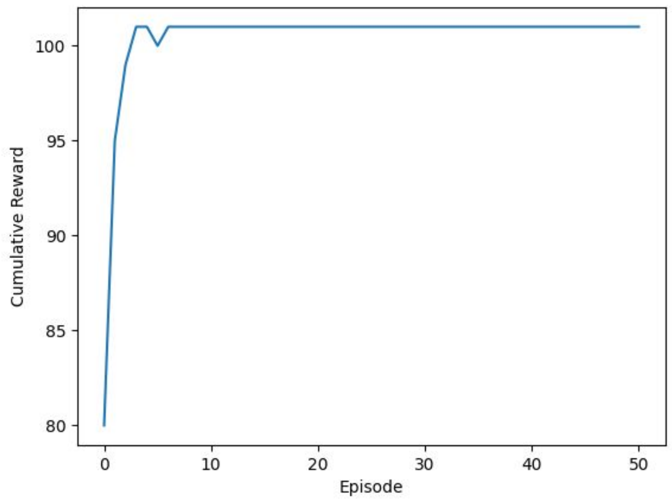
Validation and Analysis



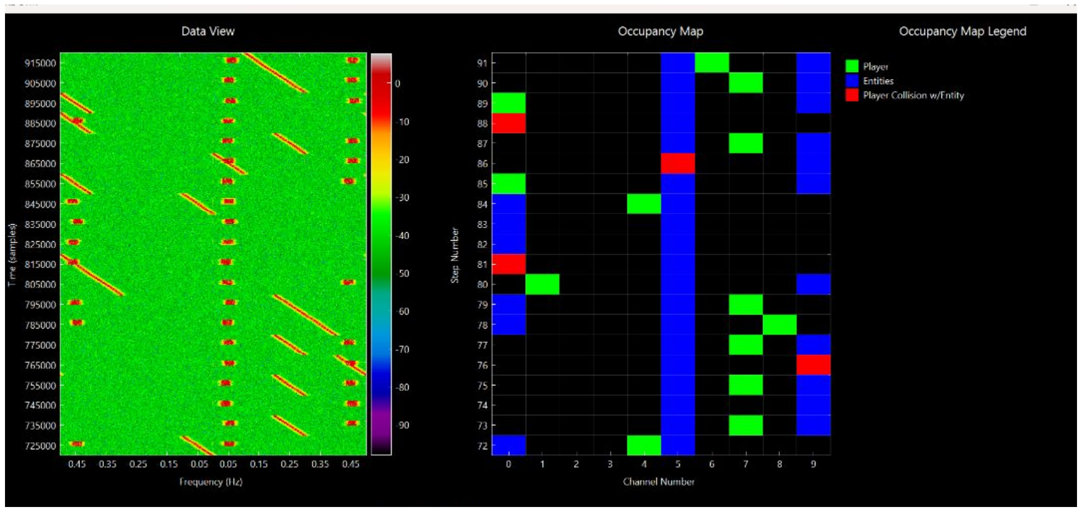
Scenario: Solvable Multi-Entity DSA

Scenario: 1 Constant Freq, 1 Stochastic Hopper

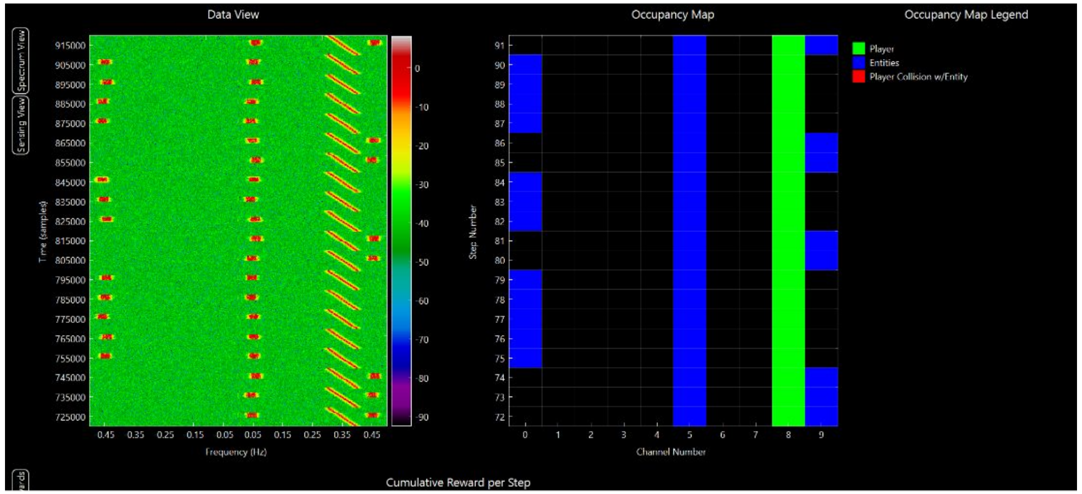
$$DSA\ Reward = \begin{cases} 1 & no\ collision \\ -1 & collision \end{cases}$$



