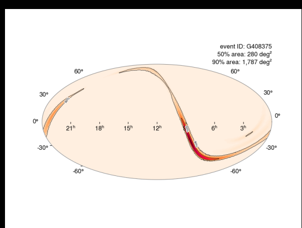
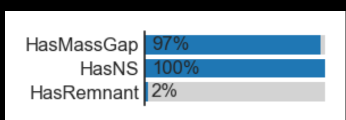
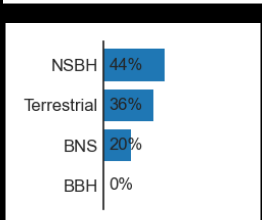
MACHINE-DIRECTED GRAVITATIONAL-WAVE COUNTERPART DISCOVERY

NIHARIKA "ARI" SRAVAN
DREXEL UNIVERSITY

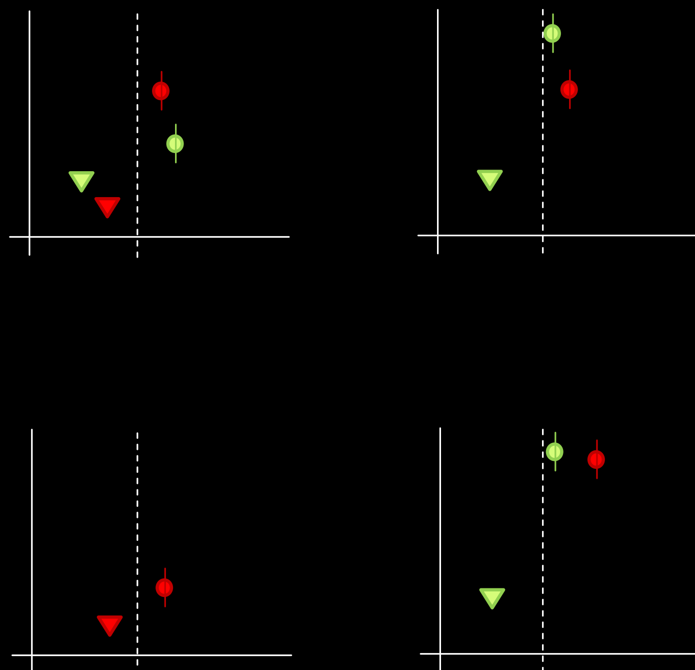
AUTONOMOUS RESOURCE ALLOCATION







FAR – 9.59 per year



30 hr per semester
g, r, i, z

1 per night
g, r

low res spec
~19mag
up to 5 per night

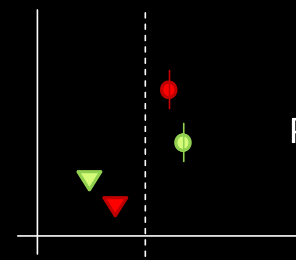
N < 500

Goal: Allocate resources optimally over all events over 7 days

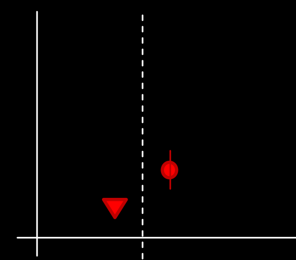
What is optimal?
How make such decisions without complete information?

NAÏVE STRATEGY

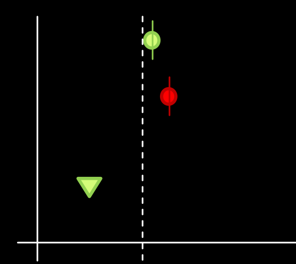
- Rank by confidence is kilonova
  - Allocate full budget to most confident kilonova?
  - Allocate one of each type in order of decreasing confidence until exhausted?
  - Allocate to improve classification scores?
  - Allocate to improve light curve constraints?
  - Other?
- Decision hard because follow-up could be misallocated and this may be apparent with a delay
  - Perhaps it was better overall to rule out borderline cases and later allocate to best guess
- Ideal strategy is optimal given all future allocations and all future outcomes



