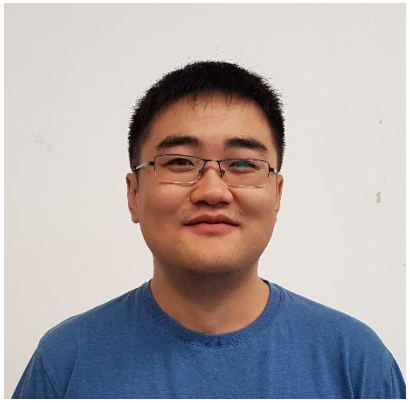
Finding the Shortest Path Using Reinforcement Learning

Group E-puck 3

Technishce Universität München

July 27, 2017

Group members

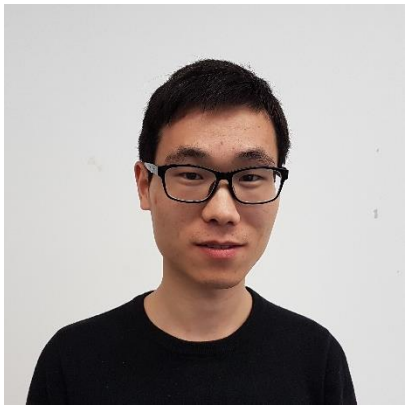


Lingfeng Zhang

- Learning Algorithm
- Modeling
- E-puck commission

Tianming Qiu

- Simulation
- Build environment
- Low-level control



Wenhan Hao

- State unit design
- Real world
- Training

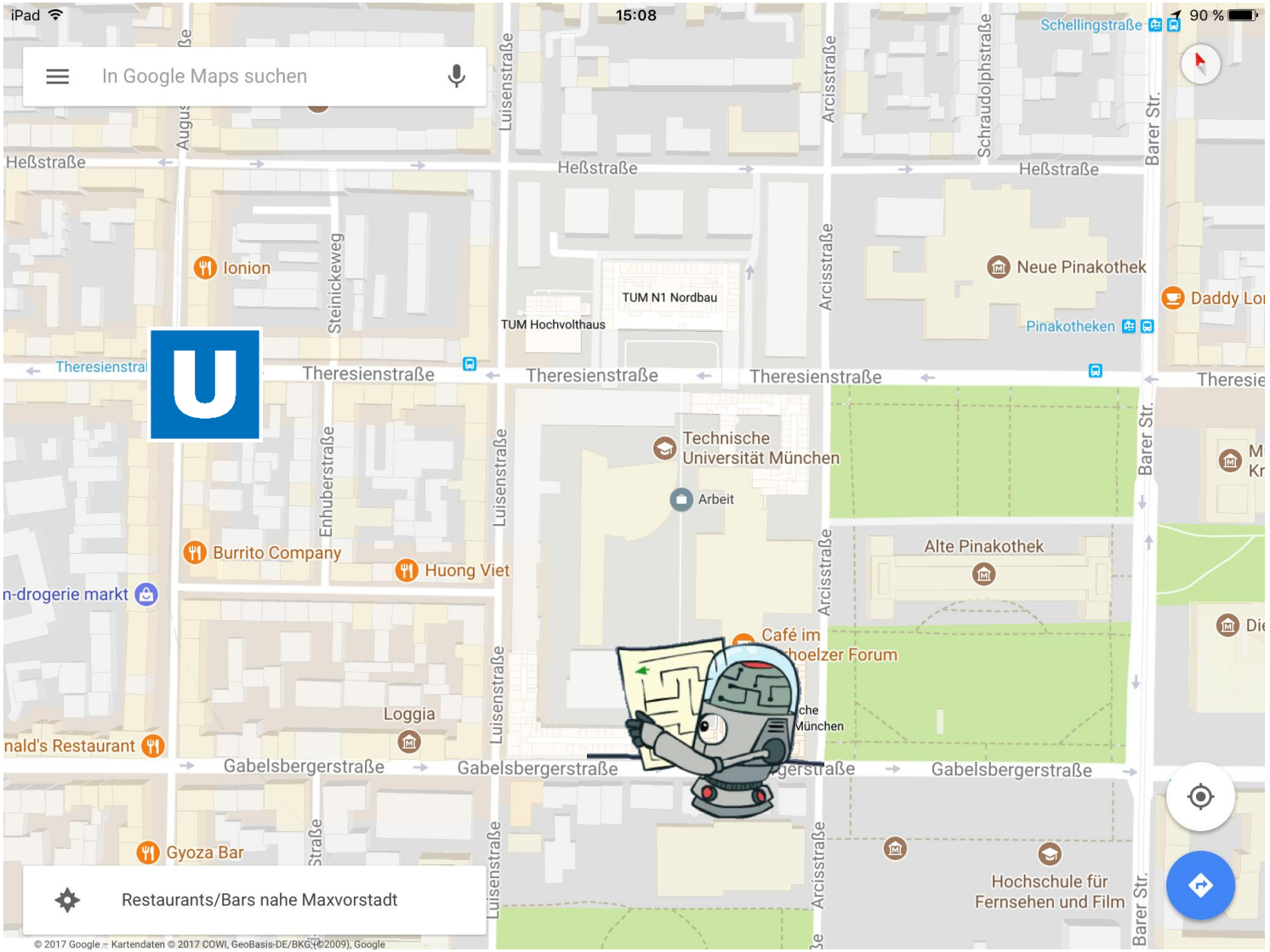
Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