Before Convergence



After Convergence

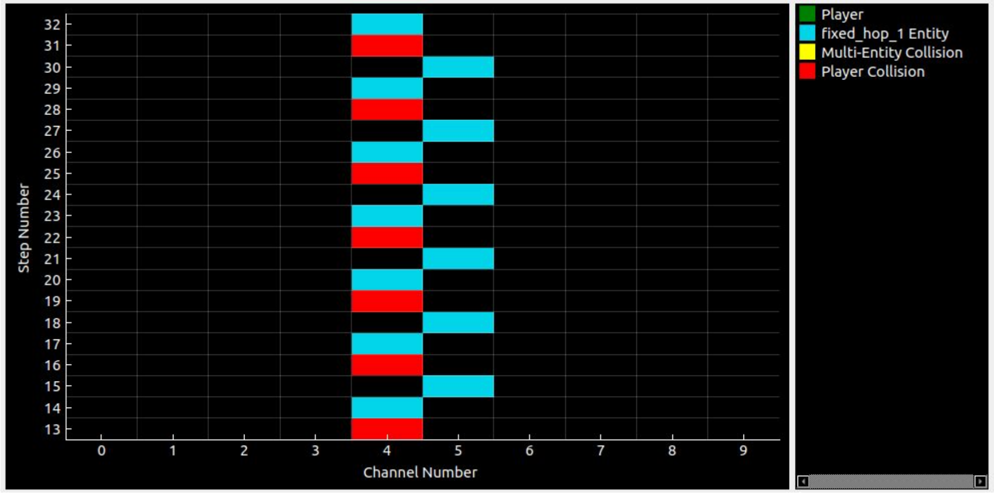


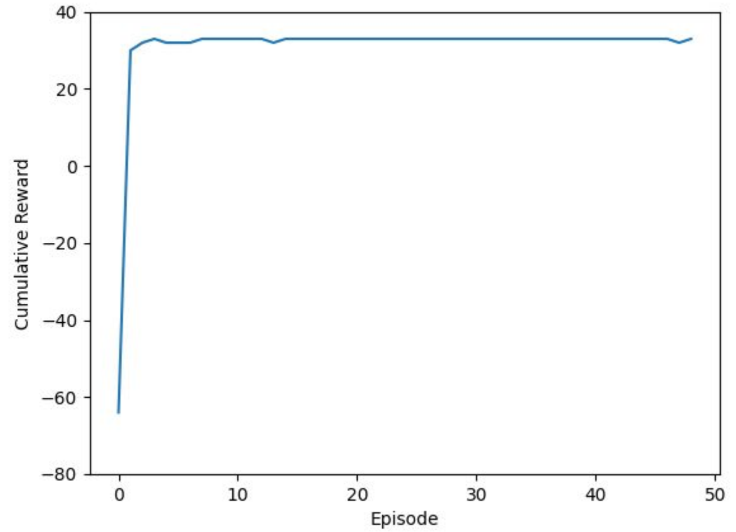
Refer to scenario 3 in the section IV of the paper

