What do Braitenberg vehicles believe?

Manuel Baltieri - 8th January 2025

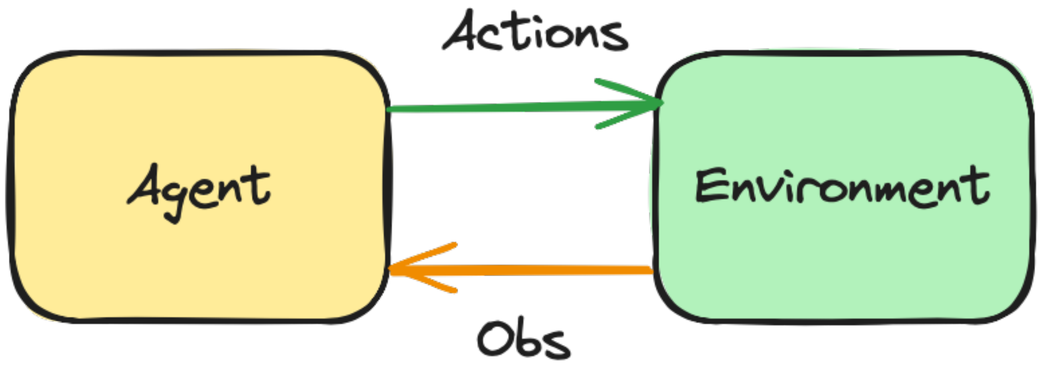
ARAYA

Contents

- Agents and (simple) beliefs
- A detour to Fristonland
- Observer-dependent vs independent agency
- Braitenberg vehicles’ beliefs

What I am interested in

Background

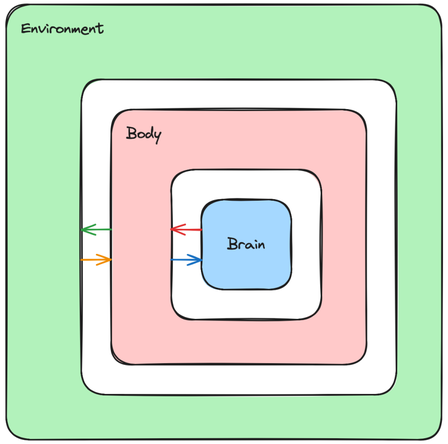


«[…] the rule “collect truth for truth’s sake” may be justified when the truth is unchanging; but when the system is not completely isolated from its surroundings, and is undergoing secular changes, the collection of truth is futile, for it will not keep.»

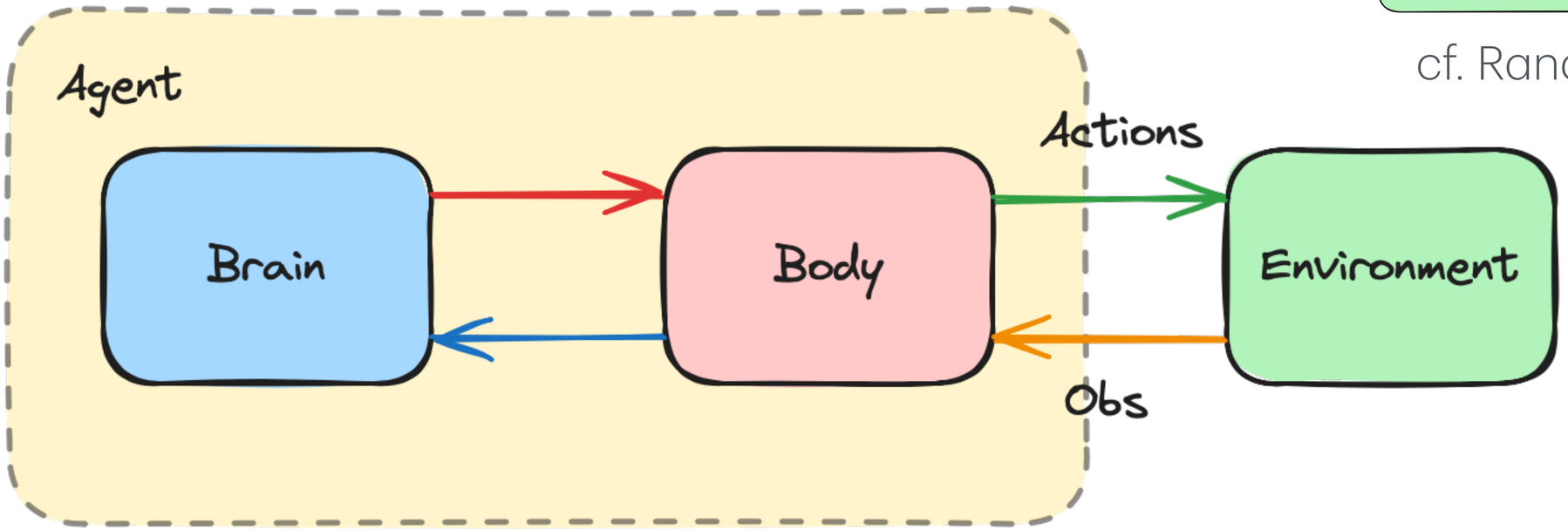
(Ashby, 1958)

