OpenMIND:
Planning and Adapting in Domains with Novelty

David Musliner, Michael Pelican, Matthew McLure,
Steven Johnston, Rick Freedman, Corey Knutson

We would like to thank DARPA and all the SAIL-ON performers for their collaborative efforts on the program, particularly the UTD, ANU, WSU, and NRL teams who have worked very hard to make the objective evaluations of OpenMIND and other agents possible.

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001120C0041. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of DARPA.

OpenMIND

SIFT

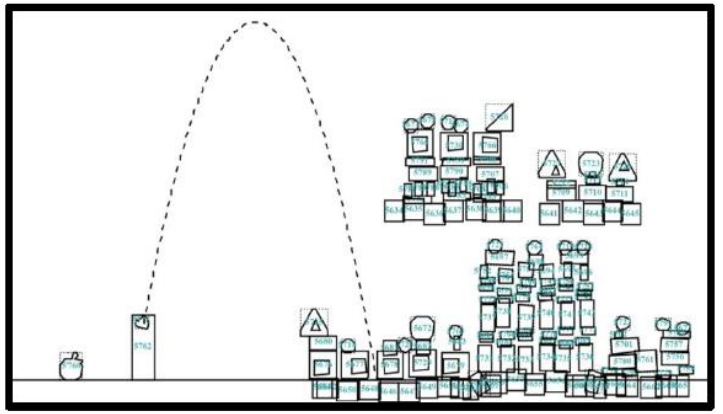
Outline

- Background: the DARPA SAIL-ON program.
- What is novelty?
- What is novel in the research?
- Evaluation approach.
- Results so far.
- Lessons.

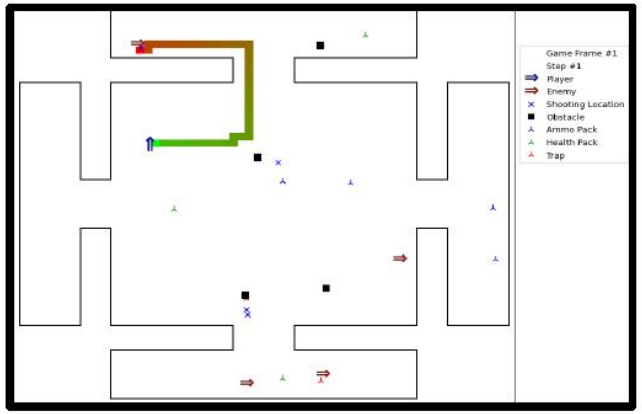
The SAIL-ON Problem

- Science of Artificial Intelligence and Learning for Open-world Novelty.
- Objective: a (mostly) domain-independent, competent agent that can adapt to novelty in its environment.
- Competent: either pre-trained or engineered to be capable of accomplishing one or more tasks in its expected environment.
  - Not a tabula rasa problem, not pure RL from scratch.
- Our version: a planning and plan-execution agent.
  - New goals are not the novelties… the novelties are in the domain.

