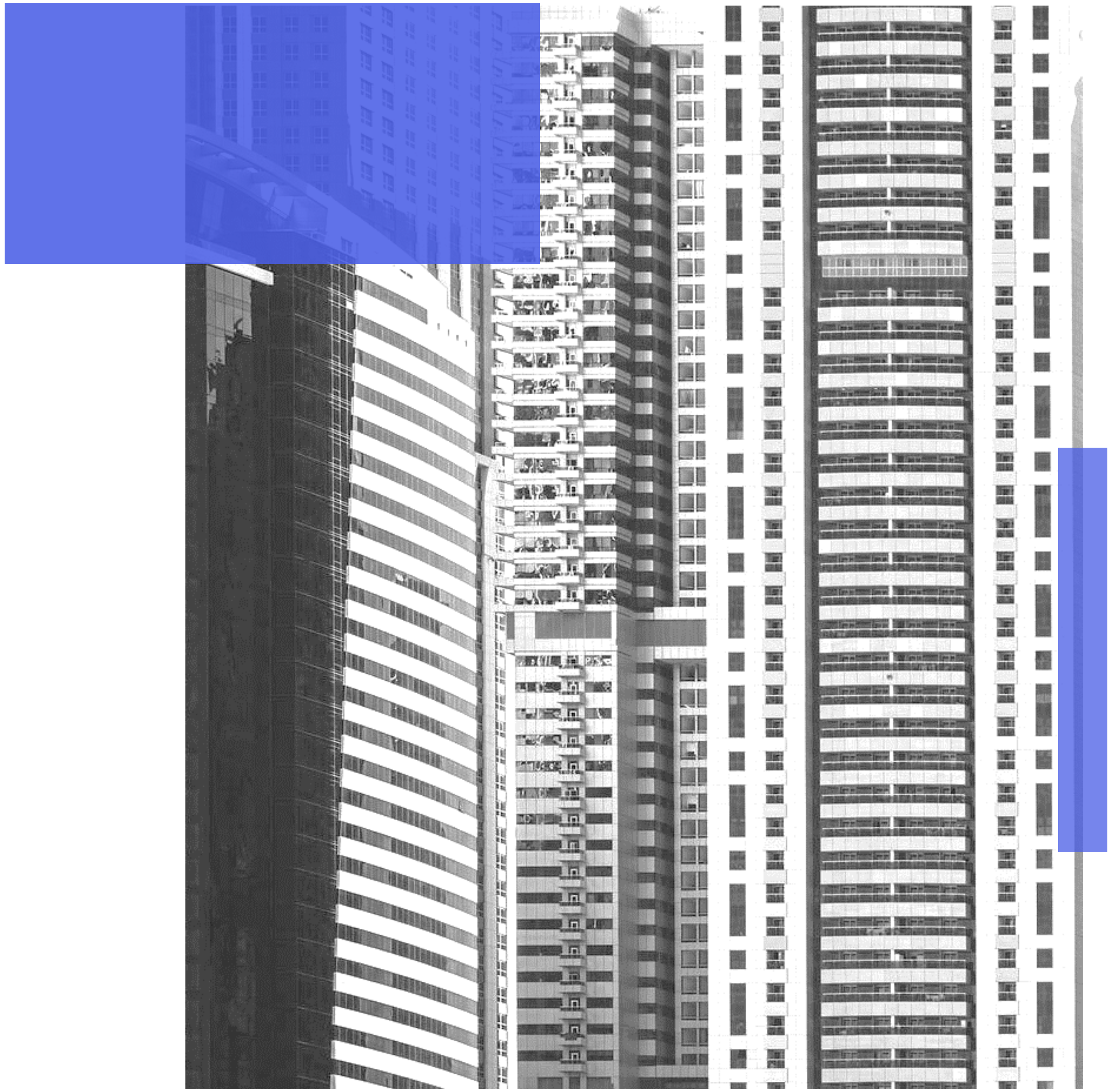
PRESENTER: MASSIMO STEFAN

WATCH-AND-HELP A CHALLENGE FOR SOCIAL PERCEPTION AND HUMAN-AI COLLABORATION



INDEX

INTRODUCTION

THE WATCH-AND-HELP CHALLENGE

THE VIRTUALHOME-SOCIAL ENVIRONMENT

THE EVAL. PROTOCOL & BENCHMARK

RESULTS

RECAP & CONCLUSIONS

QUESTIONS

INTRO

They **main contributions** are:

- The Watch-and-Help (WAH) challenge
- The VirtualHome-Social environment
- The evalutation protocol and benchmark for the challenge

❓ **Problem**: AI systems lack at recognizing others'goals and assisting them

💡 **Idea**: develop AI agents with social perception and collaborative planning abilities

🎯 **Goal**: test the social intelligence of AI agents

📄 **Social perception**: the ability to use the behavior, appearance, and communication of others to form opinions or make inferences about them, especially regarding their motives, attitudes, or values

📄 **Human-AI collaboration**: the study and practice how humans and AI agents work together to accomplish a shared goal

THE WATCH-AND-HELP CHALLENGE

WATCH-AND-HELP CHALLENGE

Goal: evaluate AI agents' social perception and their ability to collaborate with others

WATCH STAGE

The AI agent (Bob) **observes** a human-like agent (Alice) performing a task once and **infers** Alice's goal from her actions.

WATCH **stage: Bob watches Alice's behaviors and infers her goal**

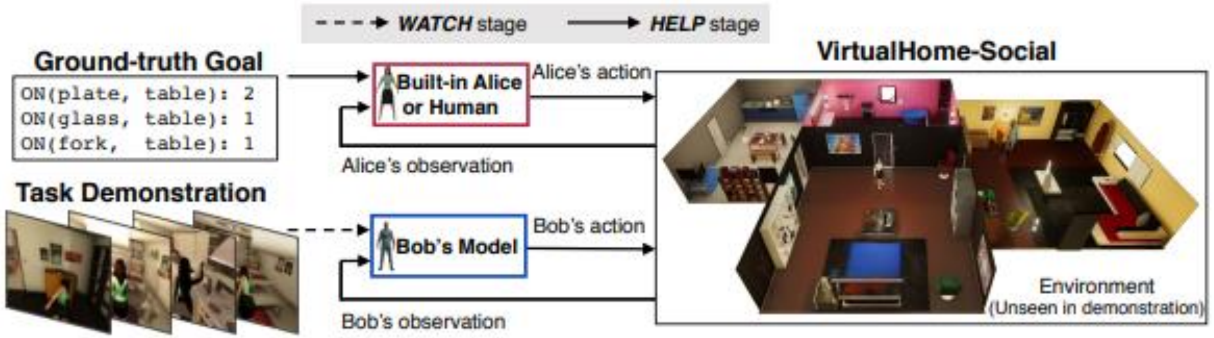
HELP STAGE

Bob **assists** Alice in achieving the same goal in a different environment as quickly as possible.

HELP **stage: Bob works with Alice to achieve her goal**

SPECIFICATIONS

- Alice is **built-in into the system** → plans her actions based on her goal and partial observation of the environment
- Bob is an **external agent** → does not know Alice's ground-truth (GT) goal and only has access to a single observation
- ONLY during training the GT goal of Alice is shown to Bob as supervision (at test time has to infer it)



PROBLEM SETUP

Each task in the challenge is defined by:

- Alice's **goal** g: a set of predicates and their count

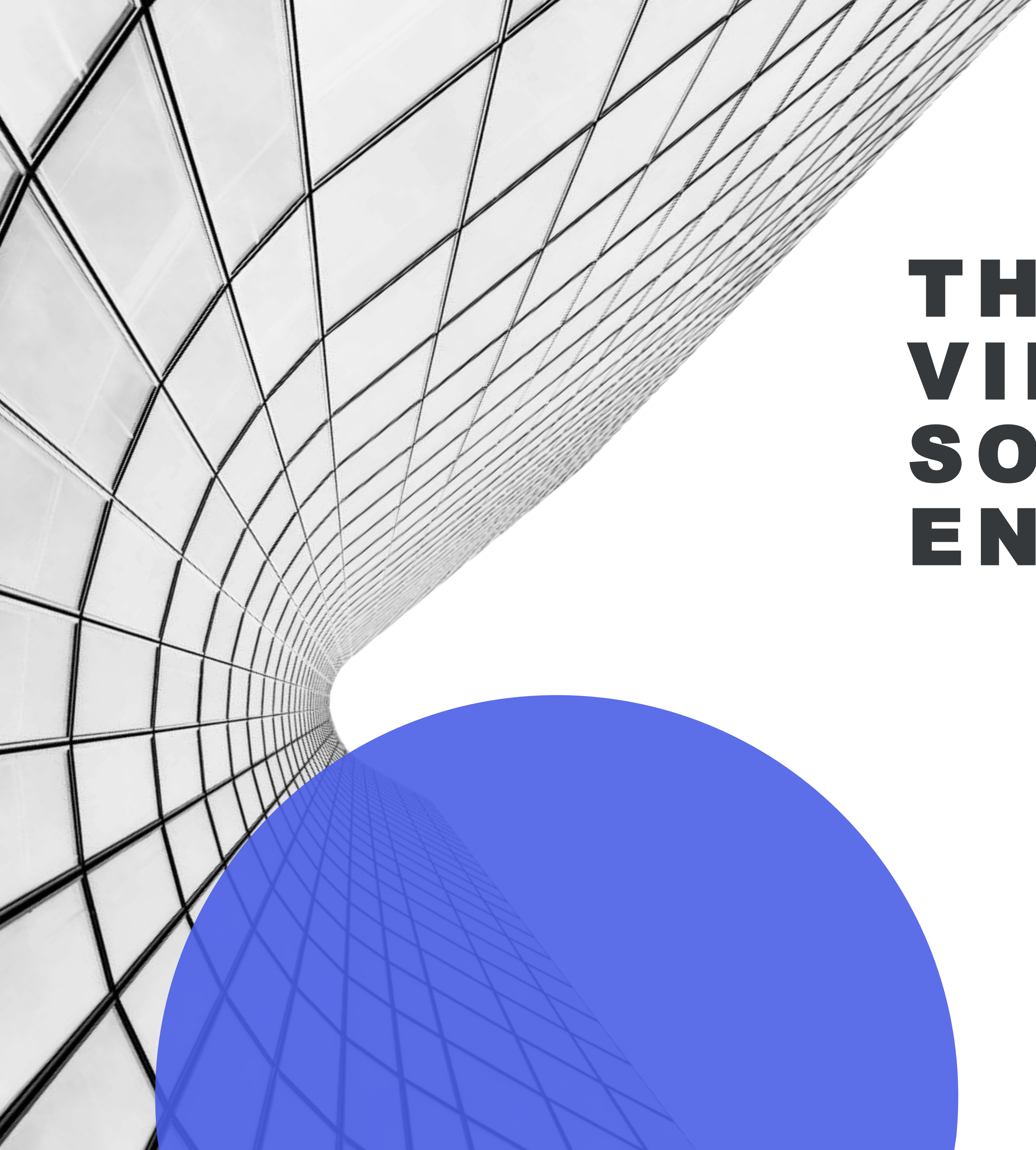
  Example: {ON(plate, dinnertable):2; ON(wineglass, dinnertable):1} → "putting 2 plates and 1 wine glass onto the dinner table"

- Alice's **actions** to achieve that goal $D = \{s_{Alice}^t, a_{Alice}^t\}_{t=1}^T$ (state, action)

- An **environment** where Bob collaborates with Alice

Note 1: the objects in a predicate refer to object classes, NOT instances

Note 2: they designed 5 predicate sets ("setting up dinner table", "wash", …)



