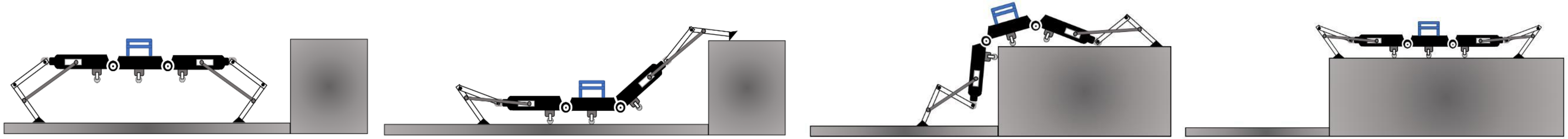
Reinforcement Learning for Autonomous Navigation of Vehicles in an Uncertain terrain : A survey



[Link for the paper]

Soham Kulkarni

--Dept. of Electrical Engineering
Indian Institute of Technology Hyderabad

31st July 2021

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Outline

1 Introduction

2 Autonomous Navigation in Uncertain Terrain

3 Autonomous Navigation in Complex Environments

4 Reinforcement Learning based methods for blind locomotion on uncertain terrain

A Approaches for Quadrupedal robots
B Approaches for Bipedal robots

5 Autonomous Navigation in complex environments and uncertain using Reinforcement Learning

6 Conclusion: Present Challenges and future perspectives

Introduction

1

For autonomous navigation of robots in challenging environments, robots need the ability to:

1. learn the environment model,
2. localize itself in the environment with respect to a global frame, and
3. move to desired locations in the environment.

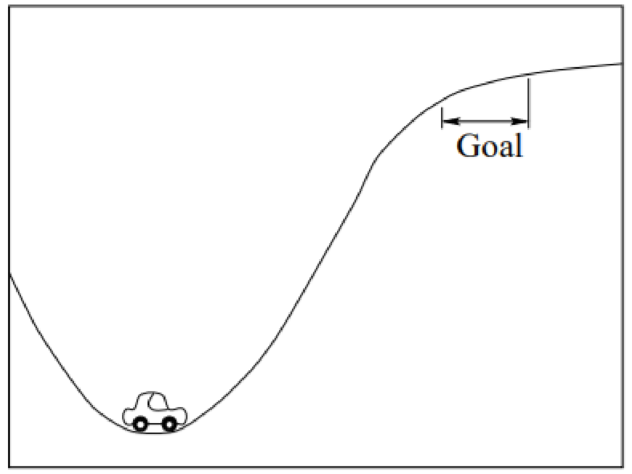
Environments:

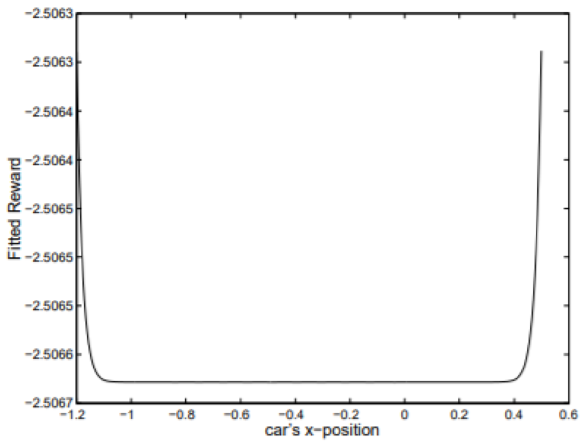
- Unknown
- Dynamic
- Unstructured

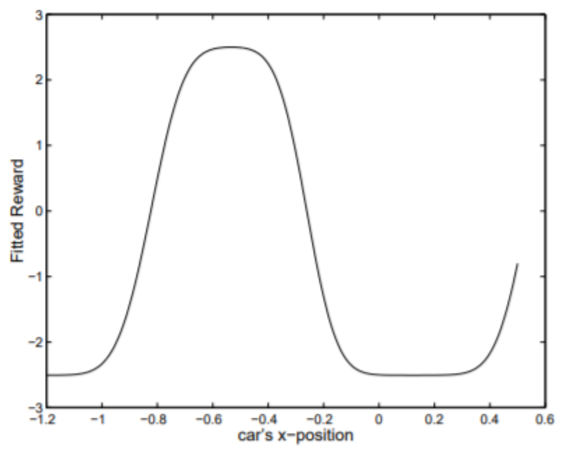
Early methods: Planning and behavior-based control

Sudden changes in the world may not reflect in our model;
quick reflex actions and immediate responding is difficult

Autonomous Navigation in uncertain terrain





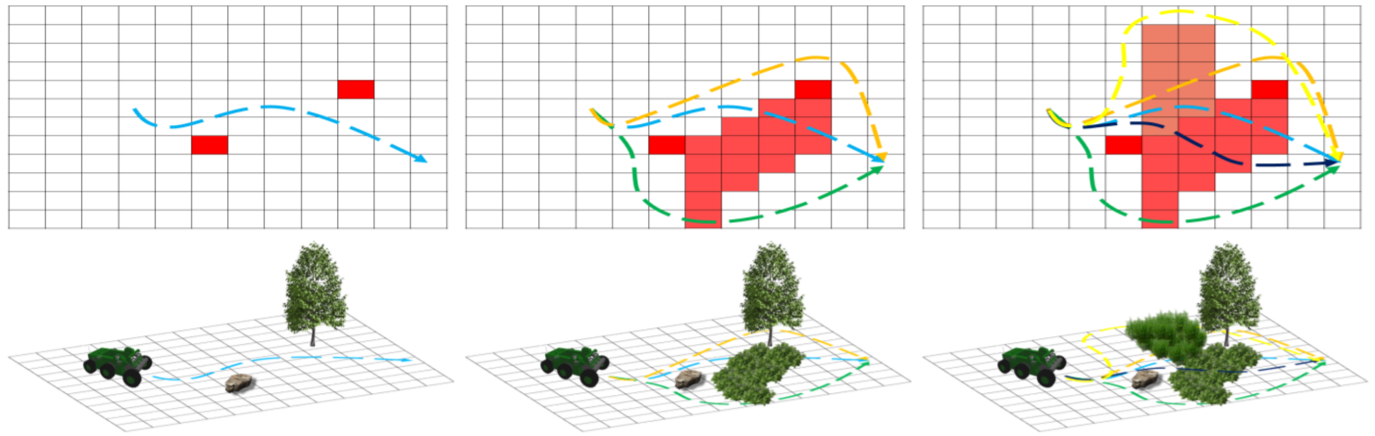


Inverse Reinforcement Learning (Ng & Russell, 2000)

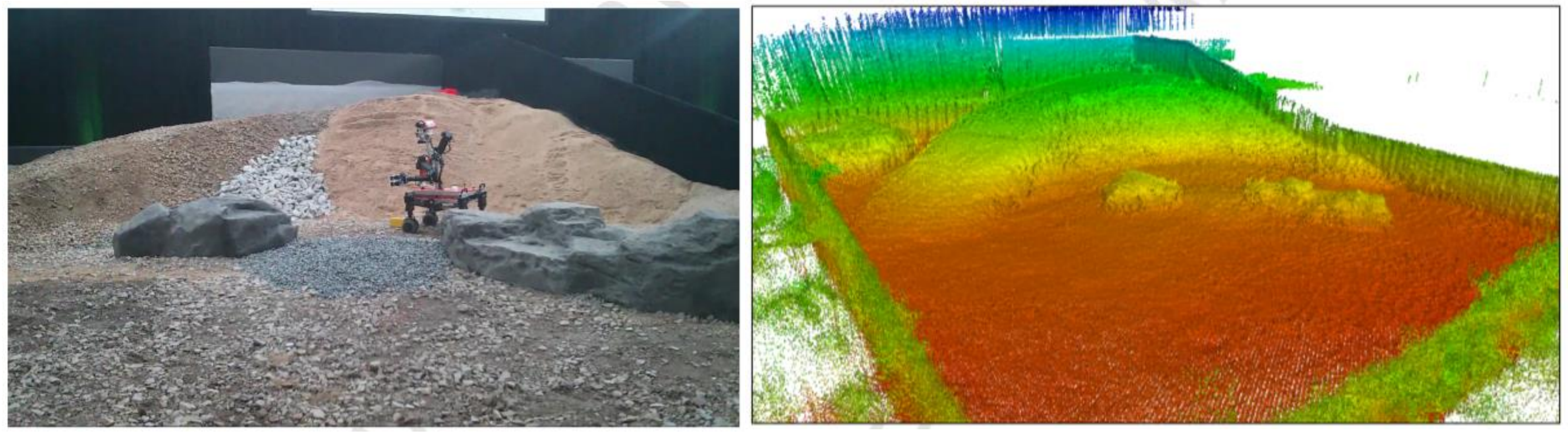
One of the first applicable methods for learning to directly map raw perceptual data to actions