Polycraft



Sciencebirds



Vizdoom

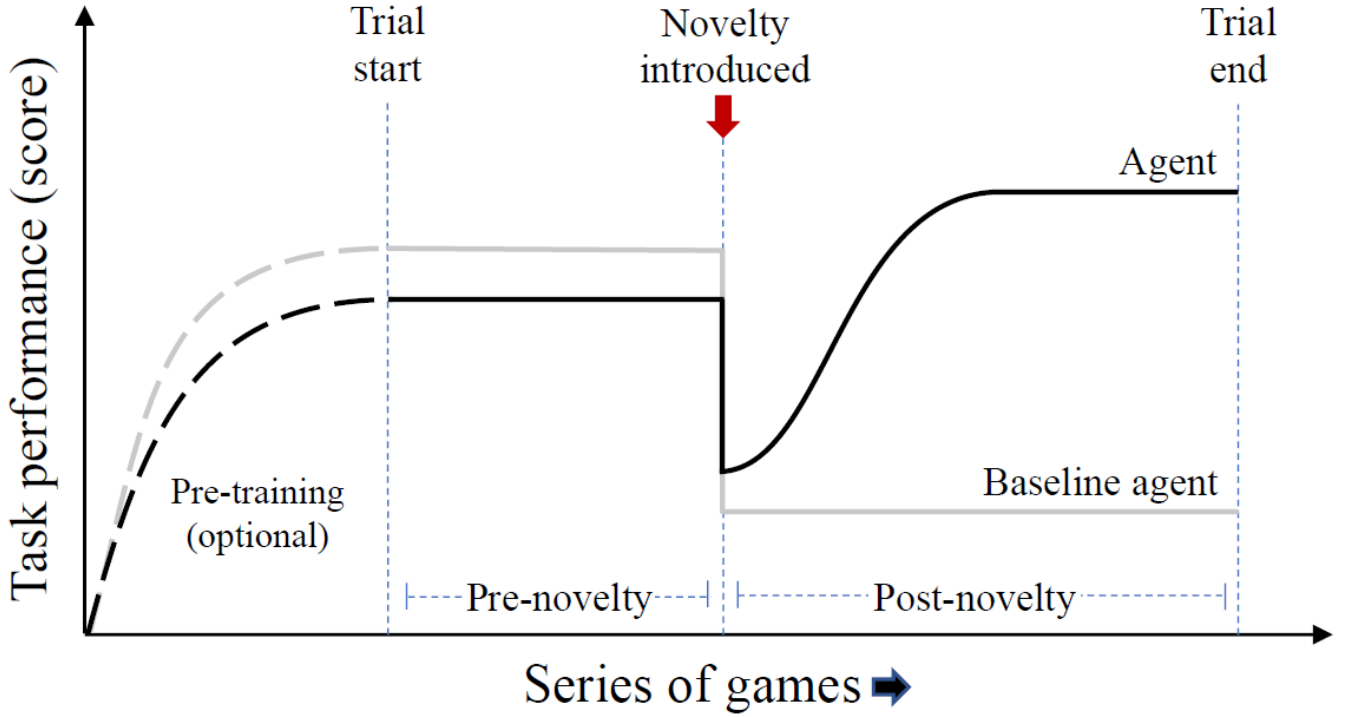
OpenMIND

SIFT

Formal Definition of Novelty

- I know it when I see it, because I haven't seen it before.
- Or, per Pat Langley's AAAI'20 paper: detectable, sudden, persistent changes to the agent's environment.
- Some novelties are opportunities, some are impediments, some are irrelevant.