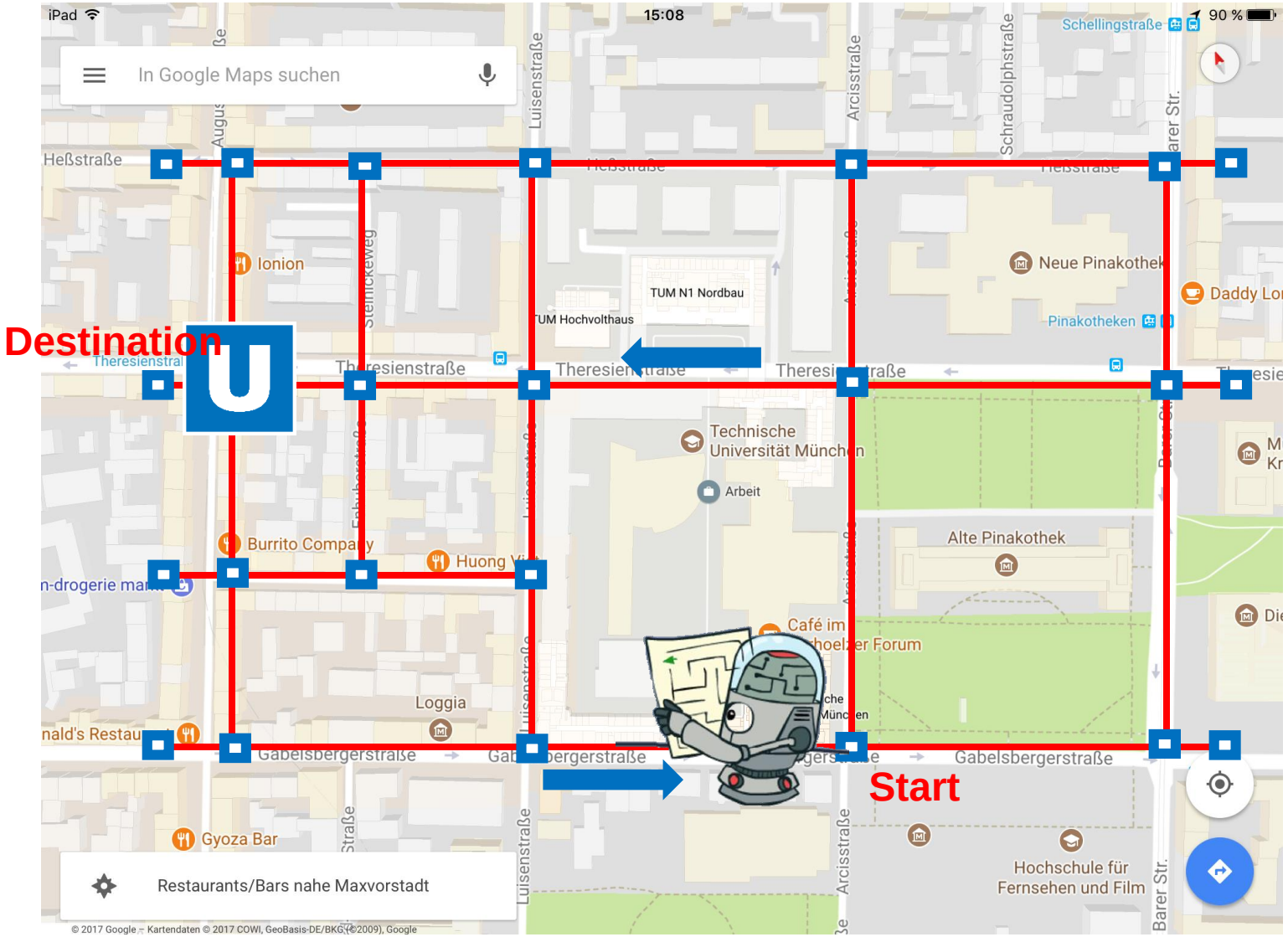
Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

