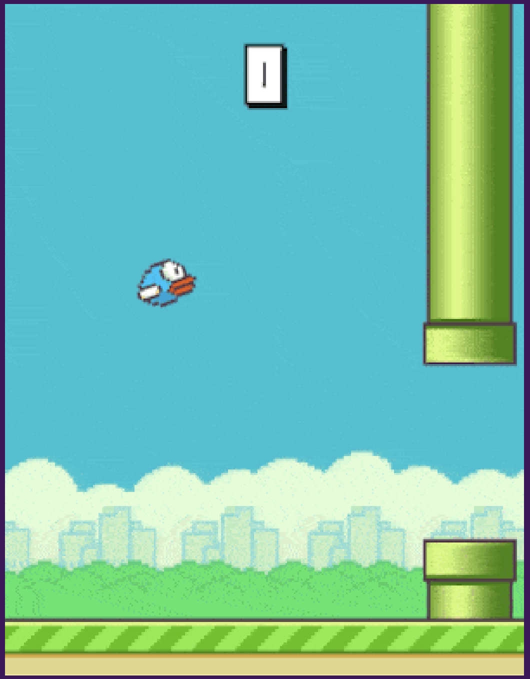
Reinforcement Learning for the Control of Microbial Communities in Bioreactors

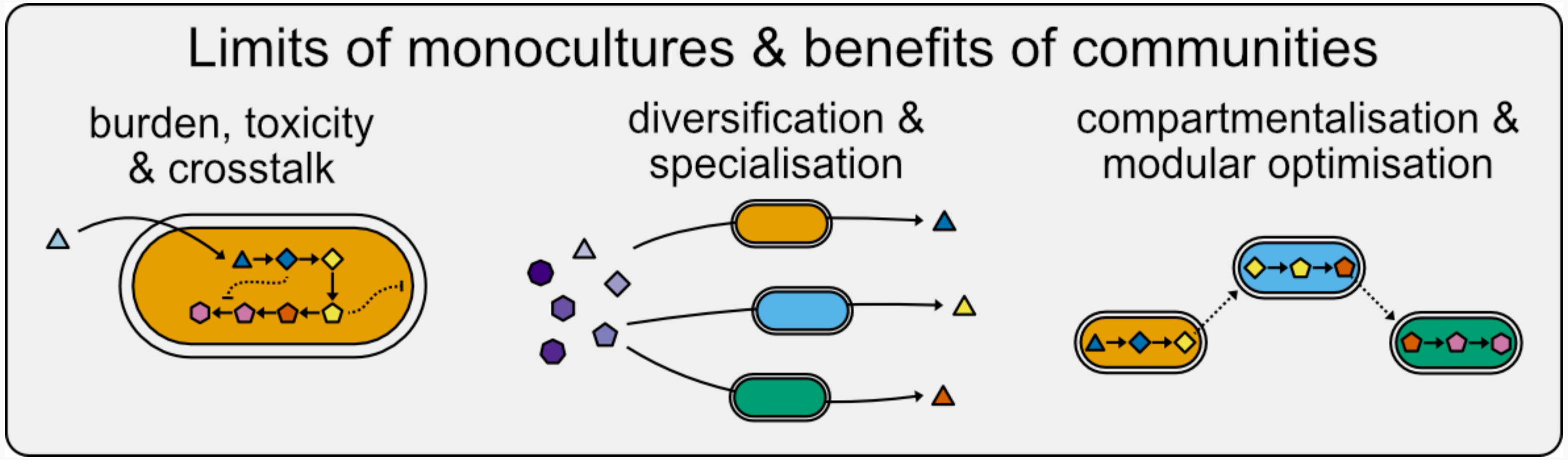
Alex J H Fedorec

30/01/2026



Division of Biosciences, Faculty of Life Sciences

Division-of-Labour in Microbial Communities



- Monocultures of microbes are metabolically limited
- Communities provide a potential solution to these limitations

Continuous Bioprocesses

- Batch processes have ephemeral, mostly sub-optimal, environmental conditions
- Continuous processes allow maintenance of optimal process conditions

Editorial: Continuous Biomanufacturing in Microbial Systems

Christoph Herwig[1,2*], Christoph Slouka[1], Peter Neubauer[3] and Frank Delvigne[4]

The Future is Continuous: Accelerating the Shift in Biomanufacturing

Cynthia A. Challener, Ph.D.
Scientific Content Director
Pharma's Almanac

Strain and process engineering toward continuous industrial fermentation

Yufei Dong[1*], Ye Zhang[1*], Dehua Liu[1,2,3], Zhen Chen (✉)[1,2,3]

Bioprocessing 4.0: a pragmatic review and future perspectives

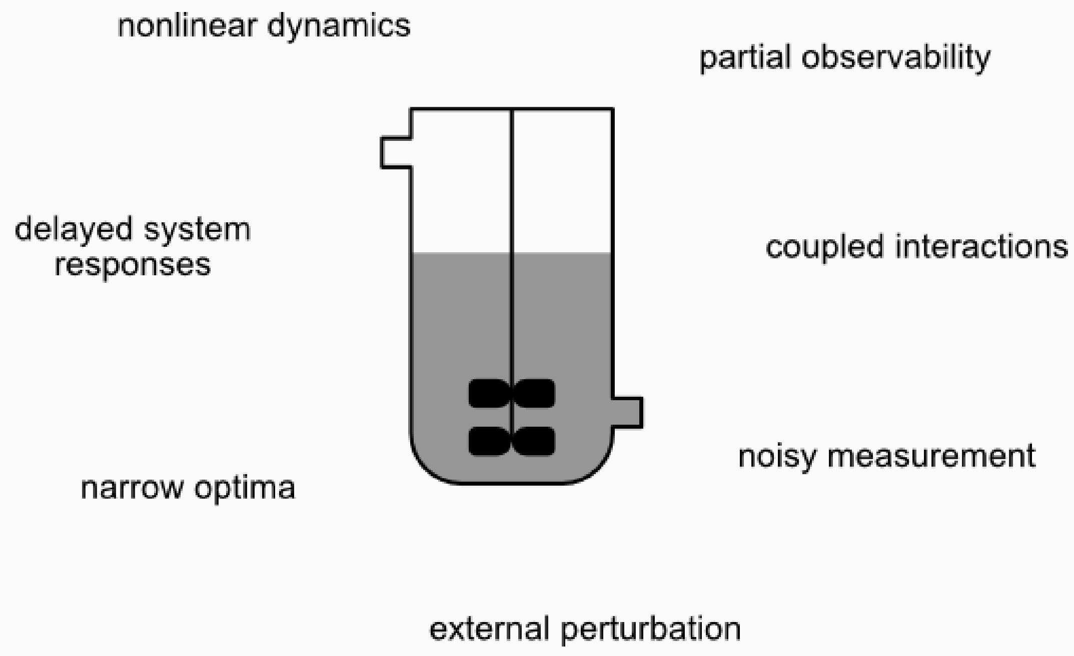
Kesler Isoko[a], Joan L. Cordiner[b], Zoltan Kis[bc] and Peyman Z. Moghadam[*a]