Evaluation Approach

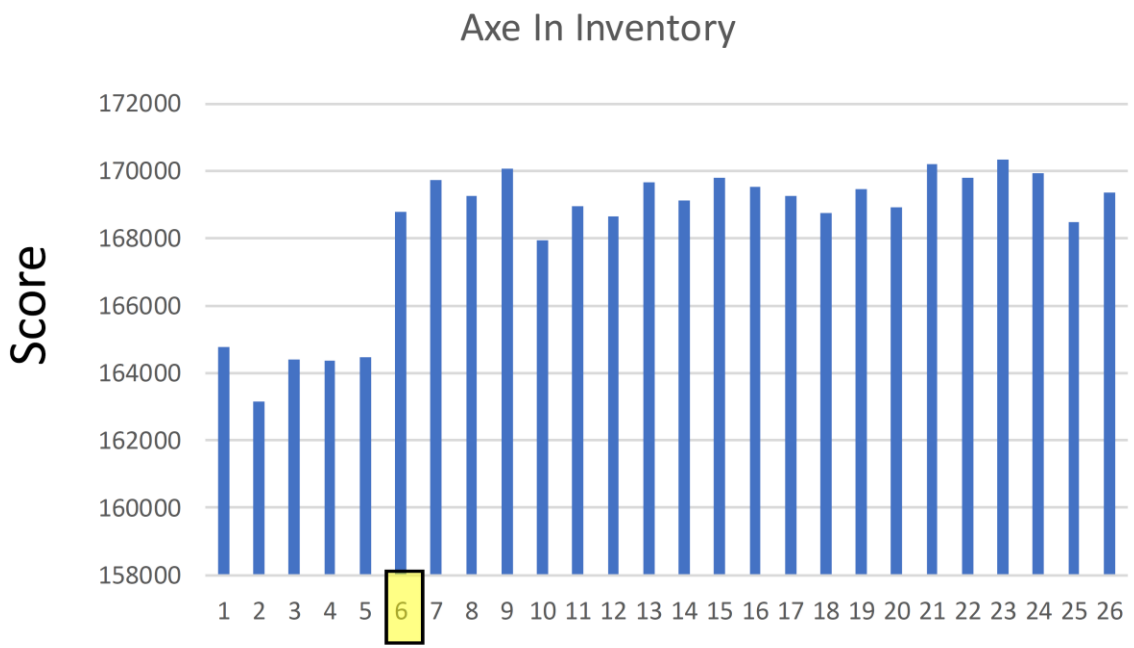


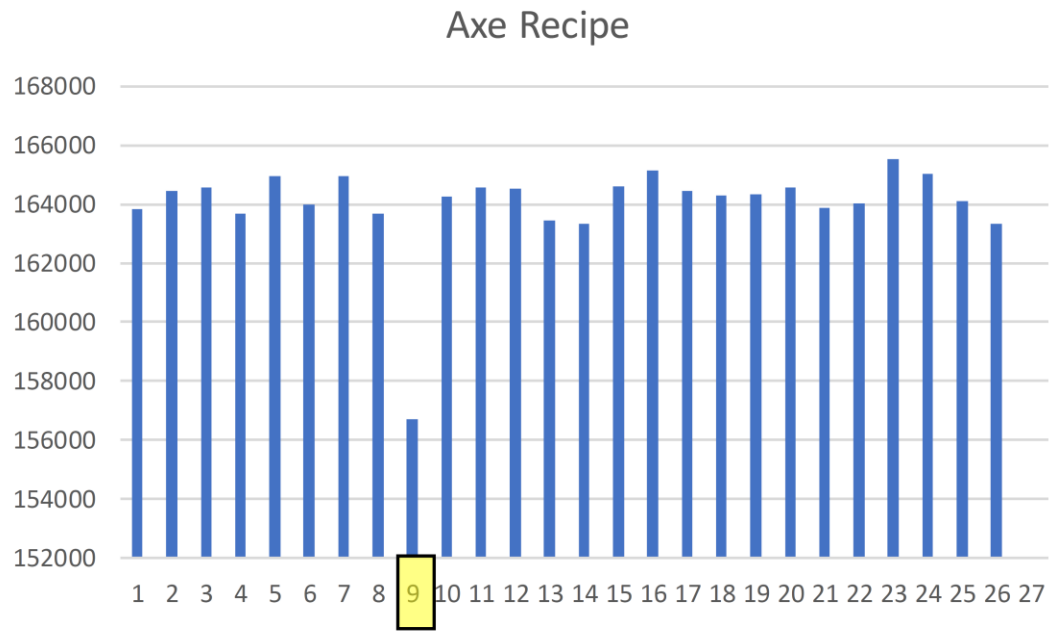
- Multiple "trials", each consisting of a series of "instances" (games).
  - Agent starts each trial fresh, without accumulating learning between trials – learning is within a trial.
- In some trials, at some point the Red Button is pushed (without telling the agent) and novelty arrives.
- Agents should report novelty as soon as detected, and adapt as needed to continue performing task.
- Metrics assess detection and performance adaptation.

Representative Trial Results

- Axes in Polycraft make it cheaper to break trees into logs.
- But the recipe for making an axe from scratch costs more than the axe can save in a game.

- If you have an axe, use it.
- If you only have an axe recipe, ignore it.





