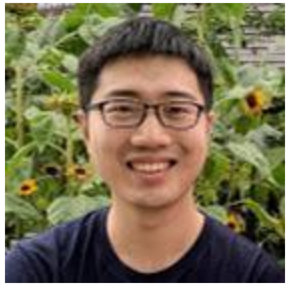
Interaction-Grounded Learning



Tengyang Xie
UIUC

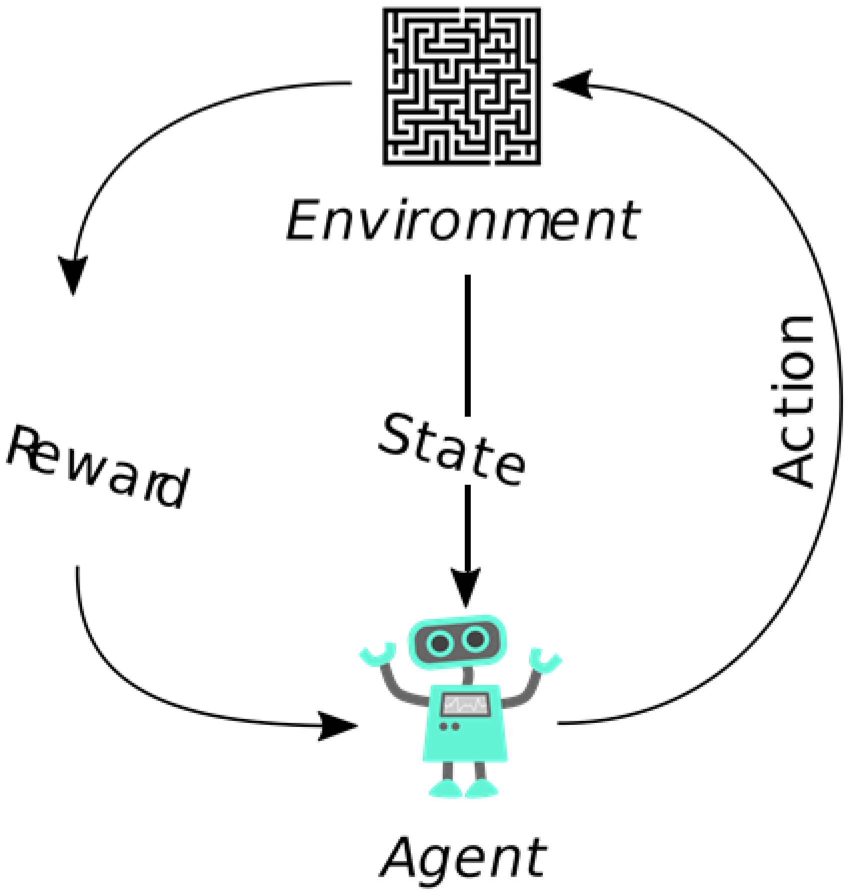
John Langford
Microsoft Research

Paul Mineiro
Microsoft Research

Ida Momennejad
Microsoft Research

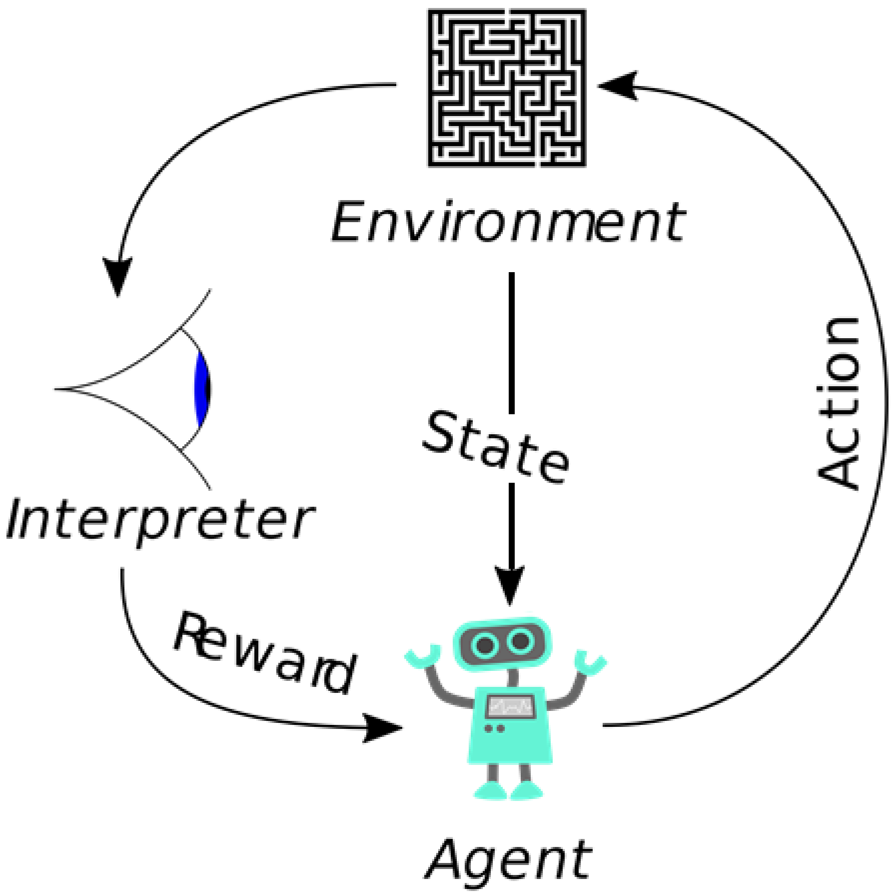
Interactive machine learning beyond *explicit reward*

