Biophysical Modeling of Satisficing Control Strategies as Derived from Quantification of Primate Brain Activity and Psychophysics

Silvia Ferrari[*], Marc Sommer[†], and Tobias Egner[#]
Pingping Zhu[*] and Hanna Oh[#]

[*]Mechanical Engineering and Materials Science
[†]Biomedical Engineering
[#]Psychology and Neuroscience
Duke University

ONR Satisficing Virtual Program Review
PM: Marc Steinberg
ONR Code 351

April 16, 2014

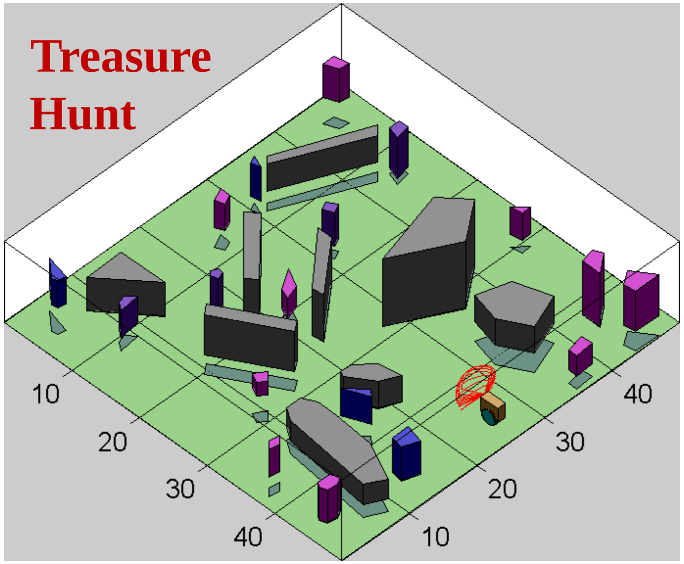
Introduction and Motivation

Satisficing in Humans and Primates

- *Bounded* rationality
- Accessing stored solutions
- Satisficing strategies for information gathering
- *Ecological* rationality – adaptive decision-making and heuristics

Satisficing Decision Strategies

- Noisy and incomplete information
- Insufficient intelligence
- Time pressure
- Utility cannot be maximized
- Consequences and probabilities may not be learned



Technical Accomplishments – Year 1

- Demonstrated systematic framework for biophysical modeling of satisficing strategies for inferential decision making
- Developed graphical model (BN) formulation of *passive* satisficing search problems
- Demonstrated integrated approach to multidisciplinary human studies and experiments via online crowdsourcing (www.mturk.com)
- Completed human and non-human primate (NHP) behavioral studies on passive satisficing task
- Developed satisficing strategies for passive search problems with time pressure / limited data / incomplete probabilistic model
- Developed human-inspired *anytime* BN learning algorithm
- Tested BN learning algorithm with empirical and virtual datasets

Technical Accomplishments – Year 2

- Developed a graphical model formulation and Webot software environment for *active and dynamic* satisficing search problems
- Started to establish the biophysical basis of satisficing using fMRI in humans and single neuron recordings in NHPs
- Developed adaptive evidence accumulation (AEA) algorithm for satisficing decision making under time pressure and with limited data
- Tested and demonstrated AEA effectiveness on empirical and virtual benchmark datasets

Outcomes and Deliverables:

- ✓ Two journal manuscripts in preparation/review
- ✓ Three conference abstracts
- ✓ Immersive environment Webot-based software

Approach

Approach: Biophysical Modeling

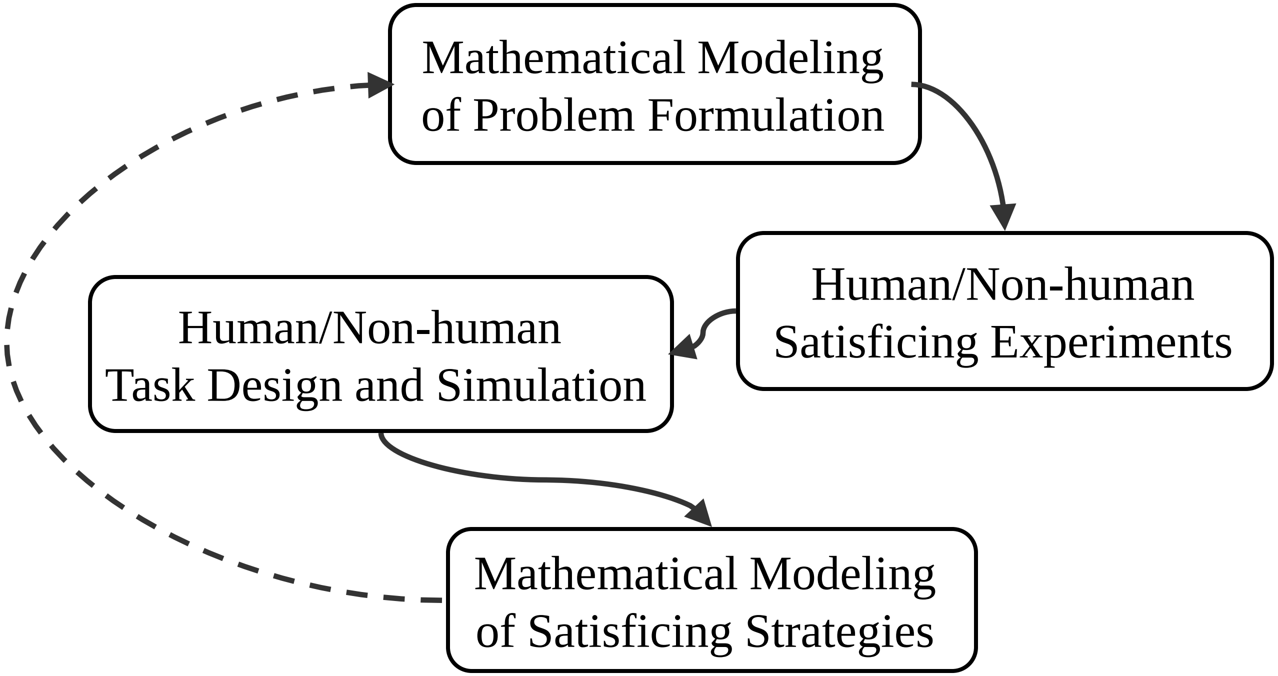


figure from Jiang & Egner, 2014

Experiments:

I. Behavioral Studies

II. Human (fMRI) Neuro-imaging and –stimulation Studies

III. Non-human Primates Neural Recording Studies

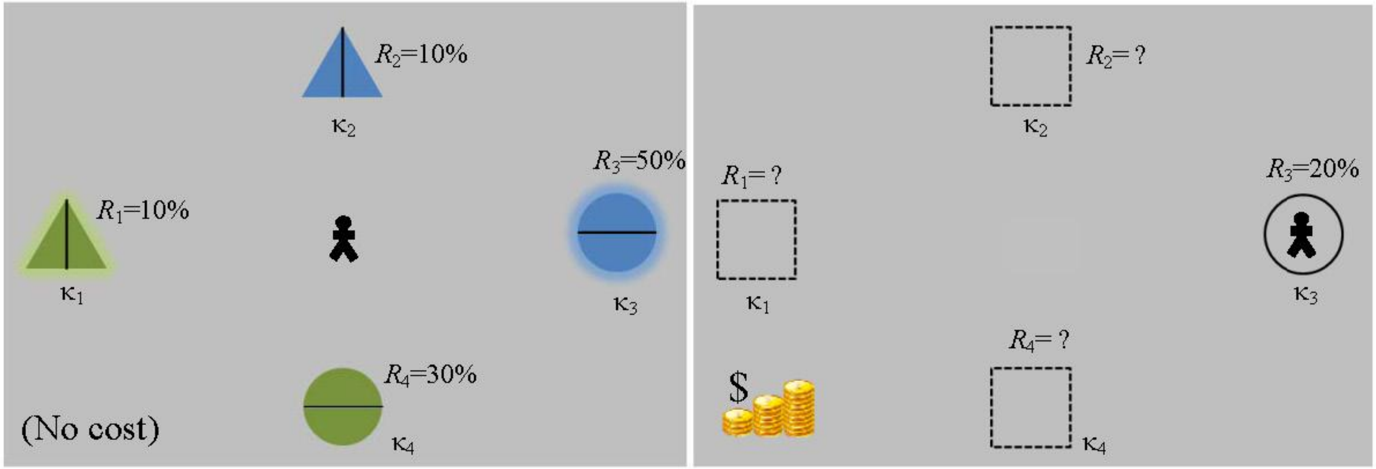
Satisficing Tasks and Experiments

1. Passive Satisficing

Learn and apply predictive information from given cues

2. Active Satisficing

Active information gathering and decision making

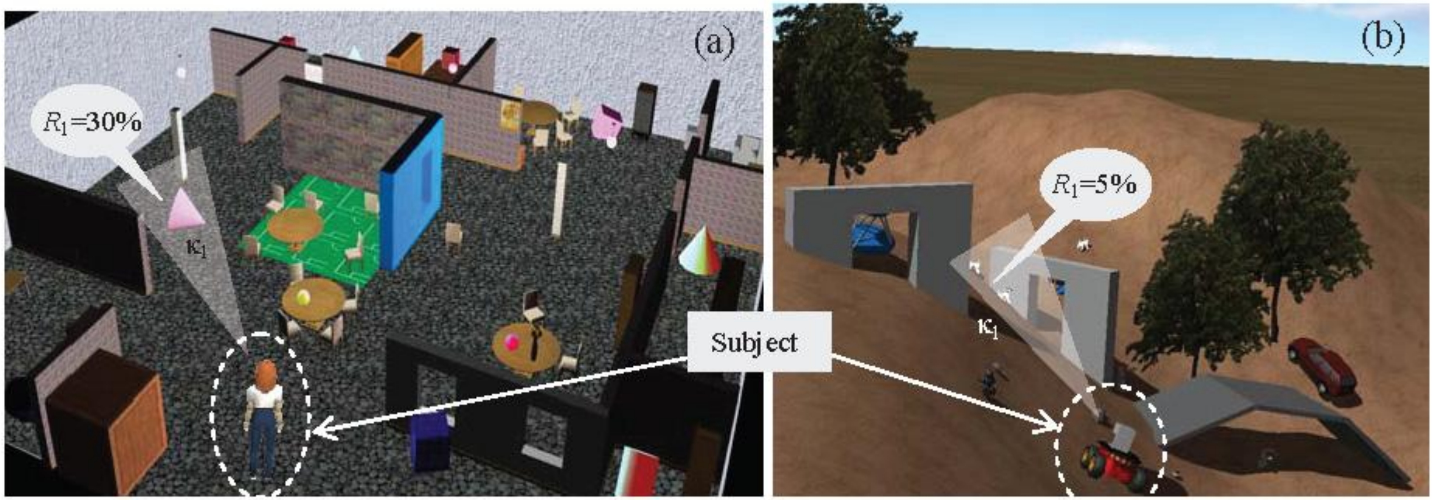


- Limited information
- Time pressure
- Cue redundancy

Satisficing Tasks and Experiments

3. Immersive Satisficing (DiVE – ML2VR)

Navigate and discover multiple treasures in complex environments



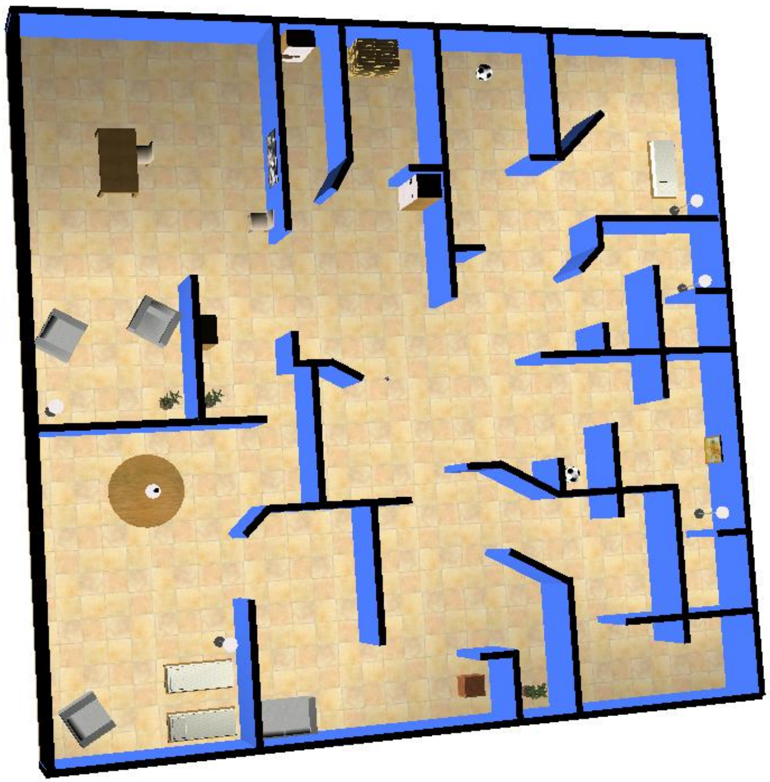
- Non-compensatory cues
- Risk and environmental pressures
- Semantically meaningful cues (context)
- Interrupted sensory signals
- Bounded controls

Immersive Satisficing Tasks and Experiments

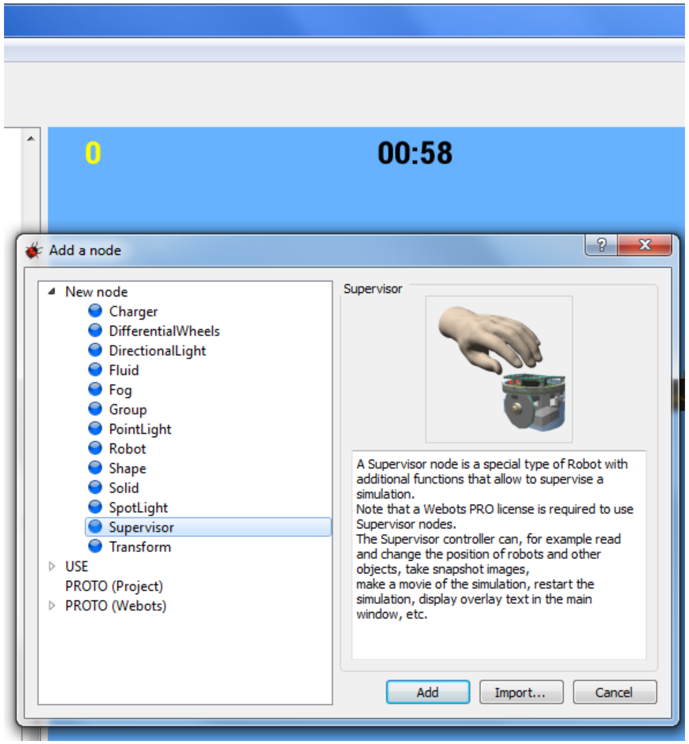
Immersive Environment

- *Active* and *dynamic* satisficing tasks: treasure hunt
- Ordering constraints on cue discovery (satisficing tree)
- Complex environment (clutter, obstacles, limited visibility…)

