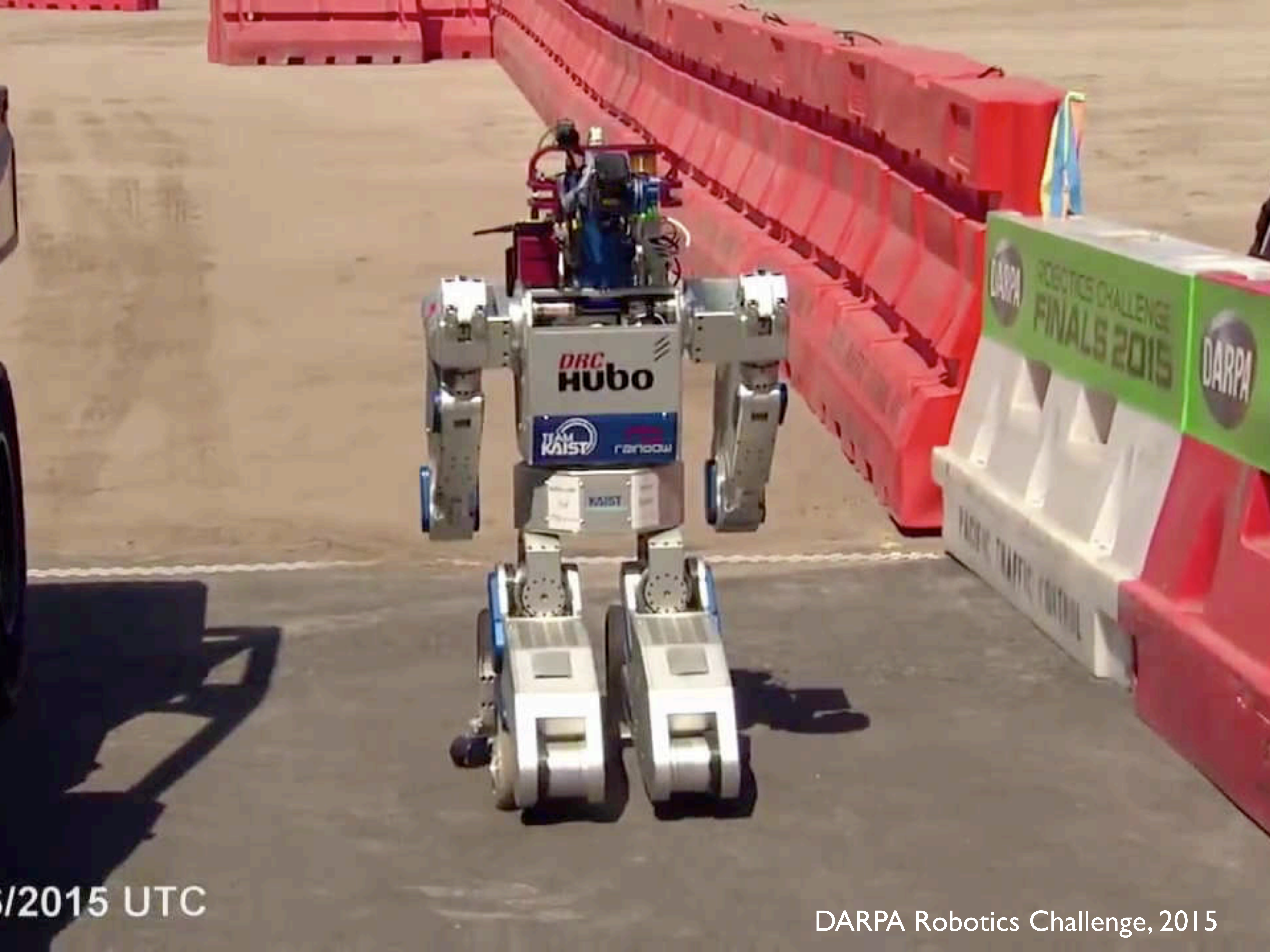
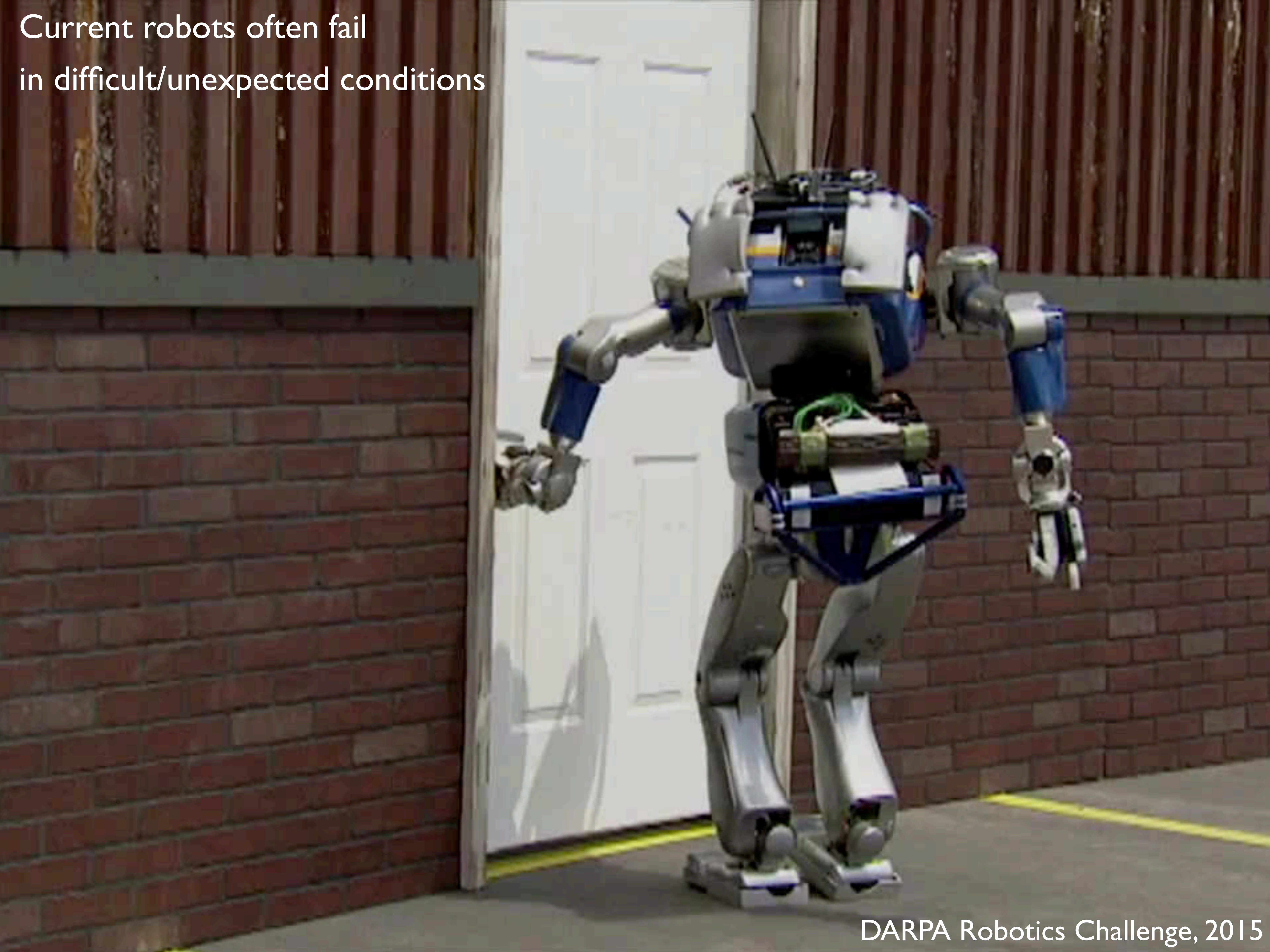
Trial-and-error damage recovery