The RFRL Gym shows expected results

Findings from Challenging Scenarios

Scenario: Non-Markovian Jamming

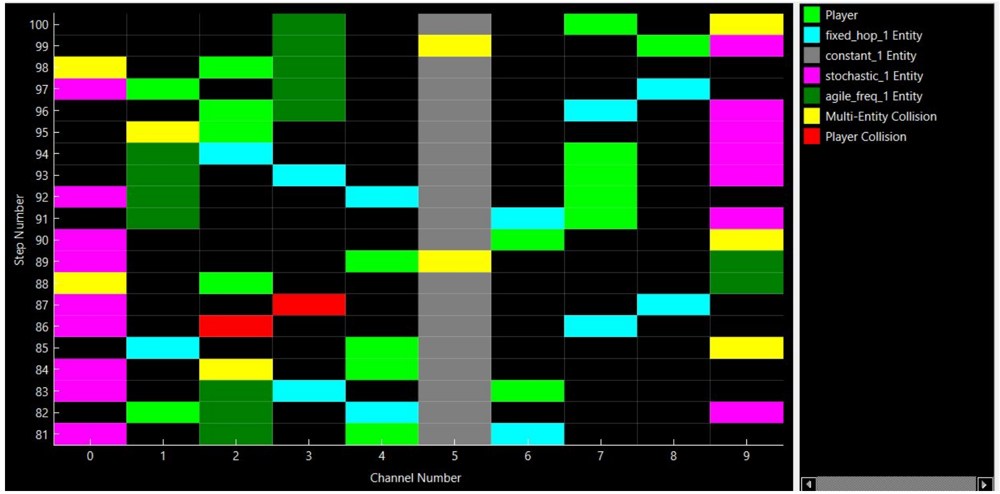




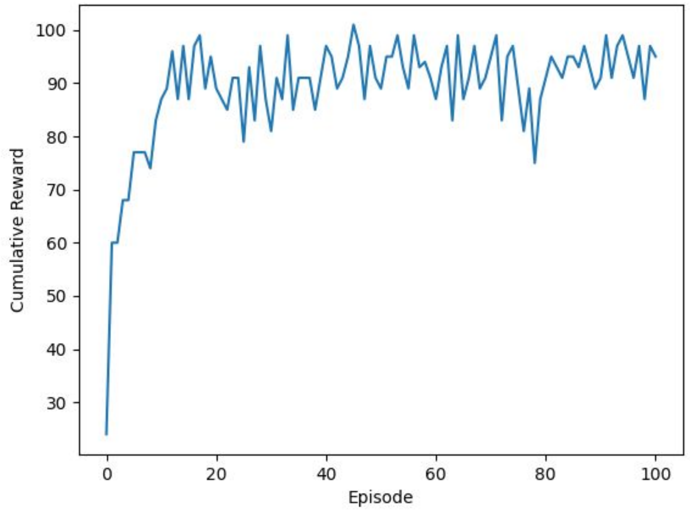
Suboptimal Convergence

$$100\left(\frac{1}{3}(1)+\frac{2}{3}(0)\right)=33$$

Scenario: Non-Deterministic Entity DSA



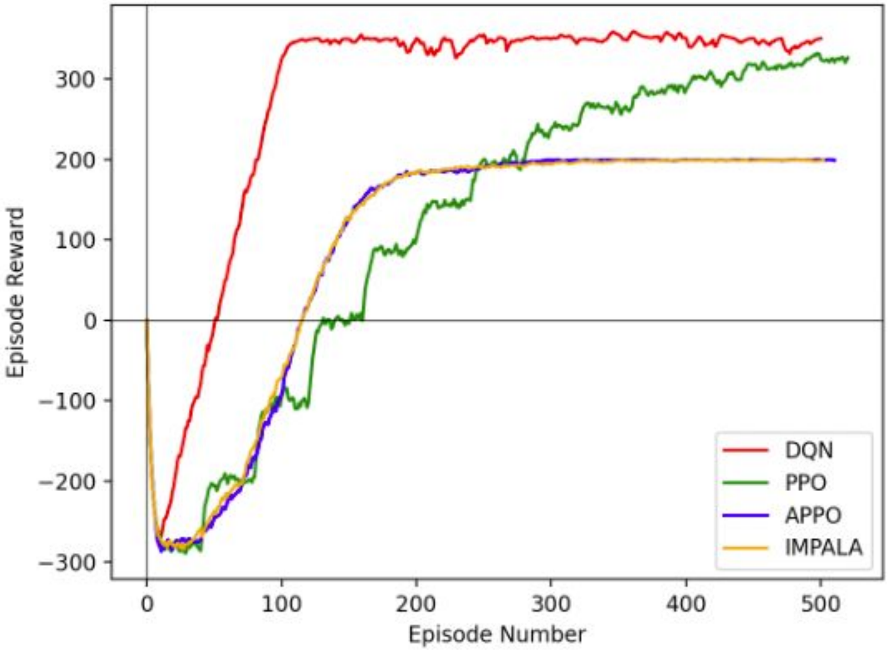
Noisy Rewards

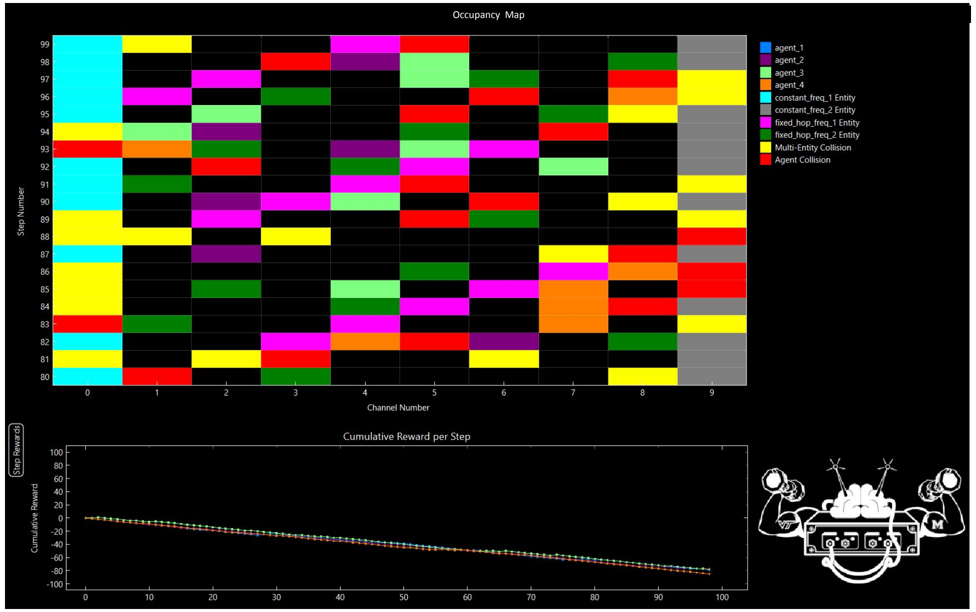


The RFRL Gym can show the limitations of Algorithms!

Future Work

- Hardware Integration
- Expansion on observation spaces: modulation classification
- GUI Expansion: end to end user interaction
- Multi-Agent Reinforcement Learning (Paper in Progress)





Codebase: **https://github.com/vtnsiSDD/rfrl-gym**

daniel20@vt.edu, alysemjones@vt.edu, cheadley@vt.edu

Thank You

Questions?

References

[1]“RF Solutions,” www.maxar.com. https://maxar.com/products/rf-solutions (accessed Dec. 13, 2023).

[2]“Summary of ctia’s annual wireless industry survey,” https://api.ctia.org/wp-content/uploads/2022/09/Summary-of-CTIAs-Wireless-Industry-Survey-2022.pdf

[3]M. Reed, “FCC Spectrum Decision Good for App Makers,” ACT | The App Association. https://actonline.org/2014/05/15/todays-fcc-spectrum-decision-great-for-app-makers/ (accessed Dec. 13, 2023).

[4]Selvaraj, Janani. (2016). INVESTIGATIVE STUDY ON ROUTING METHODS FOR CRSN. International Journal of Latest Trends in Engineering and Technology. 10. 51-60. 10.21172/1.102.08.

[5]D. Rodrigo, Pedro Ponce Cruz, and A. Molina, “CODA Algorithm: An Immune Algorithm for Reinforcement Learning Tasks,” InTech eBooks, Sep. 2016, doi: https://doi.org/10.5772/63570.

[6]“Phase Modulation of Atomic Signal,” NIST, Accessed: Dec. 13, 2023. [Online]. Available: https://www.nist.gov/image/phase-modulationjpg-0