Current challenges and recent advances on the path towards continuous biomanufacturing

Milos Drobnjakovic, Roger Hart, Boonserm (Serm) Kulvatunyou, Nenad Ivezic, Vijay Srinivasan

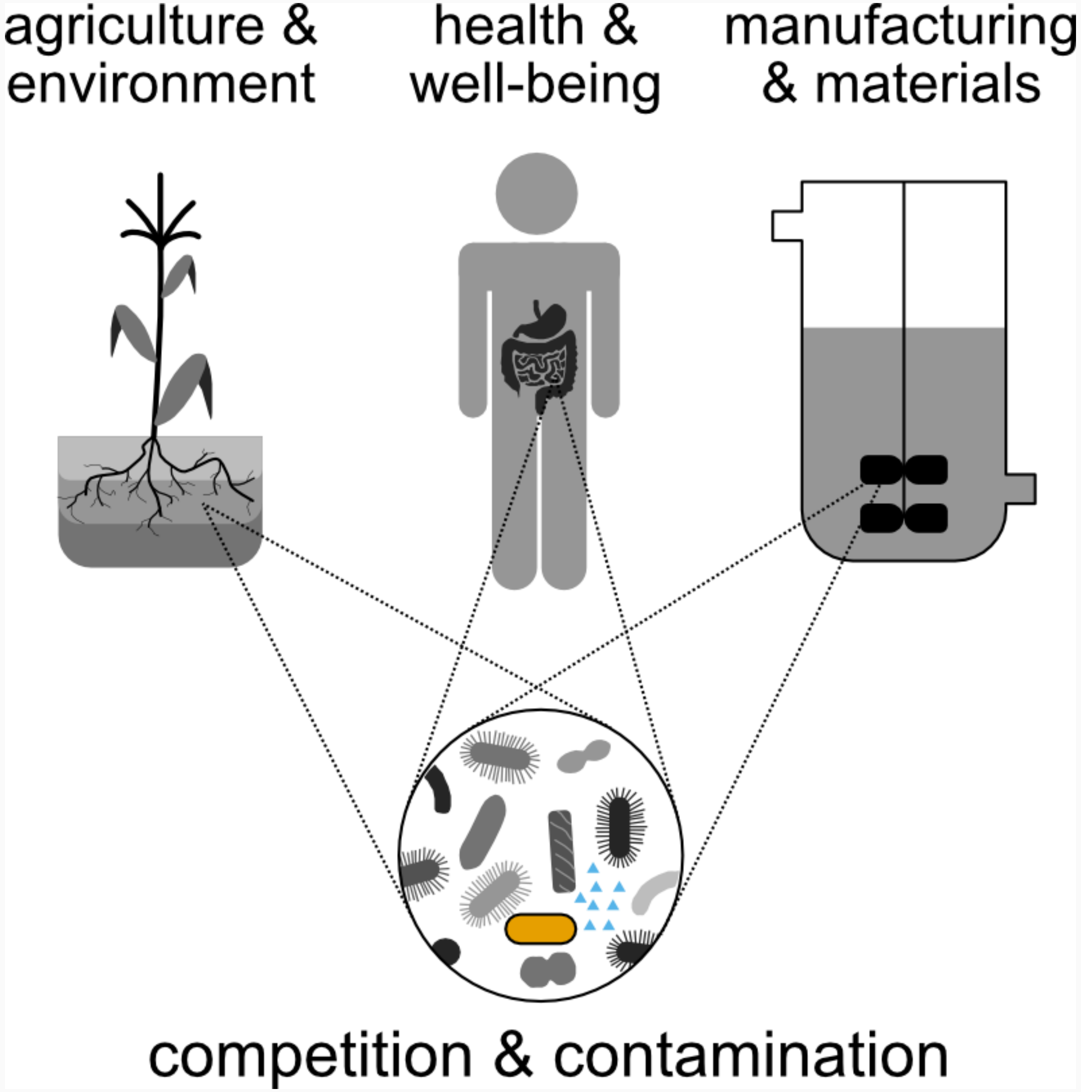
The Control Challenge

- Optimal operating point is narrow and can shift with perturbations
- Community composition drifts; competitive exclusion is a risk
- Classical controllers (PID, MPC) struggle with: nonlinearity, partial observability, coupled dynamics