OpenMIND

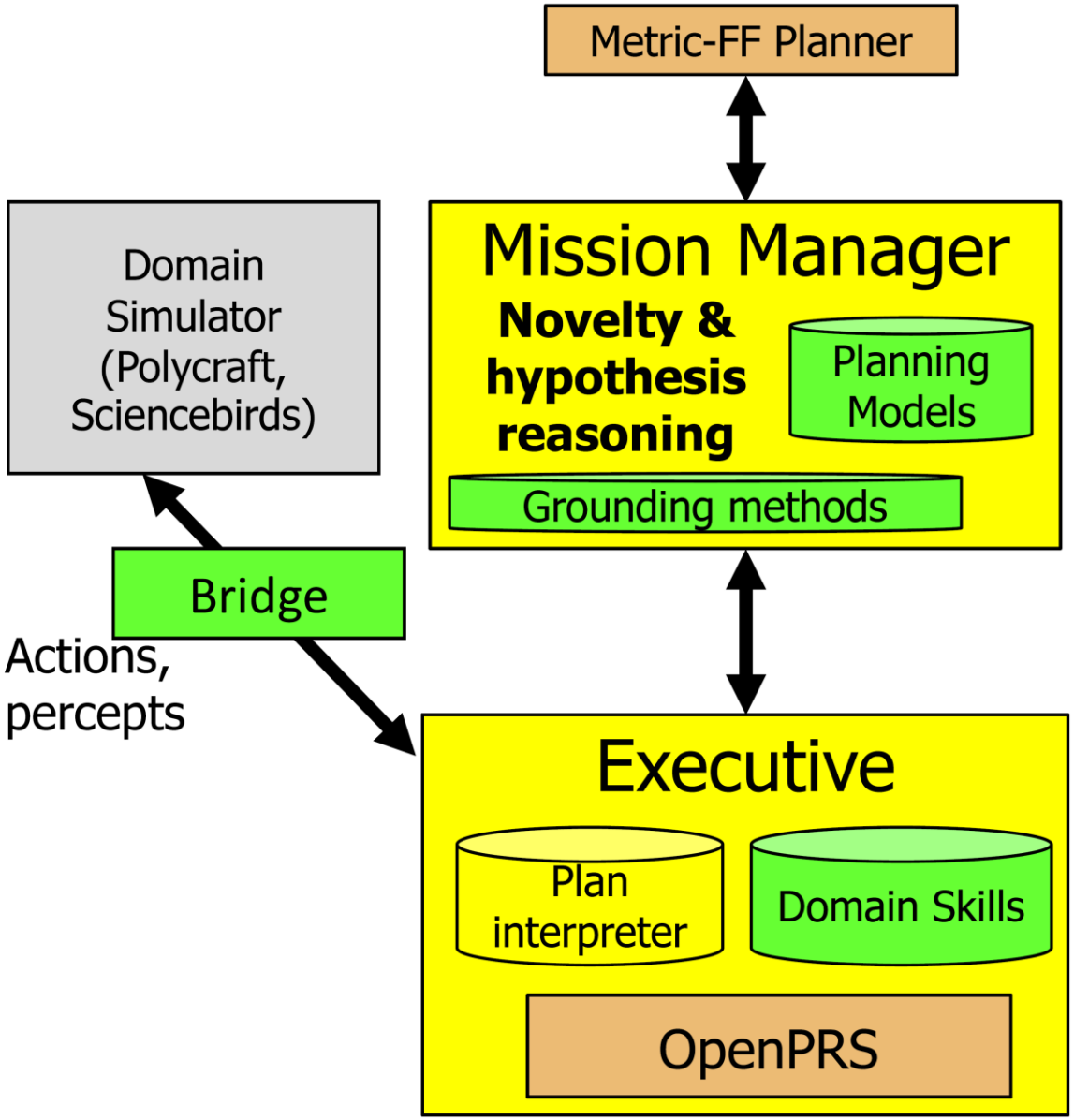
SIFT

What's Novel in the Research?

