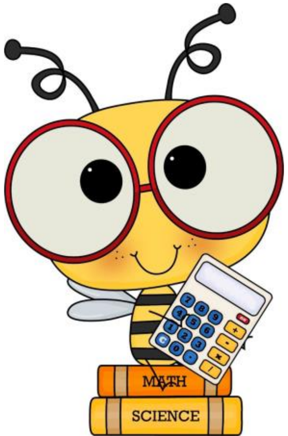
A Bee Model

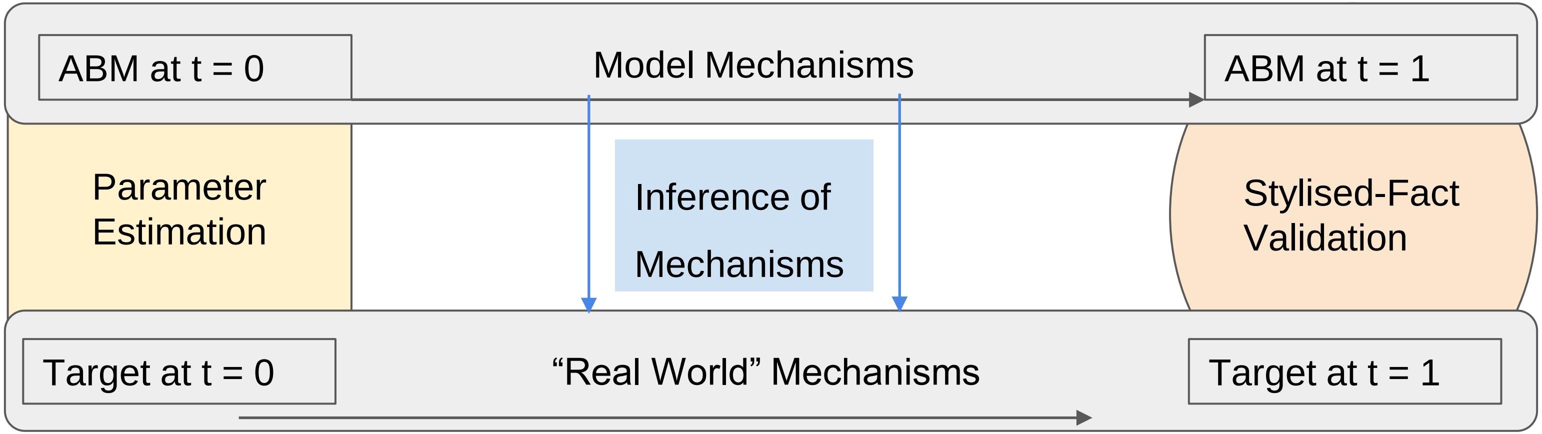
Learning agent-based models from data



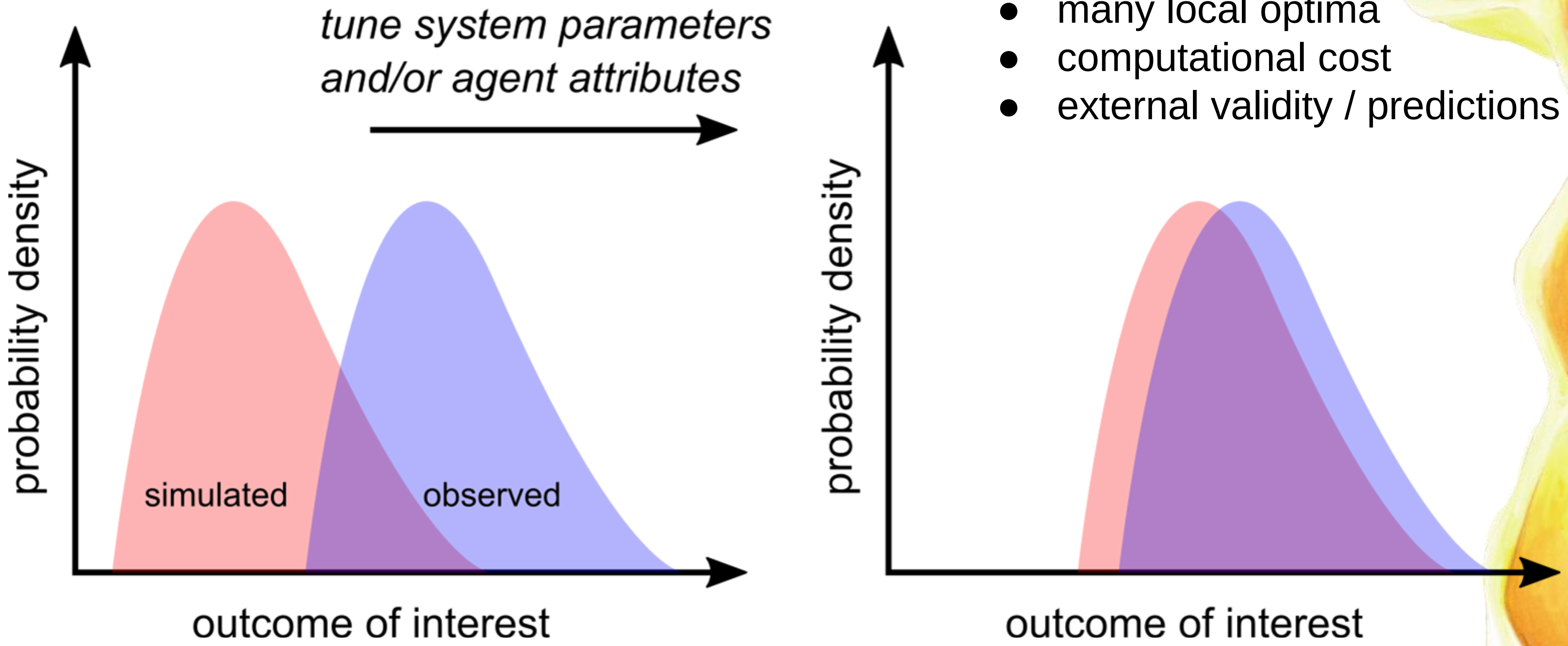
Agent based models

ABMs enable the study of a phenomenon via simulation of rule-based interaction of agents.

How to choose: i) Micro parameters ii) Mechanisms



Calibrating parameters to match moments of distributions



Challenges

- many local optima
- computational cost
- external validity / predictions

Overview of survey of IRL, shortcomings

Method	$\hat{R}_E$ params	Optimization objective	Notable aspect
Max margin methods - maximize the margin between value of observed behavior and the hypothesis			
MMP	$\boldsymbol{w}$	value of obs. τ - max of values from all other τ (Eq. 8)	provable convergence
MAX-MARGIN		feature exp. of policy - empirical feature exp. (Eq. 9)	sample bounds
MWAL		min diff. in value of policy and observed τ across features	first bound on iteration complexity
HYBRID-IRL		empirical stochastic policy - computed policy of expert (Eq. 10)	natural gradients and efficient optimization
LEARCH	$R(\boldsymbol{\phi})$	value of obs. τ - max of values from all other τ (Eq. 8)	nonlinear reward with suboptimal input
Silver et al. [29]			normalization of outlier inputs
Max entropy methods - maximize the entropy of the distribution over behaviors			