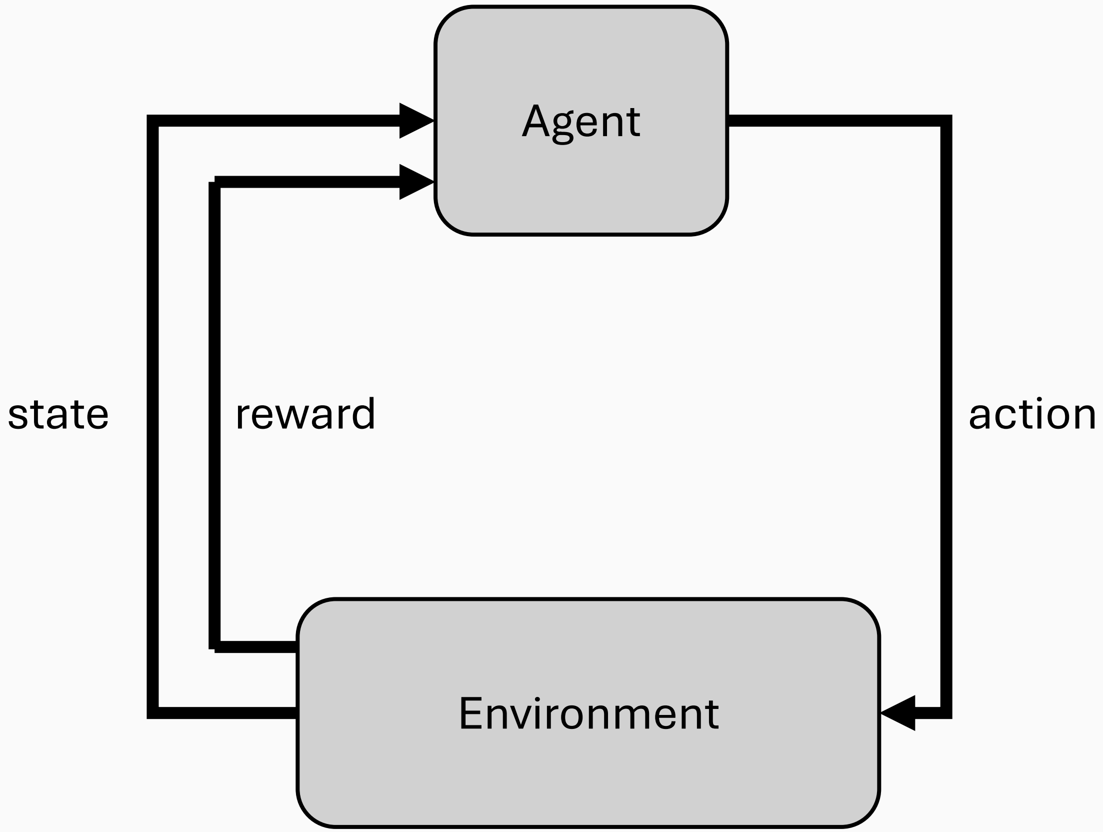


Microbes Beyond Confinement

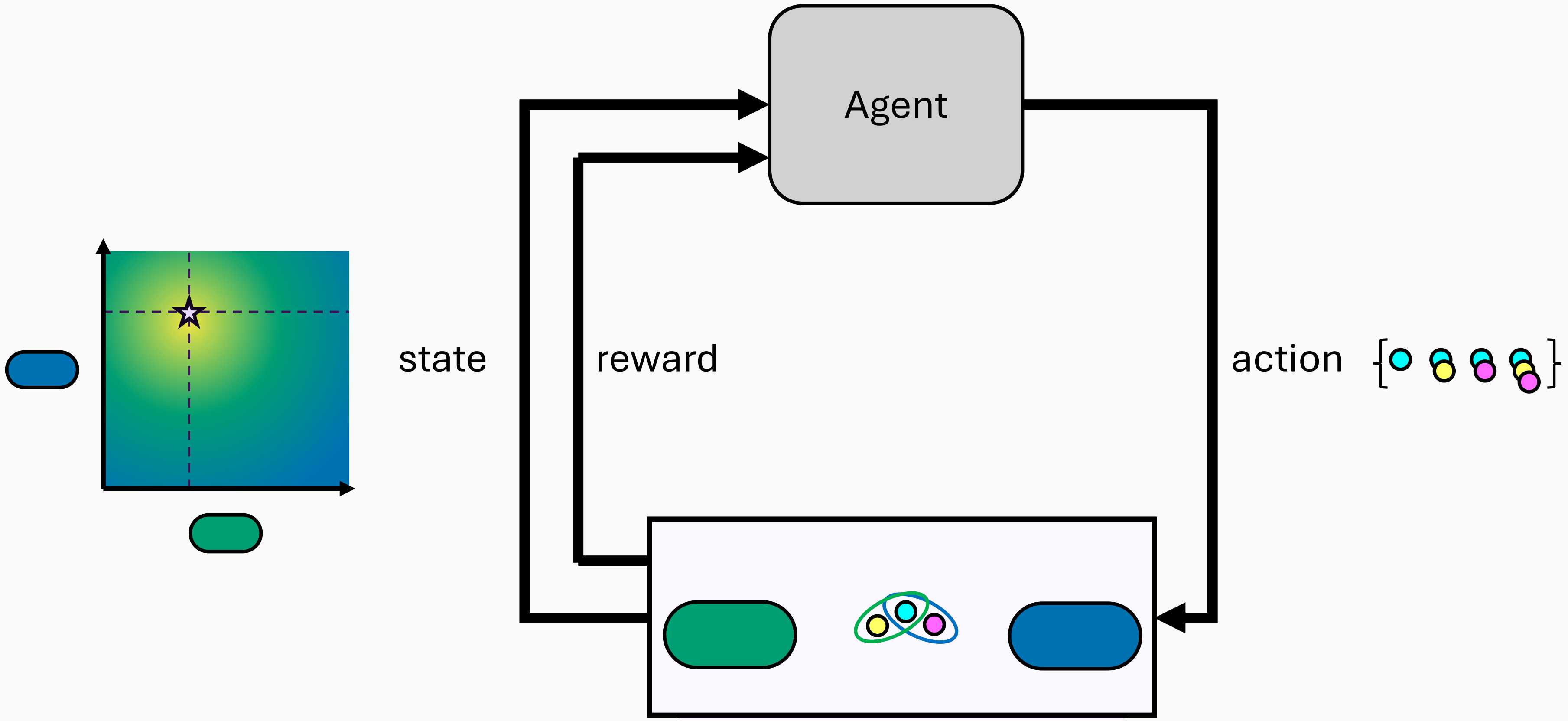
- Dysbiosis can have negative impacts beyond the lab and plant
- An ability to control arbitrary communities could have impacts in health, agriculture, manufacturing and more