Interactive machine learning in <u>Textbook</u>

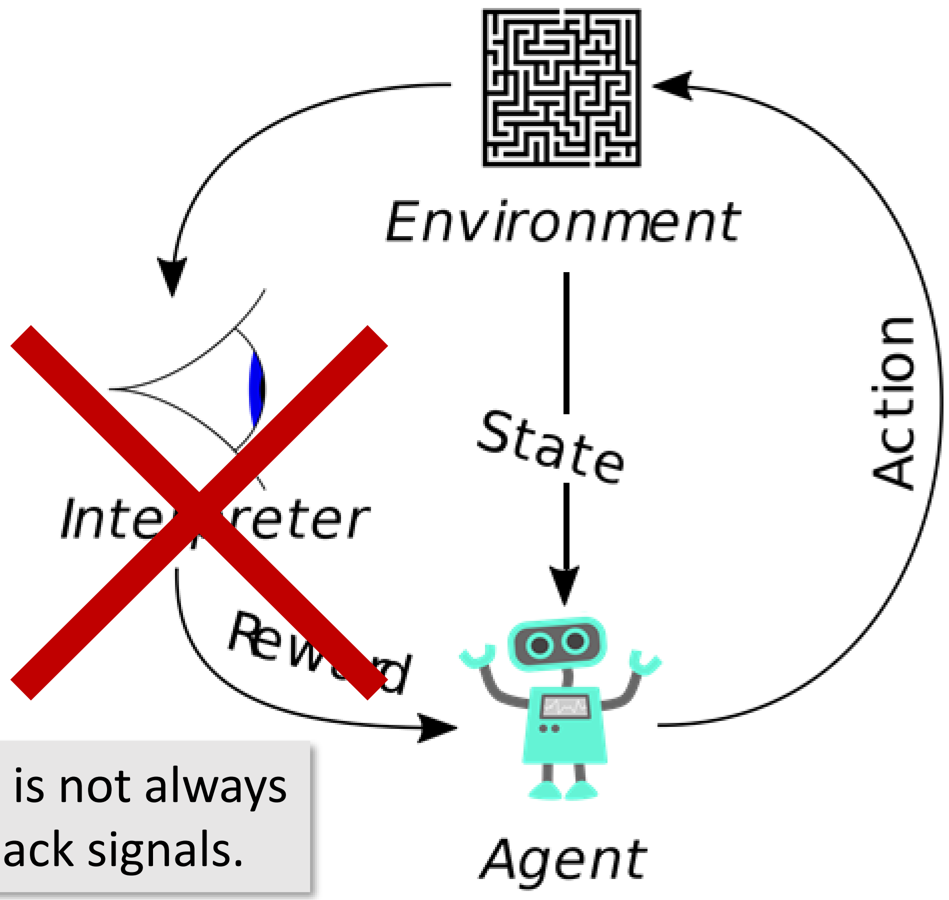


Interactive machine learning beyond *explicit reward*

Interactive machine learning in Real World