Jean-Baptiste Mouret
Inria Nancy - Grand Est



DARPA Robotics Challenge, 2015

Current robots often fail
in difficult/unexpected conditions

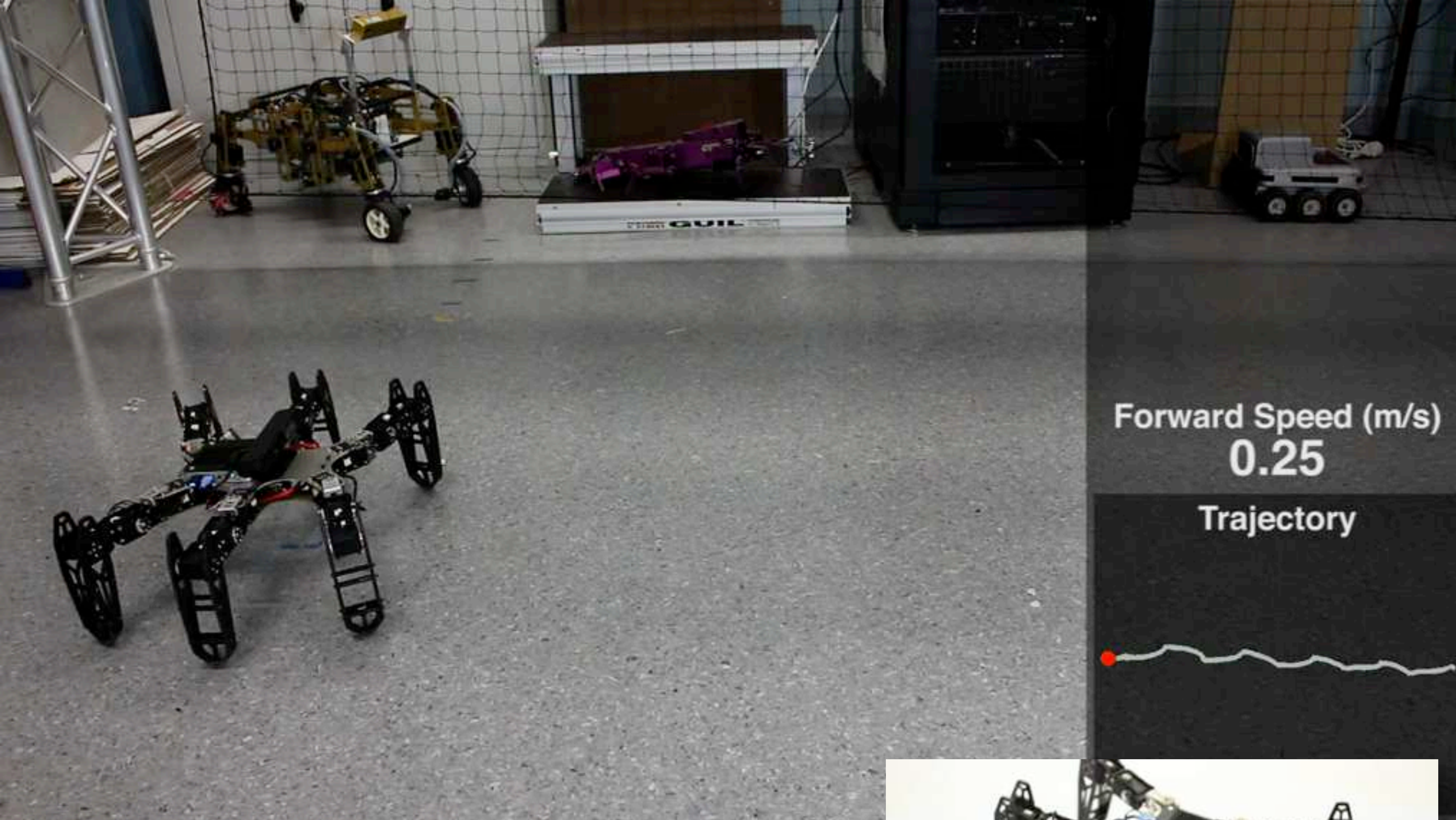


DARPA Robotics Challenge, 2015

The issue with robots is not that they fail…

… it is that they do not get back on their feet and try again

- they do not learn from their mistakes
- if broken, they give up



- Controller : periodical signals (36 parameters)
- Performance: covered distance in 5 seconds
- Performance evaluated onboard (RGB-D visual odometry)

Diagnostics is hard

- **The diagnosis-based approach:**
  - diagnose the problem
  - find a fix

➡

- sensor aliasing (similar data, different cause)
- indirect observation
- need to place the sensors "at the right place" = anticipate

… especially in unstructured environments

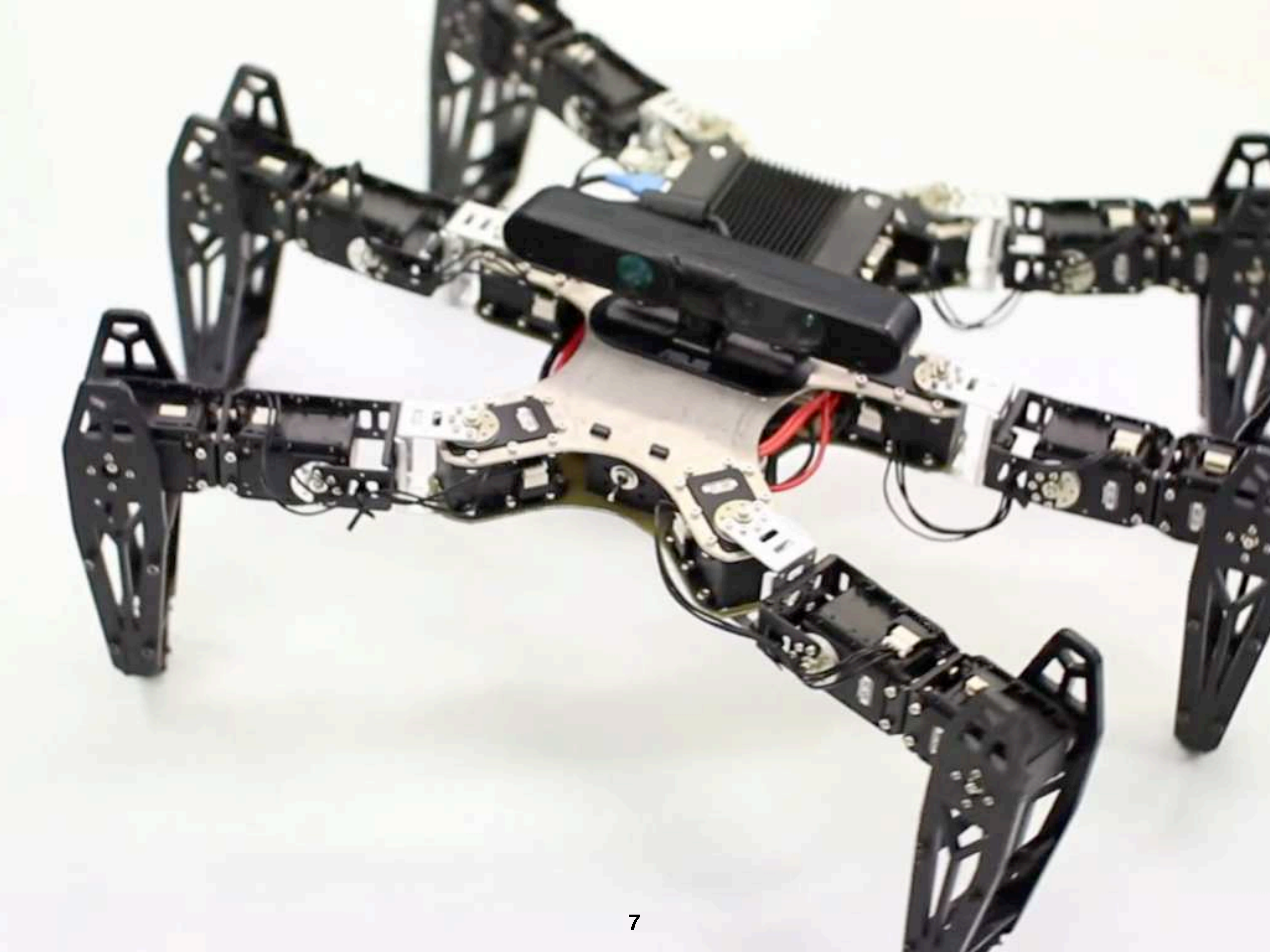
- almost infinite number of possible situations

➡ we cannot anticipate everything

➡ e.g. forests, street, damaged nuclear plant, …

… even more with low-cost robots

- not many sensors
- low-quality sensors
- e.g. Baxter, Poppy



Could we use some kind of trial-and-error learning for damage recovery?

Trial and error learning… in minutes!

(they do not « understand» the injury)

Micro-data learning

AlphaGo

30 million positions
+ self-play

« Big Data »

38 days
of learning

Deep learning ?

Amount of data

« Micro data »

1-20 trials

very few available methods

Damage recovery

The 4 precepts of micro-data learning

1. Choose wisely what to test next (active learning)
   ➡ OK to trade data resources for computational resources
2. Exploit every bit of information from each test
   ➡ e.g., use all the points of a trajectory
3. Only learn what is necessary
   ➡ e.g, do not reinvent control theory
4. Use prior knowledge
   i. use the "right" search space (possibly, design it automatically)
   ii. make prior knowledge explicit
   iii. use everything we know (e.g. simulator)

All the precepts should be combined

JB Mouret. *Micro-data learning: the other end of the spectrum.* ERCIM News. 2016

Guiding learning

Trial & error in animals:

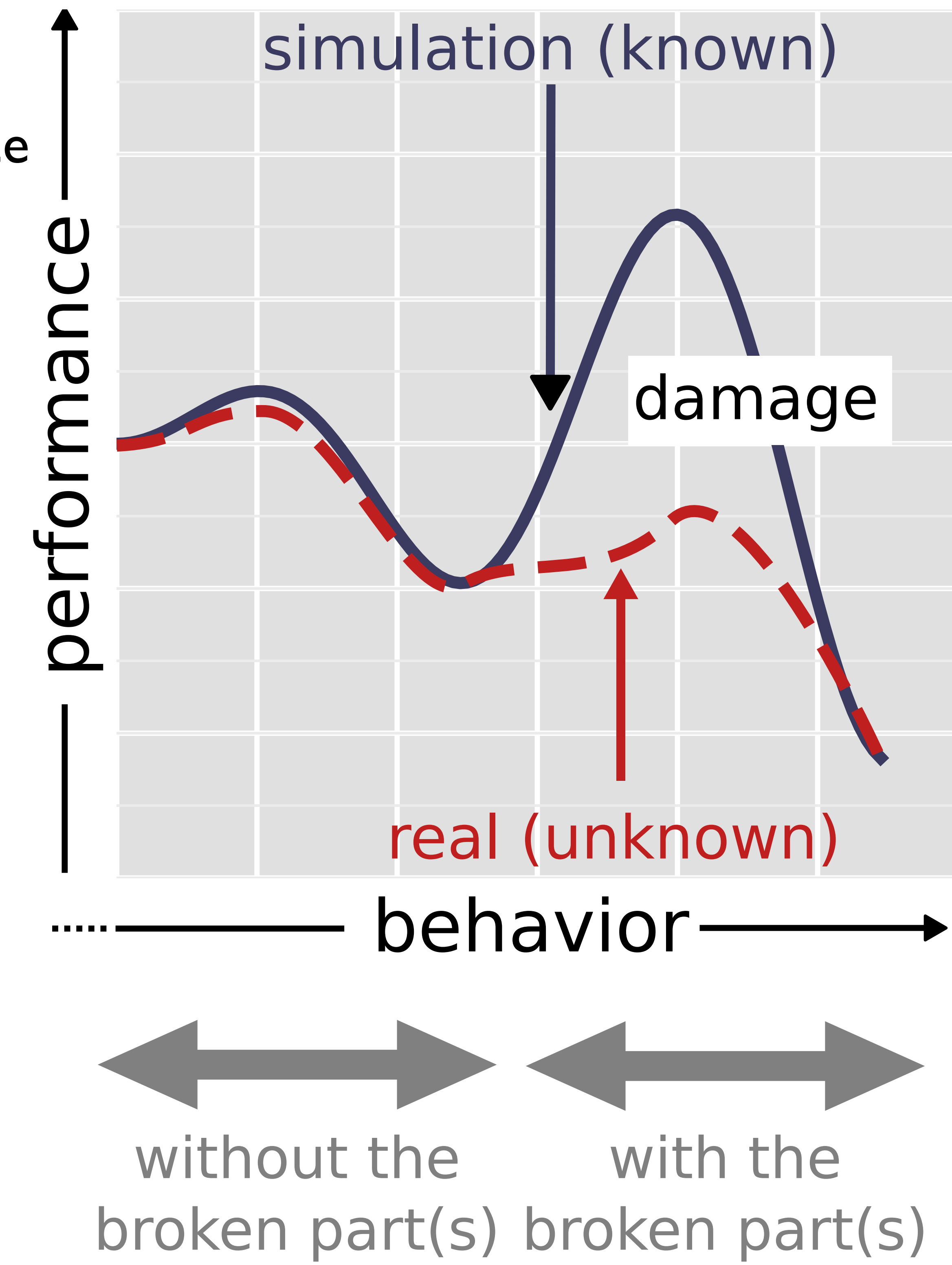
- guided by instincts (evolution) & experience
- constrained by their body (evolution)