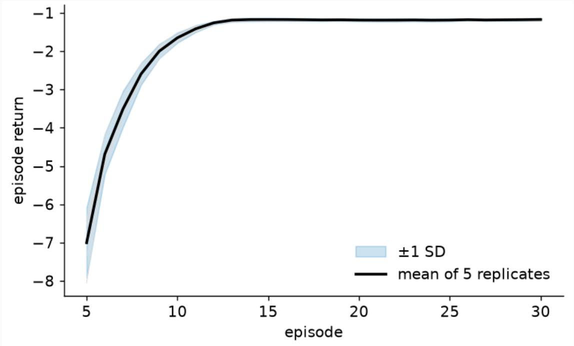
Reinforcement Learning: A Quick Introduction

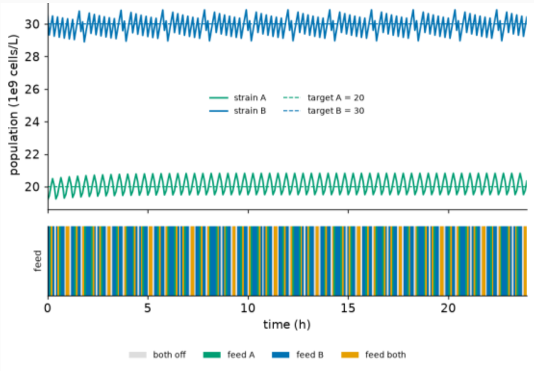


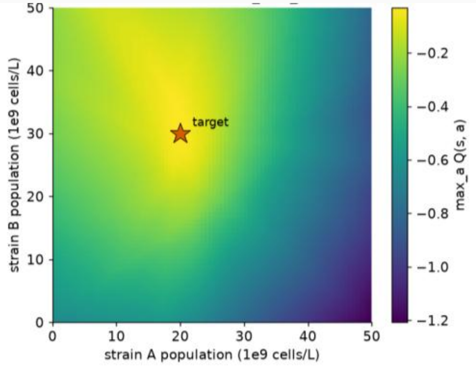
Reinforcement Learning For Bioreactor Control

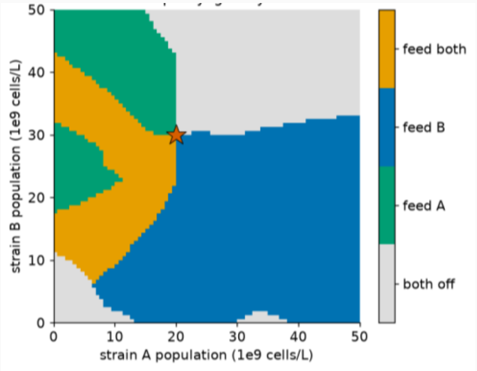


Learning to Control Two Auxotrophs



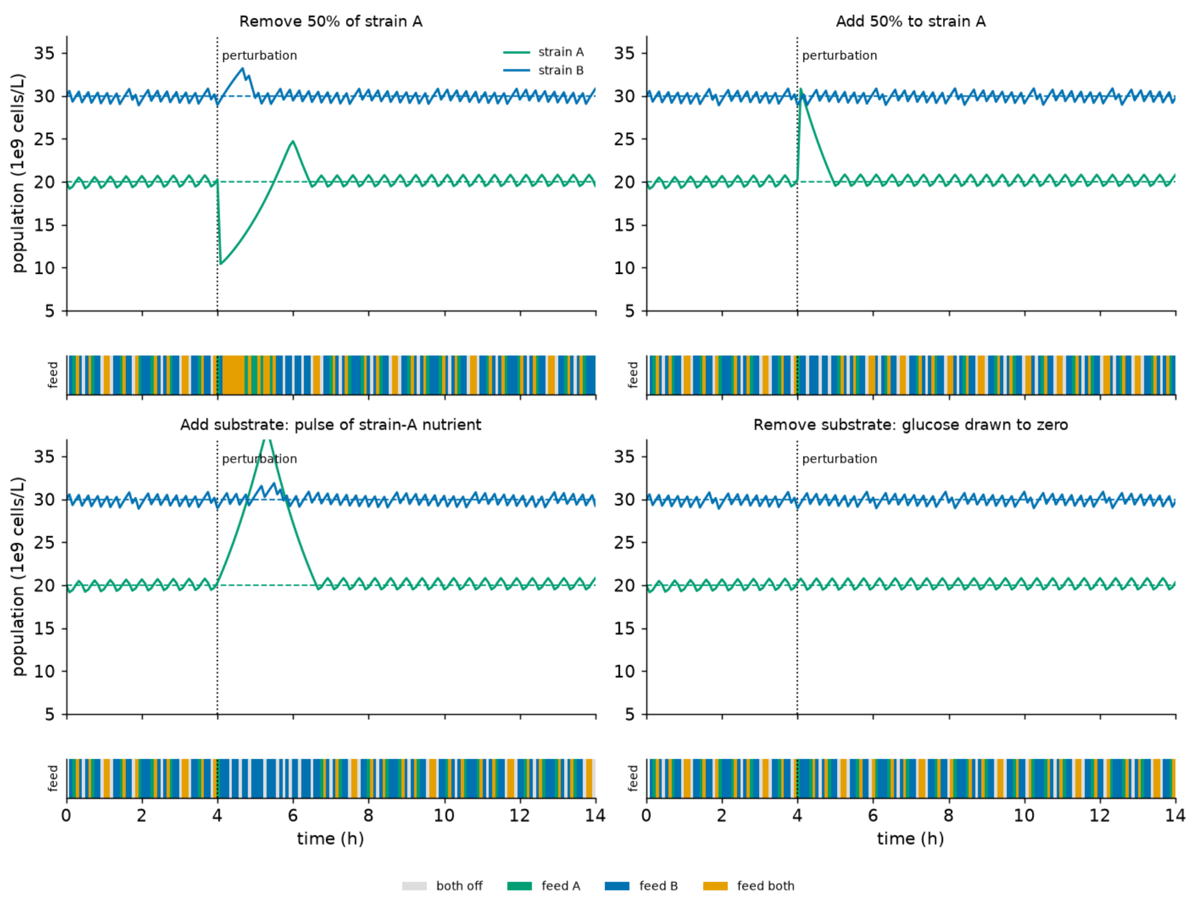






- Start with naïve exploration; exploitation as knowledge improves
- Trained agent maintains tight control
- The agent learns the "value" of different positions in state space
- The agent's policy can be interrogated; it makes intuitive sense