3D Indoor Navigation:

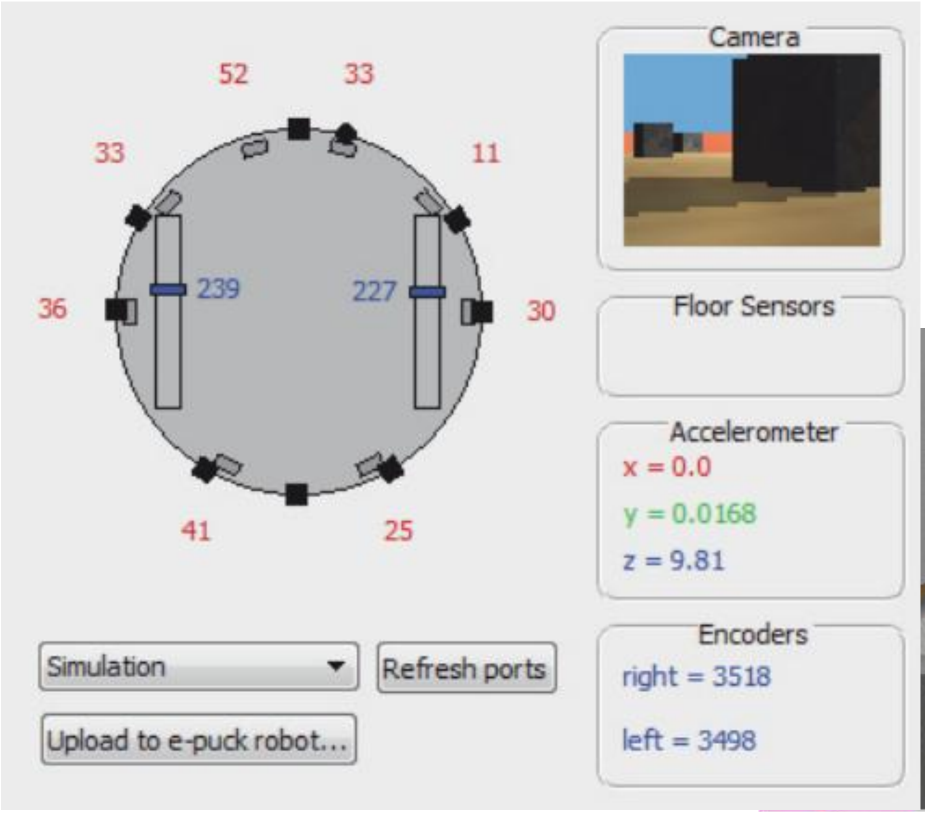


Supervisor Node for Exploration:



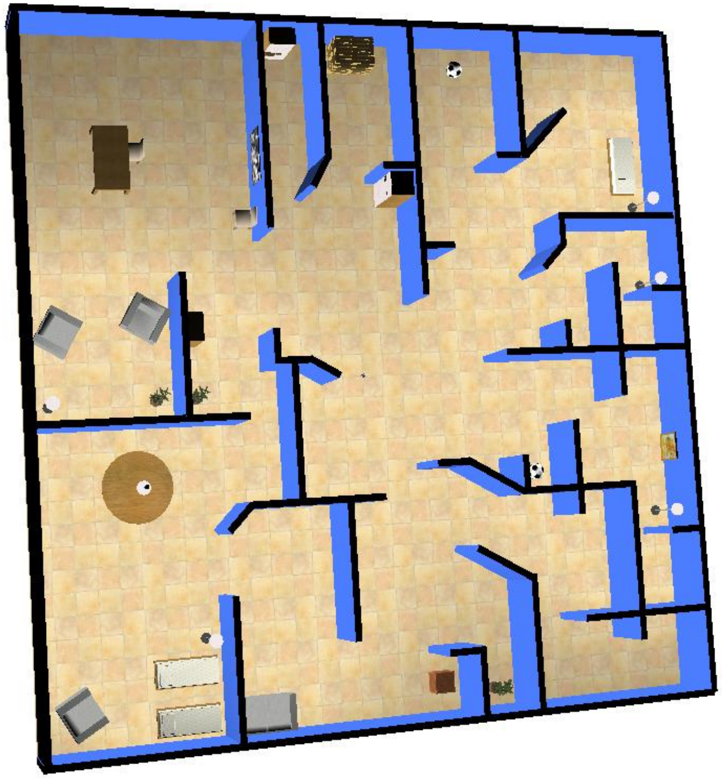
Immersive Satisficing Environment via Webot-software

Immersive Environment Robotic Capabilities



- **MATLAB® Interface**
- **Robots and sensors simulations**
- **3D navigation and perception**

Immersive Environment

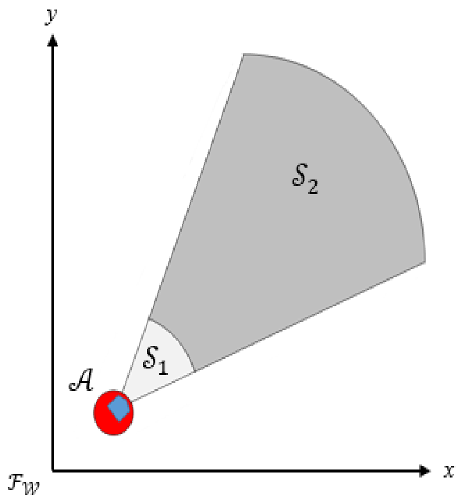


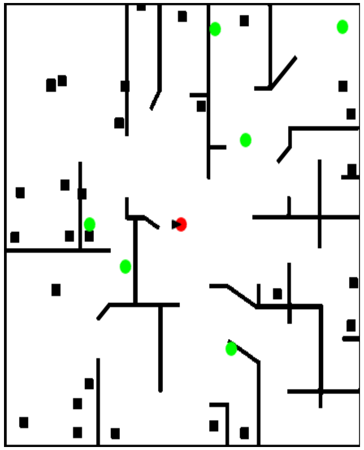
- Immersive environment is created in Webots software
- Obstacles are located at the fixed locations in the environment including table, sofa, chairs, lights and so on

Problem Formulation

Problem assumption:

- Assume the immersive environment as workspace $W \subset \mathbf{R}^2$
- Obstacles $\mathbf{B} = \{B_1, \ldots, B_m\}$ are placed at fixed locations
- Targets $\mathbf{T} = \{T_1, \ldots, T_n\}$ are placed at random locations at each trial
- A robotic sensor with platform geometry A, camera FOV geometry $S_2 \subset \mathbf{R}^2$, and measurement FOV geometry $S_1 \subset S_2 \subset \mathbf{R}^2$
- Assume the map is known to the robotic sensor a priori, including estimated geometries and locations of the obstacles





Objectives:

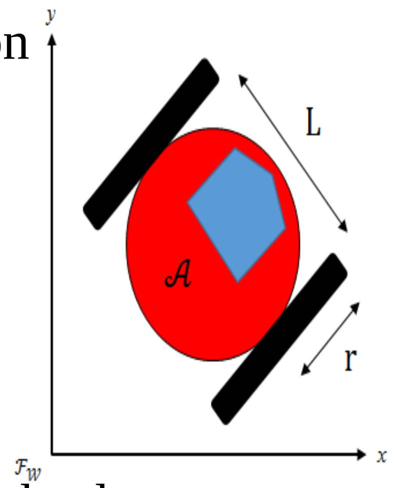
- Avoid all obstacles
- Classify all targets while:
  - Minimizing distance of robot sensor travelling
  - Maximize reward by making as many as correct classification

Robot Dynamics and Configuration Space

Robot Dynamics

- Let $\mathbf{q} = [x \quad y \quad \theta] \in W$ be a configuration vector with the position and orientation
- Let the configuration space C denote the space of all possible robot configures
- Let CB denote the set of configuration that cause collisions between the platform A and obstacles in B
- Let $C_{free} = C \setminus CB$ denote the set of configurations where no such collisions occur

The configuration vector $\mathbf{q}$ must satisfy the dynamic equation of the robotic sensor given by differetial drive kinematics



$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \frac{r}{2}(u_l + u_r)\cos\theta \\ \frac{r}{2}(u_l + u_r)\sin\theta \\ \frac{r}{L}(u_r - u_l) \end{bmatrix} = f(\mathbf{q}, \mathbf{u})$$

where r is the wheel radius and L is the distance between wheels.

The control vector $\mathbf{u} = [u_l \quad u_r]^T \in \mathbf{U}$ and control input space $\mathbf{U}$

Target Classification

Classification model assumption:

- Decision Rule:

Treasure: $P(Z_i = '\mathit{Treasure}' \mid X_{i1} \ldots X_{ik}) > P(Z_i = '\mathit{NotTreasure}' \mid X_{i1} \ldots X_{ik})$

Not Treasure: $P(Z_i = '\mathit{Treasure}' \mid X_{i1} \ldots X_{ik}) < P(Z_i = '\mathit{NotTreasure}' \mid X_{i1} \ldots X_{ik})$

- The measurements are obtained in a fixed order (sequential measurements)

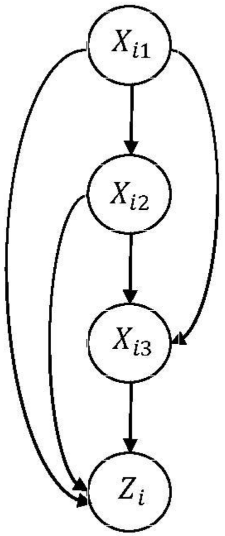
$$P(Z_i = '\mathit{Treasure}' \mid X_{i1} \ldots X_{ik})$$

$$= \frac{P(X_{i1}) P(X_{i2} \mid X_{i1}) \ldots P(X_{ik} \mid X_{i1} \ldots X_{i(k-1)}) P(Z_i = '\mathit{Treasure}' \mid X_{i1} \ldots X_{ik})}{P(X_{i1} \ldots X_{ik})}$$

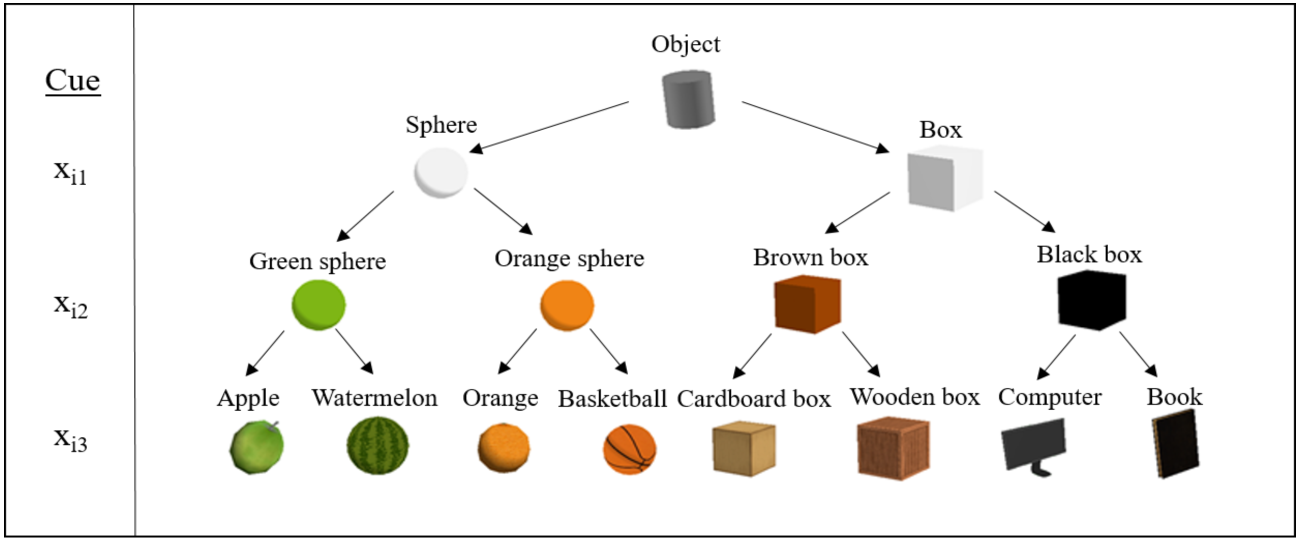
- The posterior probabilities of class are updated by more measurements
- Assume that all conditional probabilities are known

Statistics Model

Bayesian Model



Cues in layers and targets



- In each trial, targets $\mathbf{T} = \{T_1, \ldots, T_n\}$ are generated from the BN model.
- Each target T_i has K=3 layers of features/cues: $\{X_{ik}\}_{k=1,2,3}$ and classification Z_i
- Generated targets are located in the Webots environment for experiments
- Locations of targets in the workspace are generated from a uniform distribution on a set of possible locations.