- Not the architecture– a fairly common planning and execution scheme, with meta-control.
- Not the planner– Metric-ff, currently.
- Not the adaptation approach: adapt planning models and goals.

- (1) New domain-independent heuristics for model and goal adaptation.
- (2) Single-blind evaluation on unrevealed novelties (a SAIL-ON program-wide innovation).
  - Simulation developers provide a very small set of example novelties of different types (novel objects, novel other-agents, etc).
  - Evaluation, performed by the sim developers, uses "unrevealed" novelties.

OpenMIND Architecture



Color key:

Domain-dependent bootstrap knowledge
Domain-independent novelty reasoning
Open source planner and executive

The Big Science Idea: Reasoning About Novelty Hypotheses

Unexpected observations

E.g.

- Planner failures.
- Plan execution failures.
- Unrecognized item class detected.
- Unrecognized item feature detected.

Novelty hypotheses

- Characterize novelty holistically.
- Assumption can produce new operators in **domain-independent** fashion.

E.g.:

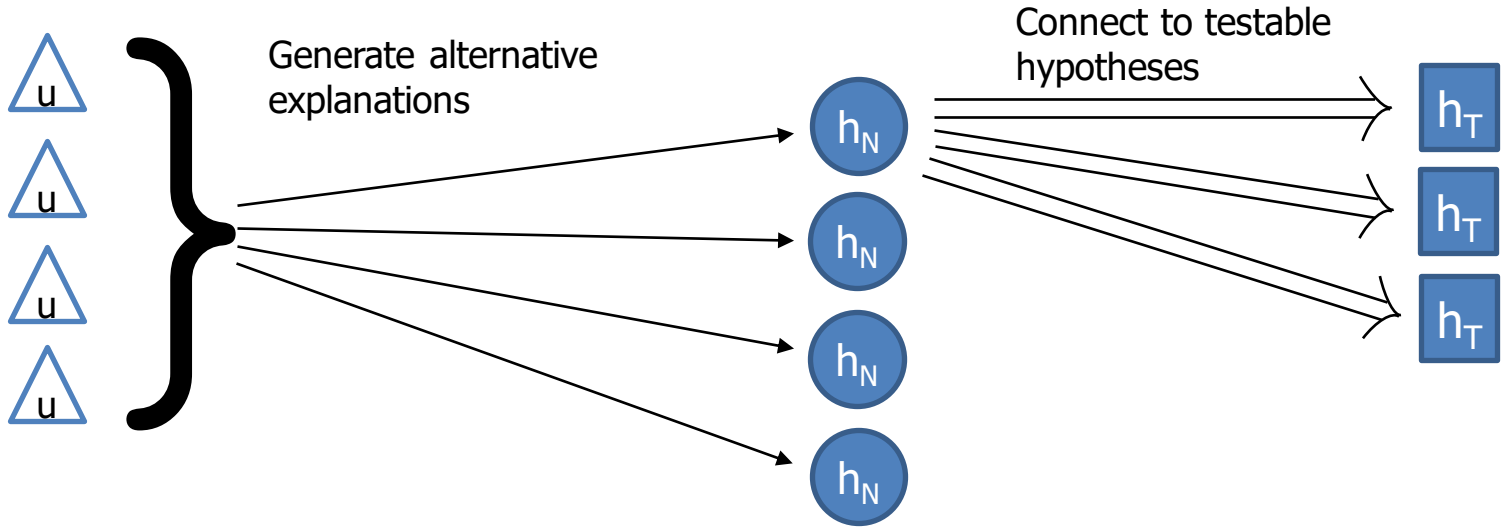
- Perception of features F on items I is transformed by T (level 3).
- Novel class C is a beneficial parameter (tool) for action A (level 1).

Testable hypotheses

- Can be validated/rejected by experimentation.

E.g.

- An operator can be executed successfully.
- One operator will have lower cost than another.
- An operator will have a particular effect.
- An operator will make it possible to create a plan, when before it was not possible.