Apprenticeship learning (Abbeel & Ng, 2004)

Maximum Margin Planning (Ratliff et al., 2006)



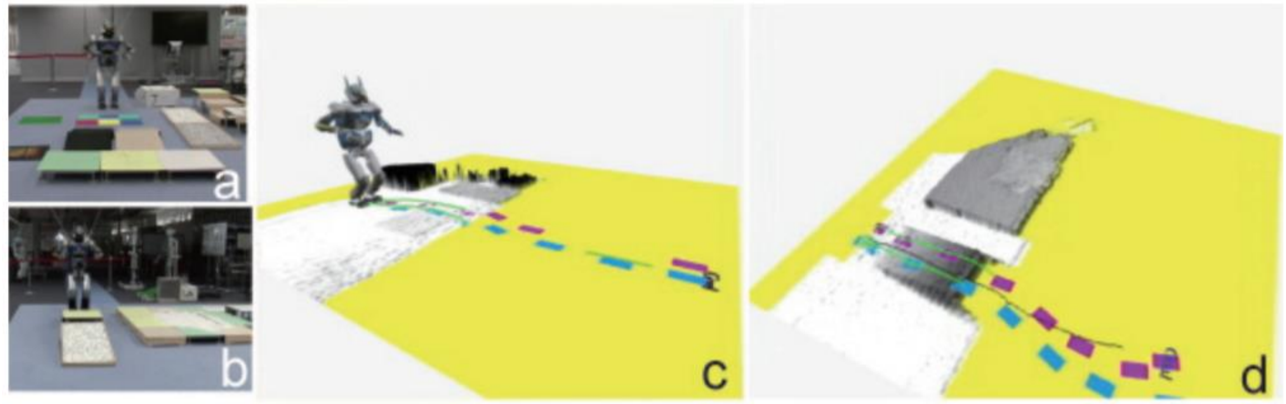
Learning from expert demonstrations for improving robustness in complex, unstructured terrain **(Silver et al., 2010)**



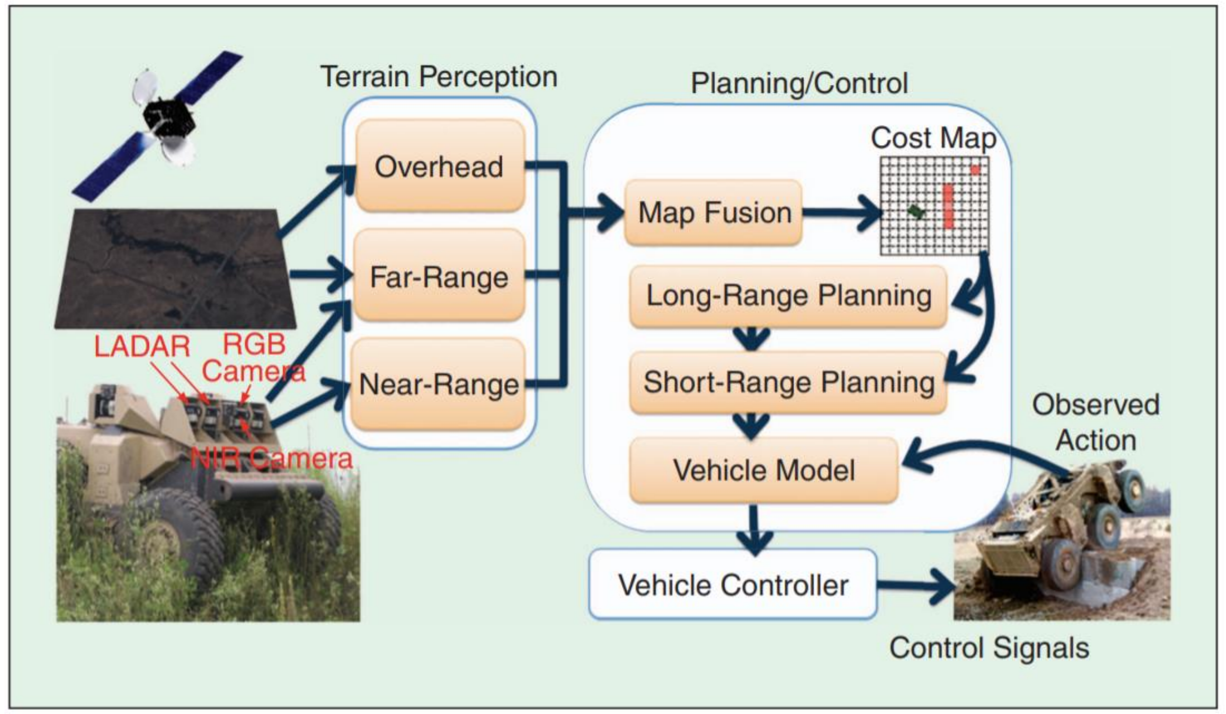
(Droeschel et al., 2017) local and allocentric mapping system that uses a multiresolution map based on the measurements of a 3D Laser Rangefinder, which is consistent in rough terrains

Autonomous Navigation in uncertain terrain

2

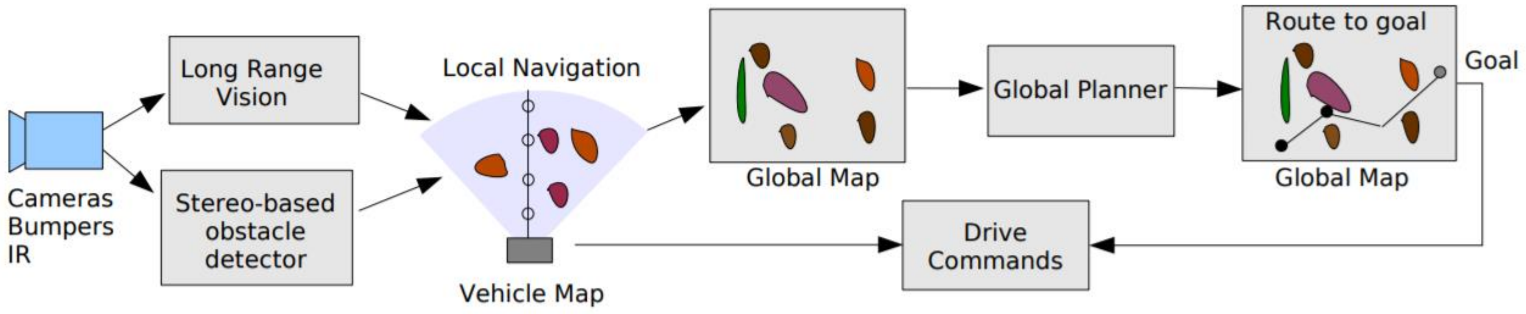


Another example of using laser based perception for autonomous navigation on an uncertain terrain; but for a humanoid robot, with the integration of perception, footstep planning and control(Nishiwaki et al., 2012)