Recovery From Perturbation

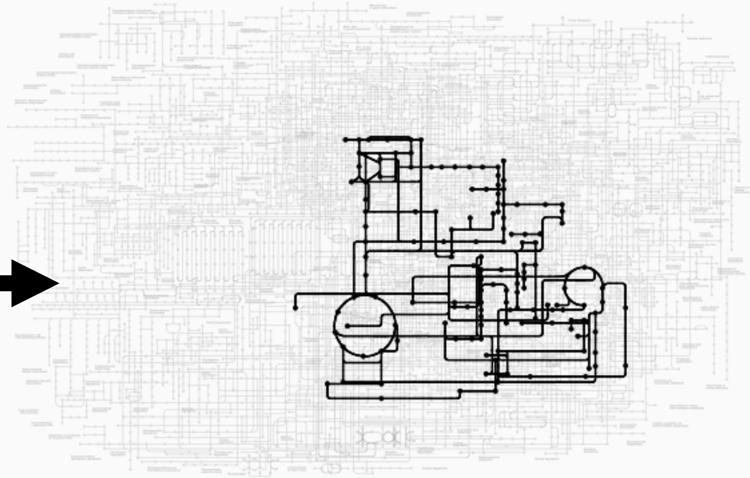


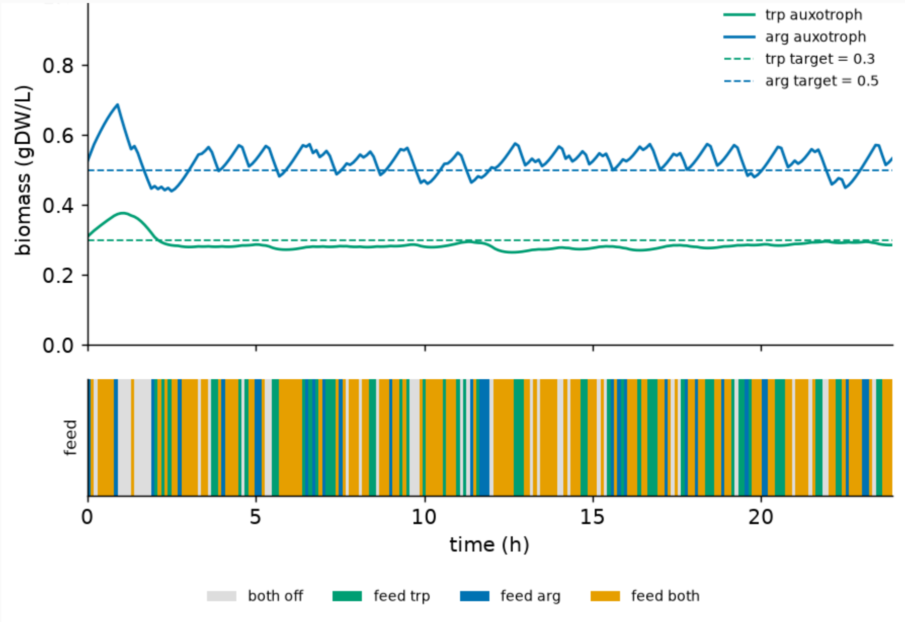
- Trained agents can respond to perturbations without further training