Unpacking that a little

Factorising the agent

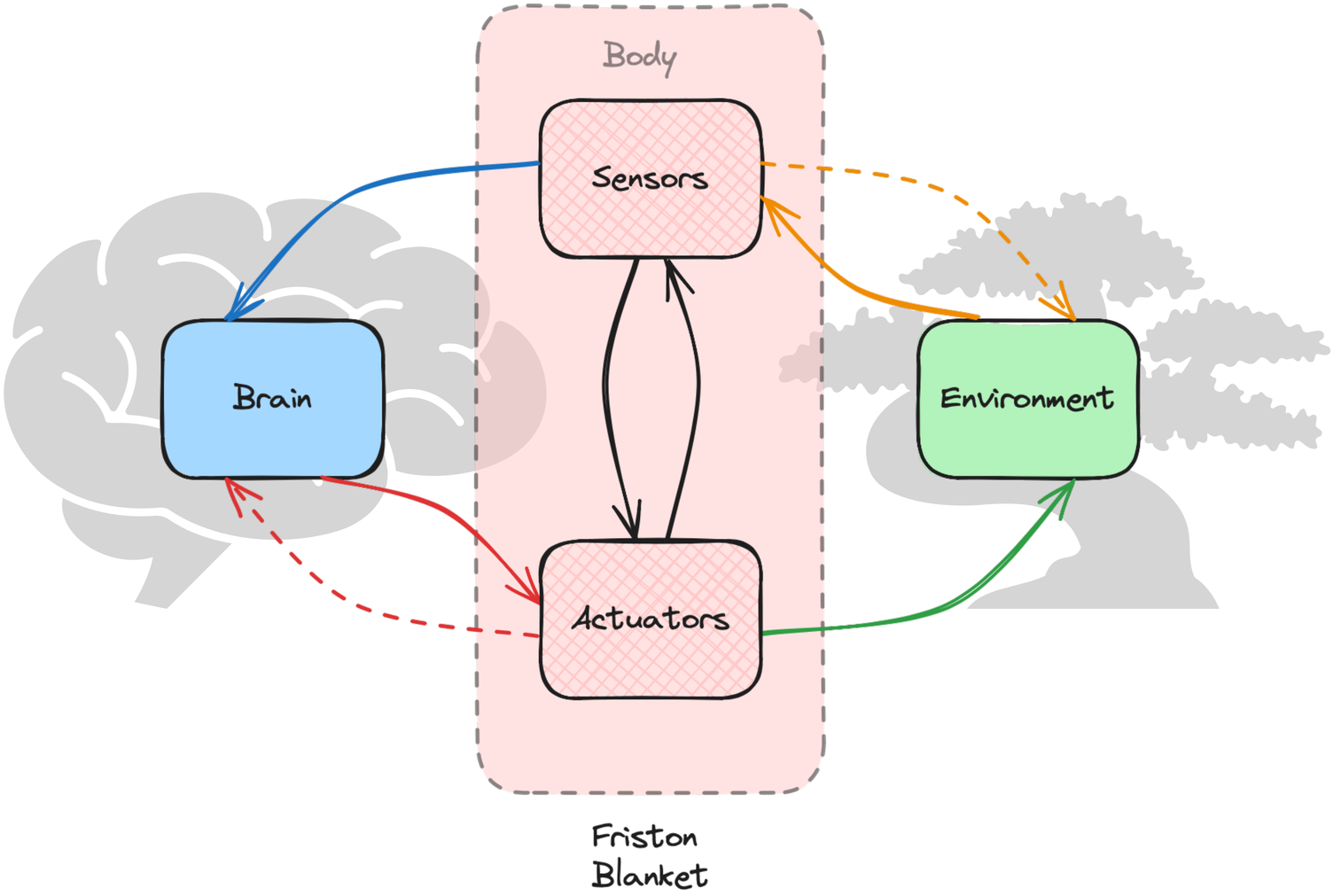


cf. Randy Beer

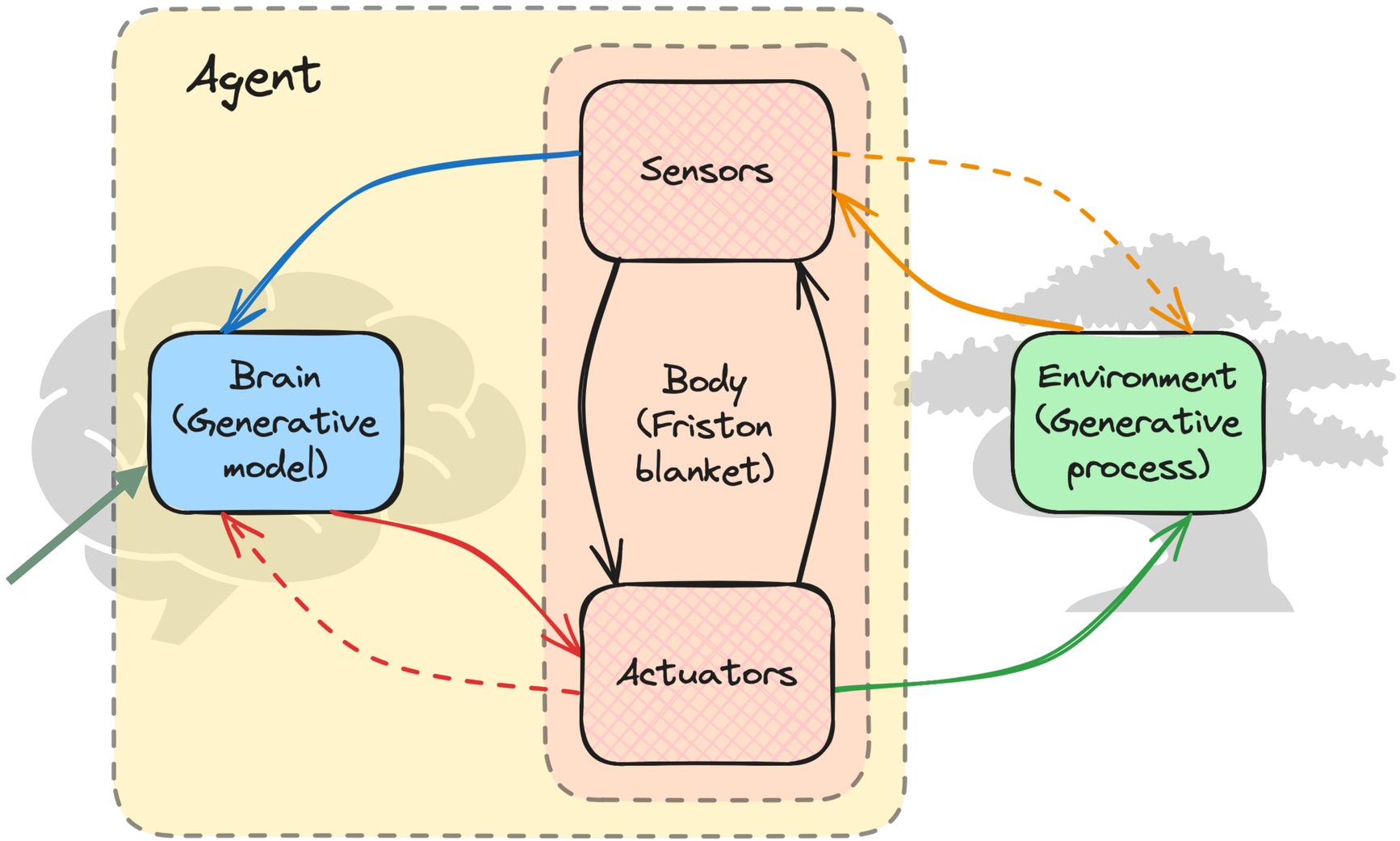


~~Alice~~ Manuel in Fristonland

Friston blankets, boundary factored into sensors and actuators



Brain states parametrise beliefs about “external” states, GM describes such beliefs