THE VIRTUALHOME-SOCIAL ENVIRONMENT

VIRTUALHOME-SOCIAL

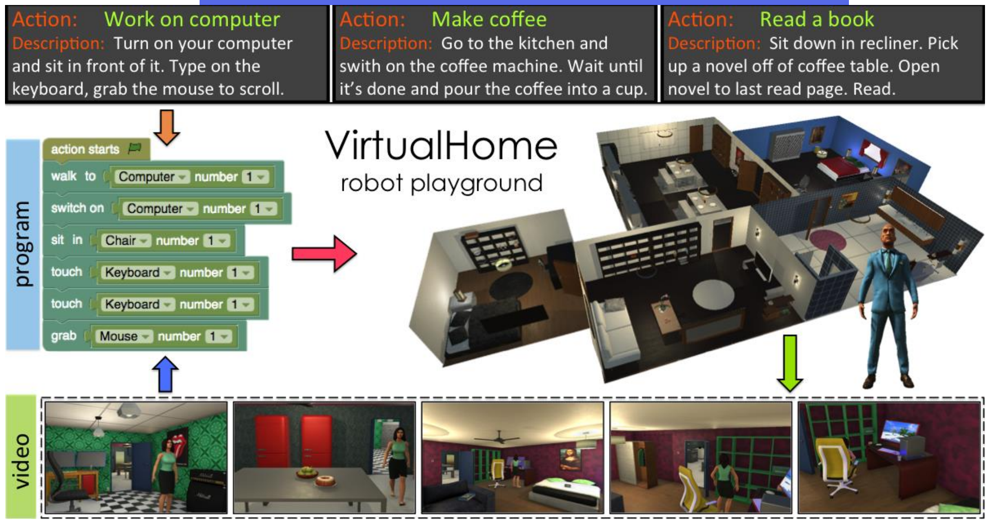
OBJECTS AND AGENTS INTERACTIONS

Simulate realistic home environments where agents can **interact** with objects and other agents.

BUILT-IN FUNCTIONS

- built-in agents that **emulate** human behaviors
- an **interface** for human players
- supports **concurrent actions** from multiple agents
- provides **observations** for the agents

Goal: allow AI agents to perform complex household tasks by interacting with objects and with built-in agents or real humans



SPECIFICATIONS

- The environment supports **symbolic** (a scene graph where nodes represent objects and edges spatial relationships between them) and **visual** observations

