



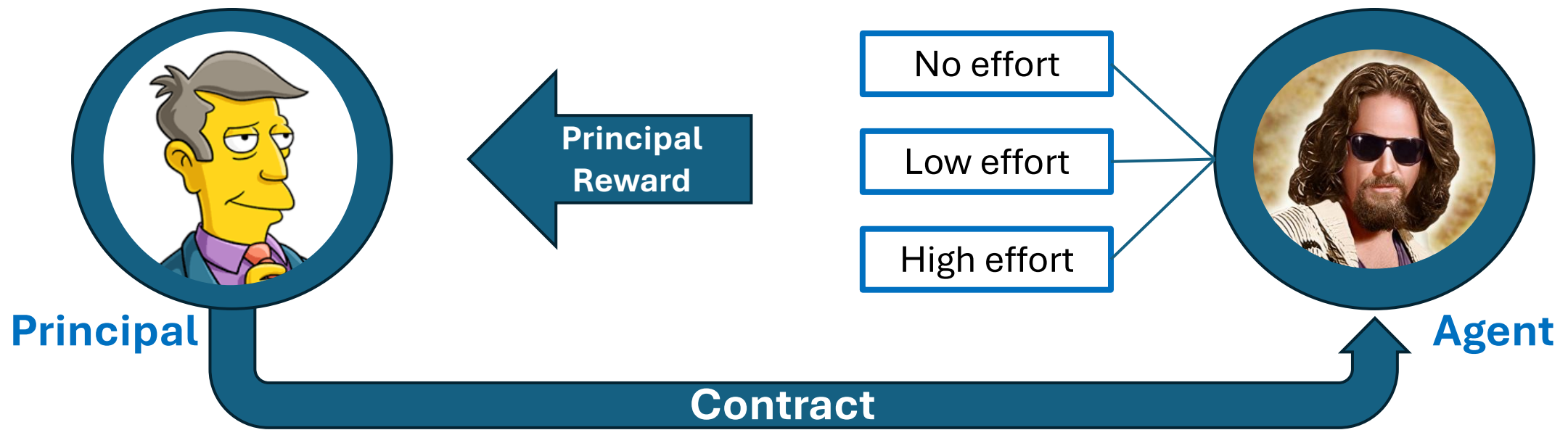
Contracting with a Learning Agent

Yoav Kolombus, NeurIPS 2024

Joint work with G. Guruganesh, J. Schneider, I. Talgam-Cohen, EV. Vlatakis-Gkaragkounis, JR. Wang, SM. Weinberg

Principals, Agents, and Contract

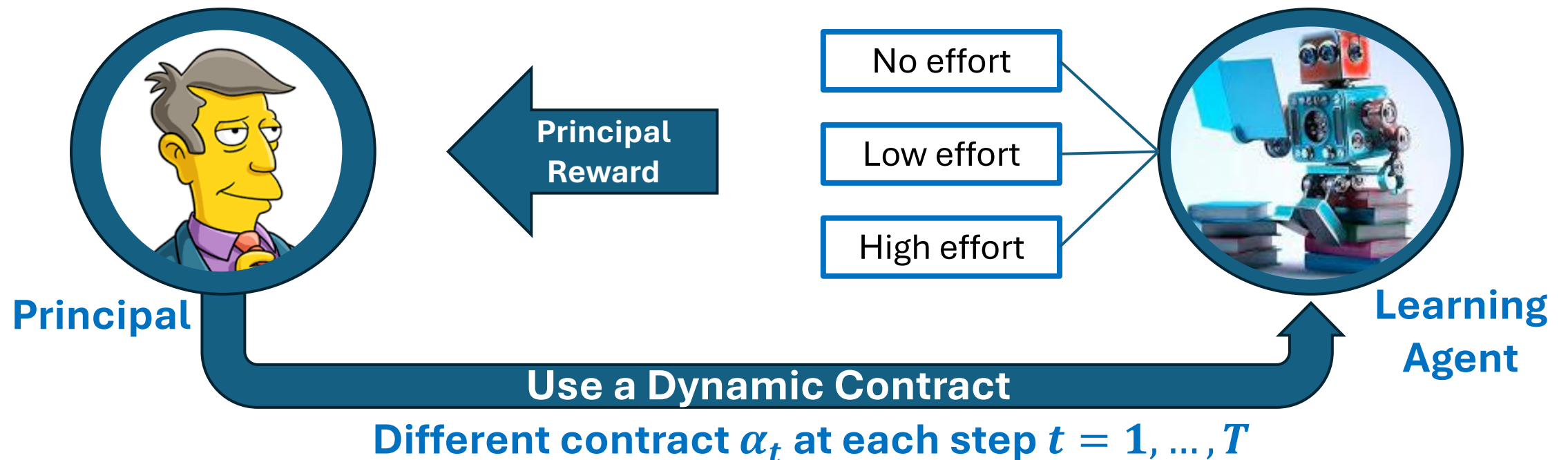
- **Agent** performs a task/service for the **principal**
- Agent selects among **costly actions**
- Principal aims to **incentivize “good” actions** (with high reward)