What agents “know” about their environment

And what we should believe agents “know”

- What beliefs can we attribute to an agent solving a task?
- What are some interesting (minimal?) classes of such beliefs?
- What goals can we attribute to an agent?
- What is the relation between goals and beliefs we attribute to a system?
- …

Observer-dependent vs independent agency

Observer-dependent AND independent agency?

- Observer-dependent
  - e.g. intentional stance (an observer wants to predict the behaviour of a system, if physical stance is too complicated, then use a more abstract stance)
  - definitions dependent on skills/limitations of an observer
- Intrinsic agency
  - e.g. autopoietic view (ask Matt)
  - definitions independent of observers' features

The intentional stance

A detour

- Physical stance: predictions from knowledge of the physical constitution of a system and the physical laws that govern its operation (mass, velocity, etc.)
- Design stance: predictions from purpose, function and design of a system (birds fly when they flap their wings, wings are for flying, etc.)
- Intentional stance: interpreting and predicting the behaviour of an entity by treating it **as if** it has beliefs, desires, and intentions (birds fly when they think they are being chased by a cat)

«Neural representations, this work has suggested, are not action neutral mirrors of the world. Instead they are in some deep sense ‘action-oriented’ (Clark 1997, Engel et al. 2013). They are geared to promoting successful, fast, fluent actions and engagements for a creature with specific needs and bodily form. Such representations will be **as minimal as possible, neither encoding nor processing information in costly ways when simpler routines, combined with world-exploiting actions, can do the job**.»

(Clark, 2015)

Different types of generative models?

- Gathering knowledge vs. achieving a goal
- Simplified generative models, encoding sensorimotor information/Umwelt

Example: Outfielder problem (Fink et al., 2009)

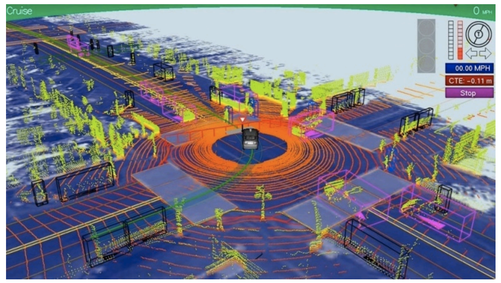
~~1) Trajectory prediction (TP)~~

2) Optical Acceleration Cancellation (OAC)

Action-oriented generative models

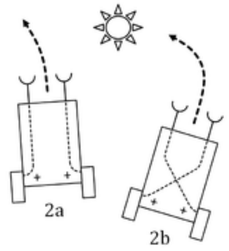
Example task: agent performing phototaxis

Perception-oriented



e.g. SLAM

Action-oriented



e.g. Braitenberg vehicles

The linebot

Some preliminary investigations

McGregor et al. (2015) look at FEP to understand what it can say about an agent's beliefs.

This agent is trying to reach a goal position when the only information available is high/low concentration of a certain chemical.