P(KN)=0.73



P(KN)=0.12

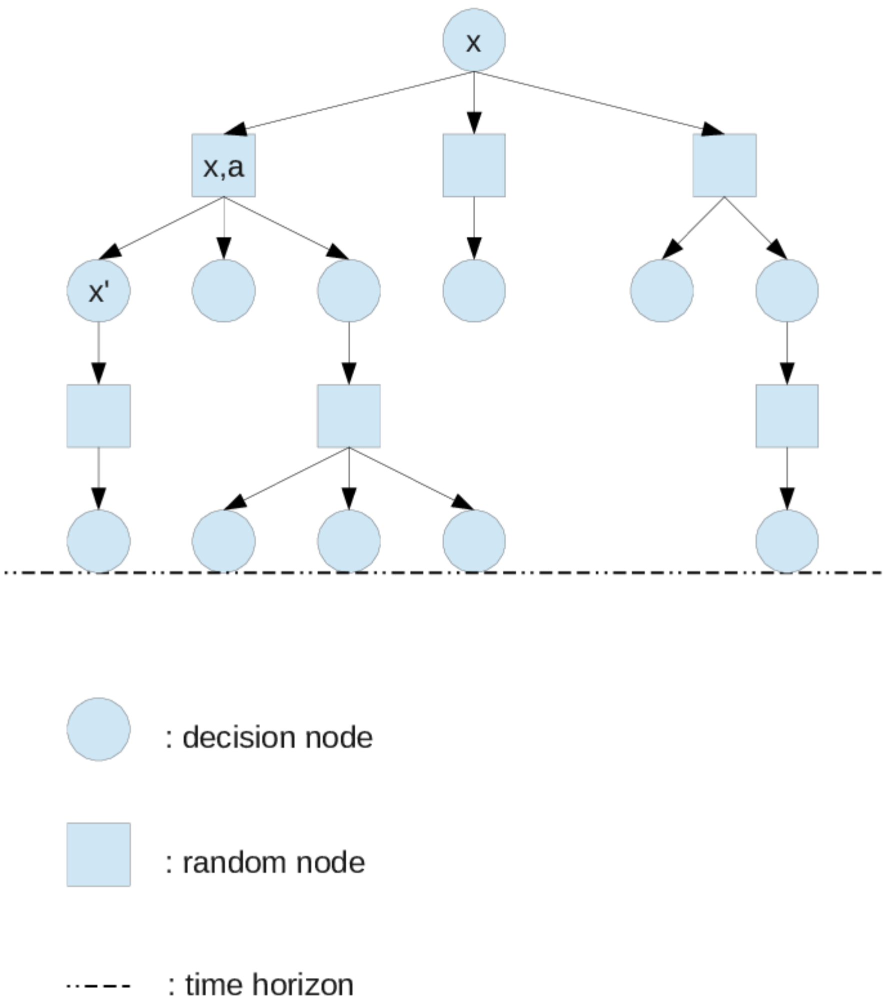


P(KN)=0.11

30 hr per semester
g, r, i, z

1 per night
g, r

low res spec
~19mag
up to 5 per night



REINFORCEMENT LEARNING

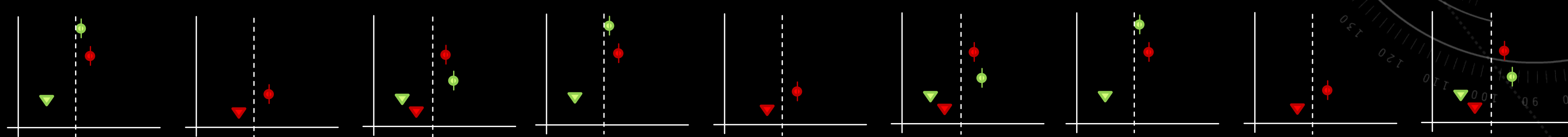
- RL agents learn through experience how, given a situation, taking an action now affects returns achieved later
  - Considers the full distribution of outcomes at every timestep, and all future actions and their outcomes, and so on, and chooses the best action right now that maximizes an overall reward

PYTHIA

TOY KILONOVA FOLLOW-UP AGENT

One per day

Episodes end day 7



9 transients, one of which (always) is the true kilonovae (min photometry = 1)

- Contaminants are SNe, unassociated GRB afterglows, shock breakout (do not include observational significance)

