Mass spectrometry control in metabolomics using reinforcement learning

Joe Wandy, Ross McBride, Stefan Weidt, Kevin Bryson, Justin J. J. van der Hooft, Vinny Davies, Rónán Daly

joe.wandy@glasgow.ac.uk

https://github.com/glasgowcompbio/vimms-gym

Glasgow Polyomics
www.polyomics.gla.ac.uk

1. Introduction

Liquid chromatography (LC) coupled to tandem mass spectrometry (MS/MS) is a widely used analytical technique in untargeted metabolomics. The acquisition strategy used to control the MS instrument plays an important role in determining the coverage of fragmented molecules.

2. Data-Dependent Acquisition

In LC-MS/MS, chemicals elute through the LC column before being fragmented in tandem MS. Data-dependent acquisition (DDA) is a popular control scheme in this setup, comprising a sequence of scans:

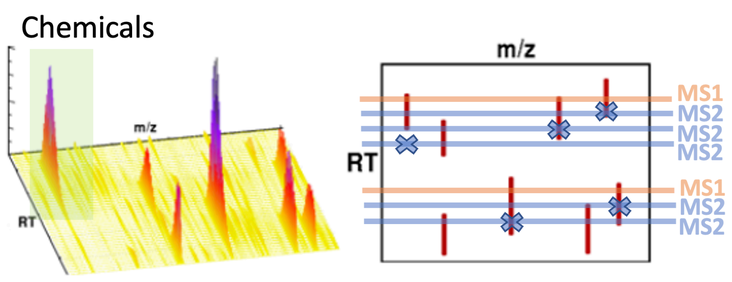

Figure 1: Chemicals elute through the chromatographic column (left) and are subjected to a series of MS1 and MS2 scans (right) using tandem mass spectrometry.

MS1 scans: explore the environment by determining the mass-to-charge (m/z) ratio and abundance of ions currently entering the MS at any given retention time (RT).

MS2 scans: exploit the environment by selecting ions in the last MS1 scan for fragmentation according to pre-determined rules. e.g. select the N most intense ions ('Top-N').

3. Learning Acquisition Method

Can we learn the optimal acquisition method from data, rather than using pre-determined rules?

Building upon our existing simulator [3], we introduce ViMMS-Gym, an Open-AI Gym environment to train reinforcement learning (RL) models and learn nearly-optimal DDA strategies.

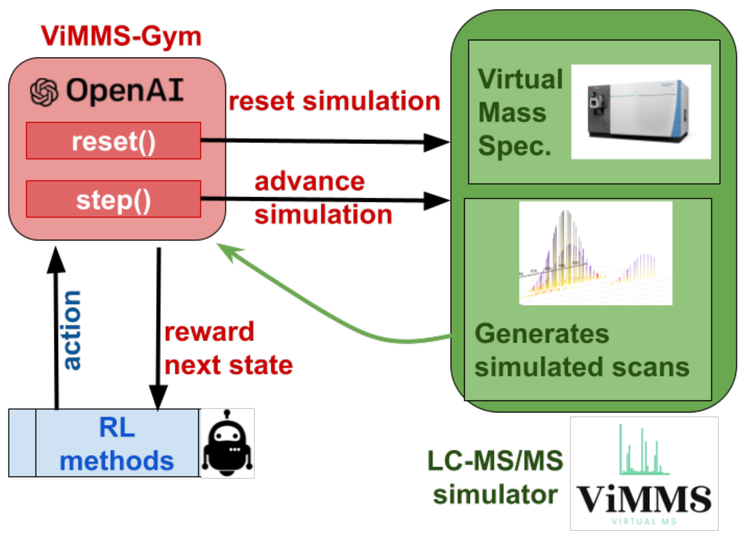

Figure 2: ViMMS-Gym implements Open-AI Gym interface and is compatible with a wide range of RL libraries.

ViMMS-Gym implements DDA as a Markov Decision Process (MDP), where at any given time t:

- The state s_t consists of features extracted from the K most intense ions in the last MS1 scan.
- The action a_t is a discrete variable that determines whether to perform an MS1 scan, or to select the k-th ion for MS2 fragmentation next.
- The reward r_t from taking action a_t is (1) a constant reward for choosing an MS1 scan, or (2) the proportion of fragmented chemical intensity from its maximum, when choosing an MS2 scan.

4. Results

- RL model was trained using Proximal Policy Optimisation (PPO) [2] with simulated chemicals from ViMMS-Gym.
- Both the actor and critic in PPO used 2-layer dense networks with 512 nodes. Training took 6 days on a 20-core machine.

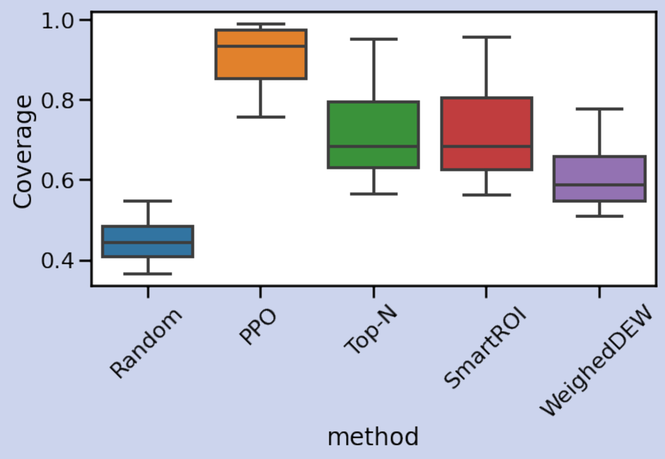

Figure 3: PPO achieves a much higher coverage compared to the conventional rule-based acquisition method (Top-N) or the advanced acquisition methods (SmartROI, WeightedDEW) in [1]

.

5. Conclusion

- Reinforcement learning can be used to learn mass spectrometry control that outperforms the baselines.
- Further research on diverse data, different network architecture, and validation on an actual instrument.

References

[1] Vinny Davies, Joe Wandy, Stefan Weidt, Justin J. J. van der Hooft, Alice Miller, Rónán Daly, and Simon Rogers. Rapid development of improved data-dependent acquisition strategies. *Analytical Chemistry*, 93(14):5676–5683, 2021.

[2] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[3] Joe Wandy, Vinny Davies, Ross McBride, Stefan Weidt, Simon Rogers, and Rónán Daly. ViMMS 2.0: A framework to develop, test and optimise fragmentation strategies in LC-MS metabolomics. *Journal of Open Source Software*, 7(71):3990, 2022.