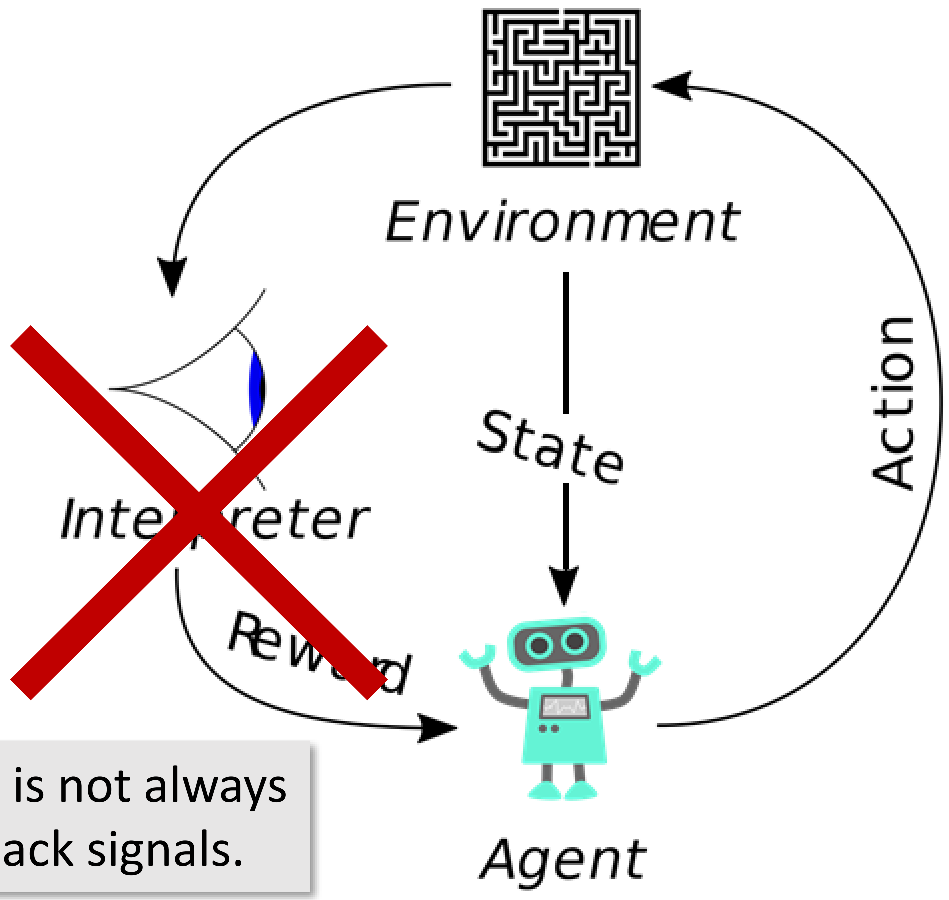
Interactive machine learning beyond *explicit reward*



Challenge: Interpretation is not always feasible with rich feedback signals.