Some solutions are more likely to be tested than others

Robots:

1. build the « right » search space (evolution)
2. use a simulation of the intact robot to guide a trial & error learning process (*instincts*)

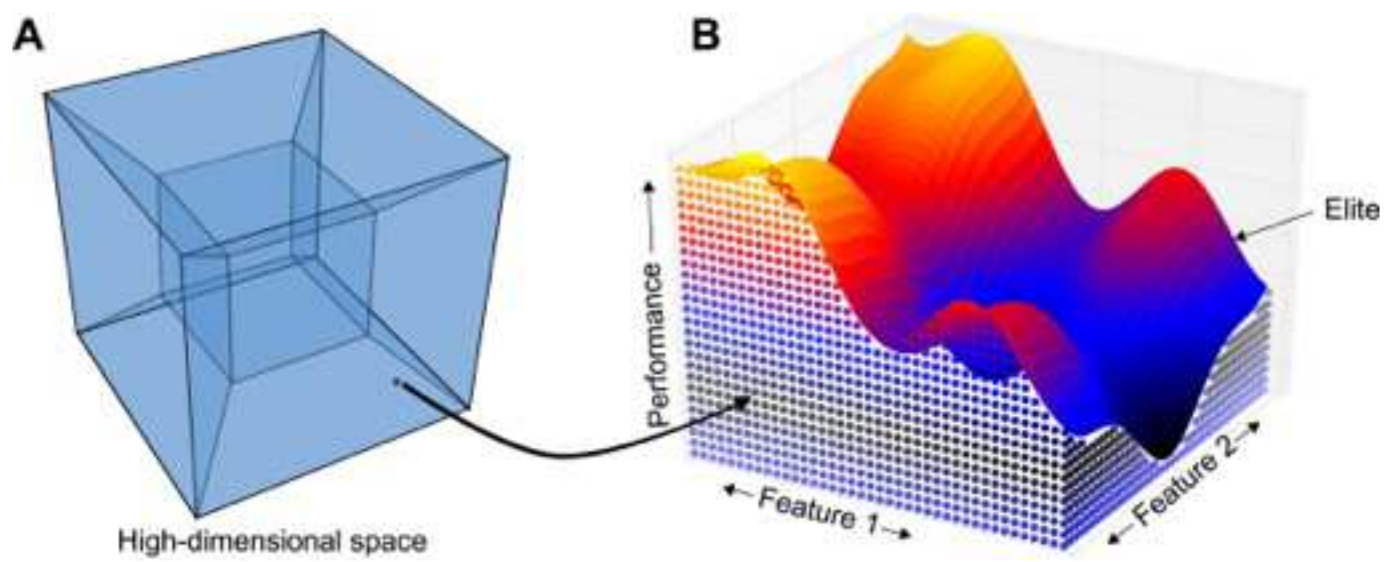
Simulation / reality: similar performance when not using the broken parts

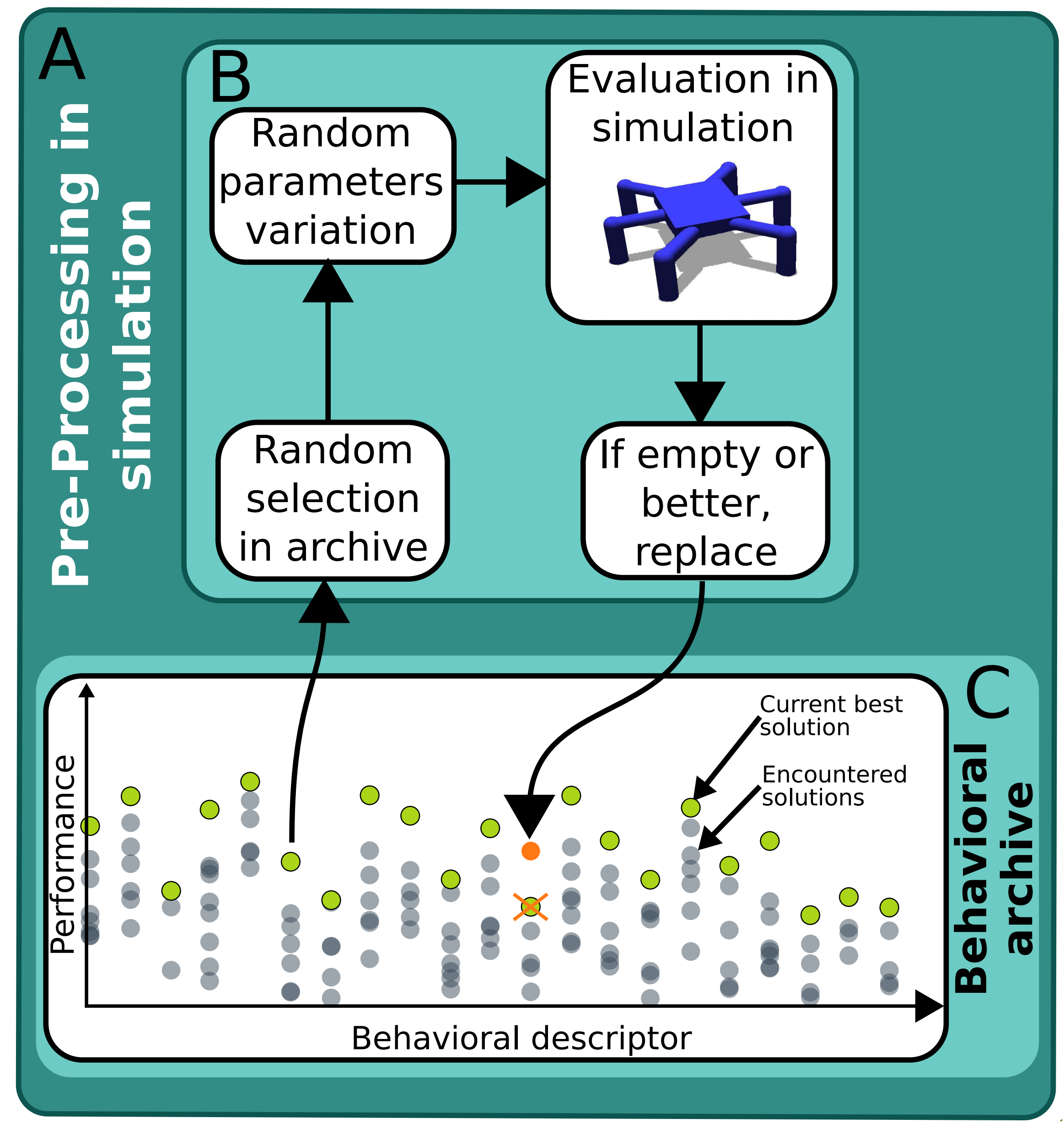


Koos, Sylvain, Jean-Baptiste Mouret, and Stéphane Doncieux. "The transferability approach: Crossing the reality gap in evolutionary robotics." IEEE Transactions on Evolutionary Computation 17.1 (2013): 122-145.