Scenario


iPad
15:08
90 %
In Google Maps suchen
Heßstraße
Luisenstraße
Arcisstraße
Schraudolphstraße
Barer Str.
Schellingstraße
Ionion
Steinickeweg
TUM N1 Nordbau
TUM Hochvolthaus
Neue Pinakothek
Pinakotheken
Theresienstraße
Technische
Universität München
Arbeit
Enhuberstraße
Burrito Company
Huong Viet
Alte Pinakothek
Café im
Loggia
Gabelsbergerstraße
Gyoza Bar
Hochschule für
Fernsehen und Film
Restaurants/Bars nahe Maxvorstadt
© 2017 Google – Kartendaten © 2017 GOWI, GeoBasis-DE/BKG (©2009), Google

Scenario

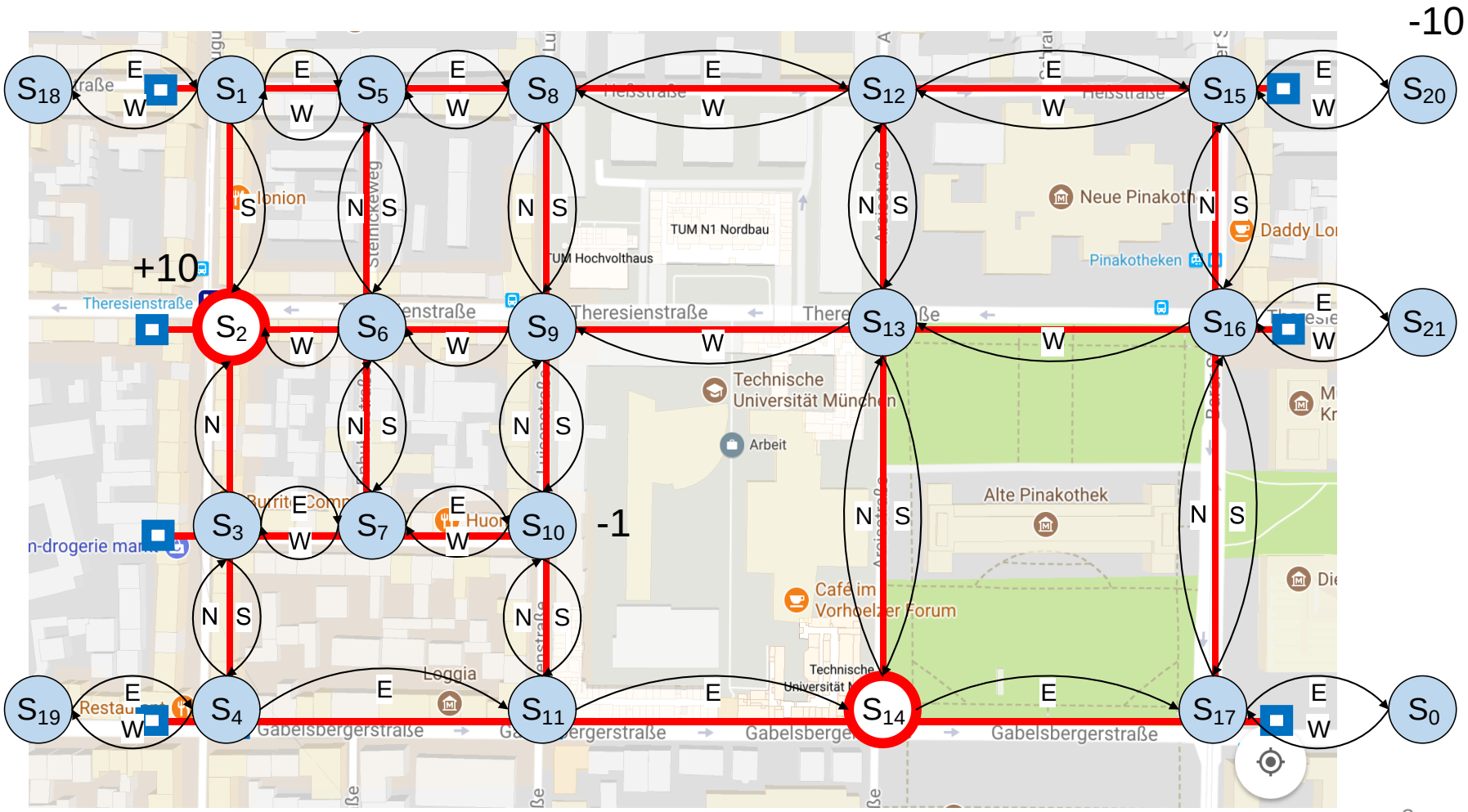

iPad
15:08
90 %
In Google Maps suchen
Destination
Start
Heßstraße
Luisenstraße
Arcisstraße
Schraudolphstraße
Schellingstraße
Barer Str.
Ionion
Neue Pinakothek
Daddy Lo
Pinakotheken
TUM N1 Nordbau
TUM Hochvolthaus
Stiglmaierplatz