Follow-up in ZTF g, r, or i (300s exposure) per day

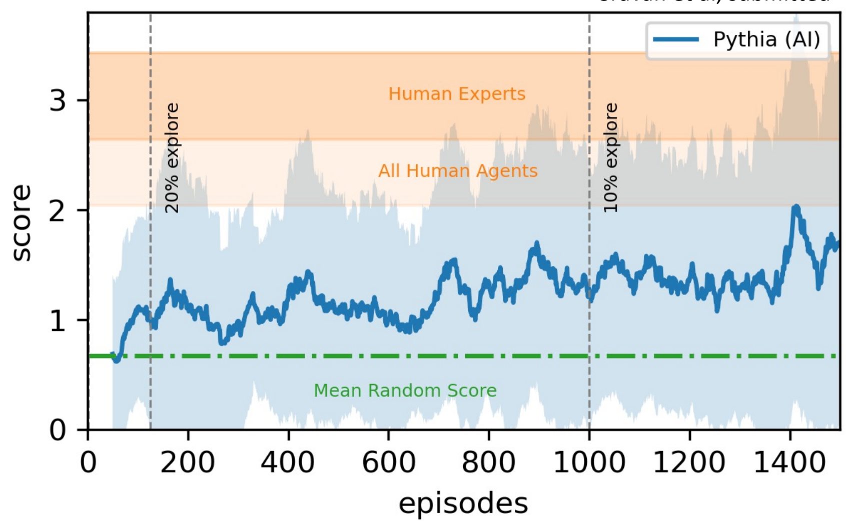
- Finite horizon – 7 days (observe on day 1)

Reward 1 if agents adds data to the kilonova else 0

- Maximize the number of follow-up to the true kilonova (non-model specific objective with the expectation that more data ~ better constraints)

Pythia

Sravan et al, submitted



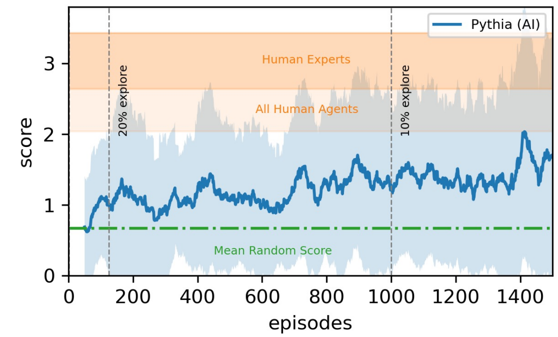
Linear VFA hypothesis class not sufficiently rich representation of true Q function

- Benefit is theoretical convergence guarantee. Demonstrates problem learnable!

Shifting to deep Q networks:

- Will remove two-step learning, one for x(s,a) in supervised/unsupervised learning and one for Q via Bellman updates in RL
- Efficient evaluation of realistic large action space, can have vector instead of scalar output

Pythia v humans



agent	score	frac KN > 1 follow-up
Pythia	**1.84**	**0.81**
Non-expert 1	2.04	0.54
Non-expert 2	3.15	0.86
Expert 1	2.64	0.76
Expert 2	2.74	0.78
Expert 3	2.94	0.72
Expert 4	**3.43**	**0.9**