Statistics Model

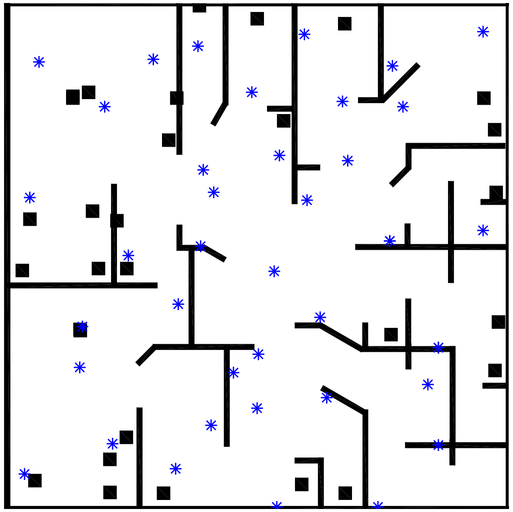
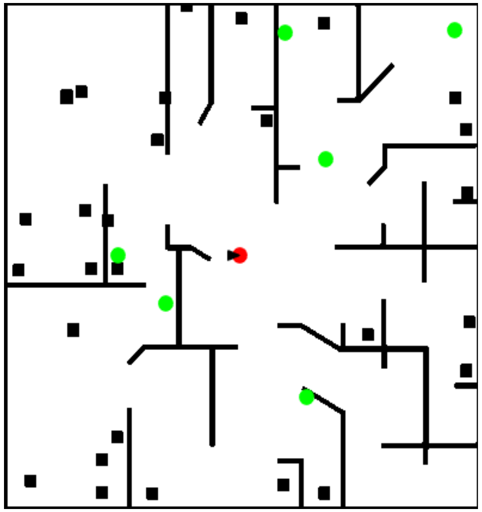
Conditional probability parameters of BN model

	Probabilities							
$P(X_{i1})$	0.49				0.51			
$P(X_{i2} \mid X_{i1})$	0.60		0.40		0.09		0.91	
$P(X_{i3} \mid X_{i1} X_{i2})$	0.98	0.02	0.65	0.35	0.46	0.54	0.86	0.14

Targets and corresponding classes

Target	Class z_i
Apple	Not Treasure
Watermelon	Treasure
Orange	Not Treasure
Basketball	Treasure
Cardboard Box	Not Treasure
Wooden Box	Not Treasure
Computer	Not Treasure
Book	Treasure

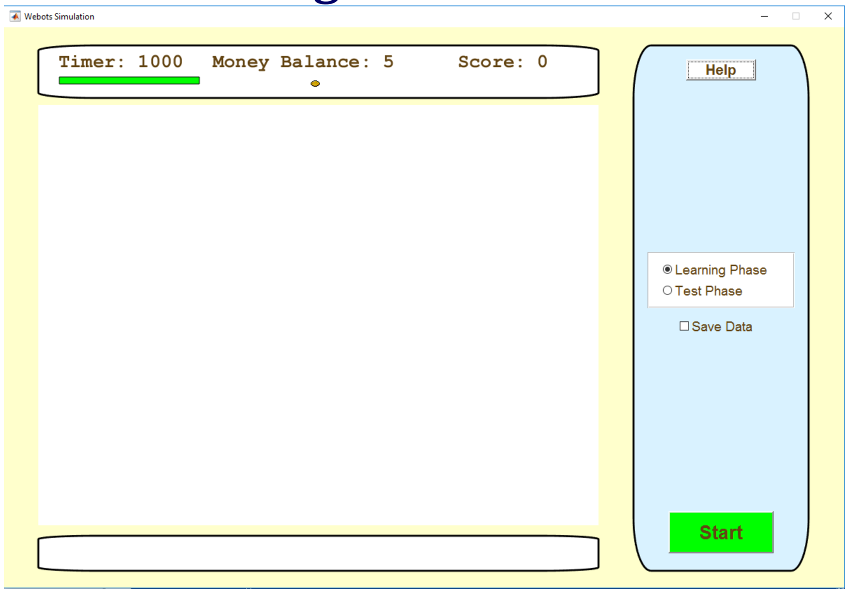
Target locations

- Possible locations of targets (blue stars) and fixed locations of obstacles (black squares) in the left figure
- Locations of targets (green circles in the right figure) in the workspace are generated from a uniform distribution on a set of possible locations.
- The robot (red circle and FOV) is initially located at the middle point in the environment.

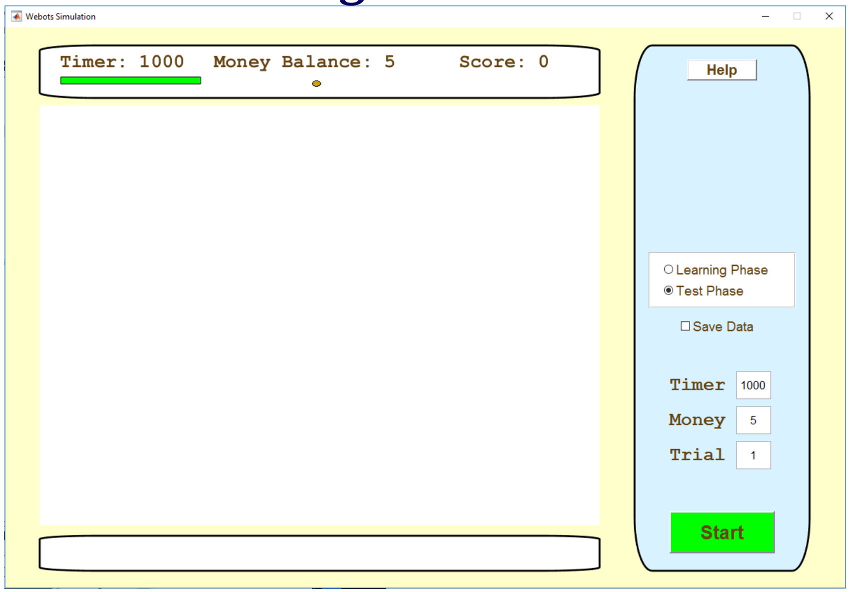
Webots GUI

Learning Phase



- No cost
- No time limitation
- No limited money balance

Testing Phase



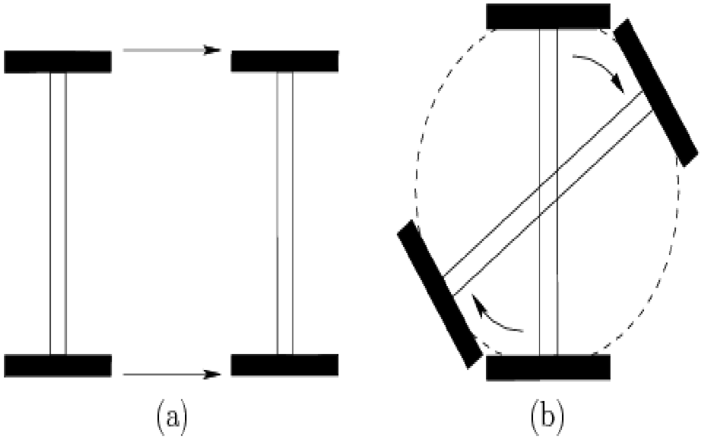
- Limited cost
- Time limitation
- Limited money balance

Robot Movements

The robotic sensor navigates the workspace w through pure translation and pure rotation. Therefore, the robot movements can be expressed by the commands “Move Forward, Move Back, Rotate Left, and Rotate Right”.