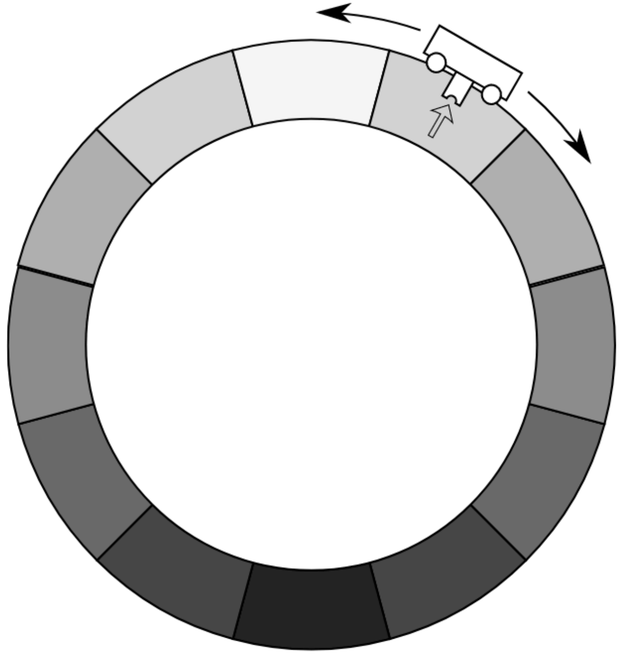


Figure 1: Illustration of agent-environment system. The agent has a sensor which reads $High$ or Low and is sensitive to chemical concentration. The agent's motor can attempt to move the agent clockwise or anticlockwise.