  ⚠ **Warning**: they do NOT specify how Bob process the visual input

- The environment provides a **large and diverse action space** for agents

  Note: it changes at every step, based on the agent's state

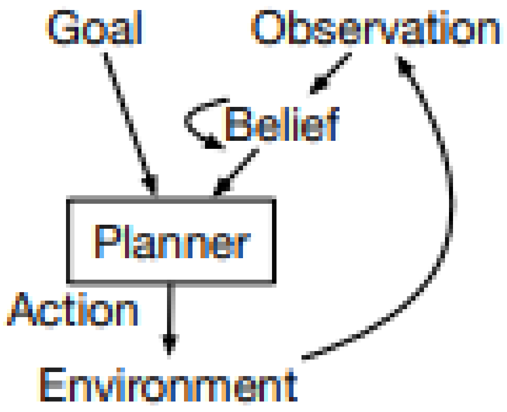
HUMAN-LIKE AGENTS

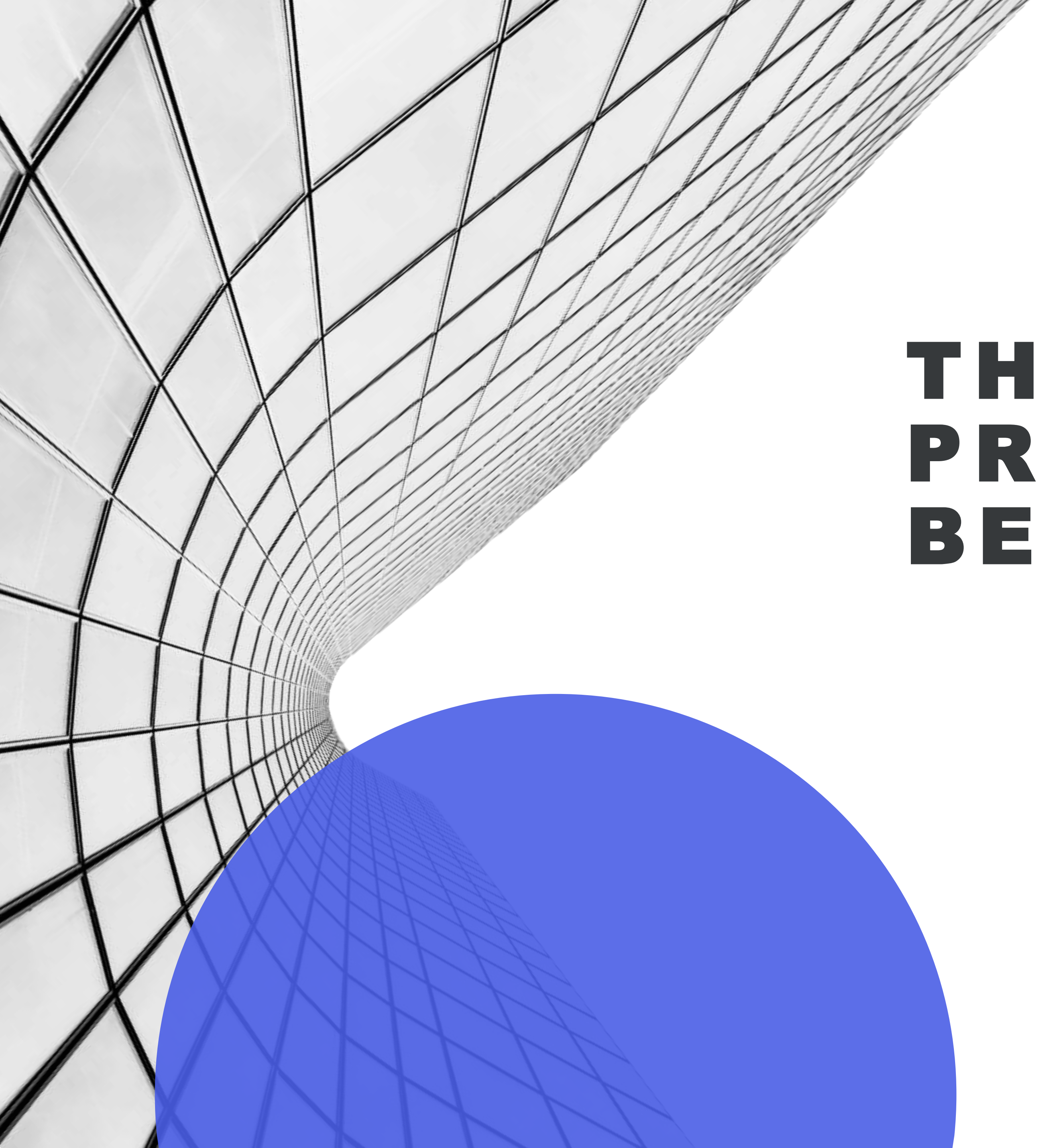
Agents operates on the symbolic representation, relying on:

1. A **belief of object locations** in the environment
2. A **hierarchical planner**, which uses Monte Carlo Tree Search (MCTS) and Regression Planning (RP)

to find a plan for a given goal based on its belief

Note: permit to plan concurrently with other agents





THE EVALUATION PROTOCOL AND BENCHMARK

EVAL. PROTOCOL AND BENCHMARK

EVALUATION PROTOCOL

Defines:

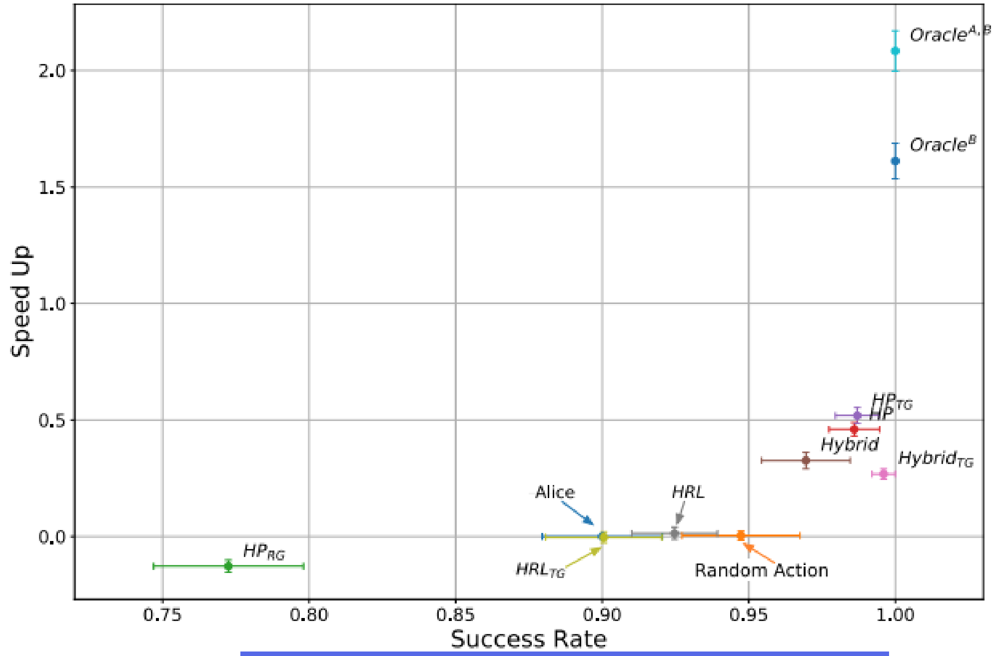
- the training and testing setup
- the goal definitions
- the metrics