The robot state equation can be rewritten as follows:

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} r u_T \cos\theta \\ r u_T \sin\theta \\ \frac{r}{L} u_\Omega \end{bmatrix} = f(\mathbf{q}, \mathbf{u})$$



translational velocity (a): $u_T = (u_r + u_l)/2$

rotational velocity (b): $u_\Omega = u_r - u_l$

control velocity: $\mathbf{u} = [u_T, u_\Omega]^T$

Both translational and rotational velocities are ranged such as

$$u_T \in [-u_{\max}, u_{\max}] \qquad u_\Omega \in [-\omega_{\max}, \omega_{\max}]$$

Information Roadmap Deployment

Probabilistic Roadmap Method (PRM)

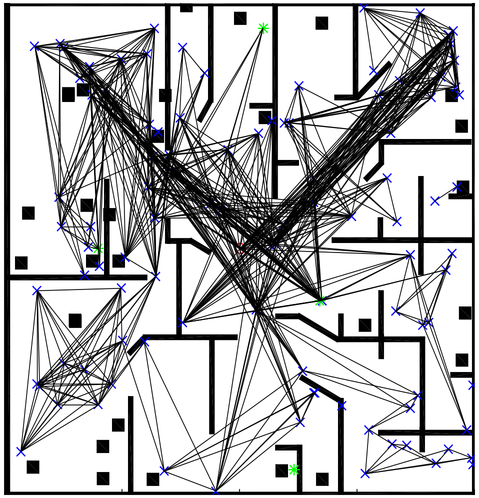
- A set of nodes or milestones $\mathbf{V} = [v_1, \ldots, v_N]$ are sampled randomly in C_{free}
- Edges $\mathbf{E} = [(v_1, v_2), \ldots, (v_{N-1}, v_N)]$ are simple local paths connecting v_i and v_j
- Nodes $\mathbf{V}$ and edges $\mathbf{E}$ form an undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ as roadmap

Nodes/milestones random sampling

Nodes are sampled using a weighted PDF π

$$\pi = w_2 \pi_V + w_1 (1 - w_2) \pi_G + (1 - w_1)(1 - w_2) \pi_U \qquad 0 \leq w_1, w_2 \leq 1$$

- A uniform PDF π_U in the workspace
- A PDF π_G for sampling narrow regions using the bridge test
- A PDF π_V for sampling regions of high information value (not applied in this project)



Path Planning

Greedy path planning

- Apply a heuristic weighted cost function

$$C(v_i, v_j) = \omega_1 D(v_i, v_j) - \omega_2 D_{avg}(v_j)$$

  - ω_1, ω_2 : cost function weights
  - $D(v_i, v_j)$: shortest path between v_i and v_j according to graph **G**
  - $D(v_j)$: node density around v_j

- Finds the path current node v_i to the next best target node v_j
  - visited nodes are remove from the node candidate list
  - shortest distance from current node
  - highest node density

Classification Strategy

Classification Strategy

- Classification confidence:

$$cc = \log \frac{P(Z_i = 'Treasure' | X_{i1} \ldots X_{ik})}{P(Z_i = 'NotTreasure' | X_{i1} \ldots X_{ik})}$$

- Classification decision conditions:
  - Sufficient confidence: $cc > Th$ Confidence threshold: Th
  - Insufficient balance: remained balance is less then the measurement cost
  - Time is up!

Immersive Satisficing Results

Approach

Experiment Setting

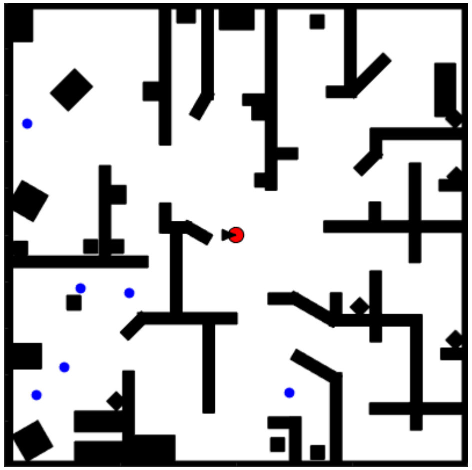
- Counting detected targets and classified targets in a fixed amount of time
- Compare the proposed path planning method with the shortest path planning method
- Vary the target configurations

Performance Parameters	
$D(\tau)$	Distance traveled along a path τ
$V(\tau)/D(\tau)$	Actual information value per distance traveled along a path τ
$R_y(\tau)$	Classification success rate along a path τ
$\eta_y(\tau)$	Number of correctly-classified targets per distance traveled
$\eta_v(\tau)$	Number of correctly-classified targets per cue purchased

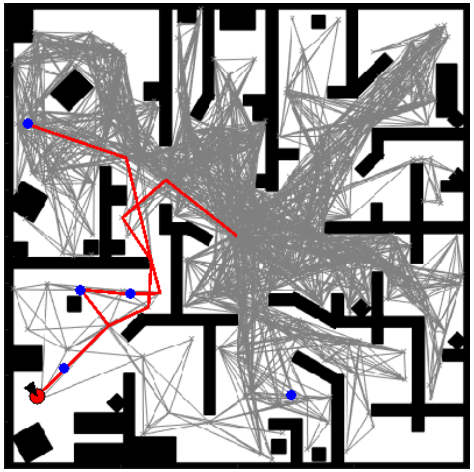
Results

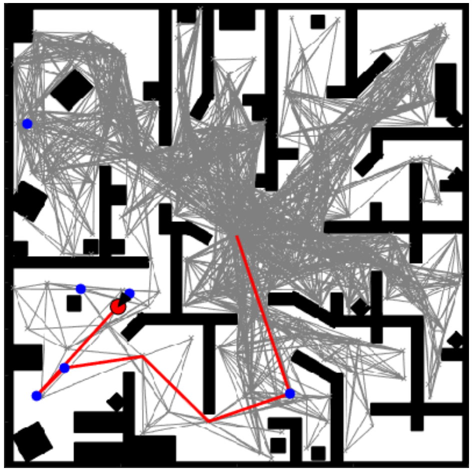
Trial #1 Cluster of targets in one region

In this setup, the proposed robot planner heads to the region of denser information value first (left), while the shortest path planner heads to the isolated target (right). Although the shortest path planner covers more distance, it is less efficient.



Performance Parameters	Proposed Planner	Shortest Path
$D(\tau)$	19.51	25.07
$V(\tau)/D(\tau)$	0.4613	0.2792
$\eta_y(\tau)$	0.2050	0.1197
$R(\tau)$	1.00	0.75
$\eta_v(\tau)$	0.4444	0.4286

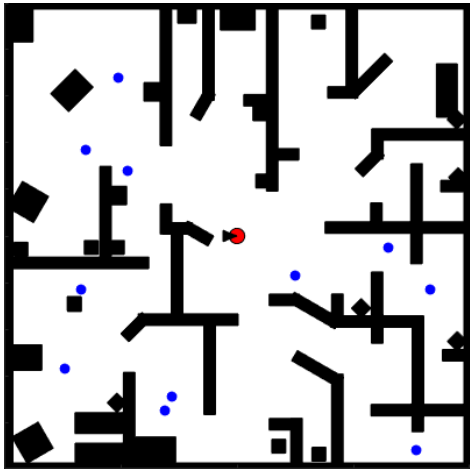




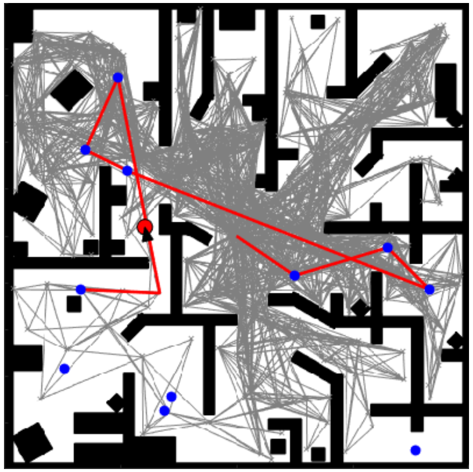
Results

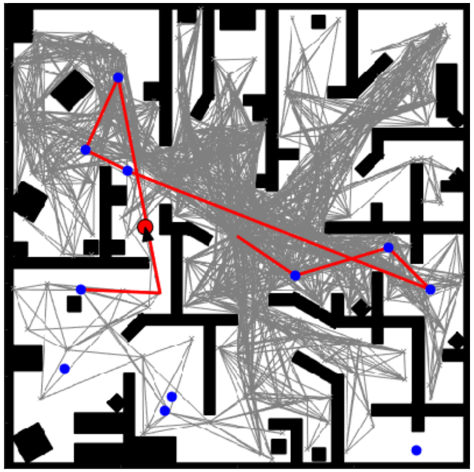
Trial #2 Small number of scattered targets