MAXENTIRL	$\boldsymbol{w}$	entropy of distribution over trajectories (Eq. 11)	low learning bias
STRUCTURED APPRENTICESHIP		entropy of distribution over policies (Eq. 12)	efficient optimization
DEEP MAXENTIRL		gradient of likelihood equivalent of MaxEnt (Eq. 13)	nonlinear reward
PI-IRL			continuous state-action spaces
REIRL		relative entropy of distribution from baseline policy (Eq. 14)	suboptimal input and unknown dynamics
Bayesian learning methods - learn posterior over hypothesis space using Bayes rule			
BIRL	$R(s)$	posterior with Boltzmann data likelihood (Eq. 16)	first Bayesian IRL formulation
Lopes et al. [43]		entropy of multinomial$(p_1(s), p_2(s), \ldots, p_{\vert A\vert-1}(s))$ derived from posterior	active learning
GP-IRL	$f(\boldsymbol{r}, \boldsymbol{\theta})$	Gaussian process posterior	nonlinear reward
MLIRL	$\boldsymbol{w}$	differentiable likelihood with Boltzmann policy (Eq. 16)	first ML approach
Classification and regression - learn a prediction model that imitates observed behavior			
SCIRL	$\boldsymbol{w}$	Q-function as classifier scoring function	actions as state labels provable convergence
CSI			unknown dynamics
FIRL	regression tree	norm of ($\hat{R}_E$ - projection of $\hat{R}_E$)	avoids manual feature engineering
AIRL	R(s)	regression error between expert demonstration $\mathcal{D}$ and $\mathcal{D}_\pi$	suitable for transfer learning