Theresienstraße
Technische
Universität München
Arbeit
Alte Pinakothek
Burrito Company
Huong
Loggia
Gabelsbergerstraße
Gyoza Bar
Hochschule für
Fernsehen und Film
Restaurants/Bars nahe Maxvorstadt
© 2017 Google – Kartendaten © 2017 GOWI, GeoBasis-DE/BKG (©2009), Google

Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

Modeling

Markov Decision Process (S, A, P, r, γ) $P(s_{t+1}) = P(s_{t+1}|s_t, a_t)$



Modeling

Model-based method: first learn a model of the system, and then optimize the policy under this model.

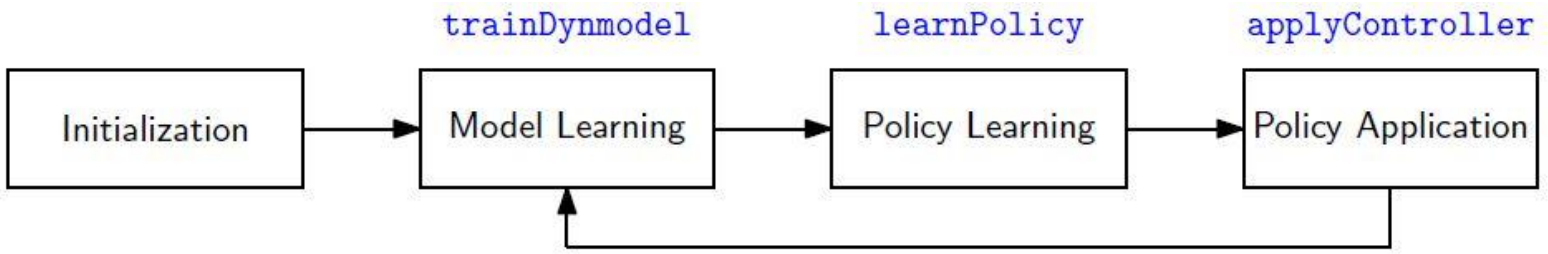


Figure: Main Modules of model-based method. [1]

Model-free method: directly learn a control policy from system interaction

[1] Marc Peter Deisenroth, 2013, Imperial College London, UK, PILCO Code Documentation

Modeling

- Problem:

- Long training time

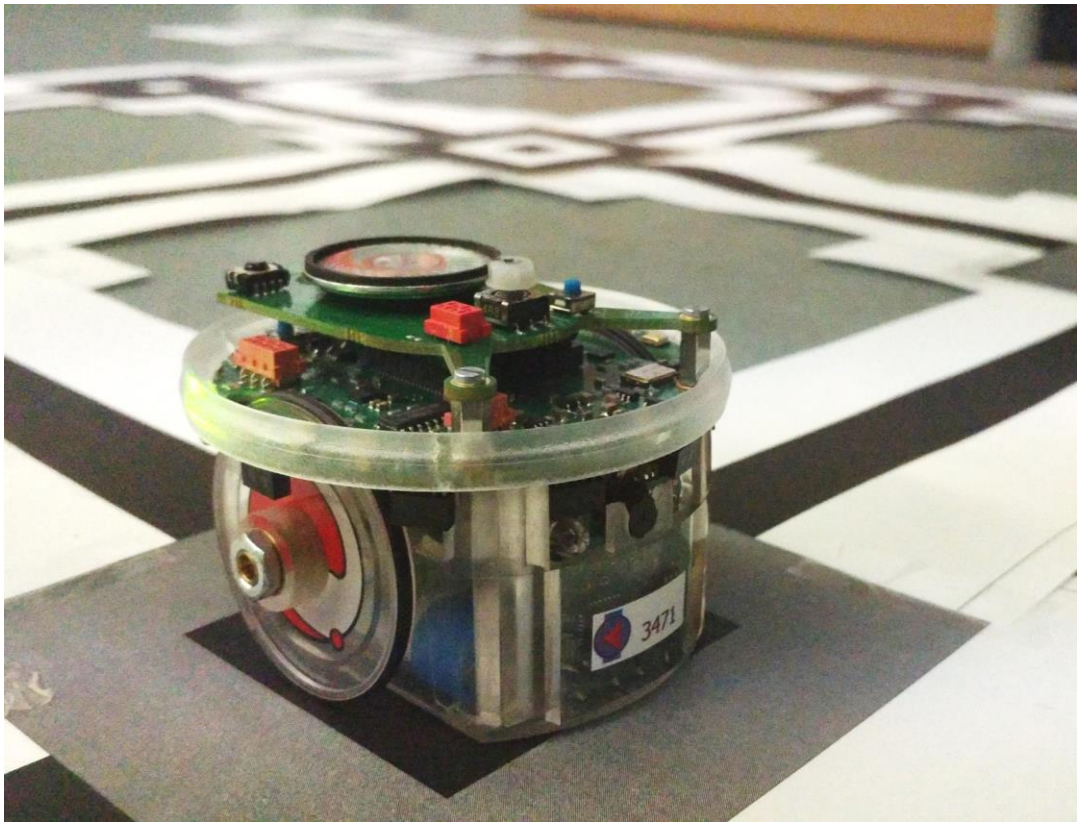
- Solution:

- Prior knowledge

Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

E-puck and its sensors



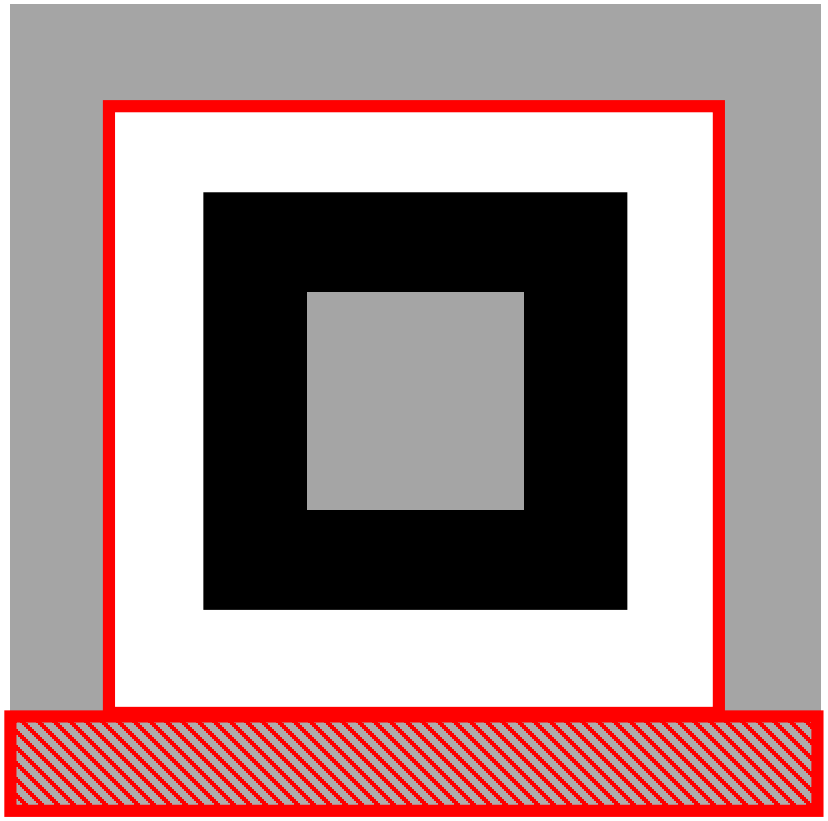
Camera

IR proximity

Microphone

Floor sensors

State detection design



Gray area for stop

Ternary detection unit

White: 2

Gray: 1

Black: 0

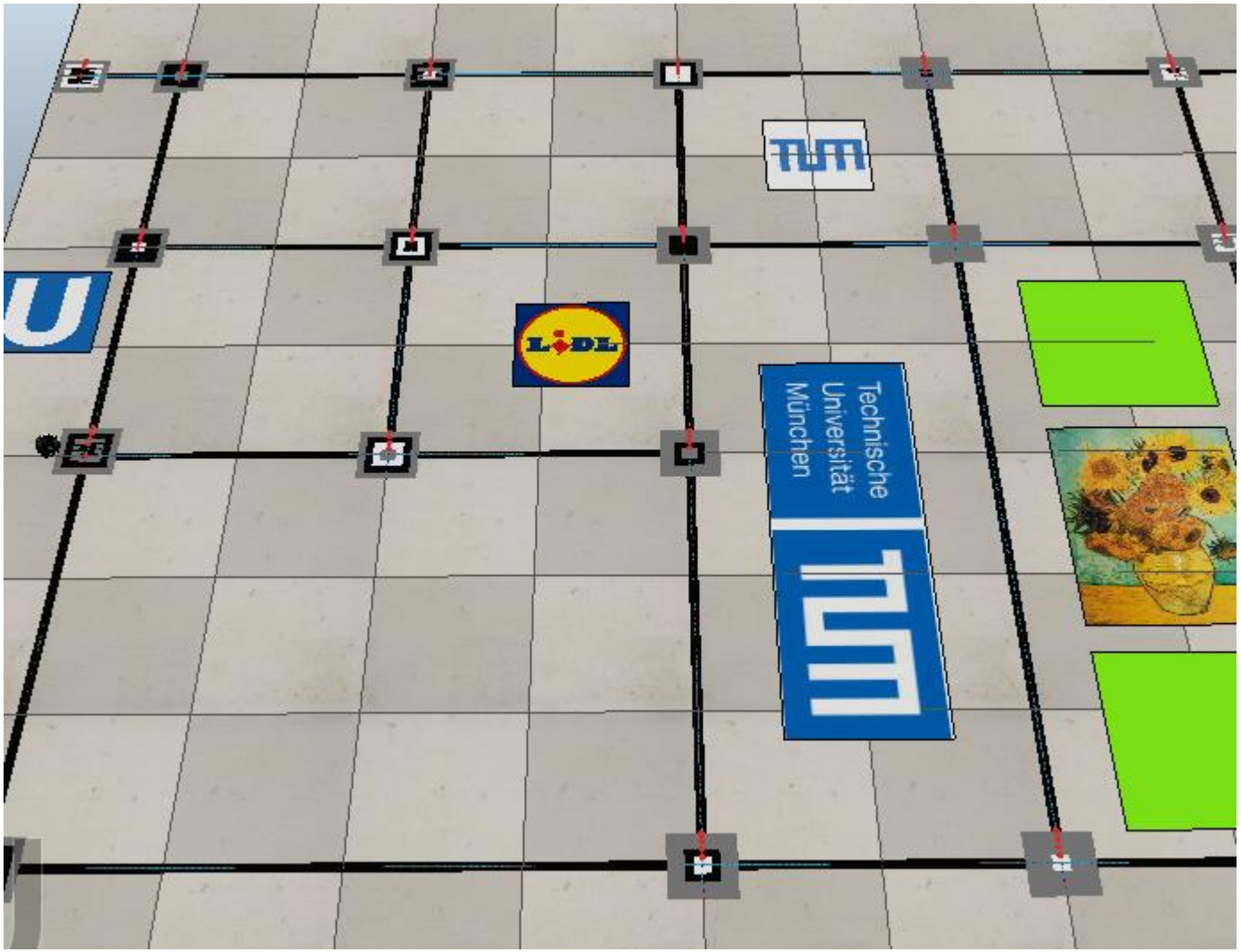
e.g. $201_{(3)} \leftrightarrow 2 \times 3^2 + 0 \times 3^1 + 1 \times 3^0 = 19_{(10)}$