of the AI agents.

BENCHMARK

Provides a set of baseline models that consist of a goal inference model and a helping policy or planner, as well as 2 oracles that show the upper bound performance.

Goal: evaluate the WAH challenge in realistic settings at scale



SPECIFICATIONS

- The training set is composed by 1011 tasks
- The test set is composed by 2 sets (test-1 and test-2) each with 100 tasks
  - With different environments (different apartments with random initial stage)
  - With different goals (predicate combinations) unseen during training
- In test-1 → all predicates in each goal are from the same predicate set

  Example: "preparing a simple meal"
- In test-2 → predicates are sampled from 2 different predicate sets

  Example: "putting groceries to the fridge and washing dishes"

EVALUATION METRICS

They used 3 types of metrics:

1. **Success rate** → the proportion of tasks where Bob successfully helps Alice achieve the goal within the given time limit

2. **Speedup** → compare the episode length when Alice and Bob are working together (L_{Help}) VS Alice working alone (L_{Alice})

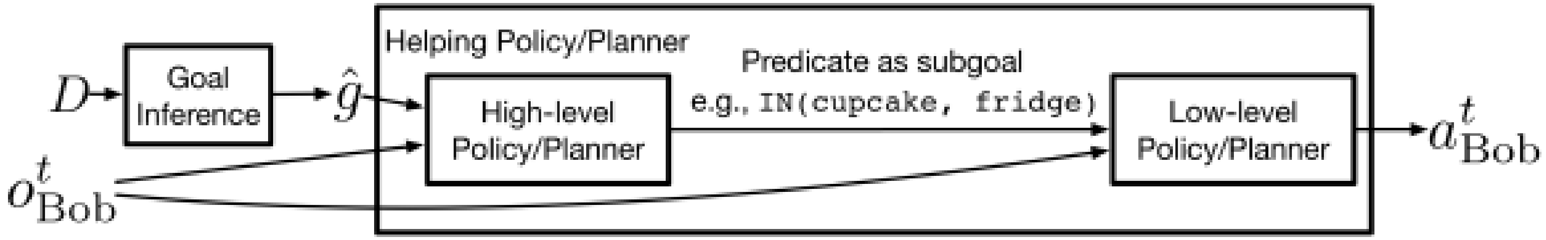
3. **Cumulative reward** of an episode with T steps → $R = \sum_{t=1}^{T}(s^t = s_g) - 0.004$ where s^t is the state at step t and s_g is the goal state

   Note: R ranges from -1 (failure) to 1 (achieving the goal in zero steps)

BASELINES

The set of baselines consist of 2 components:

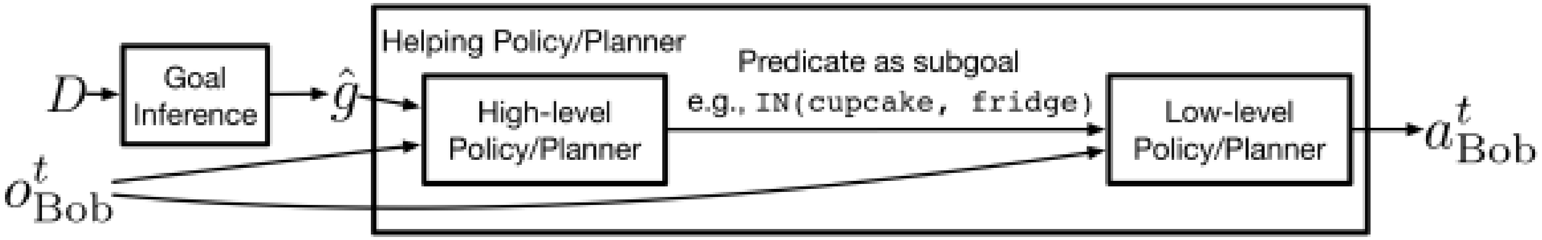
1. **Goal inference model**: trained based on the symbolic representation of states in the demonstration. At each step:
    1. Encode the state using a Transformer (over visible objects)
    2. Feed the encoded state into a LSTM
    3. Use average pooling to aggregate the latent states from the LSTM over time
    4. Build a classifier for each predicate to infer its count (**30 classifiers**)