The Map-Elites algorithm

Multi-dimensional Archive of Phenotypic Elites





Goal: find many good alternatives

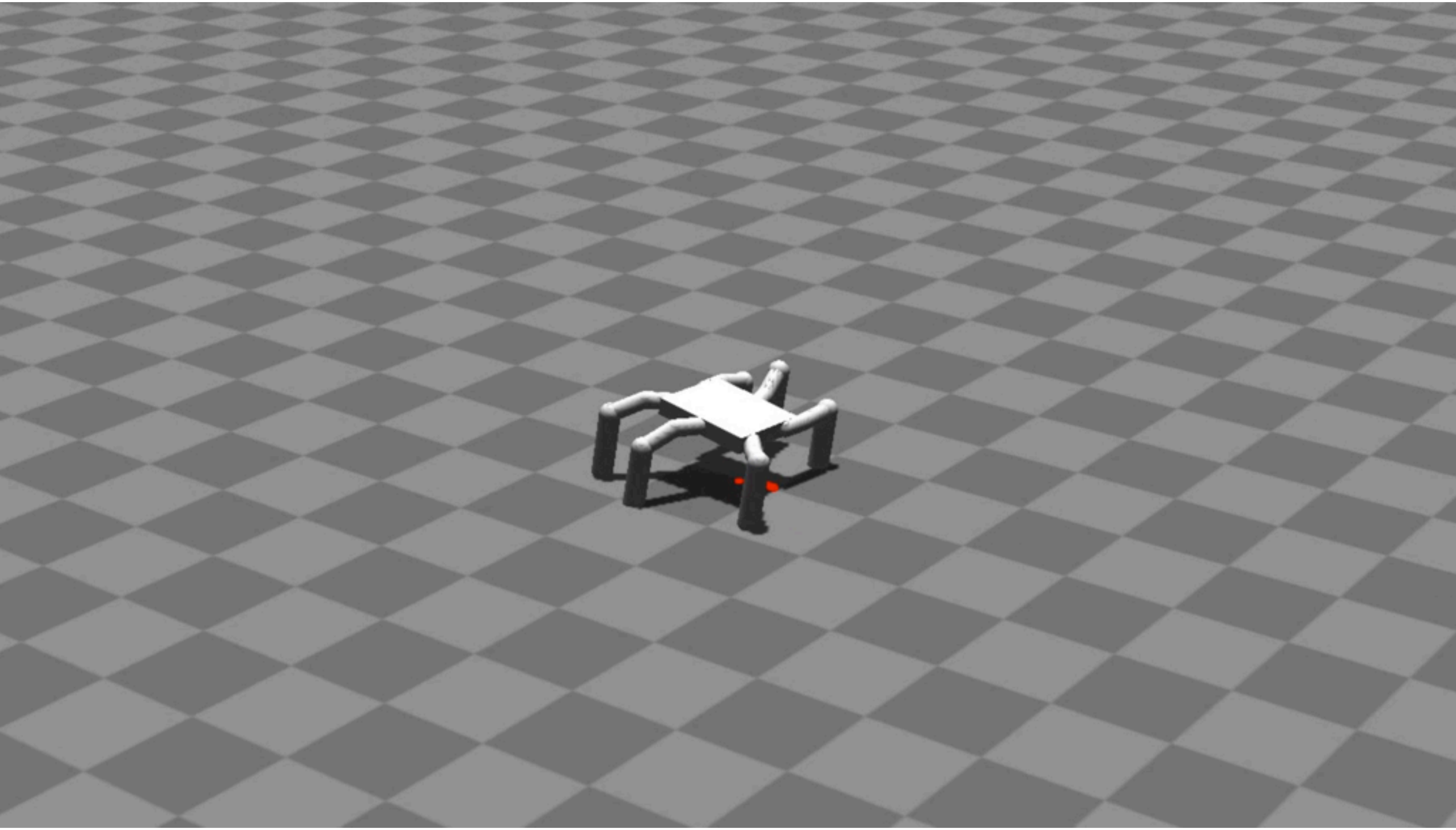
➠ The **elites** of the search space

Elite = best of the family

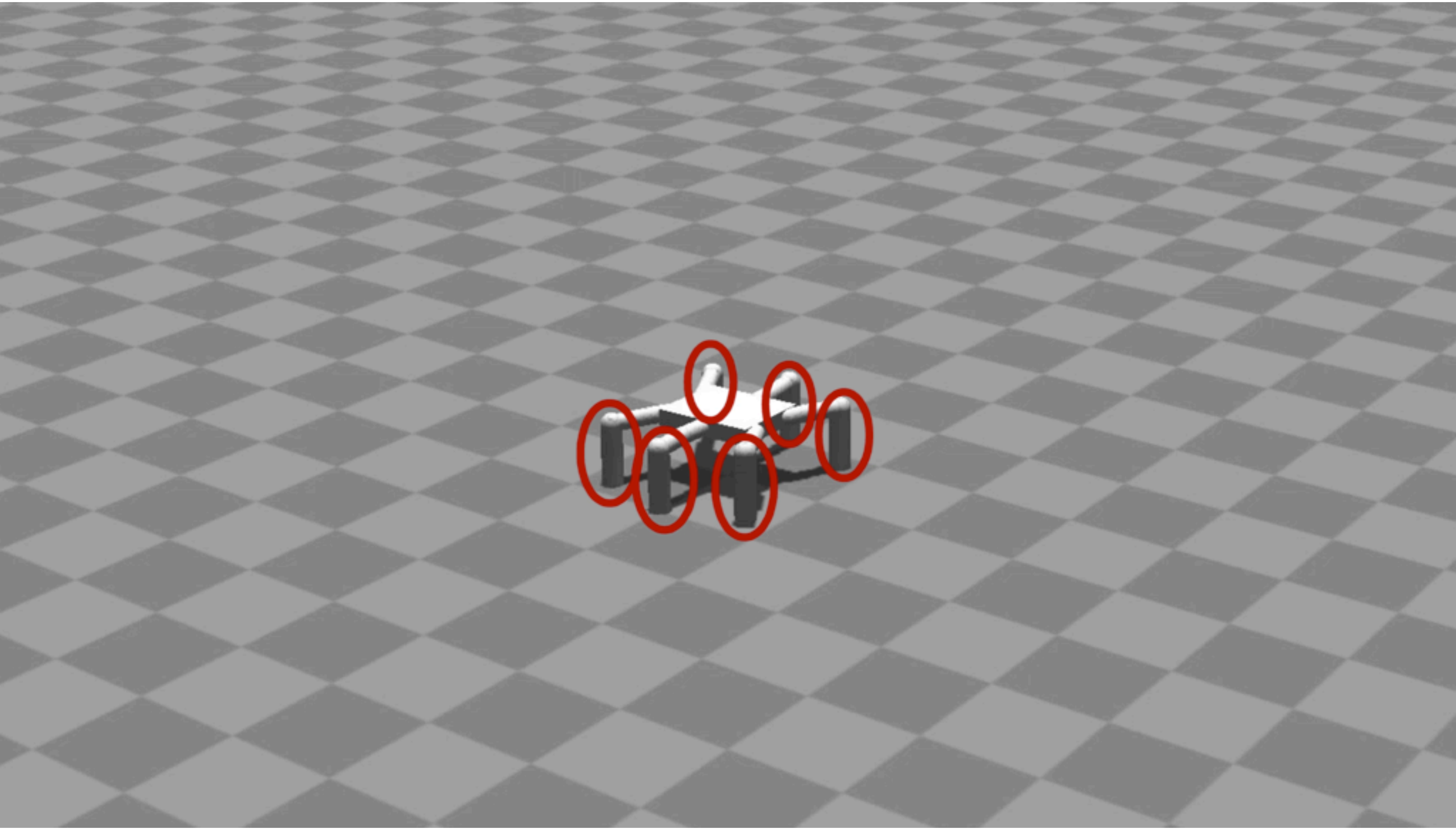
Family = solution with similar features (niche)

Mouret, J.-B., and J. Clune. "Illuminating search spaces by mapping elites." arXiv preprint arXiv:1504.04909 (2015).

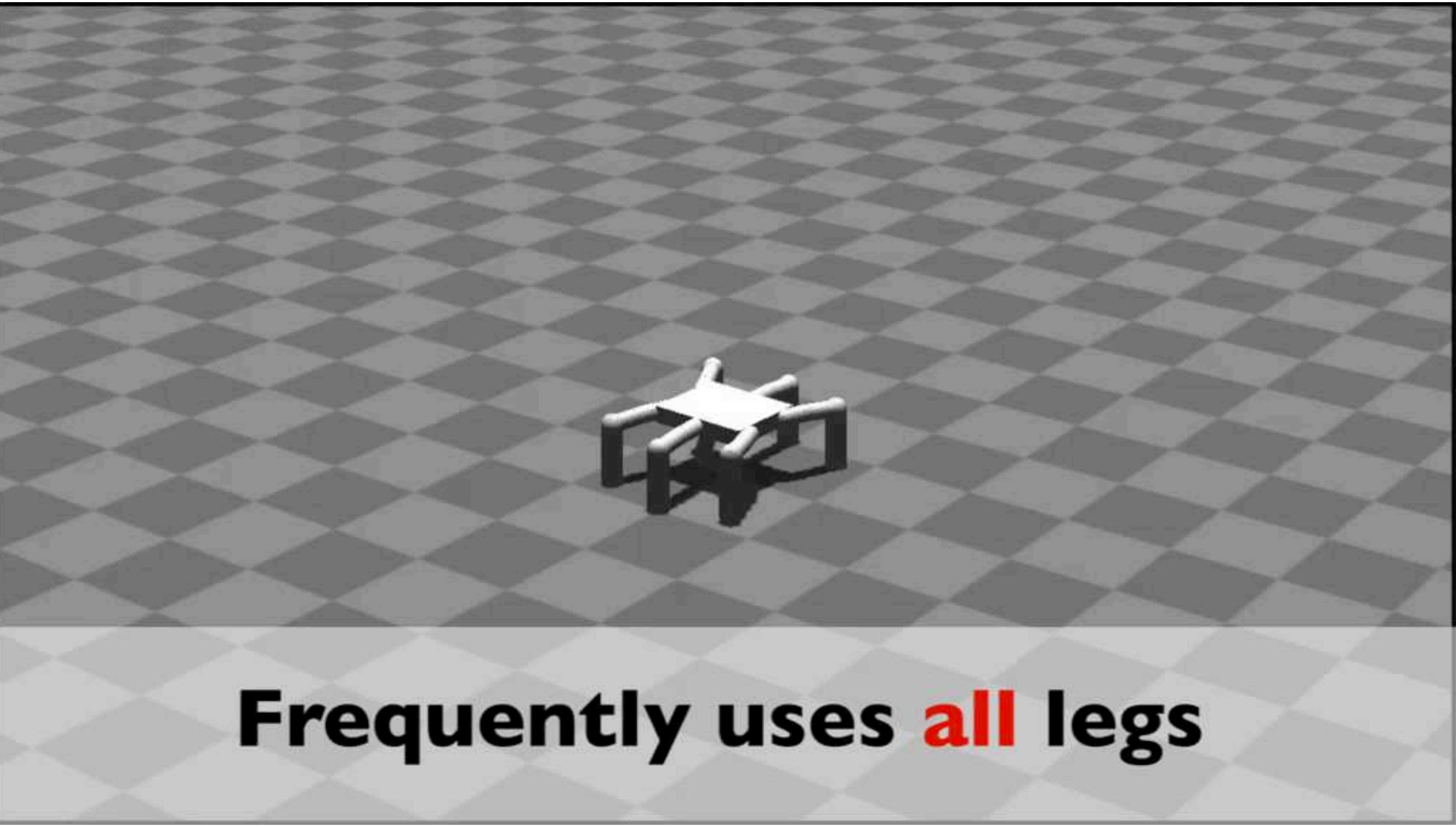
MAP-Elites: 6-legged locomotion



MAP-Elites: 6-legged locomotion



MAP-Elites: 6-legged locomotion



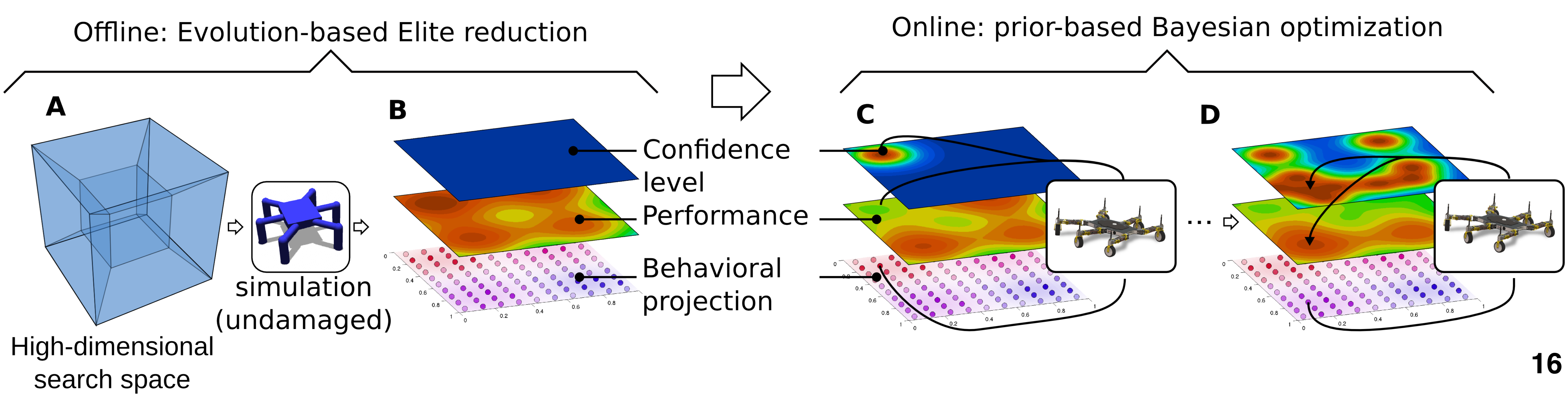
Intelligent Trial & error

An illumination algorithm generates the prior

- in simulation
- with an <u>intact</u> robot
- <u>many evaluations [simulation]</u>
- " take the needles out of the haystack"