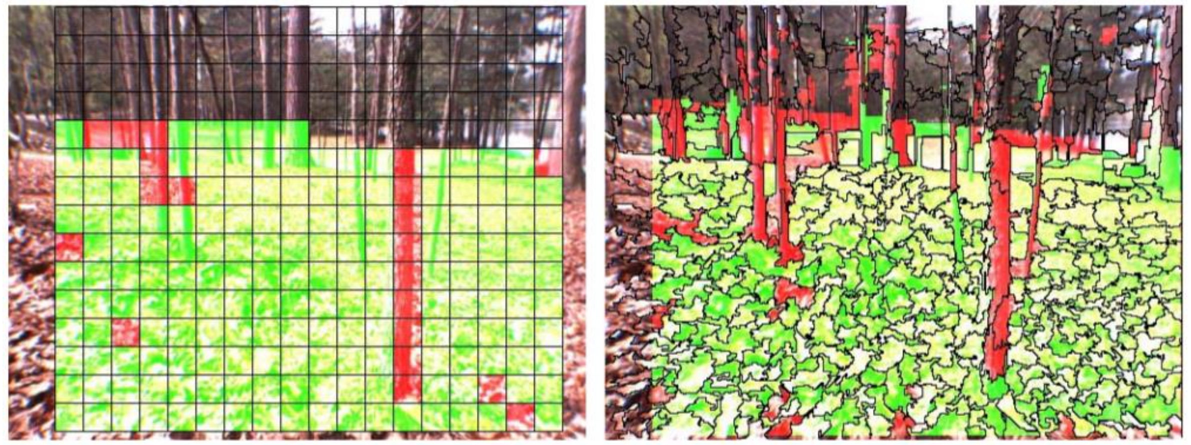


Learning for Autonomous navigation (Bagnell et al., 2010) shows extensive use of Machine Learning for mobility and robust performance of a mobile rover for field robotics

Longrange traversability in complex environments for autonomous mobile robots:



(Hadsell et al., 2007)



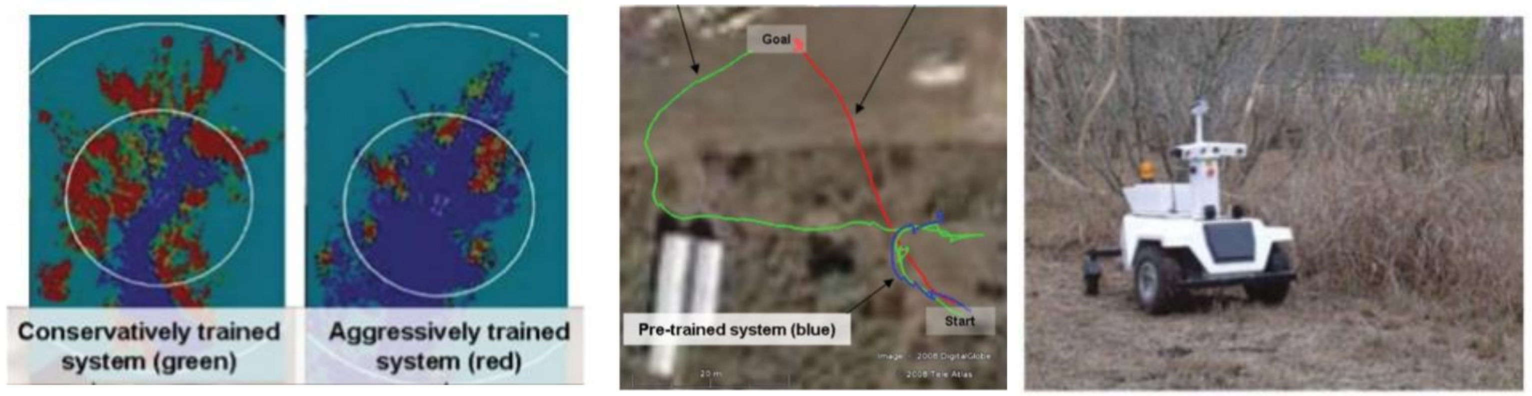
(Kim et al., 2007)

- Here, the terrain information available is spare (from sensory data) and uncertain.

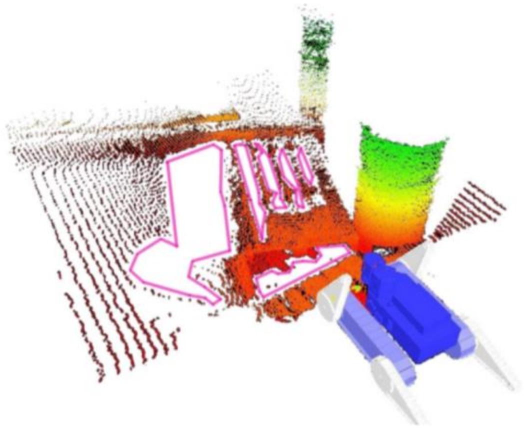
Autonomous Navigation in uncertain terrain

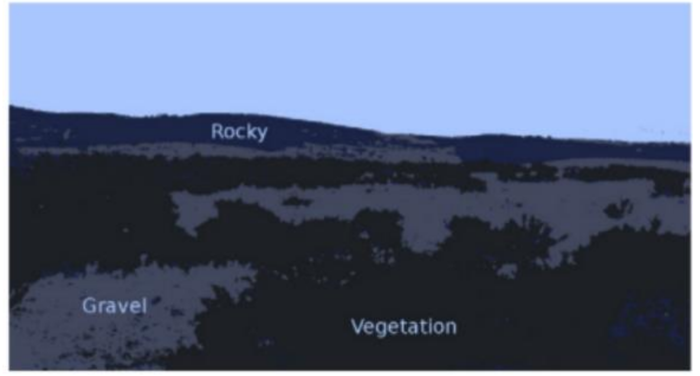
2

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad



(Bajracharya et al., 2009)
shows another end-to-end learning system for autonomous off-road navigation that adapts to local terrain and classify terrain robustly in many off-road environments using onboard proprioception and stereo cameras



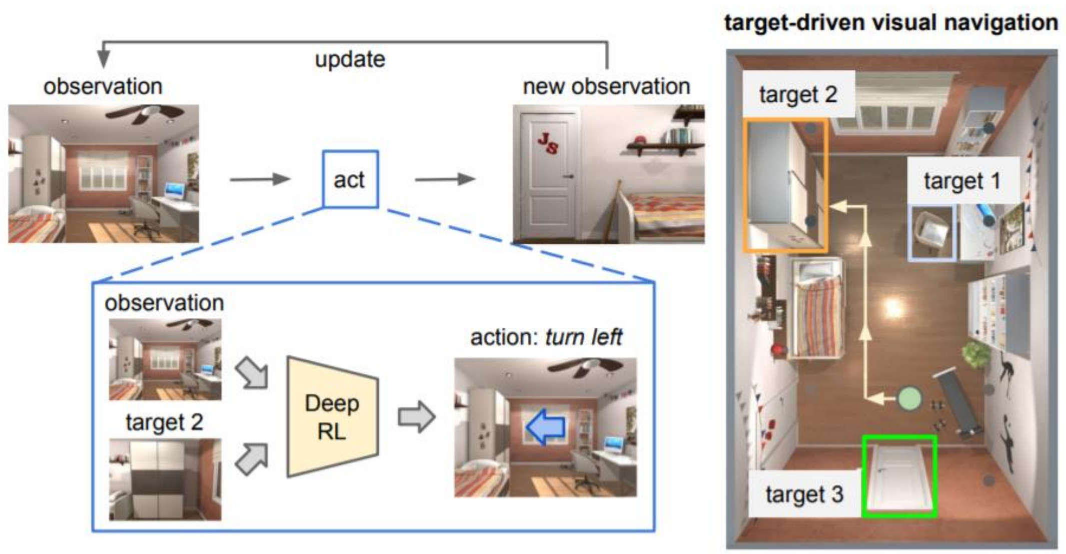


(Papadakis, 2013) is an early survey for terrain traversability analysis methods for UGVs

