Week 10 · Act 4 — Closed-loop honeypot RL

Goal: Train a multi-step honeypot agent and compare to a myopic max-bait baseline.

Thesis: When the environment reacts over a session, RL matches the problem shape.

Training may take ~10–20 seconds. Watch early vs late return on the figure.

In [ ]:
```
import seas8414_teach as teach
teach.setup()
print(teach.__version__)
teach.story_title_card(
    "Act 4 · RL earns its keep",
    "Interactive · closed-loop · multi-step",
    ["Q-learning", "Myopic max-bait baseline", "Variance is a cost"],
)
```

In [ ]:
```
curve = teach.train_chameleon()
print(curve.early, curve.late, curve.baseline_mean)
```

In [ ]:
```
teach.story_act4(curve)
```

In [ ]:
```
# Optional: needs OPENAI_API_KEY in ~/.env — skip if unset
teach.ask(
    "Why does pacing bait beat always-max-bait in a closed-loop honeypot?",
    early=curve.early, late=curve.late, baseline=curve.baseline_mean,
)
```

Exit ticket

Name three properties that make multi-step RL appropriate here.