BASELINES

The set of baselines consist of 2 components:

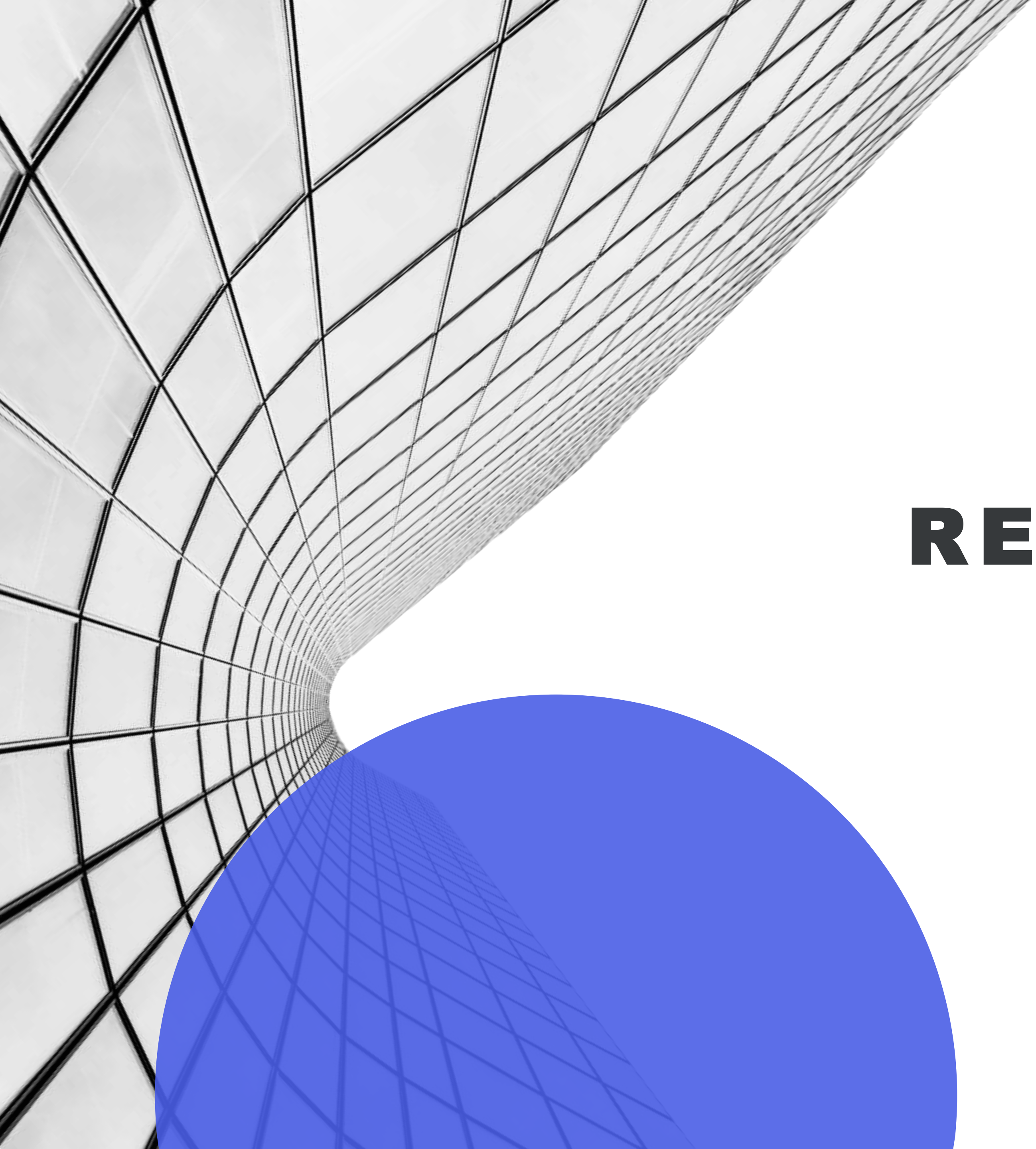
2. **Goal-conditioned helping planner/policy**: a hierarchical architecture with 2 modules for both planning and RL-based approaches. At every step:
    1. Given the Bob's observation o_{Bob}^{t} and the goal inferred from the demonstration $\hat{g}$
    2. Output a predicate s the best subgoal to pursue the current step
    3. Which will be fed to the Low-level planner/policy
    4. Resulting in the Bob's action a_{Bob}^{t}