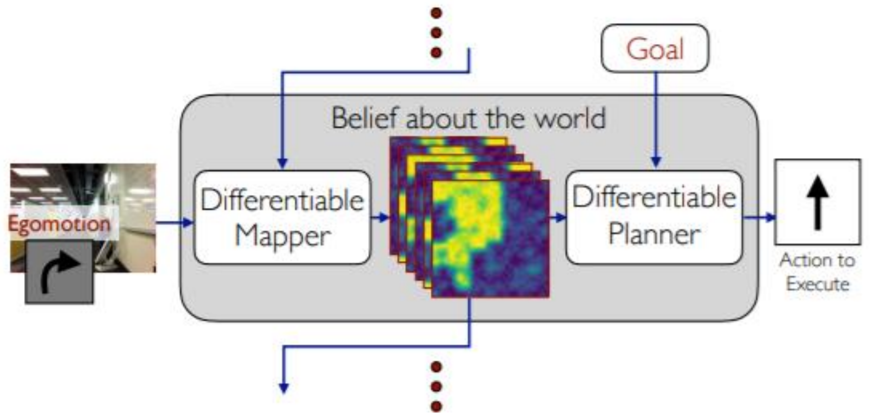
Autonomous Navigation in complex environments

3

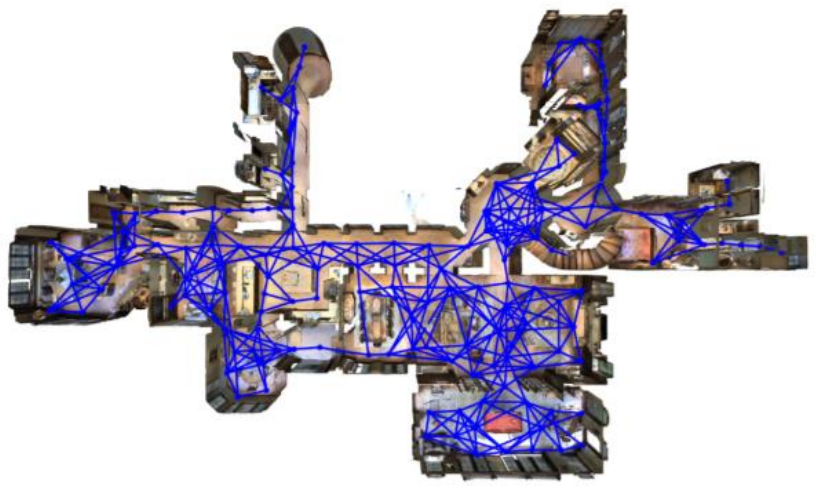
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad



(Zhu et al., 2017) shows target-driven visual navigation in indoor scenes using Deep Reinforcement Learning. This method generalizes across tasks and scenes with very little fine tuning required.



(Gupta et al., 2019) shows a joint architecture for mapping and planning which creates a top-down belief map of the environment and plans a sequence of actions with the help of a differentiable neural net planner.



- With the advances in Deep Learning in the last decade and due to the availability of huge datasets, there have been many works recently on the creation of embodied agents (e.g. mobile robots) which learn through interactions and explorations to solve challenging tasks within the environment

(Anderson, Chang et al., 2018)

(Anderson, Wu et al., 2018)

(Das et al., 2018)

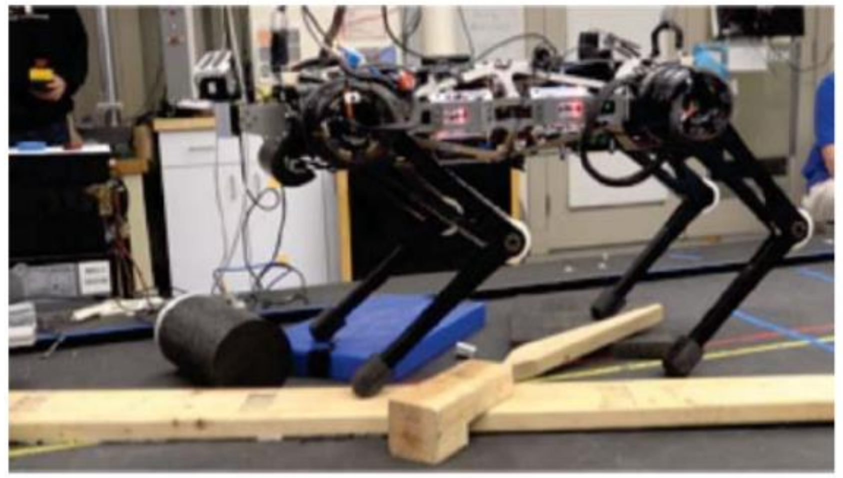
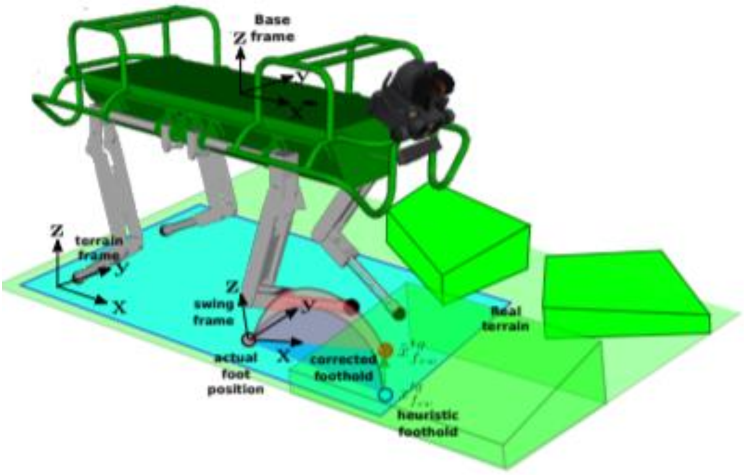
Note: The ground is considered to be flat and training scenes are either indoor or know environments with objects. The main focus is on the mapping and planning pipelines and their integration in the overall architecture.

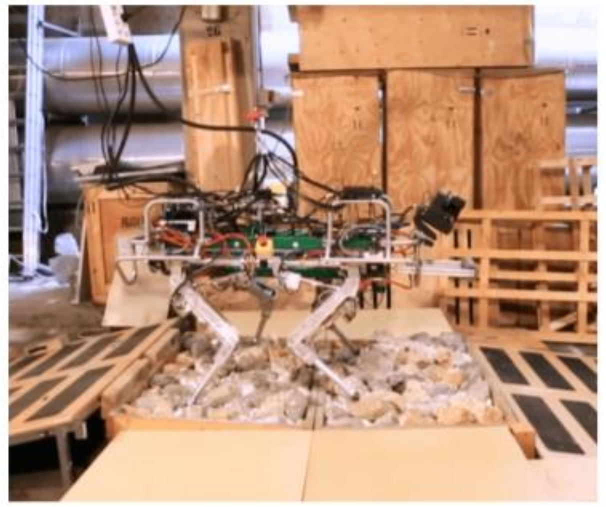
Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

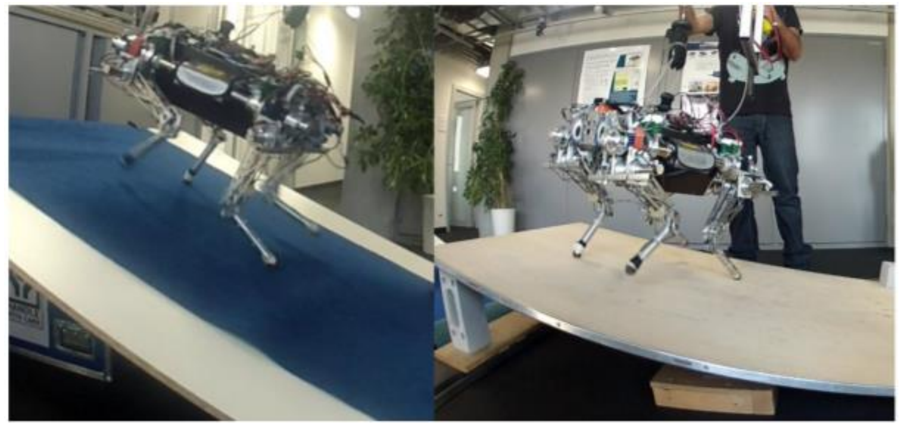
- Blind Locomotion on uncertain terrain requires precise state estimation and quick reflex strategies for external changes encountered, thereby aiding navigation in these complex environments.