Overview of survey of IRL, shortcomings

- Accuracy of inference
- Generalisability
- Sensitivity to correctness of prior knowledge
- Disproportionate growth in solution complexity with problem size

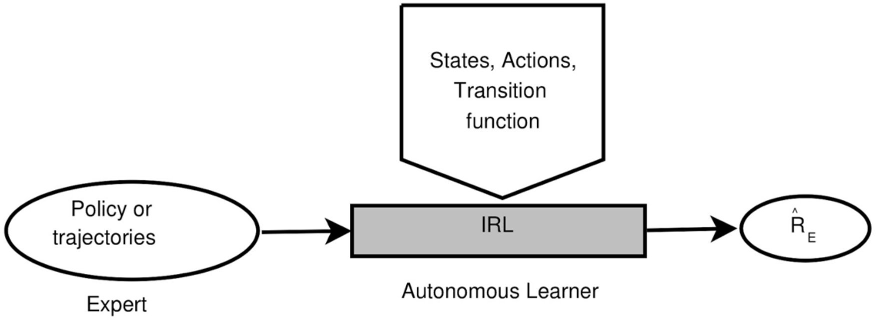


Figure 4: Pipeline for a classical IRL process. The learner receives an optimal policy or trajectories as input. The prior domain knowledge (shown here as a pentagon) include the completely observable state space, action space, and fully known transition probabilities.

A (classical) Bee Model

Basic Idea: Make as simple of a model as is useful to test data fitting strategies

The Environment: A grid with some amount of flowers randomly distributed

The Agent: A single bee that moves randomly around the ground. When it comes upon a flower, it has some probability of eating the flower and some probability of flying away