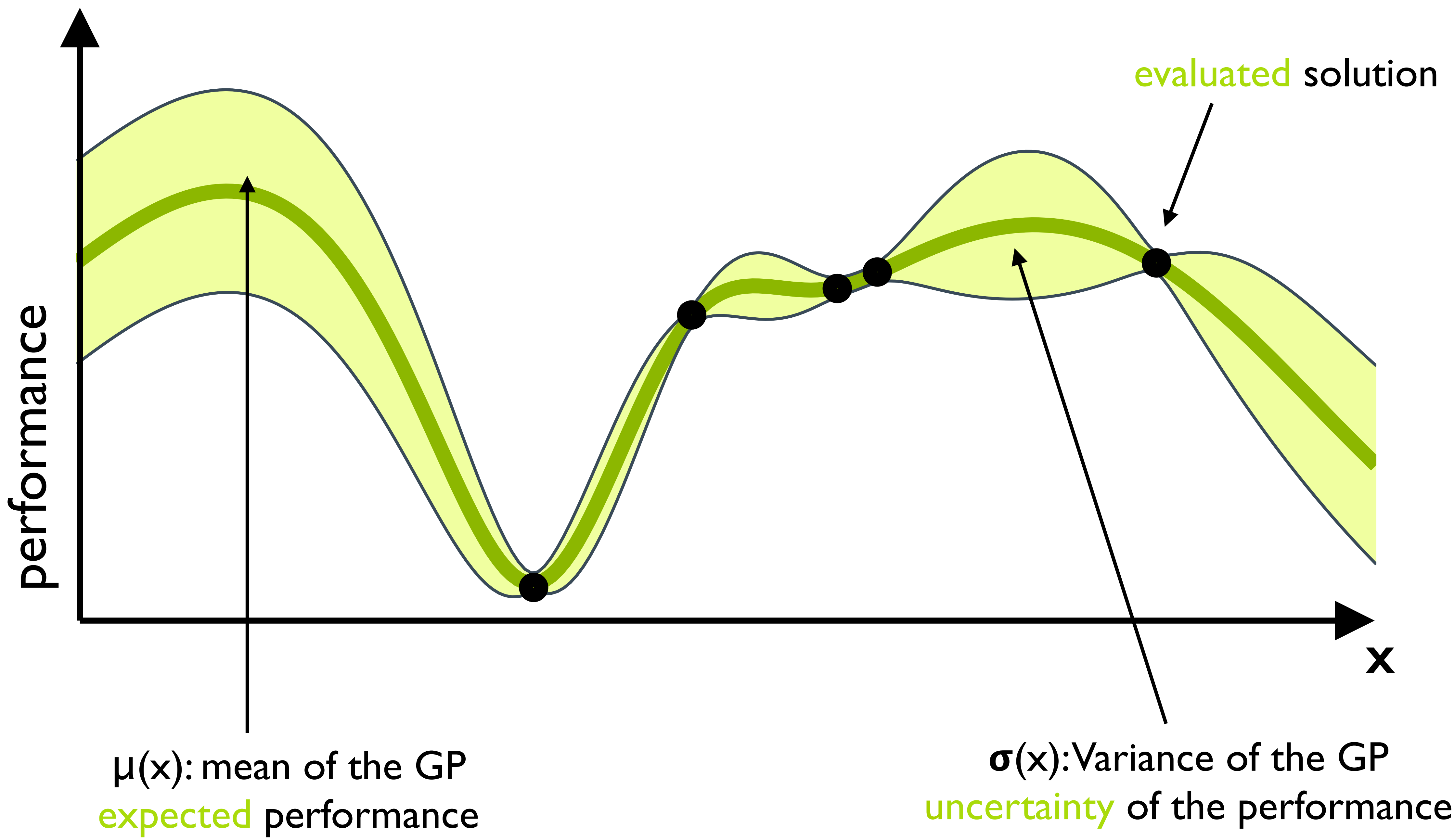
Prior-based Bayesian optimization do the online learning

- trial-and-error
- <u>few evaluations [real robot]</u>



Bayesian optimization: online adaptation

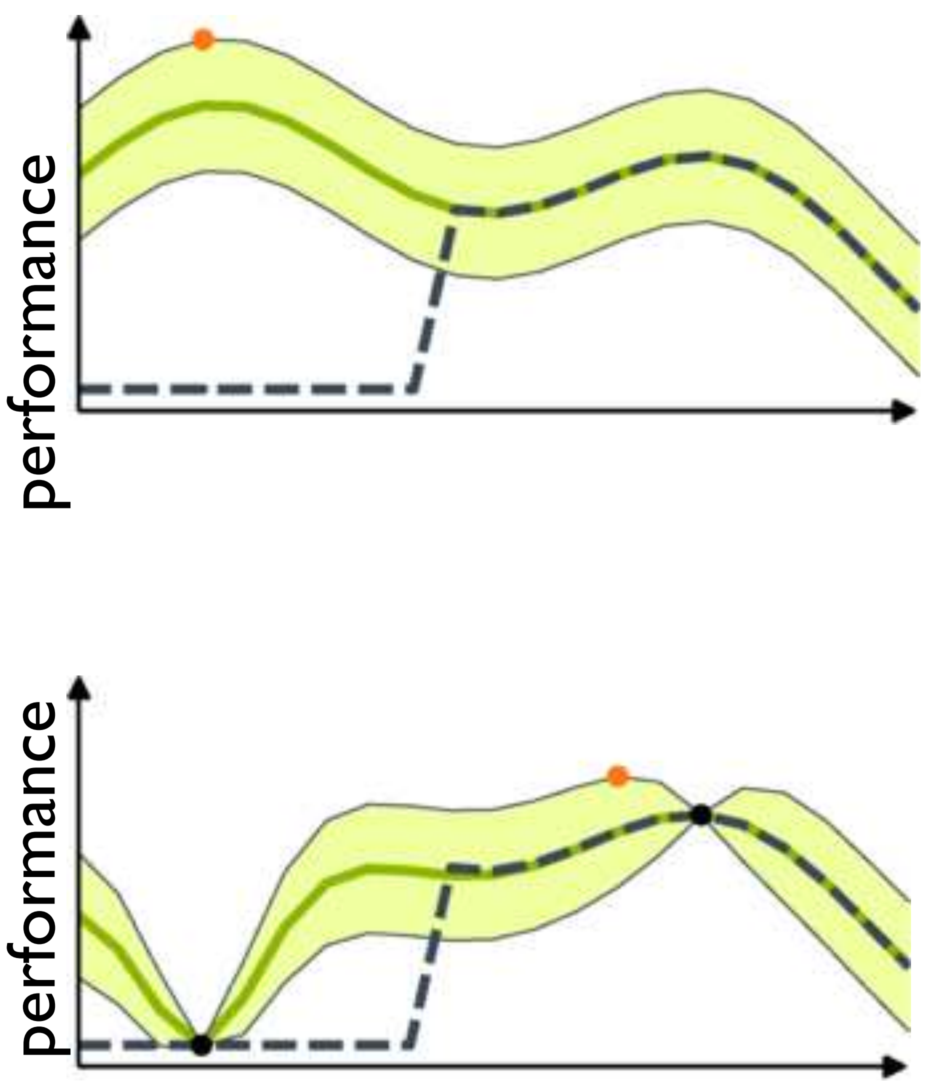
(with Gaussian processes)



Tutorial: Brochu, Cora & De Freitas. arXiv 2010 — **4-legged locomotion:** Lizotte, Wang, Bowling & Schuurmans. IJCAI 2007 — **2-legged locomotion:** Calandra et al. ICRA 2014