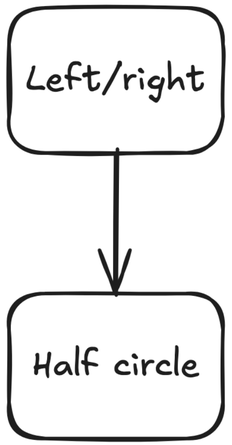
The linebot

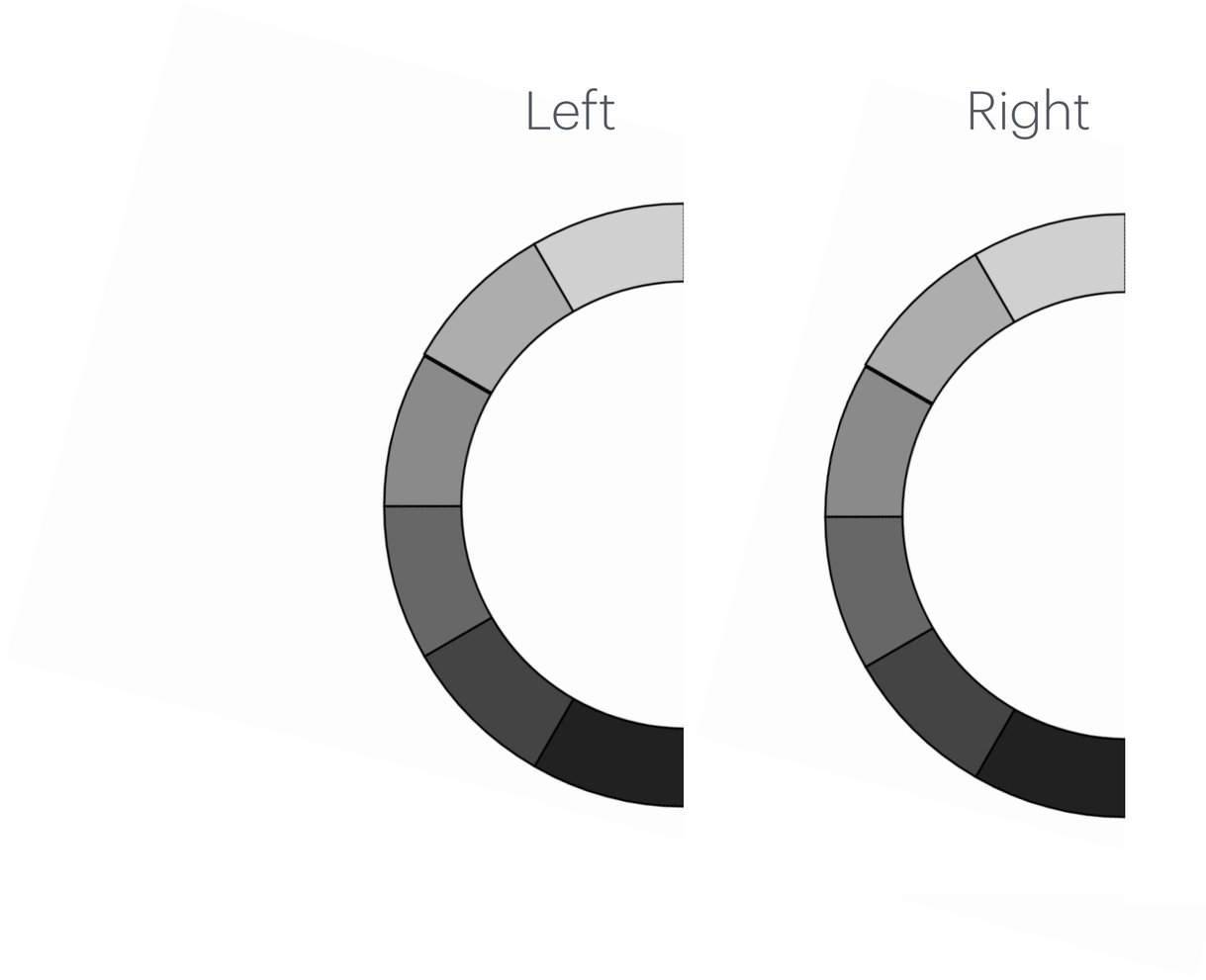
...with simplified beliefs

My master dissertation: what if the agent had some constraints, e.g. memory.

How would that affect its beliefs?

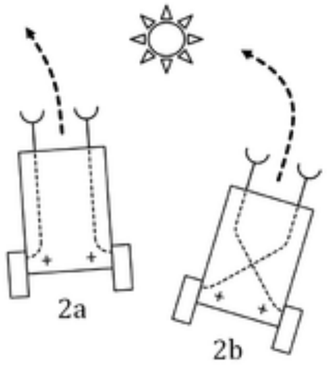
Simplified beliefs: hierarchical model with two levels: half circle + left/right.

