

RAGEN + A*PO

Evaluation on WebShop & WebArena

Team 4Cores • INFO 7375 • Professor
Suhabe Bugrara

Team Members

Jing Cao

Dhanashree Atul Nerkar

Nai-Cih Syu

Lakshminarayanan Ravi



System Overview



RAGEN: Retriever +
Generator + Policy
Network



A*PO: Advantage-
weighted PPO



Model Size: ~847K
parameters

Our Extended WebShop Environment (Last Assignment)

- ~497 items, 8 categories
- search / click / buy actions
- Shaped rewards: +1 correct buy, +0.1 helpful search, -0.05 irrelevant click

Official Benchmark Setup



WebShop: product catalog, structured search



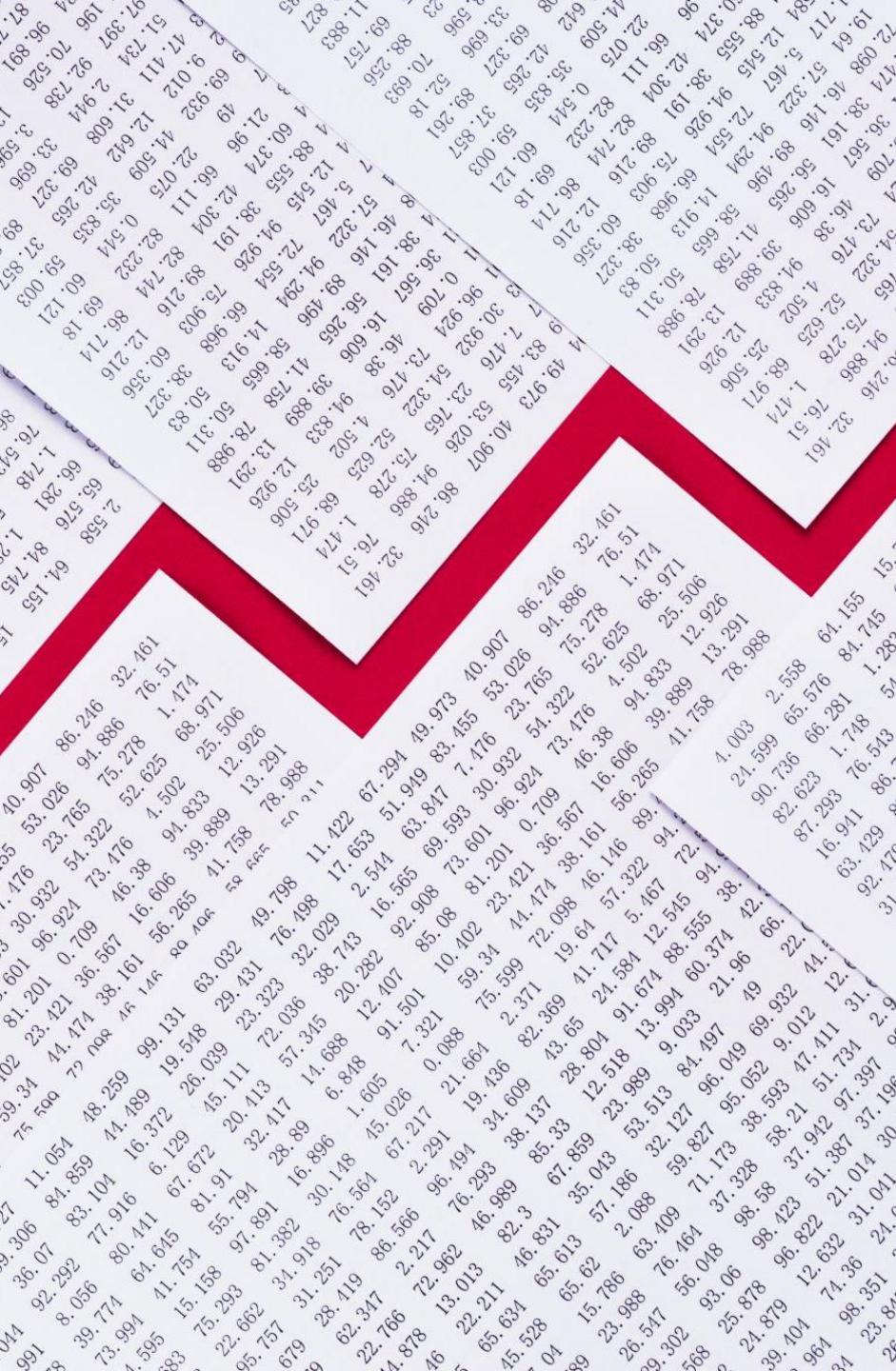
WebArena: real browser DOM + complex UI planning



Zero-shot transfer tested

Performance Comparison

- WebShop (Our Extended Version): 67% success
- WebShop (Official Benchmark): 51.2% success
- WebArena: 9.8% success



Why Extended WebShop Achieves 67%

Deterministic
environment

Reward shaping
helps stable
learning

Search vocabulary
aligned with
product attributes

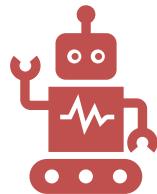
WebShop Failure Patterns

Search loops → safe reward exploitation

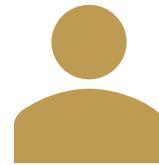
Wrong buy due to attribute confusion

Unreachable target IDs in baseline dataset

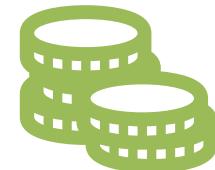
Generalization Challenge: WebArena



Large DOM
(~1000+ elements)



Structured browser
actions required



Sparse long-
horizon rewards

Why RAGEN Fails on WebArena

LSTM cannot encode full
webpage state

Fixed action vocabulary
cannot express UI
interactions

Credit assignment breaks in
30+ step tasks



Failure Example (WebArena)

Task: Sort product list → choose cheapest

Agent fails to perform multi-step UI navigation

Breakdown: cannot translate plan into browser actions

Conclusion

- Implementation correct and replicates expected patterns
- Strong performance on structured WebShop tasks
- Large performance drop in real web environment
- Future: Transformer encoder + structured action space

Thank You!