Sravan et al, submitted

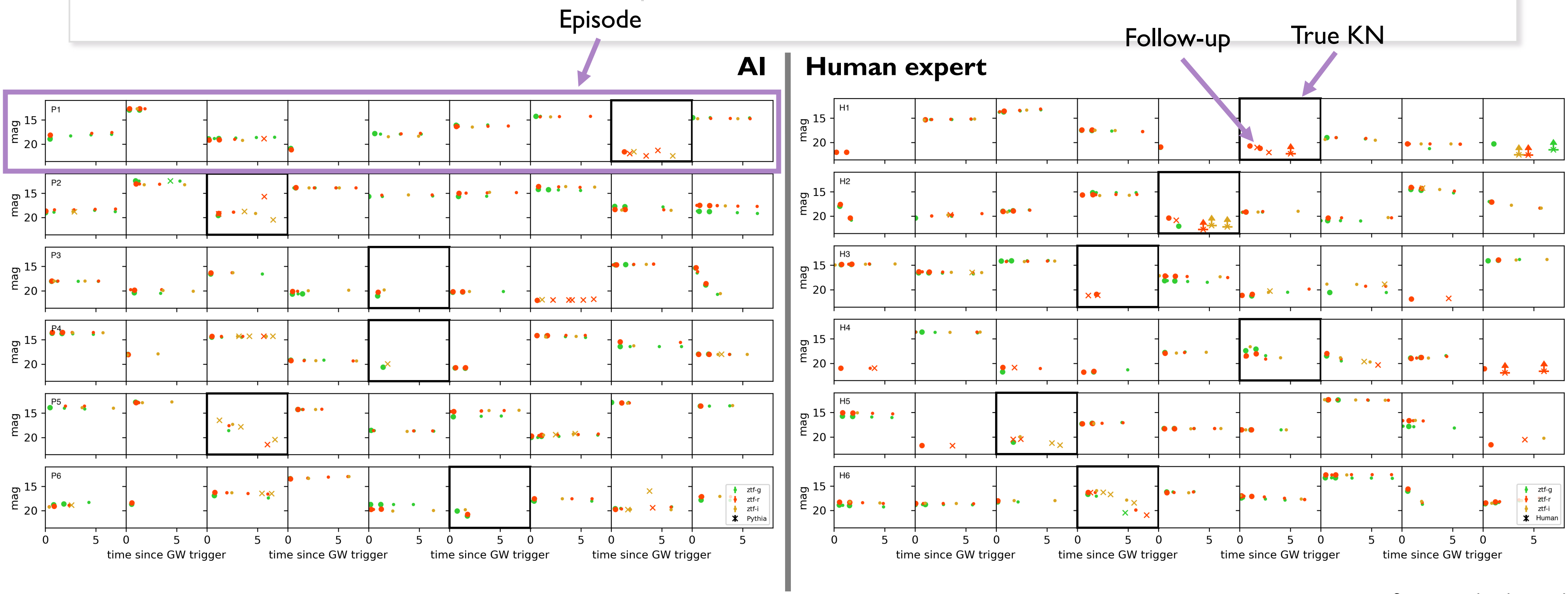
Pythia v humans

Random test episodes

Goodhart's law: "When a measure becomes a target, it ceases to be a good measure"

- AI prefers greedy when it fails since no benefit for exploring

Human experts have blindspots too



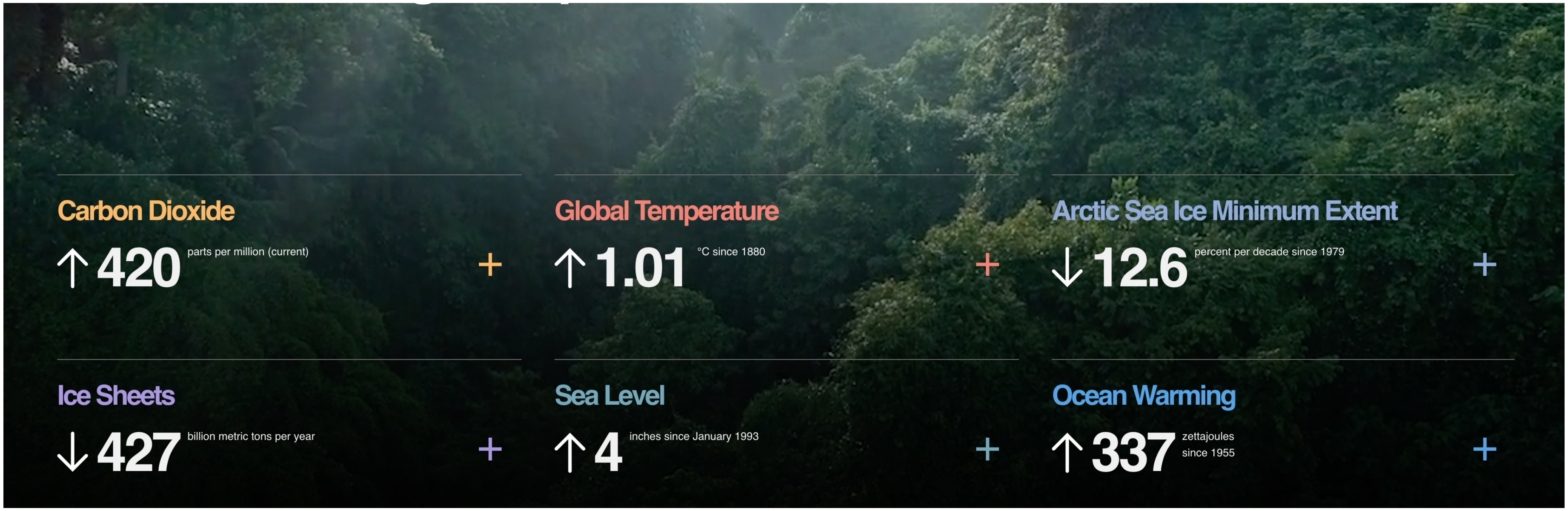
Sravan et al, submitted

Carbon Footprint

Estimated emissions: 1210 kg of CO_2eq. assuming carbon efficiency of 0.432 kgCO_2eq/kWh

Approximately equal to:

- One round trip LAX-JFK (1180 Kg CO_2)
- 4900 km driven in an average combustion engine car



climate.nasa.gov