Bayesian Optimization + MAP-Elites

"Intelligent Trial and Error"



$$P(f(\mathbf{x})|\mathbf{P}_{1:t+1}, \mathbf{x}) = \mathcal{N}(\mu_{t+1}(\mathbf{x}), \sigma^2_{t+1}(\mathbf{x}))$$

where

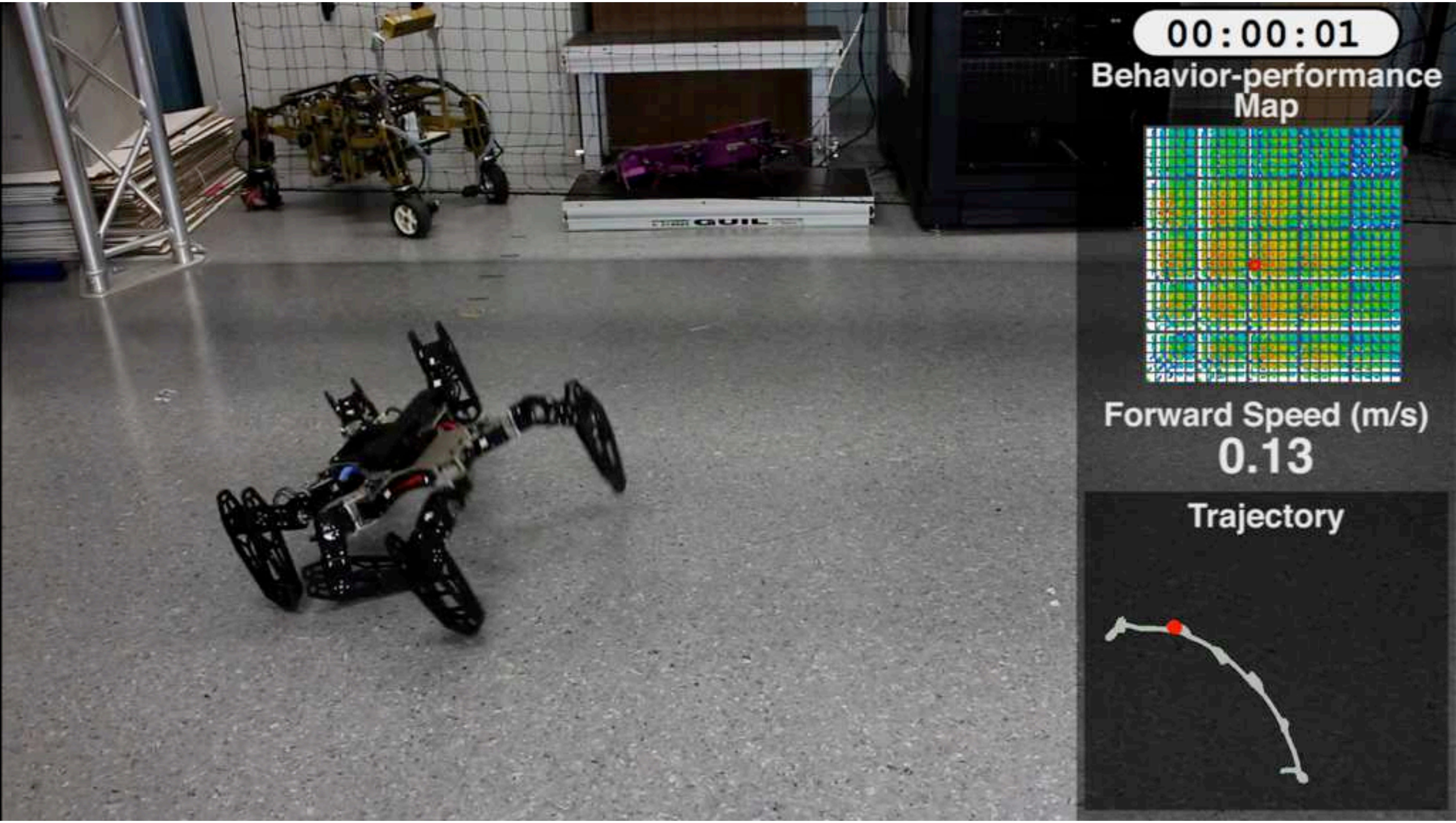
$$\mu_{t+1}(\mathbf{x}) = \mathcal{A}(\mathbf{x}) + \mathbf{k}^t\mathbf{K}^{-1}(\mathbf{P}_{1:t+1} - \mathcal{A}(\mathbf{y}_{1:t+1}))$$

$$\sigma^2_{t+1}(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \mathbf{k}^t\mathbf{K}^{-1}\mathbf{k}$$

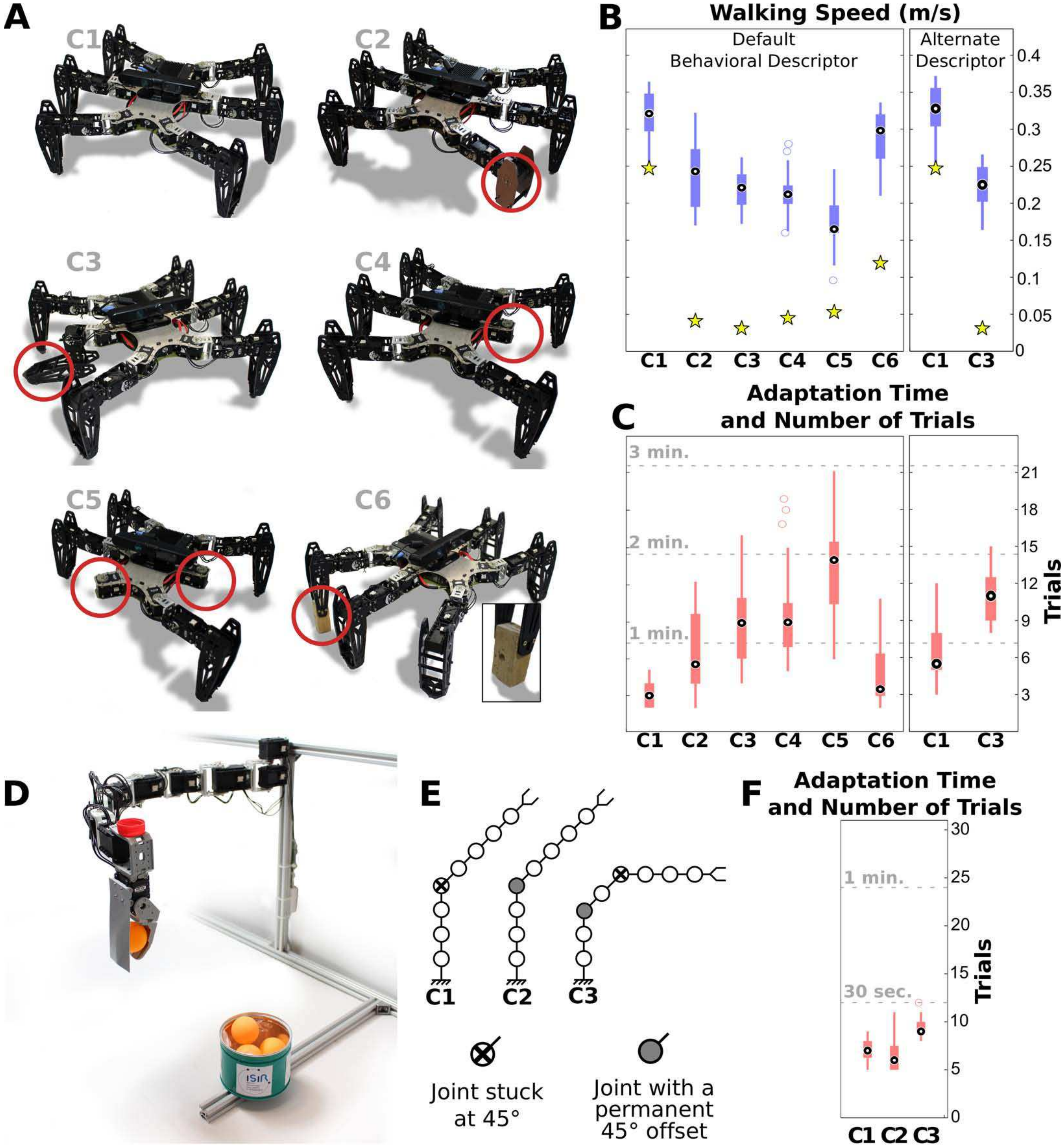
$$\mathbf{K} = \begin{bmatrix} k(\mathbf{y}_1, \mathbf{y}_1) + \sigma^2_{noise} & \cdots & k(\mathbf{y}_1, \mathbf{y}_t) \\ \vdots & \ddots & \vdots \\ k(\mathbf{y}_t, \mathbf{y}_1) & \cdots & k(\mathbf{y}_t, \mathbf{y}_t) + \sigma^2_{noise} \end{bmatrix}$$

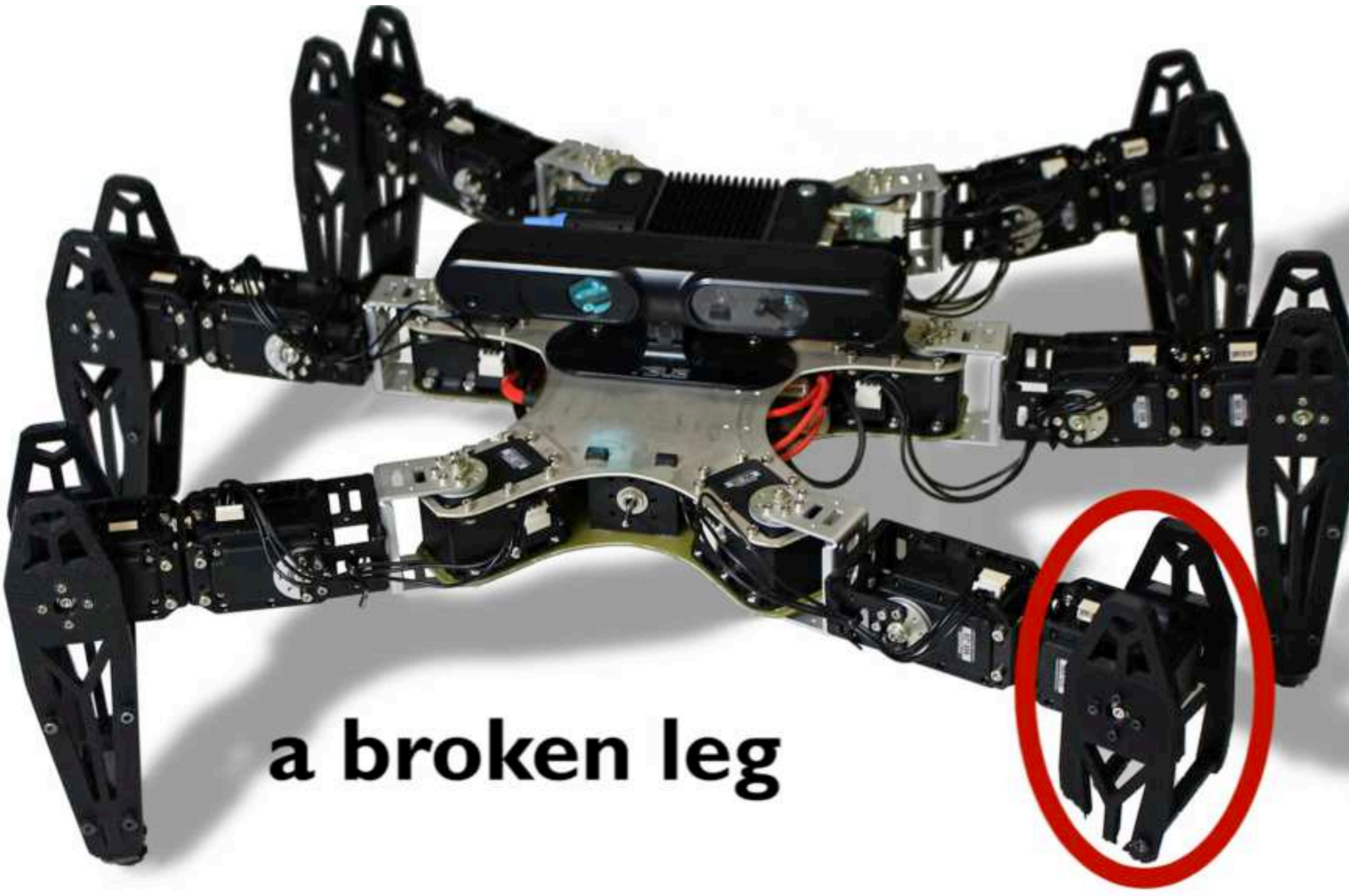
$$\mathbf{k} = \begin{bmatrix} k(\mathbf{x}, \mathbf{y}_1) & k(\mathbf{x}, \mathbf{y}_2) & \cdots & k(\mathbf{x}, \mathbf{y}_t) \end{bmatrix}$$