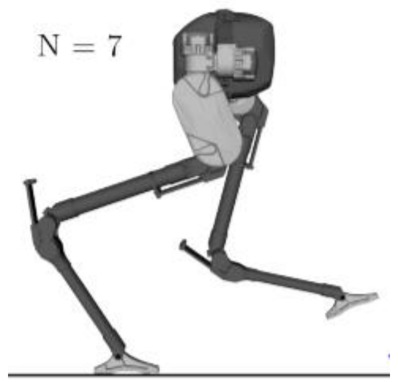






- Conventional state machines based complex architectures for legged locomotion on uneven and uncertain terrain are shown in : (Jenelten et al., 2019; Bledt et al., 2018; Gong et al., 2019; Focchi et al., 2020; Reher et al., 2019)

- Approaches with contact sensors integrated at the feet are shown in : (Bloesch et al. 2013; Gehring et al., 2015; Hartley et al. 2018)





Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Overall, conventional approaches mentioned earlier are very complex and unreliable in real-world scenarios.

A. Reinforcement Learning based Approaches for Quadruped Robots

RL based methods:
- Give promising results
- Explore various possibilities

(Lee et al., 2020) shows:
- A robust controller design for legged locomotion in challenging environments replacing the conventional state machines-based complex controller architectures.
- Only proprioceptive sensor data (from an IMU and joint encoders) is used as feedback in locomotion control at a high temporal resolution as measuring robot induced changes and physical characteristics of the environment such as friction and compliance might be difficult for exteroceptive sensors.
- Zero shot generalization and no tuning required

Fig. Deployments on various terrain types (Lee et al., 2020)

Note: Here, vehicles refer to legged and mobile robot platforms only.

Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

(Hwangbo et al., 2019) is another model-free RL based approach. shows:

- For the Sim2Real transfer, stress is given on modelling the physical system accurately, including the actuators.
- Dynamic controller, improves with experience
- Locomotion skills and recovery behaviours (on flat ground only)

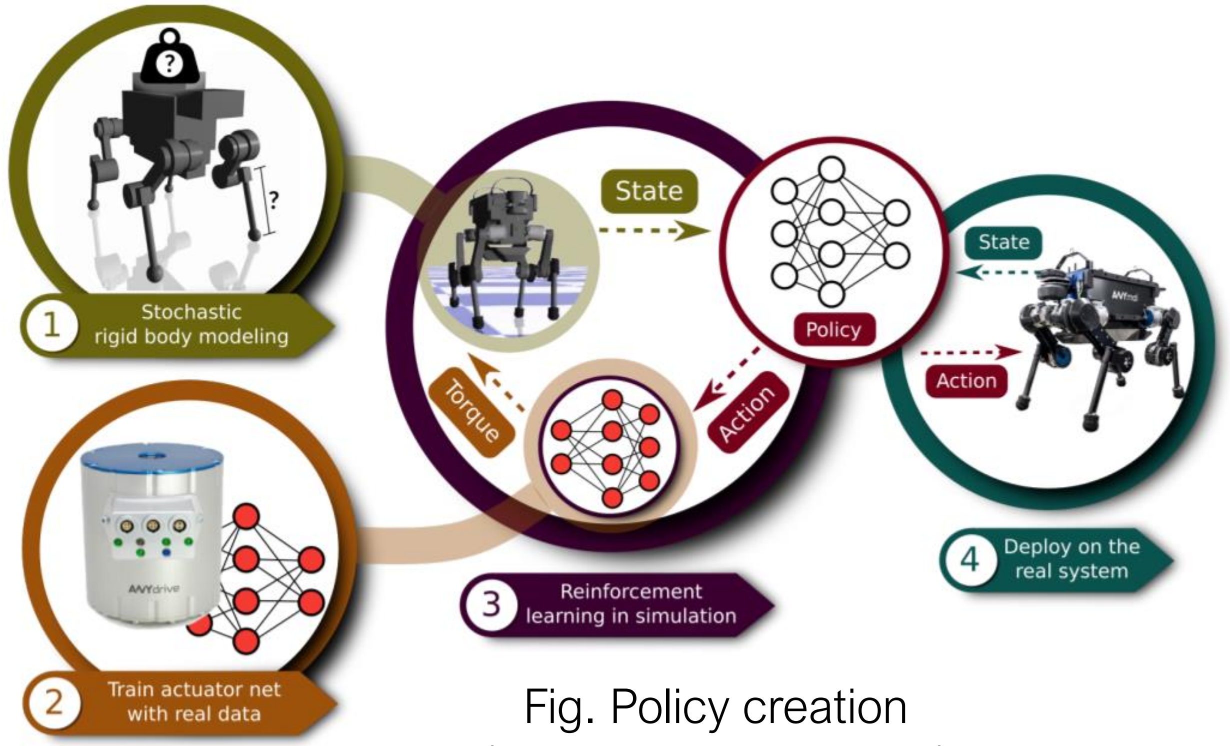


Fig. Policy creation (Hwangbo et al., 2019)

(Haarnoja et al., 2019) is yet another model-free RL based approach. shows:

- End-to-end learning of locomotion controller from scratch
- No training in simulation beforehand
- Deep RL is directly implemented
- No exhaustive pre-tuning and trials requirement

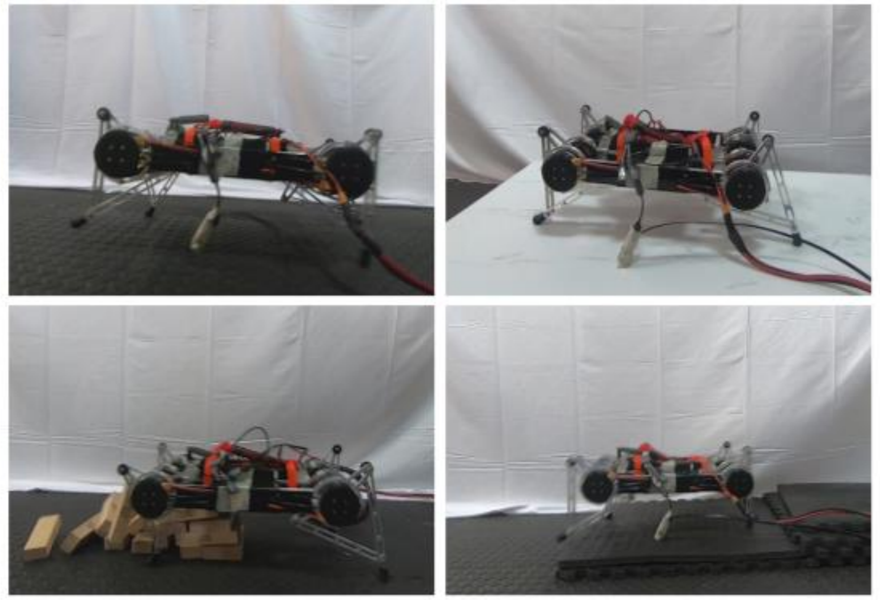


Fig. Learning of a walking gait (on flat terrain) (Haarnoja et al., 2019)

The above work is an extension of **(Haarnoja et al., 2018)** (Soft actor-critic algorithm), to remove the need of manual tuning of hyperparameters

Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Video Link:
https://www.youtube.com/watch?v=nBy1piJrq1A&t=109s

Video Paper: A.Kumar,Z.Fu,D.Pathak,andJ.Malik,"RMA:RapidMotor Adaptation for Legged Robots," 7 2021. [Online]. Available:http://arxiv.org/abs/2107.04034

Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

- (Tan et al., 2018) focuses on leveraging Deep Reinforcement learning techniques to learn agile locomotion for quadruped robots

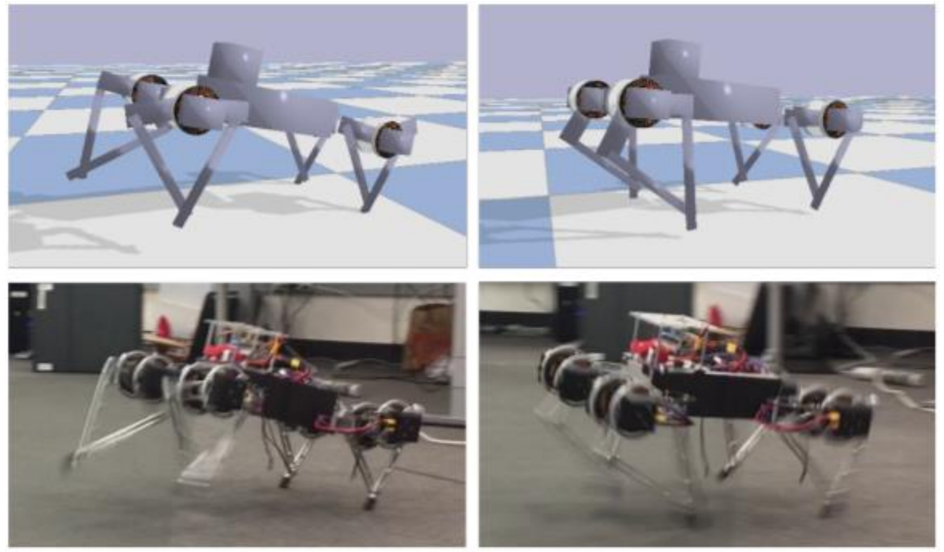


Fig. A still image of galloping gaits learned via Deep RL (in simulation and rea world env.) (Tan et al., 2018)

B. Reinforcement Learning based Approaches for Biped Robots

- (Tedrake & Seung, 2005) presents an approach for learning dynamic bipedal walking from scratch with adaptation to various terrains over experience.
  Goal: stabilizing the limit cycle trajectory for walking using a Poincare´ map[1]

- (Seikmann et al., 2021) shows robust blind bipedal locomotion for the Cassie Robot using sim-to-real RL for stair-like terrain using only proprioceptive feedback

- (Xie et al., 2018) shows:
  - ❖ Robust walking imitating reference motion
  - ❖ Feedback control: Optimal policy for a MDP
  - ❖ Trained blindly on flat terrain
  - ❖ Transition between speeds can be achieved via interpolation between policies

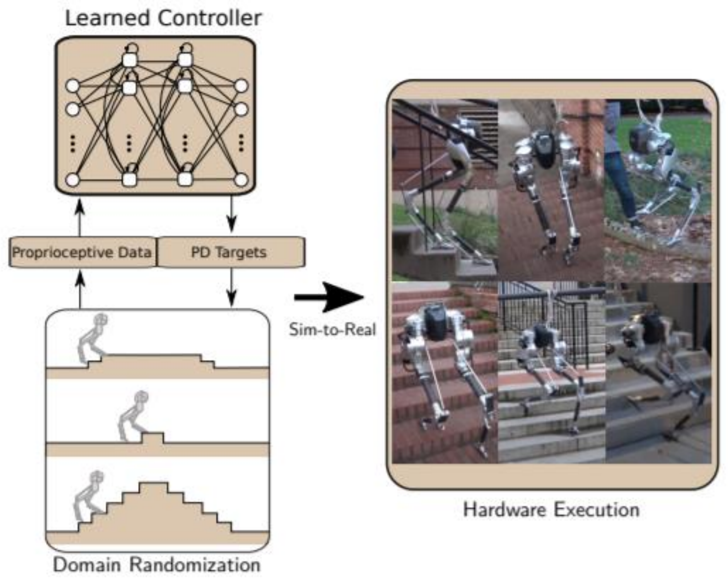


Fig. Blind bipedal stair walking via Sim2Real RL (Seikmann et al., 2021)

References: [1] S. H. S. H. a. Strogatz, Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering. Second edition. Boulder, CO : Westview Press, a member of the Perseus Books Group, [2015]

Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- **(Heess et al., 2017)** shows training (in simulation) of agents in diverse environmental conditions to learn robust behaviors that perform well across certain tasks.

- **(Peng et al., 2017)** developed a hierarchical framework for simulated robots to perform challenging tasks in complex environments.

❖ Deep reinforcement learning is used to work directly with high dimensional sensory information.
❖ Further, a network is trained for phase representation of cyclic motions.

- **(Peng et al., 2018)** shows the use of dynamics randomization to train recurrent policies capable of adapting to unfamiliar dynamics at a run.

- **(Chen et al., 2018)** shows an approach for wheeled-legged robots to learn navigation skills through a number of navigation behaviours by training action policies.