Maximal state space: $3^3 = 27$

Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

Simulation



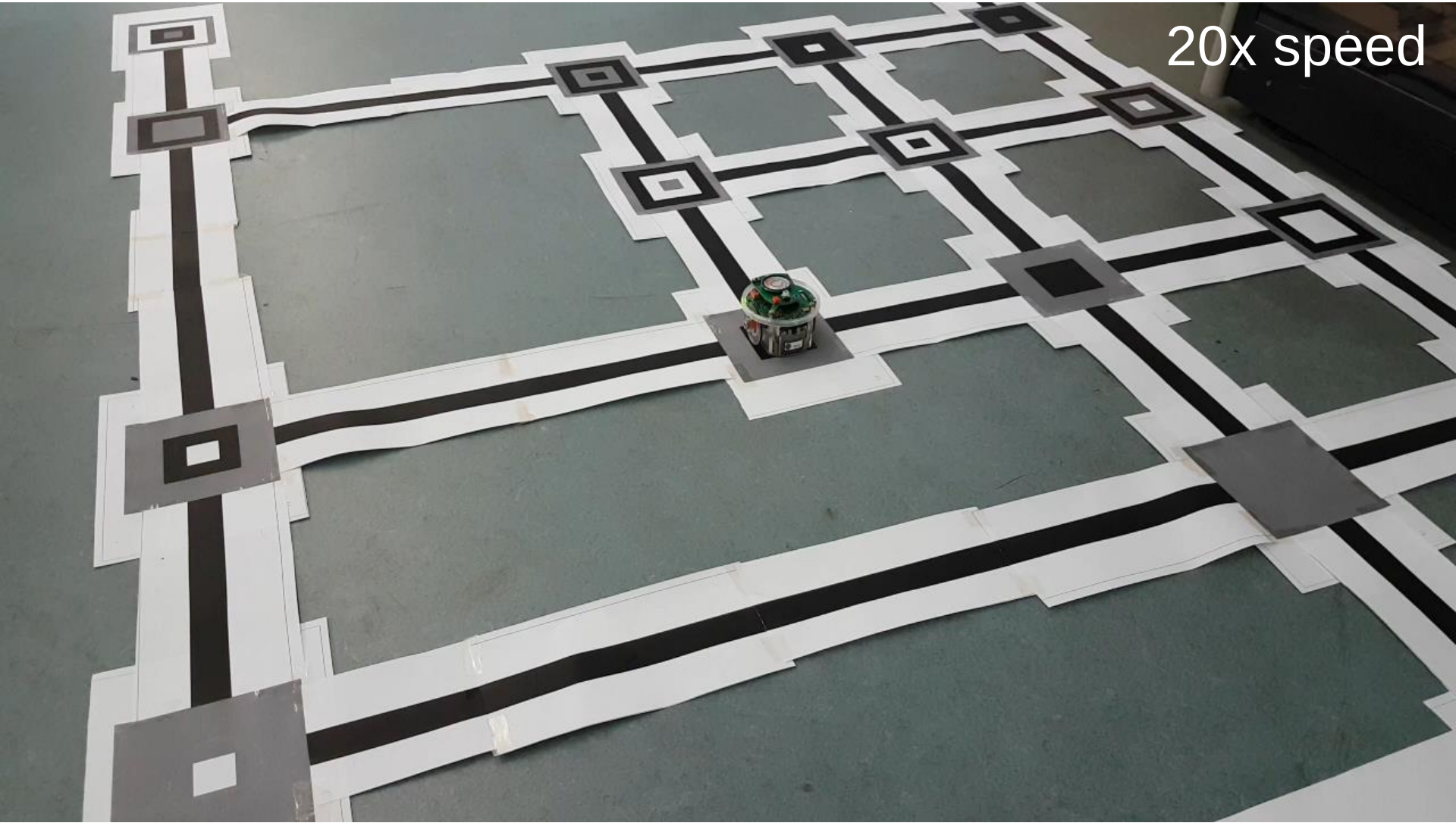
Simulation

8x speed

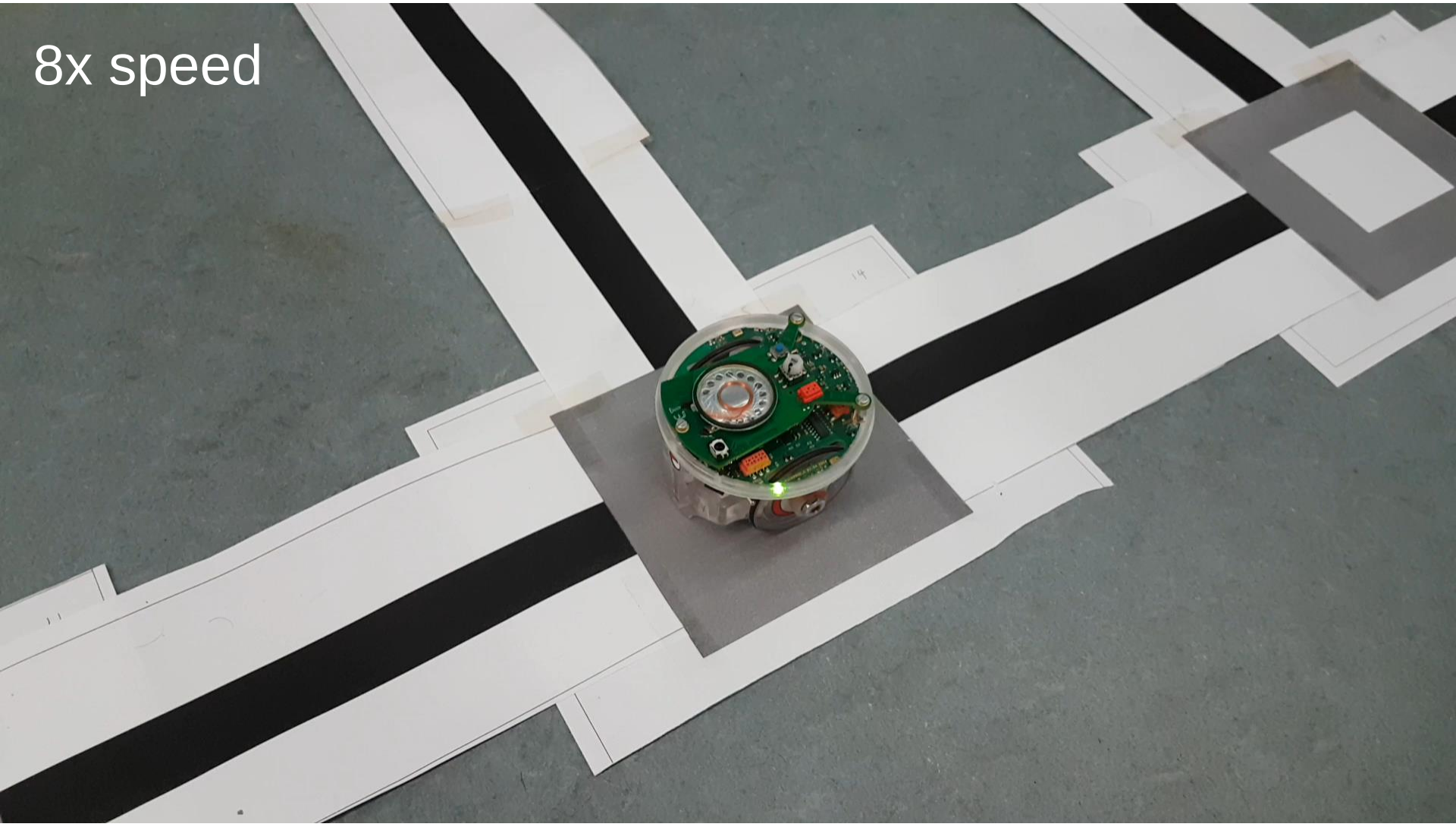
Guideline

- Introduction
- Modeling
- State detection
- Simulation
- Real world demonstration

Real world demonstration: Training



Real world demonstration: After training



Questions?