In this setup, there are no significant differences in the density of targets. Therefore, both planners give the same paths and have the same performance.

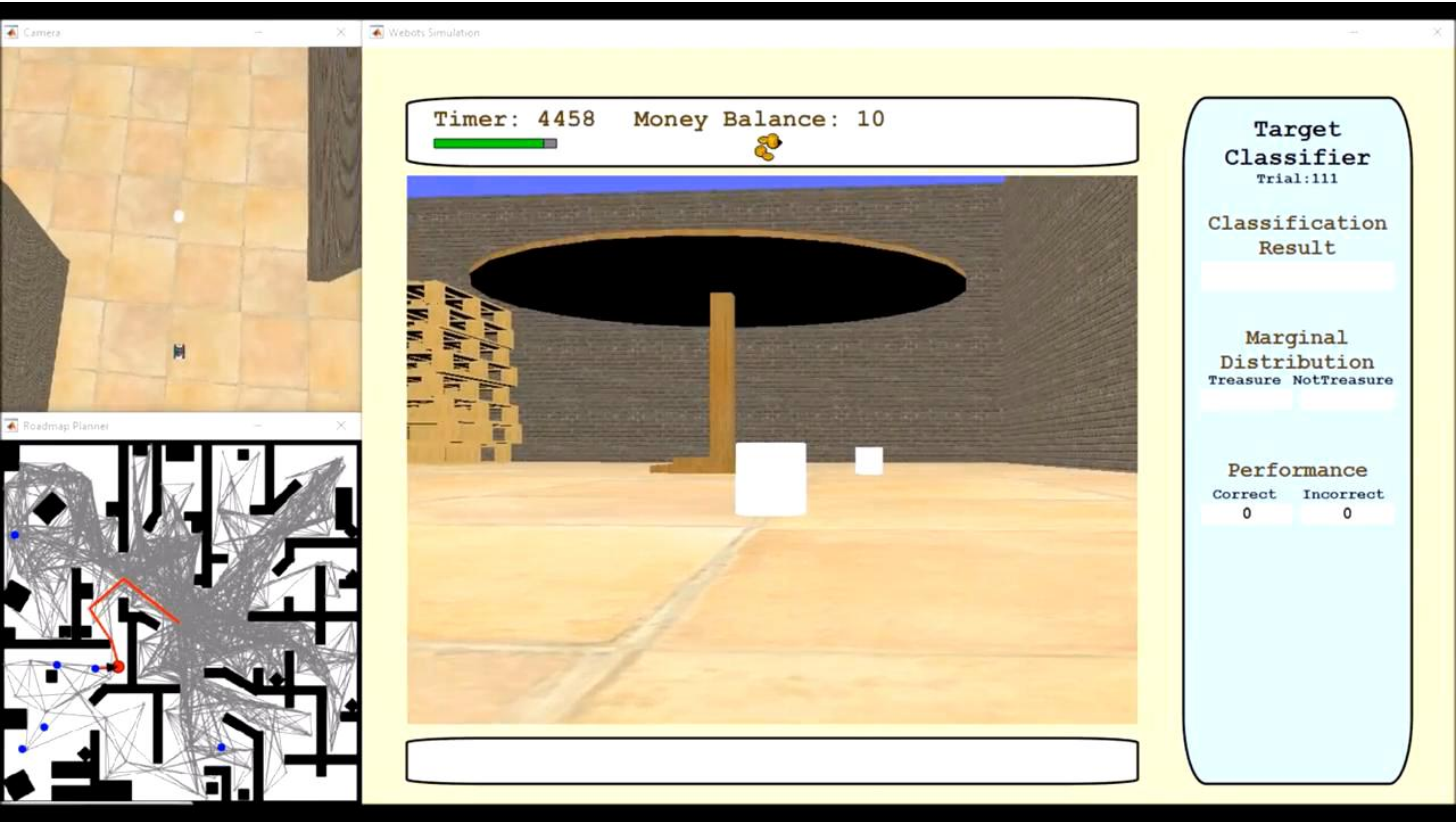


Performance Parameters	Proposed Planner	Shortest Path
$D(\tau)$	35.38	35.38
$V(\tau)/D(\tau)$	0.3675	0.3675
$\eta_y(\tau)$	0.1696	0.1696
$R(\tau)$	1.00	1.00
$\eta_v(\tau)$	0.4615	0.4615





Movie Demo


Camera
Webots Simulation
Roadmap Planner
Timer: 4458
Money Balance: 10
Target Classifier
Trial:111
Classification Result
Marginal Distribution
Treasure
NotTreasure
Performance
Correct
Incorrect
0
0

Summary (Year 2) and Future Work

- Biophysical modeling of satisficing strategies for inferential decision making
- Graphical model (BN) problem formulation
- AEA BN learning algorithm for satisficing, inferential decision making
- Time pressure / limited data / incomplete probabilistic model
- Semantic cues
- Integrate (passive satisficing) behavioral studies with:

  - Human (fMRI) neuroimaging

  - Non-human primates (NHPs) neural recordings
- Immersive environment and Webot-based software

Future Work:

- Active and dynamic satisficing task problem formulation
- Human and NHP behavioral studies (Webot - MATLAB® - DiVE)
- Mathematical modeling of active and dynamic satisficing control strategies
- Comparative studies and virtual experiments

Acknowledgements

Marc Sommer, Associate Professor
Tobias Egner, Assistant Professor
Jeff Beck, Assistant Professor
Dr. Pingping Zhu, Postdoctoral Associate
Hanna Oh, Ph.D. Student,
Vincent Prevosto, Postdoctoral Associate
Manu Raghavan, Ph.D. Student
Kimi Rafie, B.S. Student
Charlie Giattino, Ph.D. Student
Minyoung Ryoo, B.S. student
Andrew Toader, B.S. student
Jeffrey Mohl, Ph.D. student
Hrishikesh Rao, Ph.D. student

Questions?

This research was funded by the ONR Science of Autonomy Program, under grant N00014-13-1-0561.