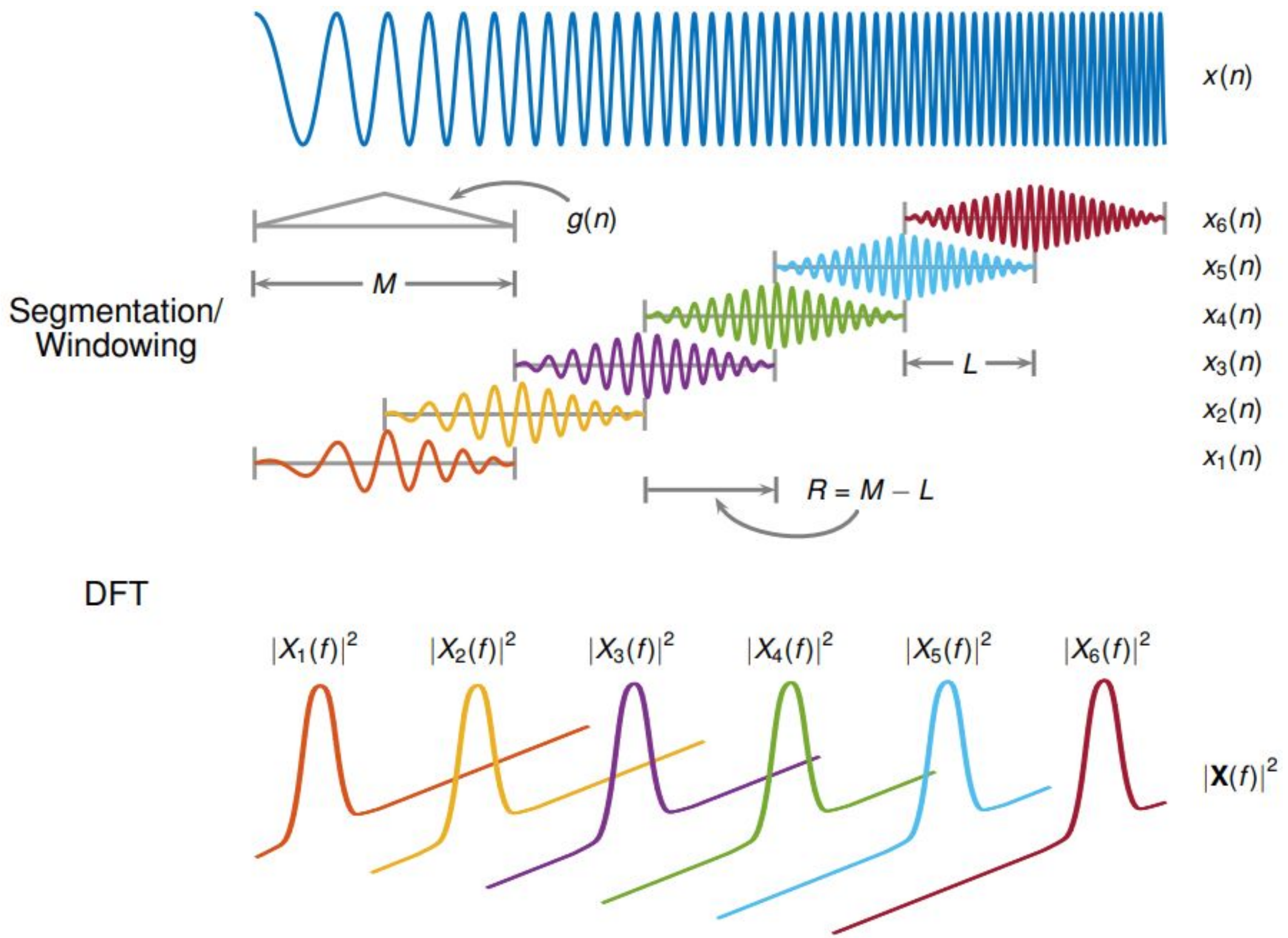
[7] Miah, Md Sipon & Yu, Heejung & Godder, Tapan & Rahman, M Mahbubur. (2015). A Cluster-Based Cooperative Spectrum Sensing in Cognitive Radio Network Using Eigenvalue Detection Technique with Superposition Approach. International Journal of Distributed Sensor Networks. 2015. 10.1155/2015/207935.

[8] P. Pandya, A. Durvesh and N. Parekh, "Energy Detection Based Spectrum Sensing for Cognitive Radio Network," *2015 Fifth International Conference on Communication Systems and Network Technologies*, Gwalior, India, 2015, pp. 201-206, doi: 10.1109/CSNT.2015.264.

Spectrogram

A series of overlapped FFTs to illustrate how the frequency domain changes over time