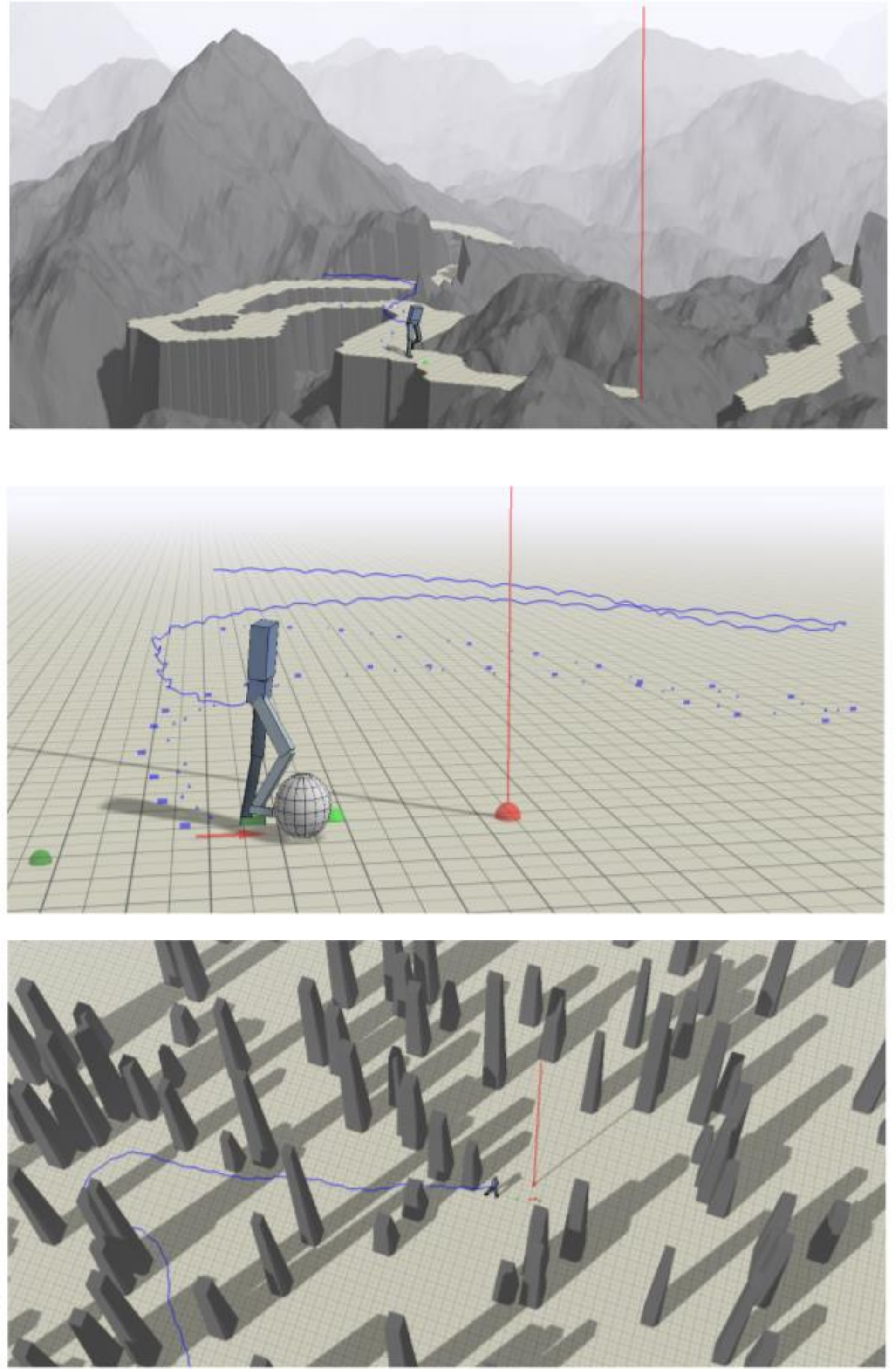


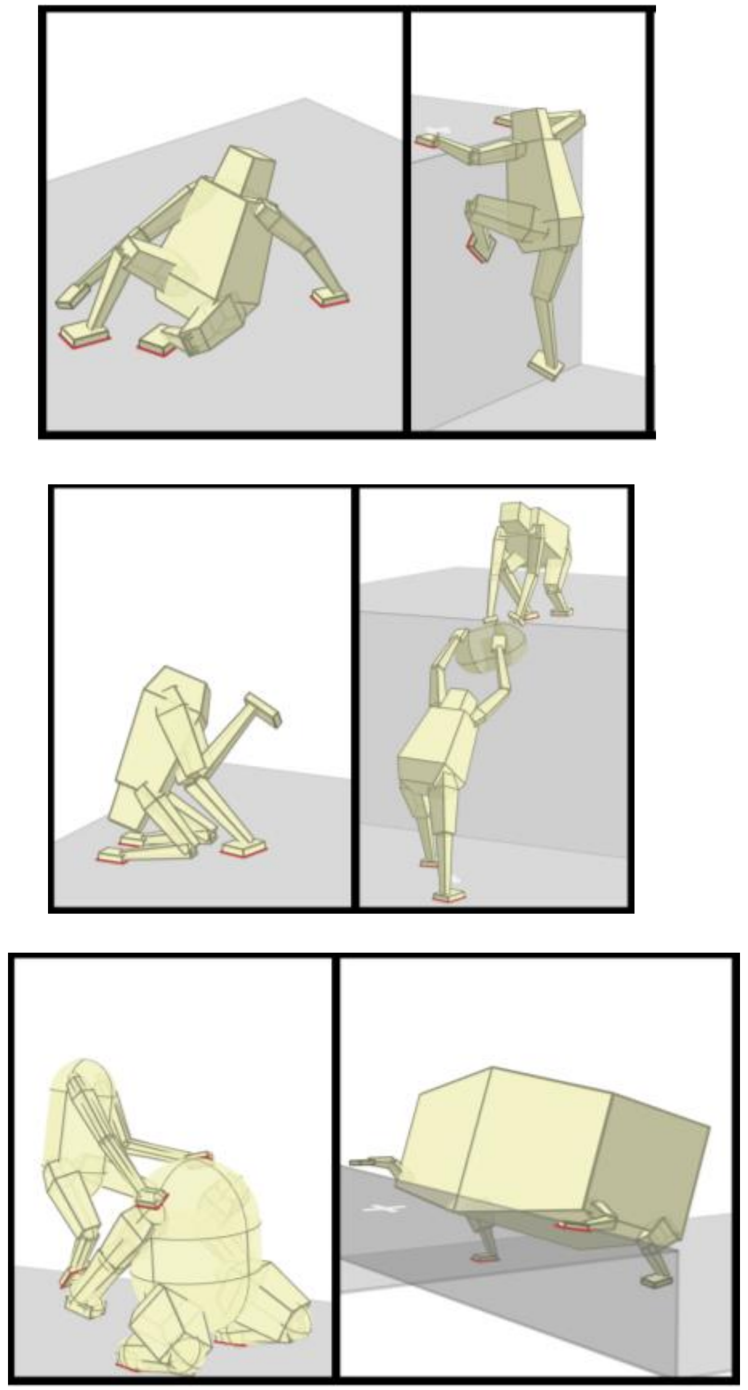
Fig. DeepLoco: Dynamic Locomotion skills using hierarchical Deep Reinforcement Learning
(Peng et al., 2017)

Reinforcement Learning based methods for Blind Locomotion on uncertain terrain

4

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- Learning of basic locomotion behaviours using RL in an end-to-end manner is shown in (Heess et al., 2015; Lillicrap et al., 2015; Schulman et al. 2015a; Schulman et al. 2015b)
- (Levine & Abbeel, 2014) shows Guided Policy Search based approach for the same.
- Terrain adaptive locomotion control has been shown by (Peng et al., 2016), but still the solution is imposed with many constraints.
- (Peng, Berseth & Panne, 2017) shows appreciable results for learning locomotion controllers for a 3D humanoid body, but these rely on domain-specific structure and motion capture data from humans to bootstrap the skills for navigating flat terrains.
- Curriculum Learning[2] is another approach which has been exploited for learned motion skills (Karpathy & Panne, 2012).
- (Mordatch et al. 2012, 2012; Wu et al. 2010) show terrain adaptive behaviours/skills emerging as a byproduct fromoptimization problems.



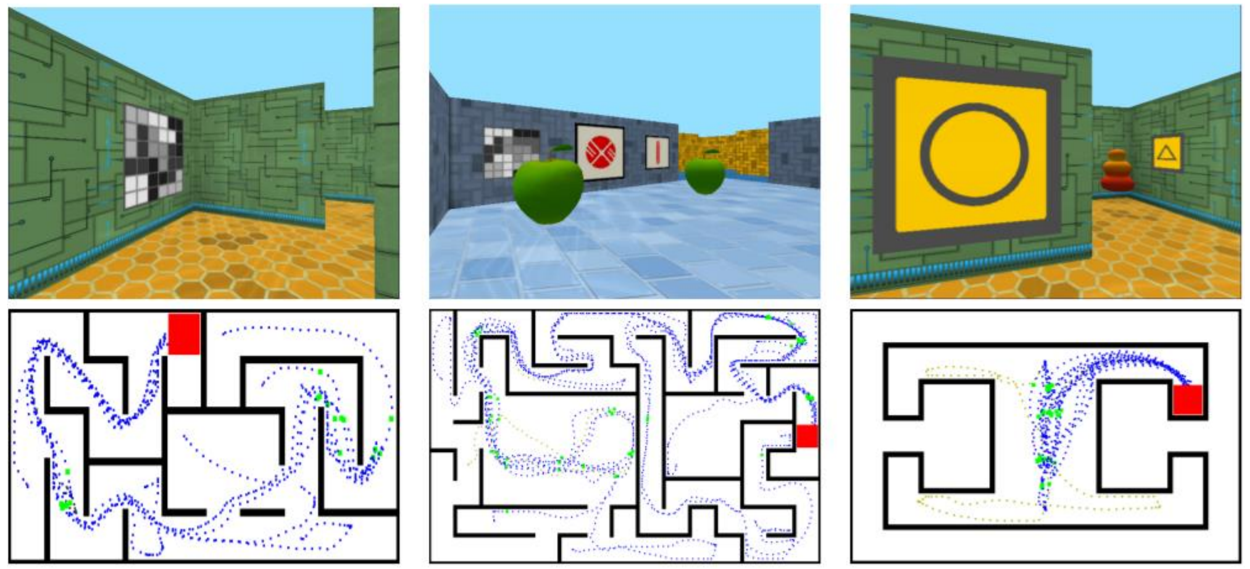
References: [2] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, “Curriculum Learning,” in Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML ’09. New York, NY, USA: Association for Computing Machinery, 2009, p. 41–48.