Observed: Location of flowers and bee in time

Target Parameter: Probability of eating a flower

A (classical) Bee Model

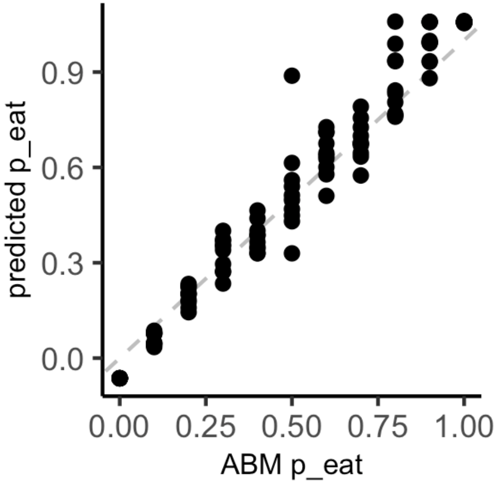
p_eat = 0.1

p_eat = 0.9

A (differentiable) Bee Model

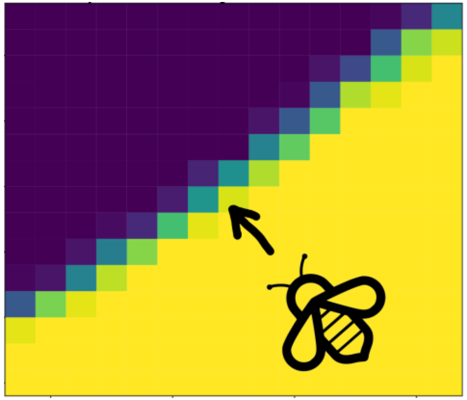
$$\frac{d}{dt}$$

- Basic idea: use differentiable operations to build ABM (Monti et al. 2023)
- Novel: keep discrete agents, and spatial structure
- Formulate discrete decisions as sampling from probability distribution – use VAE “reparameterization trick” (Kingma et al. 2013, Rezende et al. 2014)



Future directions for ABMs

- Exploration of phase transitions in parameter space
  - Computing directions as eigenvectors of Hessian (Naumann-Woleske et al. 2022)
  - Detection of different parameter regions



- Further ideas for learning initial values of latent variables (Monti et al. 2023)
  - Does the algorithm depend on the dynamic properties of the system? (Stationary distribution / cyclic attractor / chaos)
  - Flower growth as logistic map (latent variable)

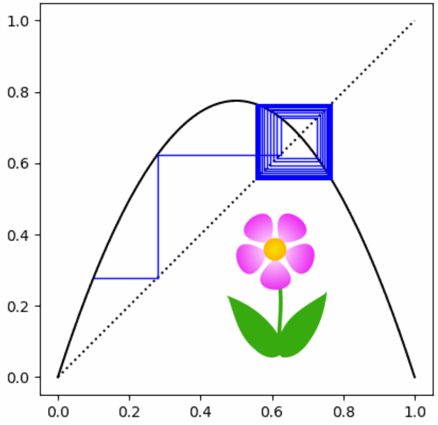
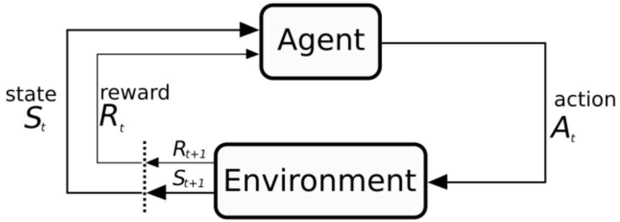


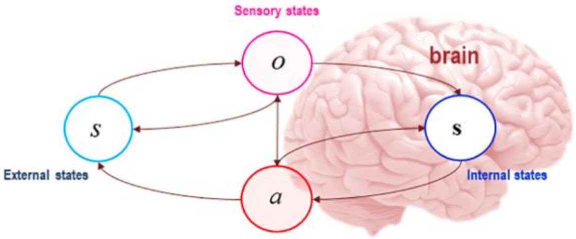
Image sources: Cobweb: https://www.johndcook.com/cobweb2z.png, Bee: https://cdn.onlinewebfonts.com/svg/img_563816.png
Flower: https://clipart-library.com/newimages/clip-art-flowers-41.png

Future directions for ABMs [Pablo 1 min]

- Multi Agent *Inverse* Reinforcement Learning (MIRL)
    - Agents are generated to optimize pursue certain reward
    - They interact, exploring/exploiting strategic possibilities
    - Agents develop their own strategies along the way
    - **Inverse**: Learning agents' reward function from their strategies



- Active Inference and Free Energy Principle
    - Agents try to minimize "free energy" (uncertainty) in environment
    - They do so by refining their models and *acting* upon the environment
    - A modular ("markov-blanket") causal model naturally emerges
    - **Inverse**: Learning agents' internal states from s,o,a



Future Directions for Project

- Join us on GitHub and Zotero to create a collective of teaching students and learning teachers.
- Overview of ABM learning resources, to be extended with examples from contemporary ABM modelling.