$$X(f,\tau) = \frac{1}{N}\sum_{t=1}^{N} x(t)w(t-\tau)e^{-2\pi jft/N}$$



Acquired from https://www.mathworks.com/help/signal/ref/stft.html

Spectrum Sensing Detector - PSD-based Detector

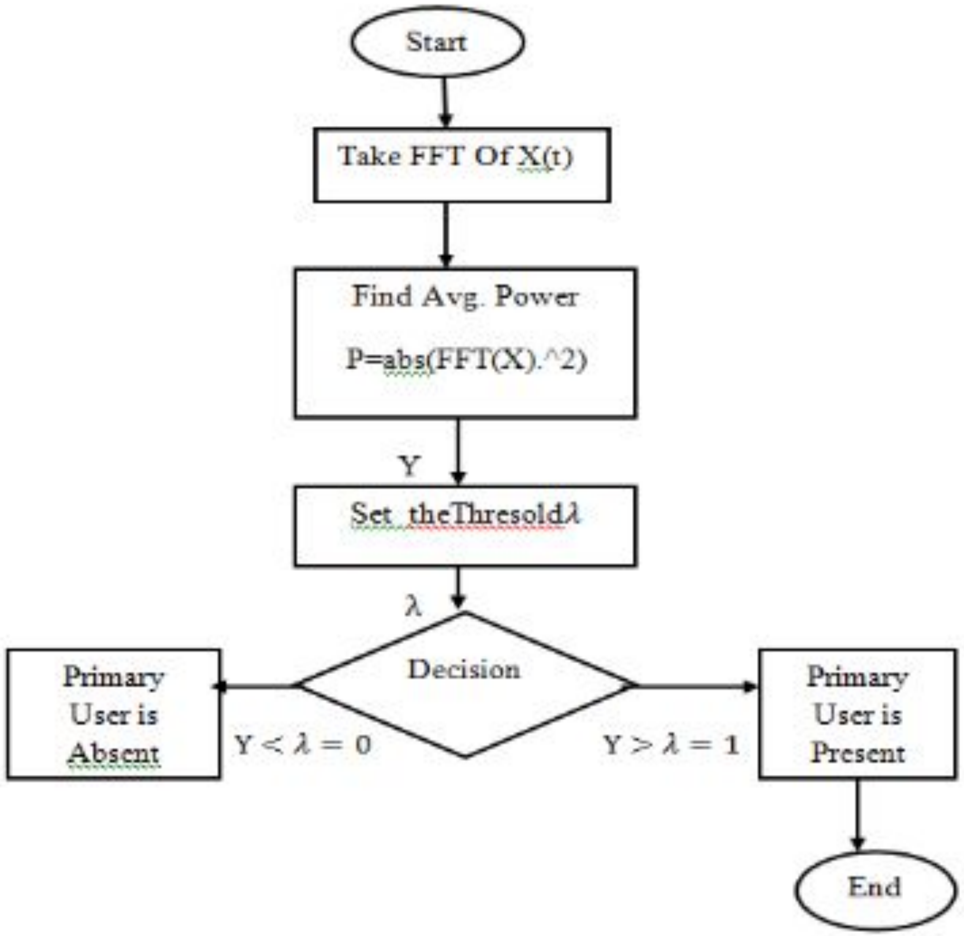
- Splits the wideband spectrum into *N* sub-channels
  - Ex: 1 MHz wide band sample rate, N = 8 filters, each sub-channel 125 kHz wide
- Each output represents the *state* of the RF spectrum
- After split into N sub-channels, the energy of one sub-channel is calculated

$$E[x_1] = \frac{1}{K}\sum_{i=1}^{K} |x_1[i]|^2$$

- Hypothesis test performed to distinguish signal energy from the noise floor

H_0	$y[n] = w[n]$	$E[x_1] < \lambda$	Unoccupied (0)
H_1	$y[n] = x[n] + w[n]$	$E[x_1] > \lambda$	Occupied (1)