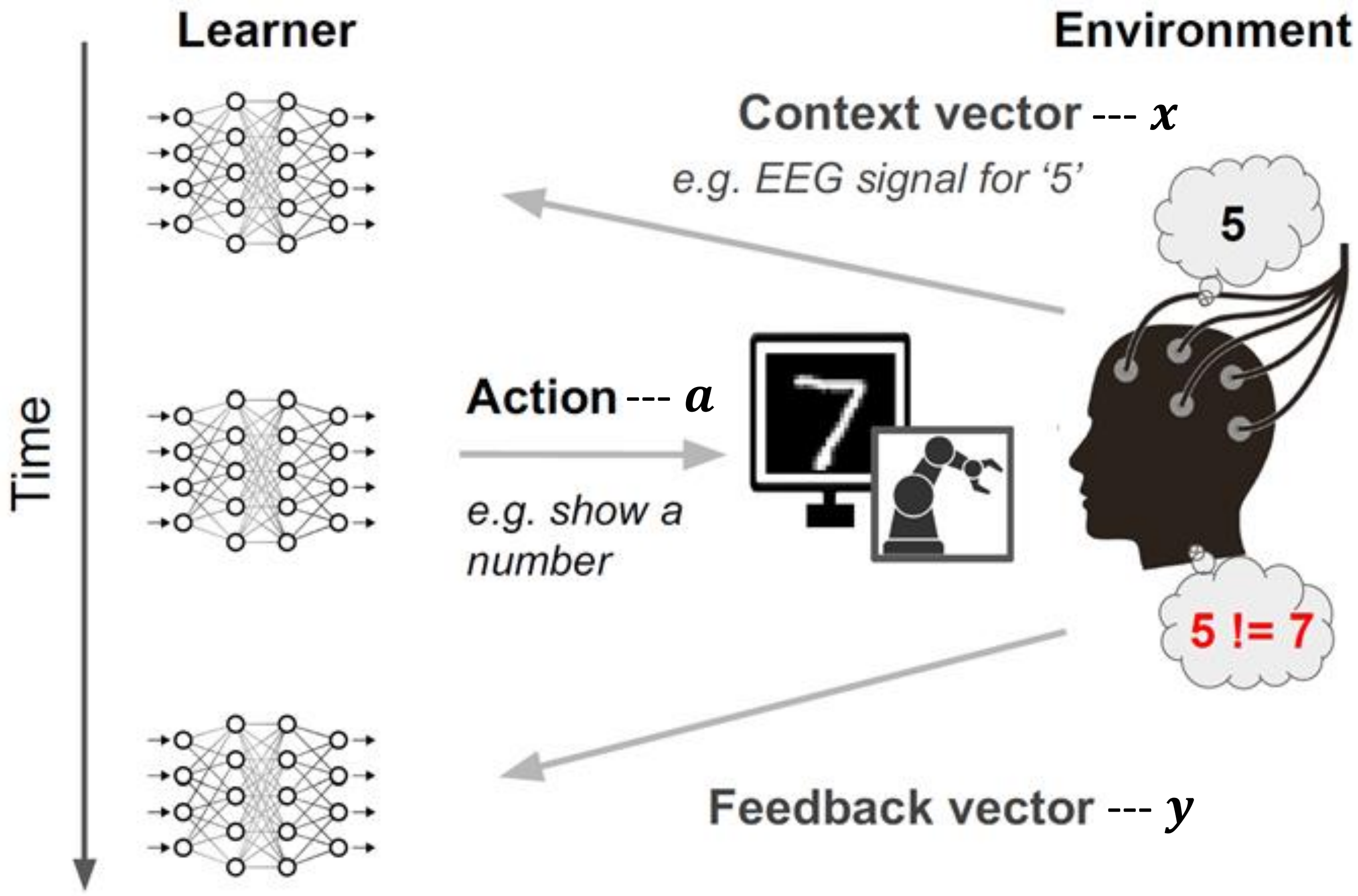
Interactive machine learning beyond *explicit reward*

How can we learn without explicit reward?



Challenge: Interpretation is not always feasible with rich feedback signals.

Interaction-Grounded Learning (IGL)



Algorithm for IGL

- Learning Goals:
  - Policy π: $\mathcal{X} \to \mathcal{A}$ (approximating optimal policy $\pi^\star$)
  - Reward decoder ψ: $\mathcal{Y} \to [0,1]$ (approximating optimal reward decoder $\psi^\star$)

- Proxy Learning Objective: (corresponding algorithm --- E2G)
  - $\mathrm{argmax}_{\pi,\psi} \mathrm{V}(\pi,\psi) - \mathrm{V}(\pi_{\mathrm{bad}},\psi)$

$\mathrm{V}(\pi,\psi) := \mathbb{E}_{x,a,y\sim\pi}[\psi(y)]$

Our Result

- E2G provably converges to $(\pi^\star, \psi^\star)$ under natural assumptions.

- Assumptions:
  - *Conditional Independence*

    --- feedback vector y only contains information about reward r
  - *Identifiability*

    --- π_{bad} is "bad" enough to be identified

Conclusion

- IGL: A novel setting that conducts interactive machine learning without explicit reward function.
- E2G: A novel algorithm provably solves IGL under natural assumptions.
- Future Directions:
    - Relaxing the assumption of conditional independence
    - Applying IGL to real-world problems (e.g., brain-computer interface)

Thank you!