BASELINES

The baseline models follows different policies:

- **HP**: a selects actions based on predefined rules and heuristics
- **Hybrid**: combines handcrafted policy with a NN policy traned with SL
- **HRL**: HL policy selects subgoals, LL policy selects actions to achieve them
- **Random**: selects actions uniformly at random
- **Oracle**^{B}: selects the best action according to the GT reward function
- **Oracle**^{A,B}: selects the best action according to the GT reward function and the GT truth user state



RESULTS

🏆 TEST-1 RESULTS 🏆

🎯 **Goal 1**: evaluate the 👀 stage (predict the goal predicates)

🔍 **How**: measuring the recognition performance of the predicates

📈 **Results**: precision of **0.85** and recall of **0.96** (👍)

🗒 **Note 1**: performs better when it sees the full demonstration rather than just the last observation of the actions taken by Alice

👁 **Remark**: the model can learn from the temporal and causal structure of the demonstration, not just from the final outcome

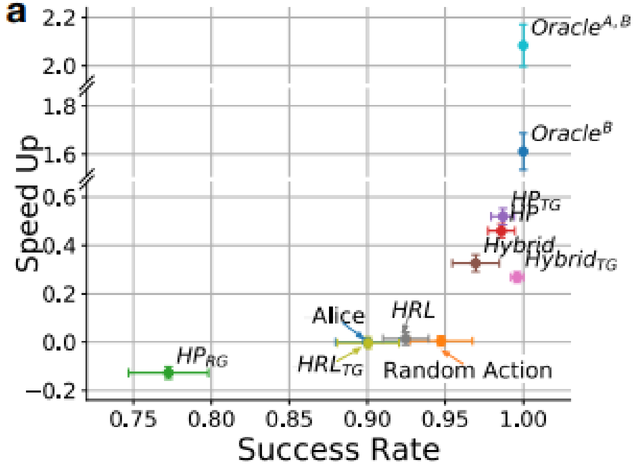
🗒 **Note 2**: the chance precision and recall is 0.08 and 0.09

TEST-1 RESULTS

Goal 2: evaluate the SOS stage (predict the goal predicates)

Results:

- Planning-based approaches (**Oracle**, **HP**, **Hybrid**) are the most effective
- DRL does not work well for Bob (fails to achieve the subgoals)
- No significant difference between **Random actions** and Alice working alone
- Using the inferred goals (**HP** and **Hybrid**) works well
- But, with a random goal inference (**HP_{RG}**) becomes counter productive



🏆 TEST-2 RESULTS 🏆

🎯 **Goal 1**: evaluate the 👀 stage (predict the goal predicates)

🔍 **How**: measuring the recognition performance of the predicates

📈 **Results**: precision of **0.68** and recall of **0.64** (👎)

📝 **Note**: the model fails to generalize to multi-activity scenarios, overfitting to prediate combinations seen during training

🏆 TEST-2 RESULTS 🏆

🎯 **Goal 2**: evaluate the 🆘 stage (predict the goal predicates)

🔍 **How**: evaluate the performance of Alice alone and **HP** (the best performing baseline)

📈 **Results**:

- Alice alone achieves a success rate of **95.40** $\pm$ **0.01**
- **HP** achieves
    - a success rate of **88.60** $\pm$ **0.02**
    - a speedup of **0.21** $\pm$ **0.04**

THANK YOU

Massimo Stefan

Paper pdf and GitHub links

Questions link