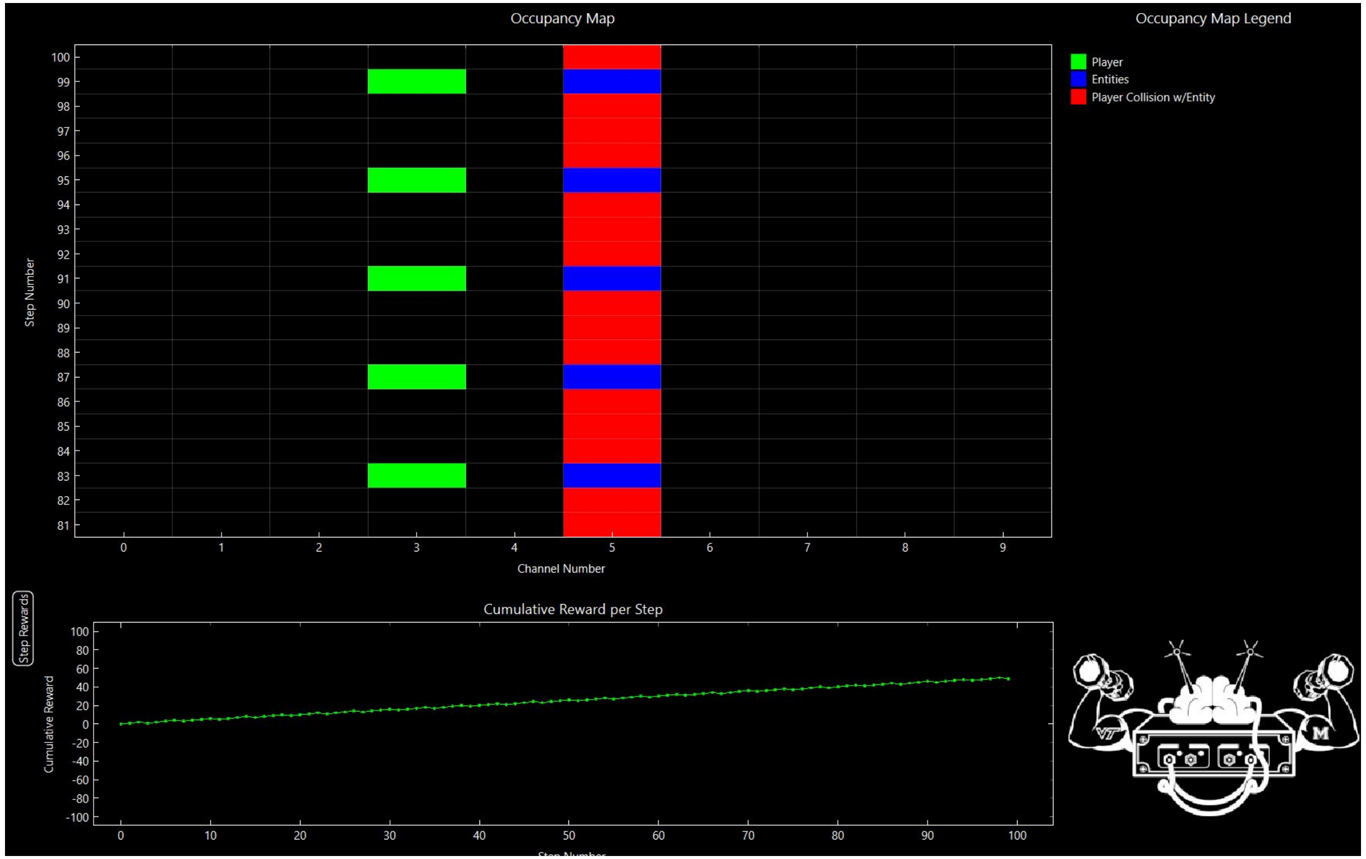
Energy Detection Algorithm



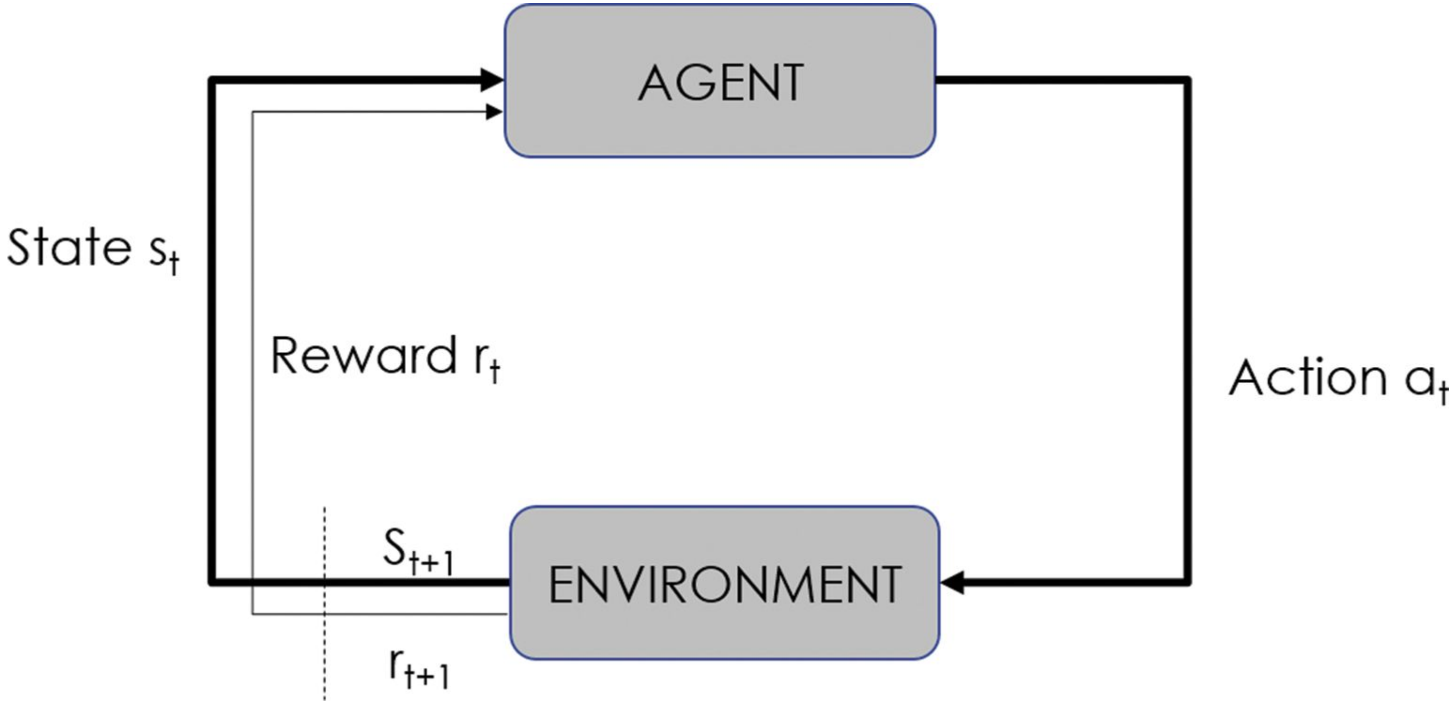
Jamming Scenario

Scenario: 1 Constant Freq

$$\text{Jamming reward} = \begin{cases} 1, & \text{collision} \\ -1, & \text{no collision} \end{cases}$$