From “Toy” Models to Metabolic Models

$$\dot{\boldsymbol{B}} = f(\boldsymbol{B}, \boldsymbol{S}, D)$$

$$\dot{\boldsymbol{S}} = g(\boldsymbol{B}, \boldsymbol{S}, D)$$

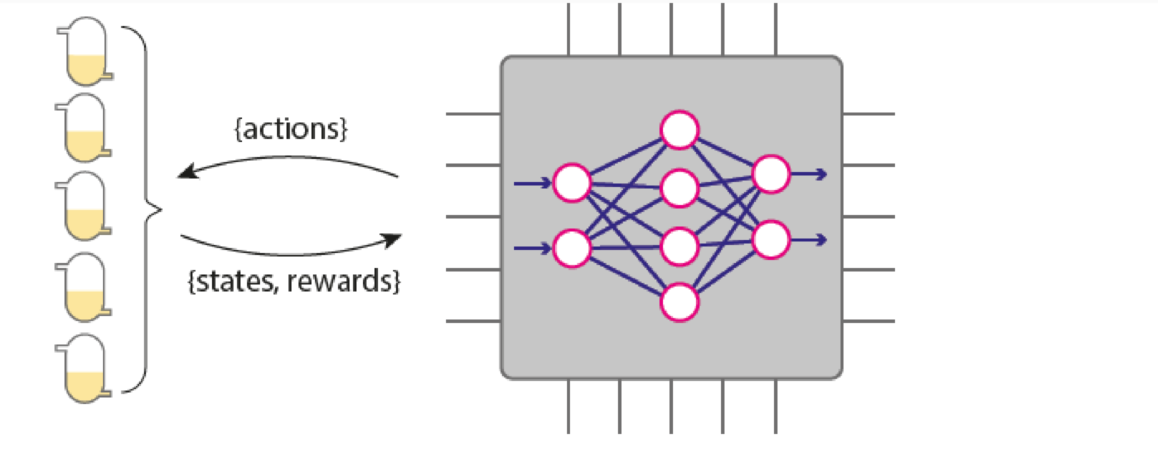




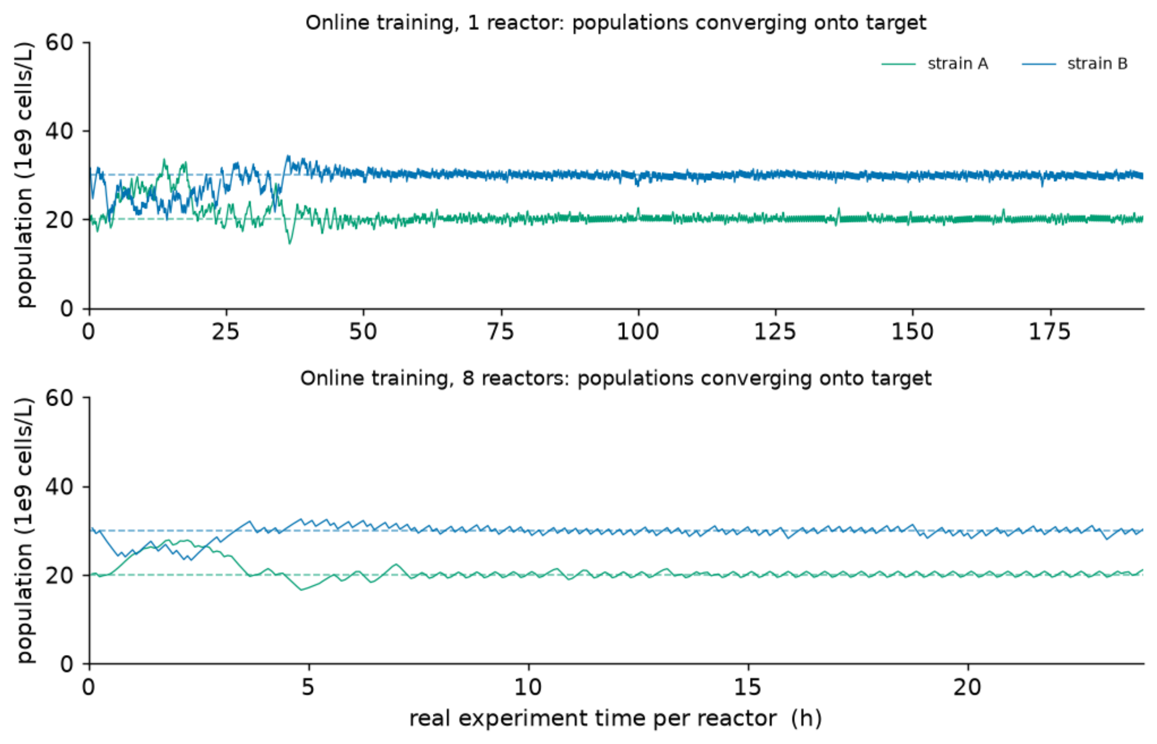
- Complexity ratchet: gLV -> ODE -> GEM -> in vitro
- Genome Scale Metabolic models capture more biological complexity than our initial ODE models

- Include mechanisms that we didn’t assume to be important
- The same RL framework can control these models; but simulation is more expensive

Learning Online in Parallel

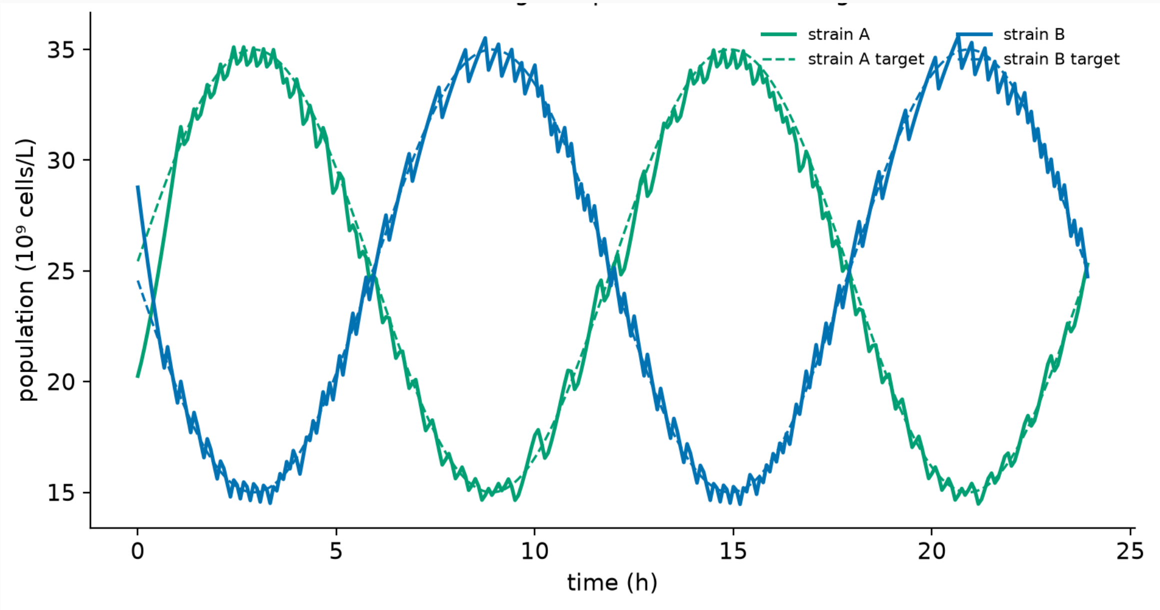


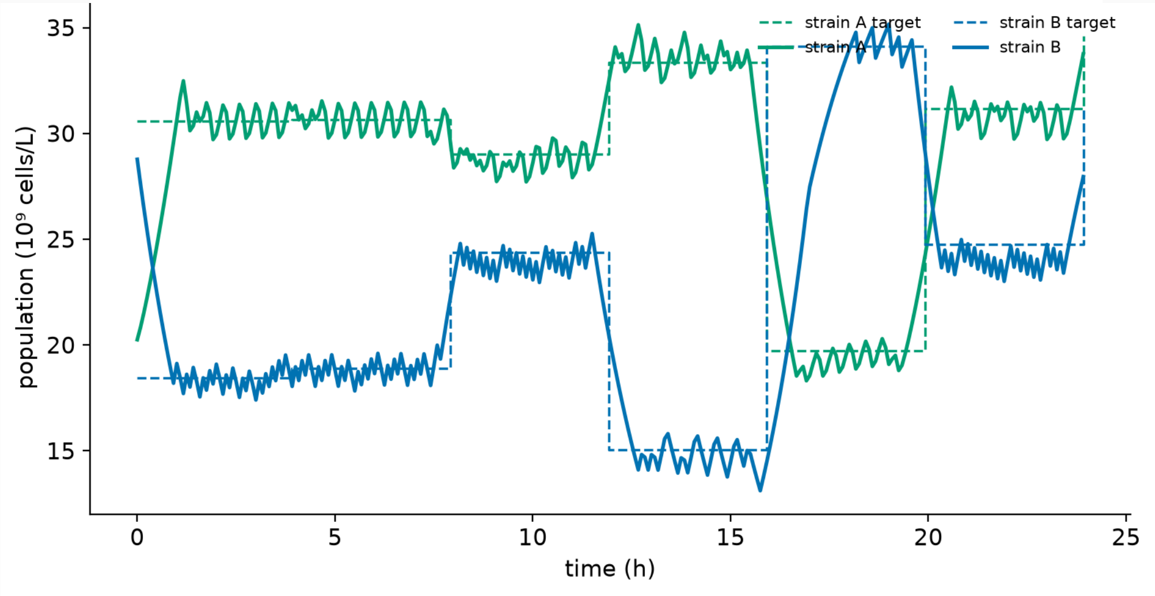
- RL is data hungry and traditionally trained sequentially
- Parallel environments allow simultaneous simulations with pooled experience