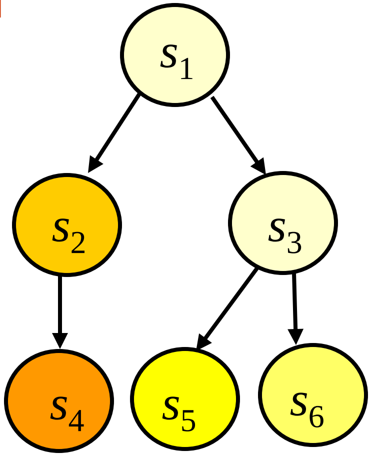
Backup Slides

Satisficing Decision and Control

Epistemic Utility Function [Curtis, Beard, Goodrich, Stirling, et al.]

- Seeks a family of control laws via model predictive control (or CLFs)
- Selectability criterion: control benefit (tracking, regulation, or stability)
- Rejectability criterion: control cost (time or control effort)

Satisficing Decision Trees [Simon, Kadane, Dieckmann, et al.]



- Inferential decision making
- Cues are predictors of performance
- Search for Spanish treasures

  - Unknown number of treasures at n sites, $s_1, \ldots, s_n$

  - Given probabilities, $p(s_i)$, and cost $\phi(s_i)$, $i = 1, \ldots, n$

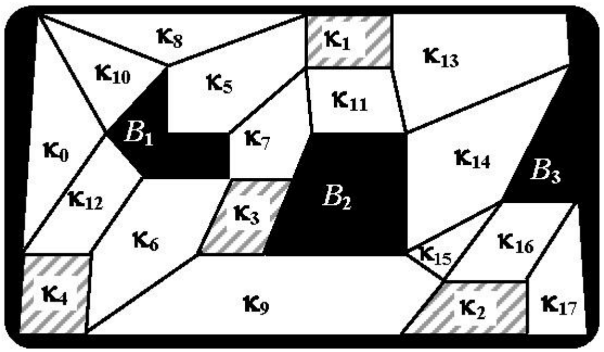
  - Find any treasure at minimum cost

- AND/OR trees, ordering constraints

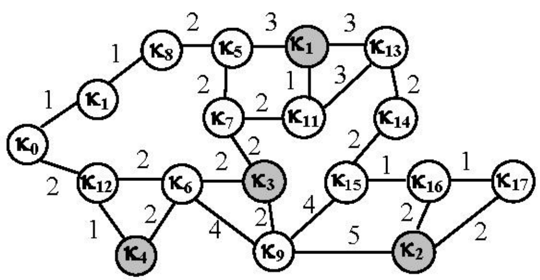
Satisficing Searches and Inferential Decision Making

Treasure Hunt Problem [Cai, Ferrari, et al.]

For a given layout $\mathcal{W} \subset \Re^3$ with r targets and n obstacles and a given joint probability mass function $p(z, \xi, \lambda)$, find the obstacle-free path that minimizes the distance traveled by a robot $\mathcal{A}$ between two configurations q_0 and q_f, and maximizes the information gain, $I(z \mid \xi, \lambda)$, for a sensor with field-of-view $\mathcal{S}$, installed on $\mathcal{A}$.

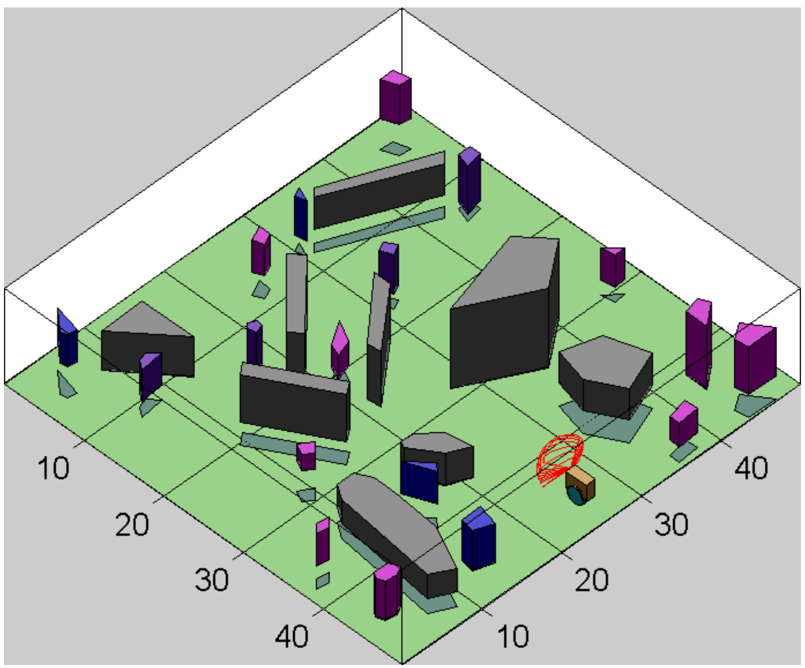


Cell Decomposition



Connectivity Graph → Graphical Model

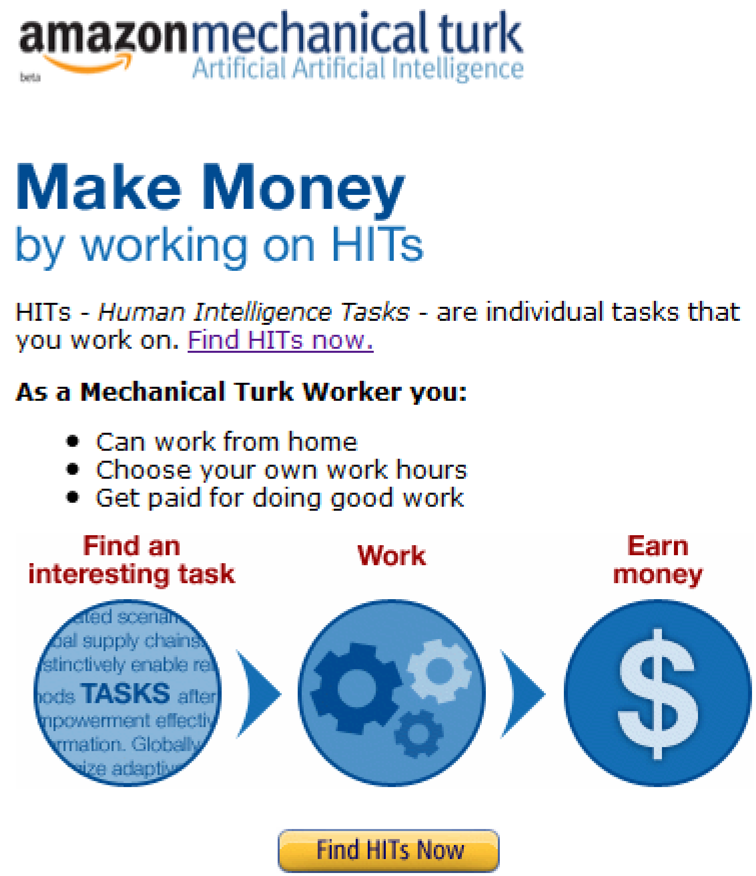
Robotic Sensor Path Planning



Human Behavioral Experiments Online

- Online crowdsourcing service
- Anonymous online subjects complete web-based tasks for small sums of money
- Advantages
  - Data can be collected quickly
  - Diverse demographics
- Empirical validation through replication using classic psychology tasks

  [Crump, McDonnell & Gureckis, 2013]



www.mturk.com