Cully, A. and Clune, J. and Tarapore, D. and Mouret, J.-B. (2015).
Robots that can adapt like animals. **Nature**. Vol 521 Pages 503-507.

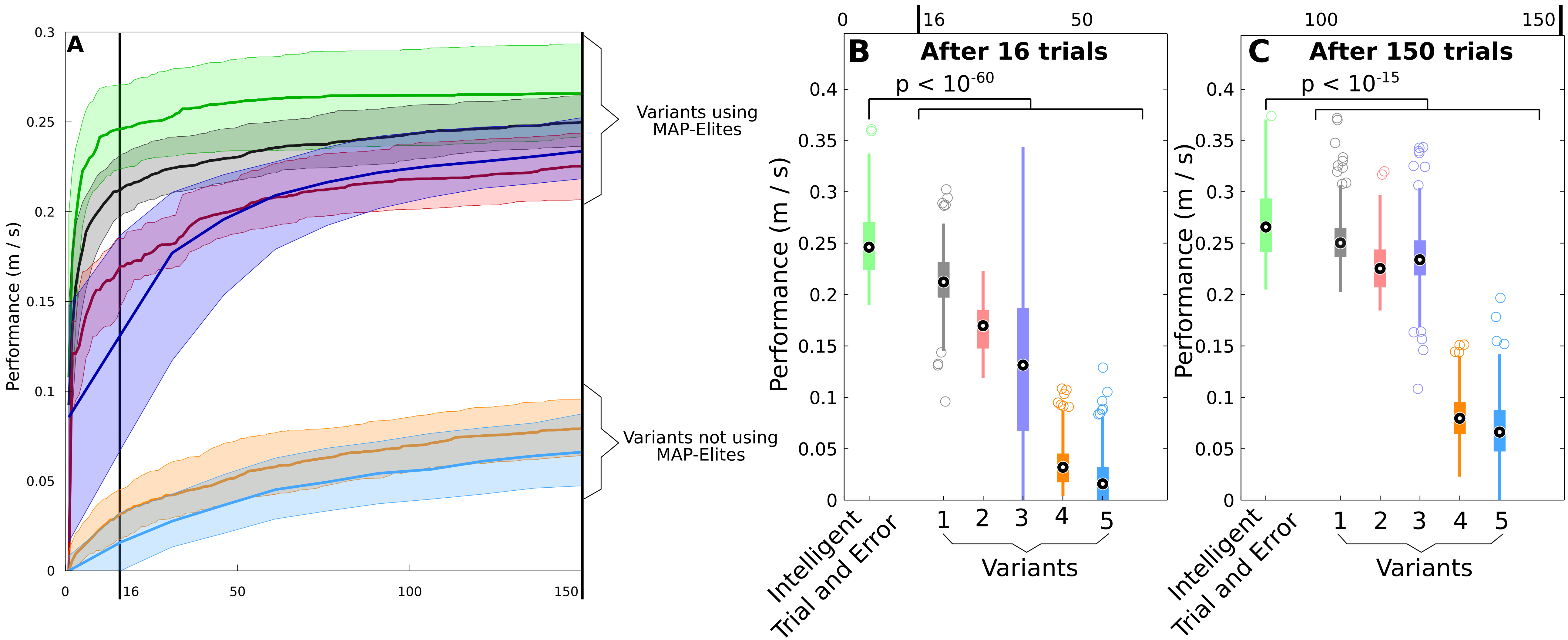


- Controller : periodical signals (36 parameters)
- Fitness: covered distance in 5 seconds
- Fitness evaluated onboard (RGB-D visual odometry)


A
C1
C2
C3
C4
C5
C6
B
Walking Speed (m/s)
Default
Behavioral Descriptor
Alternate
Descriptor
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
C1 C2 C3 C4 C5 C6 C1 C3
C
Adaptation Time
and Number of Trials
3 min.
2 min.
1 min.
21
18
15
12
9
6
3
Trials
C1 C2 C3 C4 C5 C6 C1 C3
D
E
C1
C2
C3
Joint stuck
at 45°
Joint with a
permanent
45° offset
F
Adaptation Time
and Number of Trials
30
25
20
15
10
5
0
1 min.
30 sec.
Trials
C1 C2 C3


a broken leg

Comp. with other approaches (simulation)



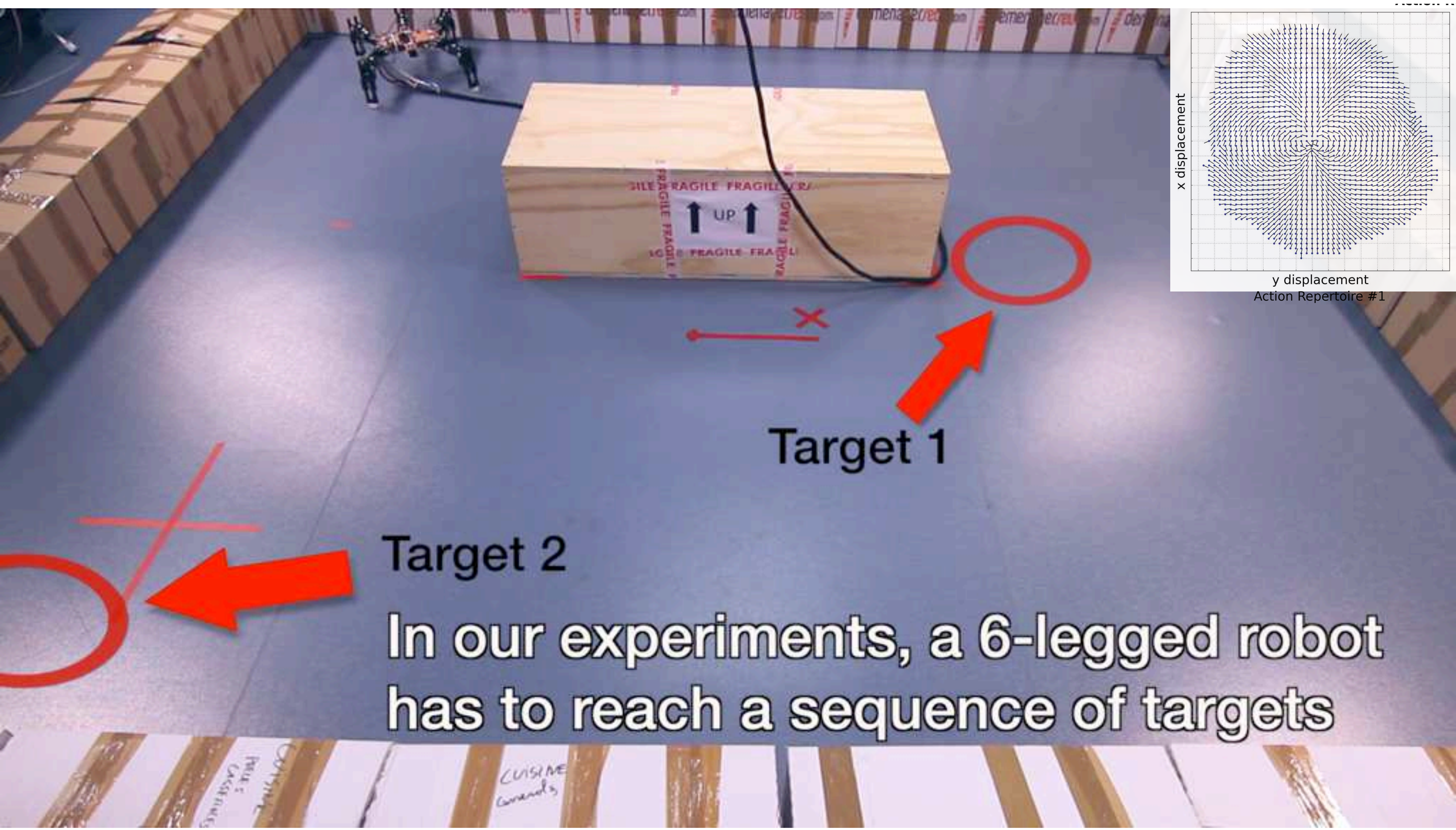
Variant	Behavioral repertoire creation	Priors on performance	Search algorithm	equivalent approach
Intelligent Trial and Error	MAP-Elite	yes	Bayesian Optimization	-
Variant 1	MAP-Elite	none	Bayesian optimization	-
Variant 2	MAP-Elite	none	policy gradient	-
Variant 3	MAP-Elite	none	random search	-
Variant 4	none	none	Bayesian optimization	Lizotte et al. (2007) (*33*)
Variant 5	none	none	policy gradient	Kohl et al. (2004) (*23*)

Kohl & P. Stone. In Proc. of the IEEE Int. Conf. on Robotics and Automation 2004.
Lizotte, Wang, Bowling, Michael & Schuurmans. In Proc. of the Int. Joint Conf. on Artificial Intelligence 2007.