OpenMIND

SIFT

A Spectrum of Behavior Modification Hypotheses

- Depending on type of unexpected observations.

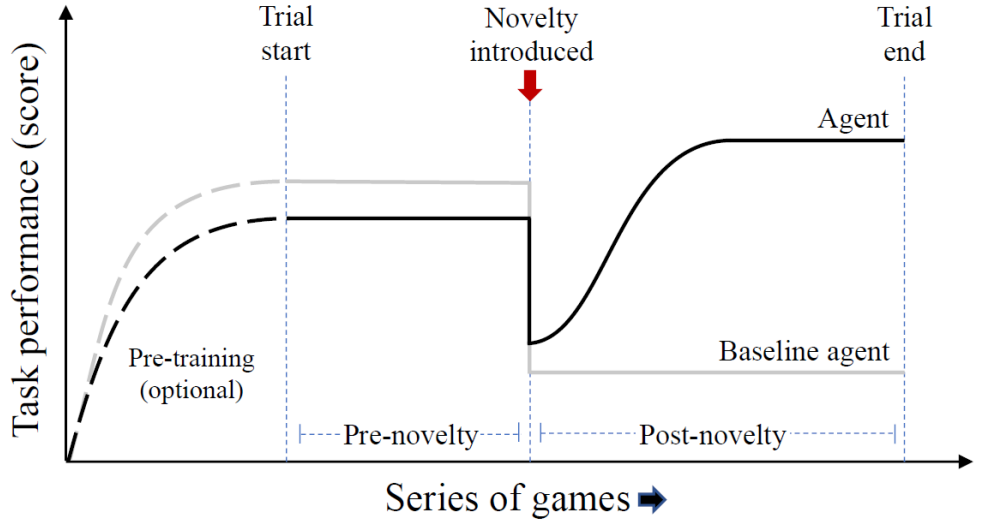
- Preferred operator – use specific parameters (e.g., an axe) in certain actions.
  - "When I break blocks with the axe, I get lower action costs".
- Prefix game goal – try to achieve an early goal (e.g., make an axe) before pursuing the usual goal.
- Bad arguments – avoid using specific parameters in certain actions.
  - "When I try to break block-27, over and over, it doesn't work…so I won't do that anymore".
- Misrepresentation – something is wrong in the percepts, try a transform (e.g., recipe rotation).
- Novelty-removal hypotheses: last-resort domain-independent tactic for handling interfering novelty.
  - "When I break fence blocks, I'm able to achieve my original goals".
- Repeat hypothesis handling – keep trying them, even if they didn't help the first time.
- Do-anything goals: last-resort domain-independent tactic for handling imperceptible novelty.
  - "I can't see anything novel, and there are no pigs to shoot…but the game isn't over… I'll shoot anything".

OpenMIND

SIFT

SAIL-ON Program-Wide Metrics

CDT = Correctly Detected Trial



Measure	Type	Definition
M1: $M_{\overline{FN}}$	Distribution Change Detection	Mean # of FNs among CDTs†
M2: $M_{\%CDT}$		% of CDTs (among all Trials)
M2.1: $M_{\%FP}$		% of Trials with at least 1 FP
M3: M_{NRP}	Performance Task	$\frac{\alpha_{post}}{\beta_{pre}}$

OpenMIND

SIFT

Months 12 and 18 Results

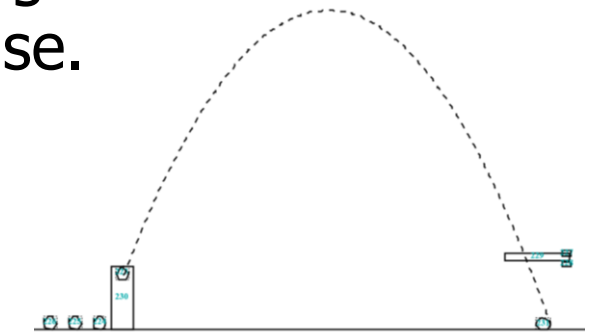
- Generally improving results.
- Removing semi-random false-positive detector in Sciencebirds reduced CDT (bad) but also FP% (good).