- This speed-up enables convergence in tractable wall-clock time

Objectives Don't Have to be Static

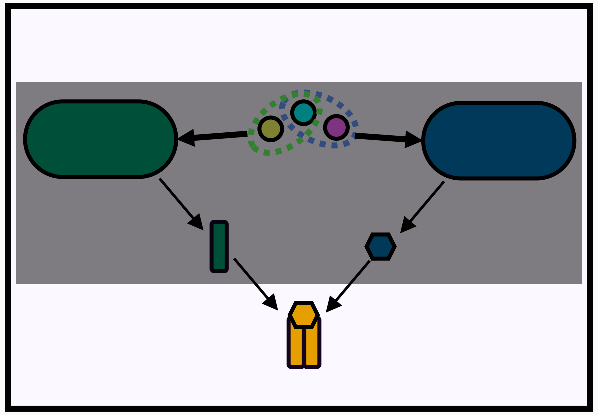




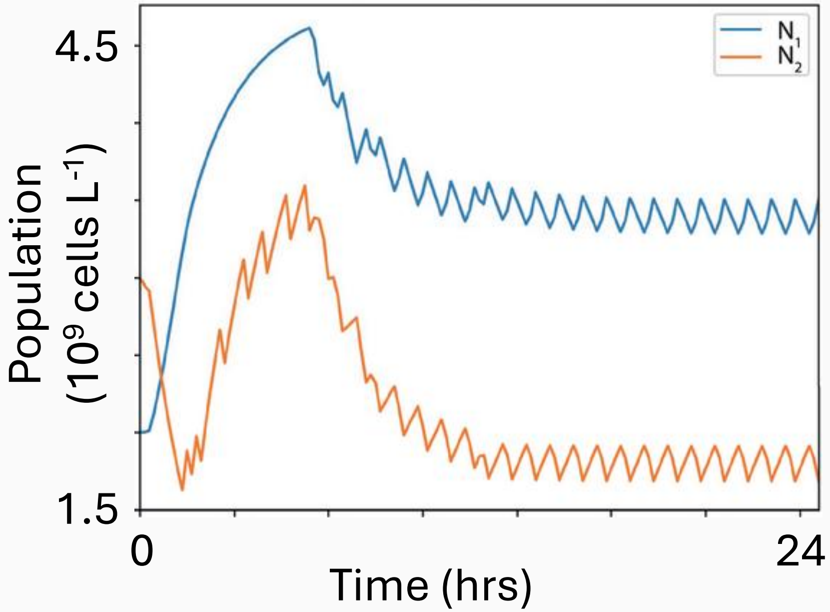
- Fixed targets may be suboptimal over a production run
- An agent can learn to anticipate and follow a smooth, periodic trajectory

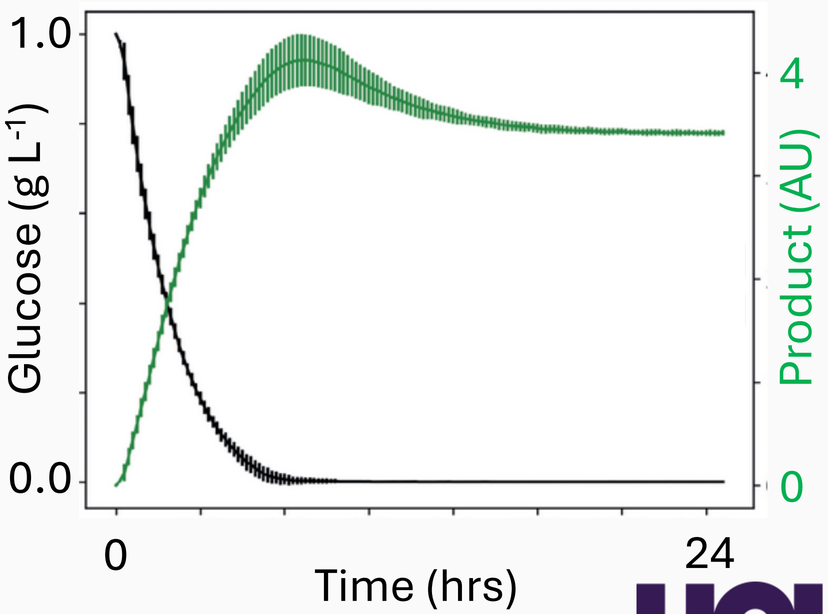
- Can also react to rapid, unpredictable changes
- Both use the same trained framework – no retraining for different types of target