Autonomous Navigation in Complex Environments and uncertain terrain using Reinforcement Learning

5

▪ **(Mirowski et al., 2017)** shows RL-based navigation in complicated and highly dynamic environments with frequently changing start and goal locations, leveraging raw multimodal sensory input.

❖ RL-based end-to-end learning framework
❖ *Primary Objective:* Maximise the cumulative reward using Asynchronous Advantage Actor-Critic (A3C) algorithm[3]
❖ Supplementing the primary goal target, depth prediction (as an additional auxiliary loss) from RGB images and loop-closures detection are also performed as **auxiliary tasks**, aiding the data-efficiency and learning process for given tasks.



▪ **(Tai, Paolo & Liu, 2017)** presents an end-to-end learning based map-less motion planner which take sparse LiDAR sensory data as input and outputs continuous steering commands as output for navigation.

References: [3] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous Methods for Deep Reinforcement Learning," 2 2016.

Autonomous Navigation in Complex Environments and uncertain terrain using Reinforcement Learning

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

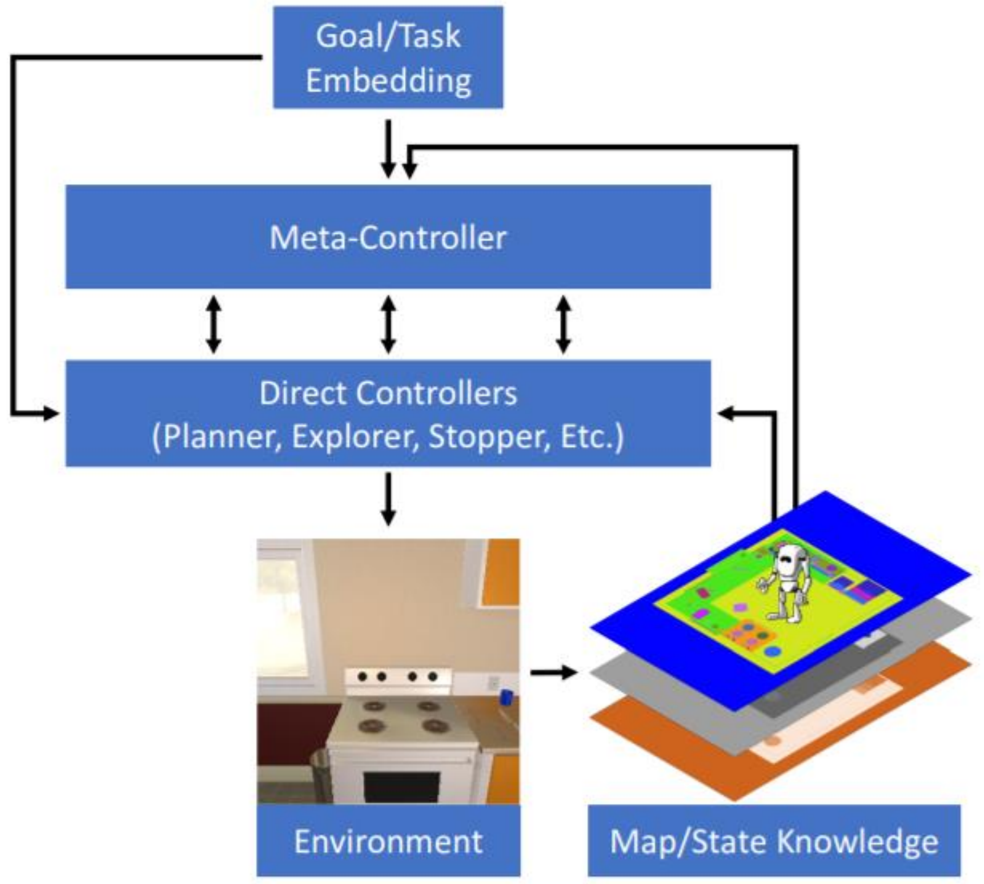
- **(Bjelonic et al., 2018)** presents Weaver, a six-legged hexapod robot for autonomous navigation on unstructured terrain.

❖ Here, stereo cameras and proprioceptive sensing based terrain perception is used for adaptive control while using VIO (Visual-Inertial Odometry) for autonomous waypoint based navigation.

❖ The autonomous Navigation and Control architecture is mentioned in [4].

- **(Gordon, Fox and Farhadi, 2019)** shows a method for Symbolic planning integrated with the learning abilities of Deep Reinforcement Learning for dynamic visual environments.

- **(Kurenkov et al., 2020)** presents a neural network architecture that uses neural message passing of vectors to reason across layers of the scene graph, while combining semantic and geometric cues.



References: [4] M. Bjelonic, T. Homberger, N. Kottege, P. Borges, M. Chli, and P. Beckerle, "Autonomous navigation of hexapod robots with vision-based controller adaptation," 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 5561–5568, 2017.

Autonomous Navigation in Complex Environments and uncertain terrain using Reinforcement Learning

5

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

❖ Latest works

- **(Graves et al., 2018)** propose learning method which efficiently learns robust driving policies in an offline manner, which are generalizable.

- **(Liang, Chen & Song, 2021)** focuses on visual semantic navigation, the task of producing actions for an active agent to navigate to a specified target object category in an unknown environment.

- **Vision-based navigation:** Dense depth algorithms (Computationally expensive: Sensor fusion, Collision checking path planning; Potentially inaccurate)

- **(Feng et al., 2021)** explores the use of passive, stereo sensing for vision-based navigation.

❖ Stixel representation: Compact and sparse for local free space

Autonomous Navigation in Complex Environments and uncertain terrain using Reinforcement Learning

5

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

❖ Latest works

- Hierarchal approaches 1. Path planning (in the HL representation) and 2. Using sub-goals derived from the plan to guide the RL policy in the source task
- Complex Dynamics, Sub-optimal performance
- (Wohlke, Schmitt & van Hoof, 2021) proposes a novel hierarchical framework that utilizes a trainable planning policy for the HL representation.

- (Huang et al., 2021a) propose a novel Deep RL system consisting of:

❖ Map Navigation
❖ Multi-view perception
❖ Multi-branch control

For mobile robots

- (Huang et al., 2021b) proposes a novel Deep RL system for local path planning.

❖ Multi-modal perception: Direct policy learning for generating flexible actions to avoid collisions with obstacles in the navigation

For UGVs

Autonomous Navigation in Complex Environments and uncertain terrain using Reinforcement Learning

5

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- The video shows full autonomous navigation and 3D semantic mapping in the given environment.

- The pose of the robot is estimated from invariant EKF odometry in the IMU frame. It is a result of pure odometry and simple navigation, without any loop closures.

Video Link:
https://www.youtube.com/watch?v=uFyT8zCg1Kk

Conclusion: Present Challenges and future perspectives

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

- This work presents a survey on Reinforcement Learning for autonomous navigation of vehicles in an uncertain terrain.
- Initially, the autonomous navigation problem in complex and dynamic environments is introduced and discussed.
- Then, Reinforcement Learning based methods for Blind Locomotion and Autonomous navigation are discussed.
- Reinforcement Learning methods for locomotion on uncertain terrain have shown impressive results in terms of robustness and complexity of the system.
- RL-based methods for navigation are still facing challenges, due to heavy reliability in the dynamic environmental changes.
- A unified framework integrating both Locomotion and Navigation pipelines based on shared objectives can be a future research area at the intersection of both these methods.

Thank you for your attention!