Example: *I'll pay you an α part of the reward (linear contract)*

Contracting with a Learning Agent

- Interaction is **repeated**
- Agent is **learning** (responds to past experience)



Example: outcomes are success/failure

- **Best static contract:** paying $\alpha = 2/3$ of reward to agent is optimal (or $1/3$)
- Principal (net) utility: **$1/3$**
- Agent (net) utility: **$1/6$**

Actions	“Failure”	“Success”
$a_1: (c_1 = 0/6)$	1	0
$a_2: (c_2 = 1/6)$	1/2	1/2
$a_3: (c_3 = 3/6)$	0	1



Agent actions
and costs

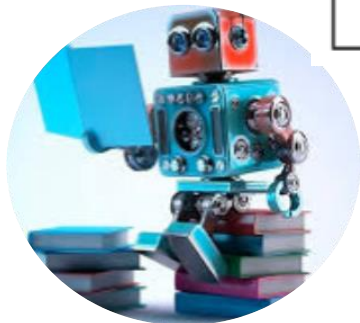
Expected
principal rewards



Example: Now with a Learning Agent

- Agent is a **mean-based no-regret learner**
- **Best dynamic contract:** pay $\alpha_t = 2/3$ until $T/2$, then pay zero until T
- Agent's response is a “**free fall**” through the actions:
 - plays a_3 (high effort) until $T/2$, then plays a_2 (low effort) until T
- Principal utility is now higher, **$5/12$** . Agent utility is **zero**.

Actions	“Failure”	“Success”
$a_1: (c_1 = 0/6)$	1	0
$a_2: (c_2 = 1/6)$	1/2	1/2
$a_3: (c_3 = 3/6)$	0	1



Agent actions
and costs

Expected
principal rewards



Results Roadmap

With mean-based learners:

- Principals would prefer to use dynamic contracts
- “Free-fall” contracts are optimal dynamic linear contracts (Theorem 3.1)
 - Use a fixed contract until some time T^* , then switch to pay zero.
 - Can be computed efficiently
- Extends to general linearly-scaled contracts (Theorem D.1)
- Optimal dynamic contracts may have win-win outcomes (Theorem 3.2)

With no-swap regret learners:

- Best static contract is optimal (observation I.2)

Uncertainty about the time horizon:

- Principal added gains from being dynamic degrade as uncertainty increases (Thm. 4.2-4.3)

Summary notes

Contracting with a Learning Agent

- Results can be **very different** with **learning agents**
- **Principals benefit** from using simple dynamic contracts
- The learning **agent can be worse off or better off**
- A rich setting, **many open questions**:
 - Algorithmic, game-theoretic, and computational

Thank You!

yoav.kolumbus@cornell.edu