Maximising Yield Without Full Process Knowledge

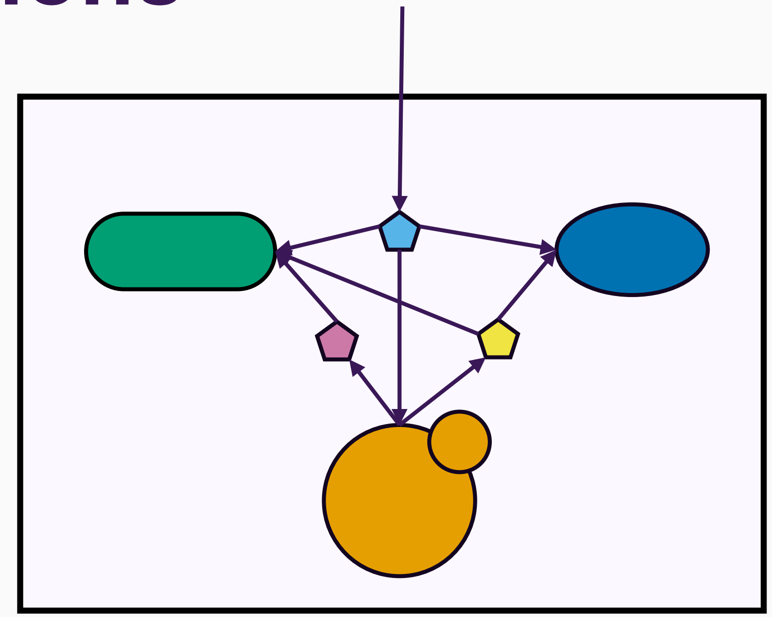


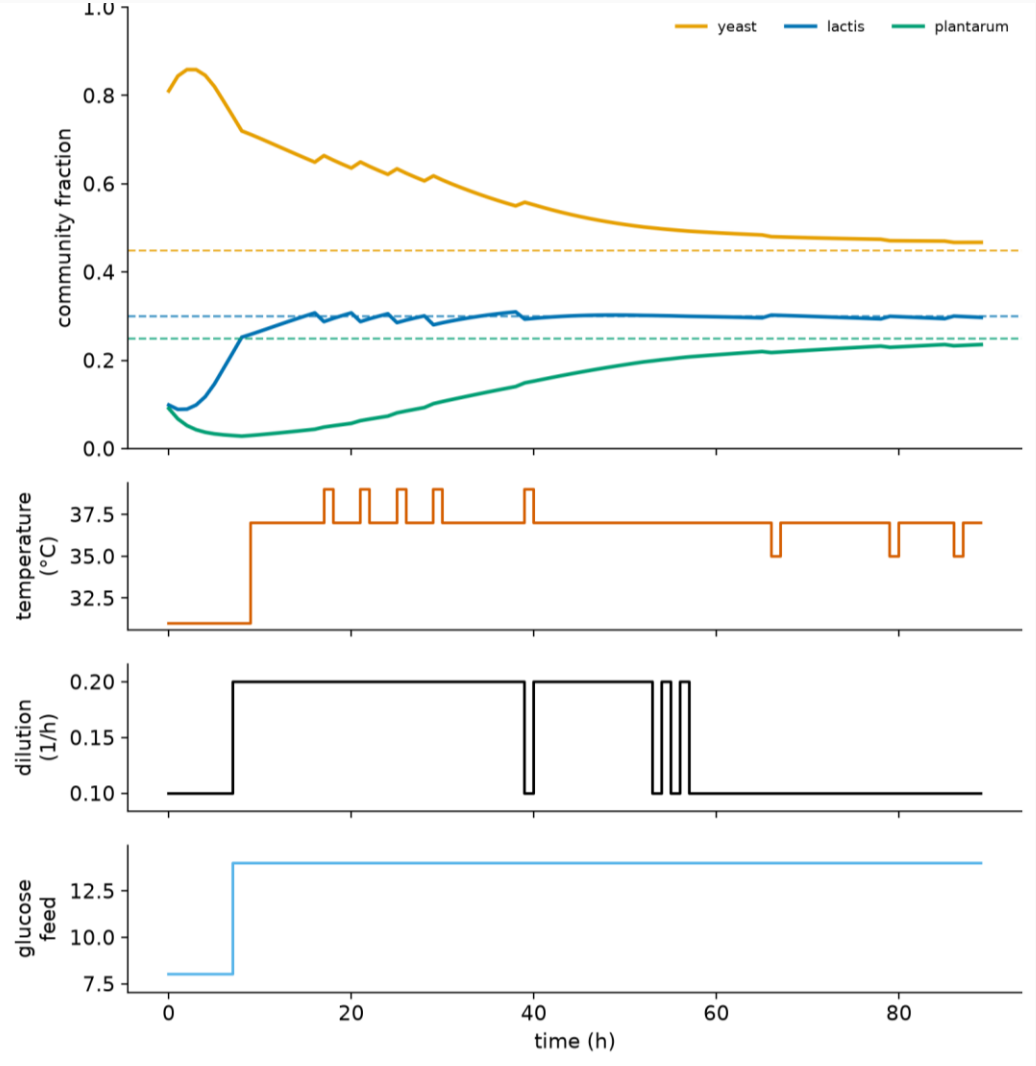
- Same two auxotrophic strains with the same control actions
- Simple division-of-labour; the agent can only measure the product
- Agent is rewarded based on product titre





Learning with Non-Orthogonal Actions





- A more complex, S. cerevisiae + L. lactis + L. plantarum community
- Cross-feeding and competition entangles community dynamics
- Control actions (temperature, dilution, glucose) affect all strains but differentially
- RL can exploit physically coupled control handles that would defeat classical MIMO controllers

Summary

- Division-of-labour in continuous bioprocesses offers efficiency — but demands control
- Reinforcement Learning provides a flexible, model-light framework for that control
- We can achieve:
    - target tracking, recovery from perturbation, ratchetting system complexity, parallel training, dynamic objectives, yield maximisation, non-orthogonal actions

Reinforcement Learning turns the control problem into a learning problem — and learning is something we can scale

Acknowledgements

Fedorec Group
Casey Chen
Alice Cook
Louie Destouches
Manuel Godoy
Kirill Sechkar

Dongning Chang
Ryan Kingston

Cell & Developmental Biology, UCL
Chris Barnes
Neythen Treloar

University of Waterloo
Brian Ingalls

Biotechnology and Biological Sciences Research Council

Engineering and Physical Sciences Research Council

European Research Council
Established by the European Commission

Treloar et al., PLOS Comp Biol, 2020