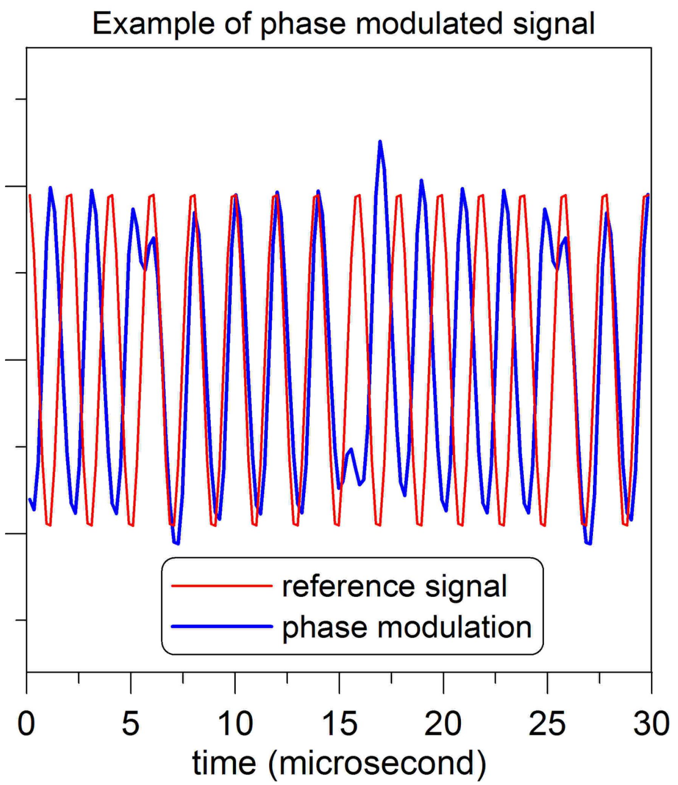
RL Cycle



Acquired from [5]

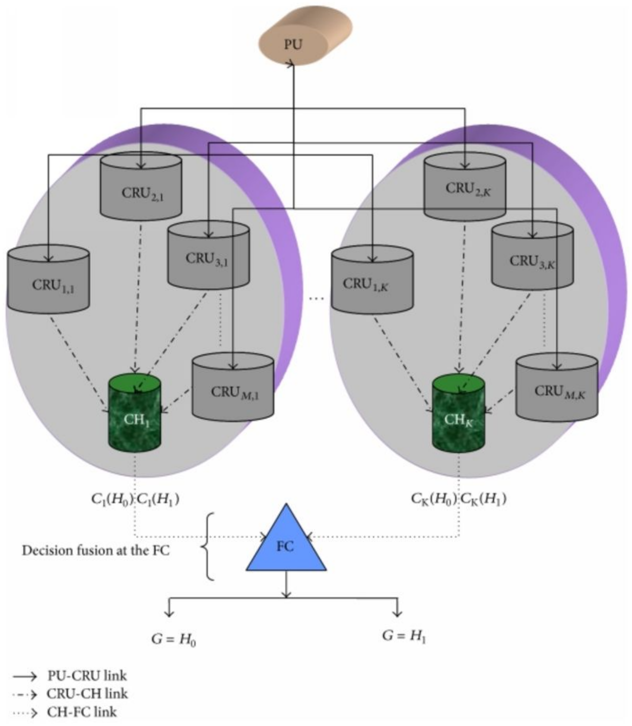
RFML Outside Spectrum Management

Signal Modulation



[6]

Collaborative Spectrum Sensing



[7]

Modulation Classification