22

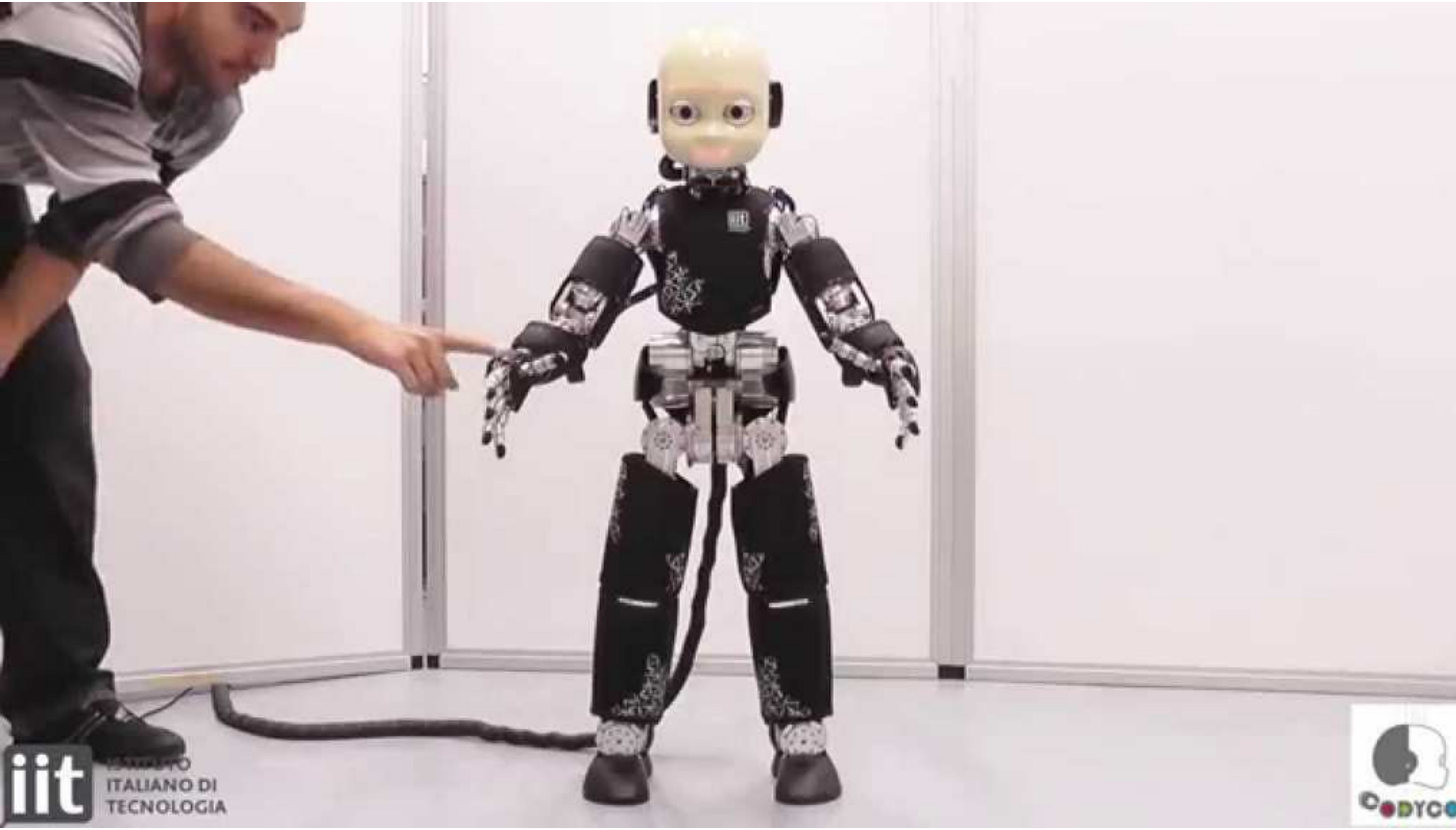
Undamaged robotic arm

Reset-free IT&E / GP + MCTS

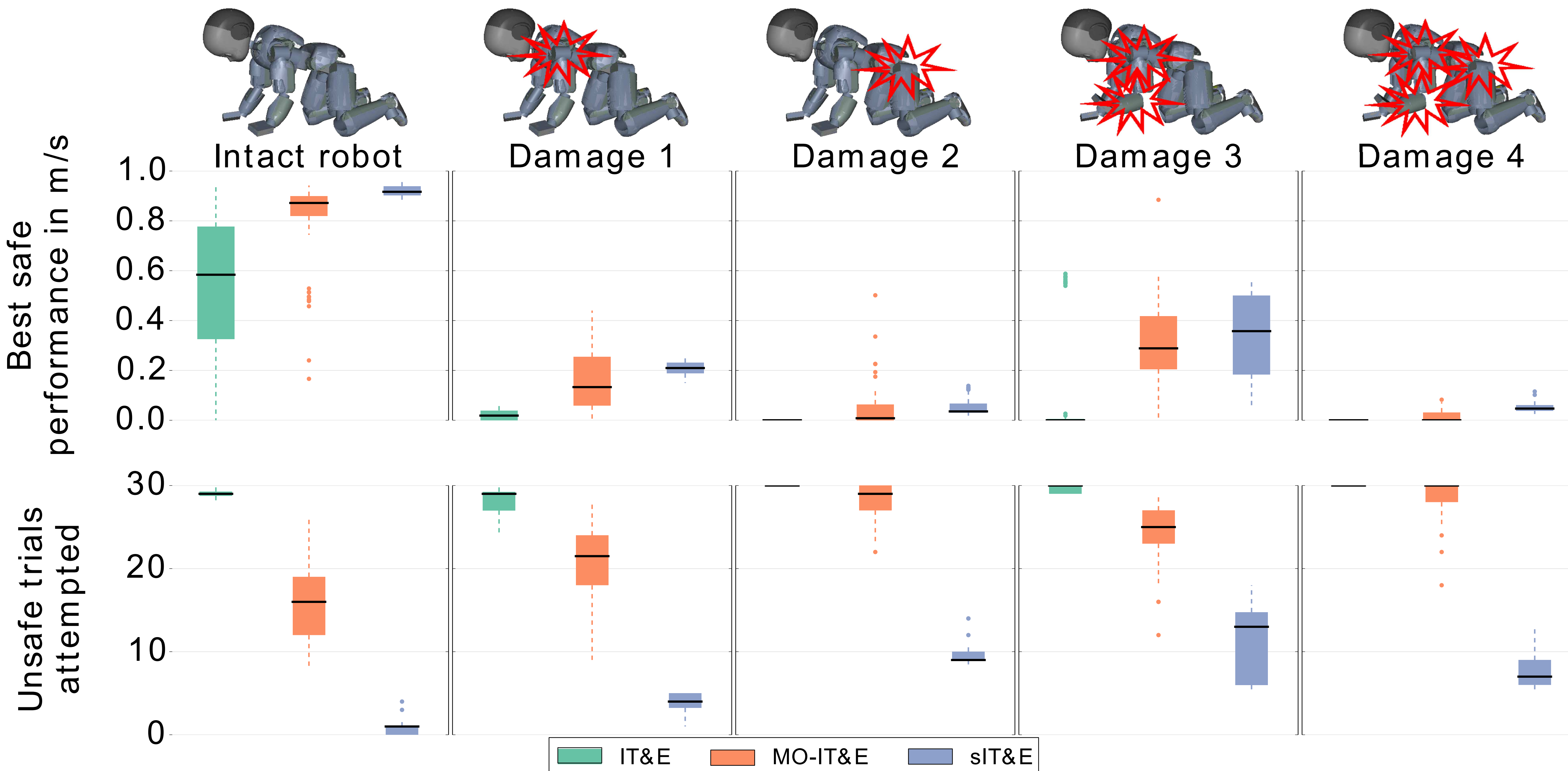


K. Chatzilygeroudis, V. Vassiliades, and J.-B. Mouret (2016). *Reset-free Trial-and-Error Learning for Data-Efficient Robot Damage Recovery*. Submitted.

Expensive robots — iCub



Safe optimization



Trial & Error

Multi-objective Trial & Error

Safe Trial & Error

V. Papaspyros. K. Chatzilygeroudis, V. Vassiliades, and J.-B. Mouret (2016). *Safety-Aware Robot Damage Recovery Using Constrained Bayesian Optimization and Simulated Priors*. Submitted.


IT&E
After 30 Trials
Crawling speed: 0.60 m/s
Measurement
Dangerous
Safety threshold
Safe
MO-IT&E
After 30 Trials
Crawling speed: 0.20 m/s
Measurement
Dangerous
Safety threshold
Safe
sIT&E
After 30 Trials
Crawling speed: 0.22 m/s
Measurement
Dangerous
Safety
Safe

Conclusion

Damage recovery

- recovery in less than 15 trials (1-2 minute)
- for the 5 damage conditions (at least) / 2 robots
- 36 parameters to learn
- reset-free extension
- safe extension

➡ It works well!

… but

- no guarantee of optimality
- need a working sensor (reward)

The next steps

1. Choose wisely what to test next (active learning)
   ➡ OK to trade data resources for computational resources
2. Exploit every bit of information from each test
   ➡ e.g., use all the points of a trajectory
3. Only learn what is necessary
   ➡ e.g, rely on QP whole body control

Future work

4. Use prior knowledge
   i. use the "right" search space (here, MAP-Elites)
   ii. make prior knowledge explicit (here, the map)
   iii. use everything we know (e.g. here, simulator of the intact robot)

All the precepts should be combined

Cully, A. and Clune, J. and Tarapore, D. and Mouret, J.-B.
Robots that can adapt like animals.
Nature. Vol 521
Pages 503-507.(2015).

UPMC SORBONNE UNIVERSITÉS
ISIR INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE
Inria INVENTEURS DU MONDE NUMÉRIQUE
erc European Research Council
resibots
ANR

Doctorants, post-doctorants & ingénieurs

								
Sylvain Koos doctorant	Paul Tonelli doctorant	Antoine Cully doctorant	Danesh Tarapore post-doctorant	J.-M. Jehanno ingénieur	Vassilis Vassiliades post-doctorant	K. Chatzilygeroudis doctorant	Dorian Goepp ingénieur	Adam Gaier doctorant