	FN_CDT		CDT_%		FP_%	
	M12	M18	M12	M18	M12	M18
Polycraft	2.4	0.6	82.0%	88.5%	4.6%	0.0%
Sciencebirds	27.8	23.0	68.3%	27.2%	12.6%	0.6%

	NRP	
	M12	M18
Polycraft	81.7%	81.0%
Sciencebirds	84.0%	182.8%

Insights, Lessons

- Fairly simple, completely domain-independent, and somewhat syntactic modifications to planning operators lead to routinely-effective changes in behavior.
  - This really is a validation of the notion of symbolic planning and the representations it uses– the whole idea is to have declarative models support goal-achieving behavior.
- Insight: very general knowledge-poor novelty-handling strategies are feasible and effective, without explanatory hypotheses or models.
- Insight: some stochastic behavior by our agent would mitigate some forms of obstructive novelties, without detection or focused response.
  - E.g., Sciencebirds "awning" instance.

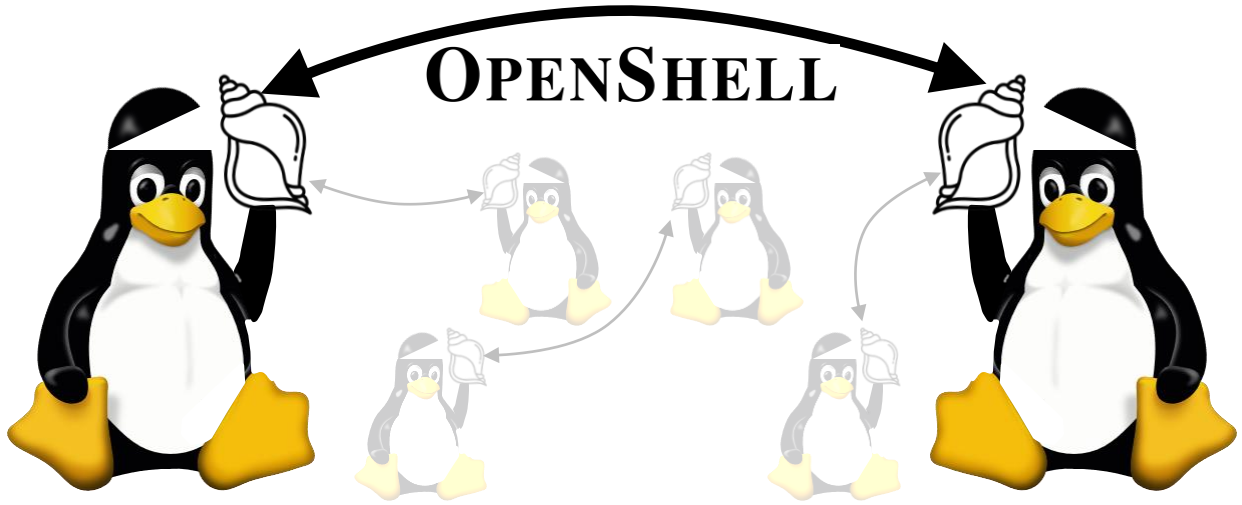


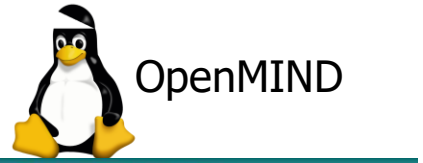
- Insight: in a novelty-biased evaluation, false positives were beneficial to CDT.
- Simulation Bug or Novelty?

Future Directions

- More structured design of experiments.
- Multi-step validation/refutation: accumulation of evidence.
- Postcondition modification hypotheses (e.g., break-block results in more than one log).
- Operator success/fail probability modeling for change detection.
- Explore/exploit tradeoff: learn from domain?
- Other possible directions: extend qualitative modeling of kinematics.

- Sharing info among lifelong learners.





Questions?