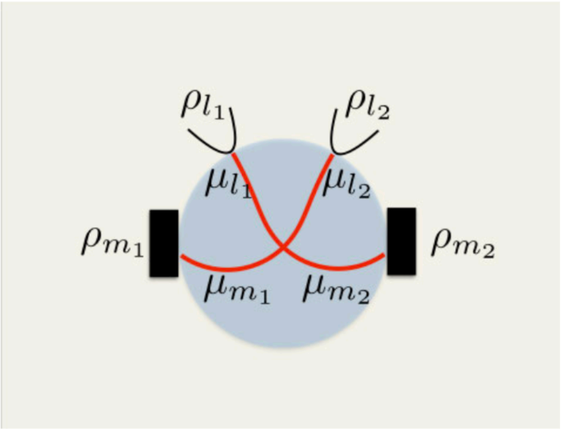




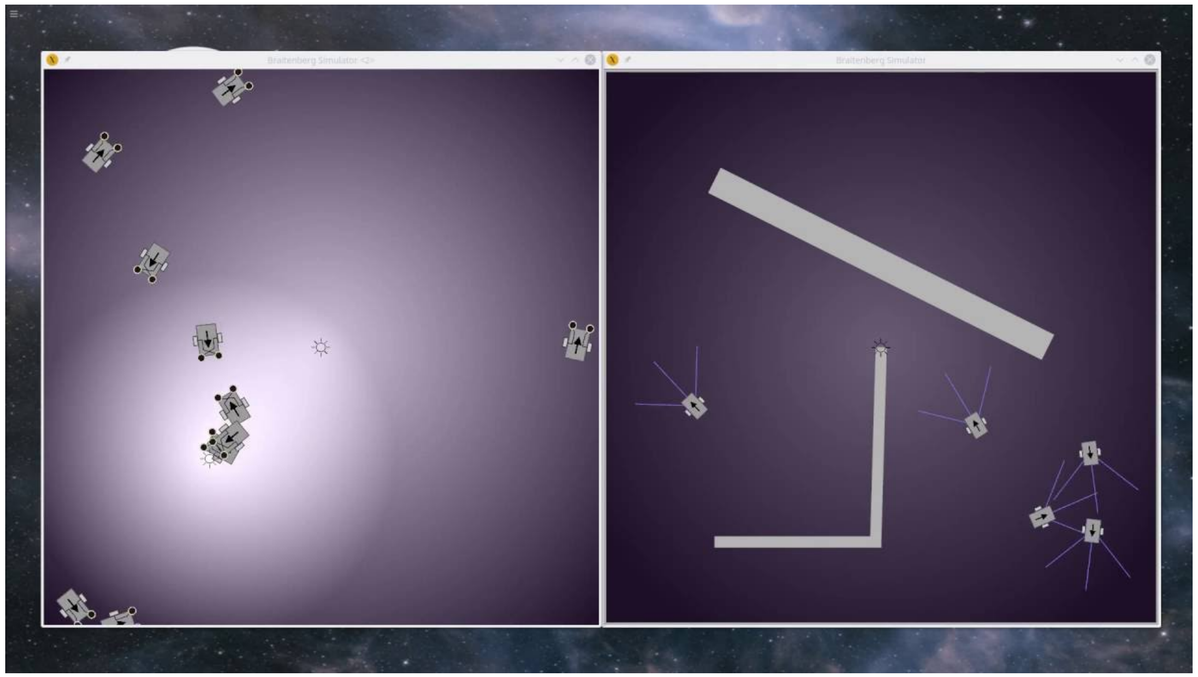
Braitenberg vehicles

Photo/chemo/rheo/tropo/... taxis





- Vehicles 2 and 3
- Agent with two sensors and two wheels
- Sensors and wheels connected by wires
- Implementation: (Left/right) Wheel rotational velocity = constant * (right/left) sensory reading



A hierarchy of models?

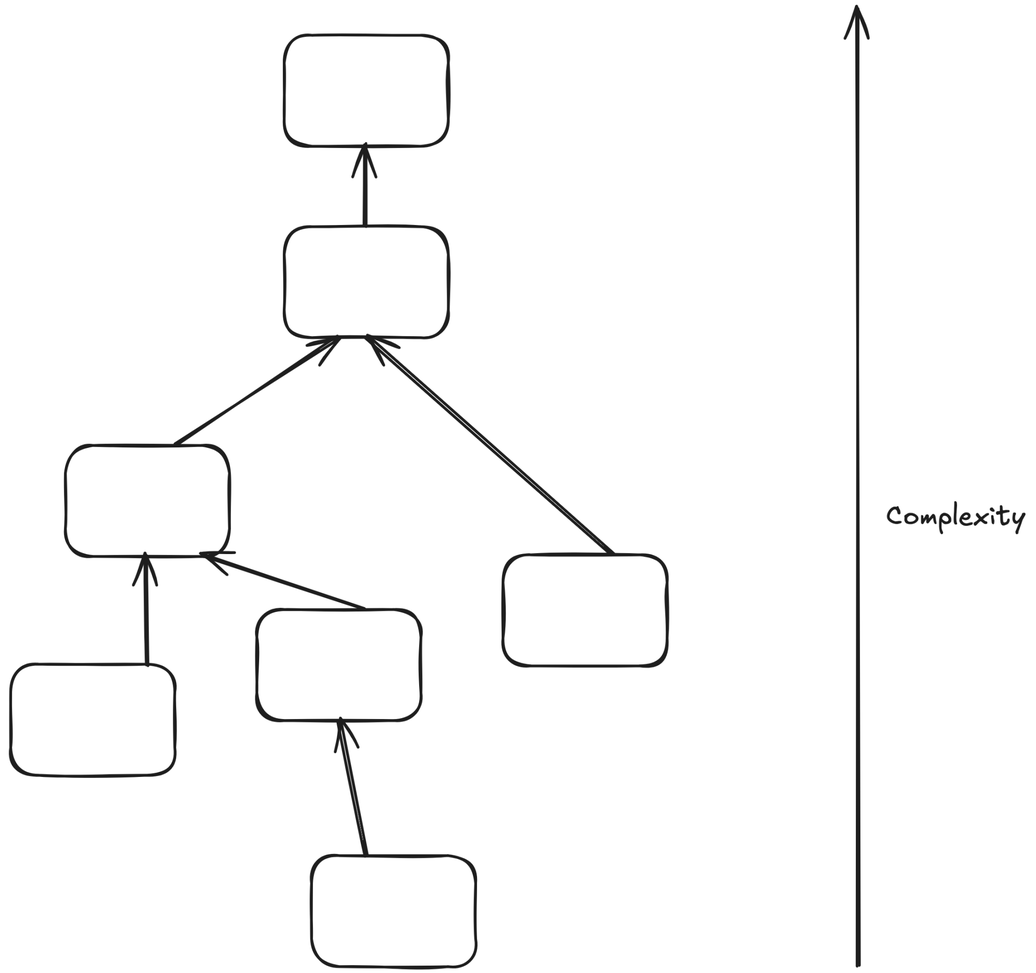
New (?) ideas

What if there is a hierarchy of models?

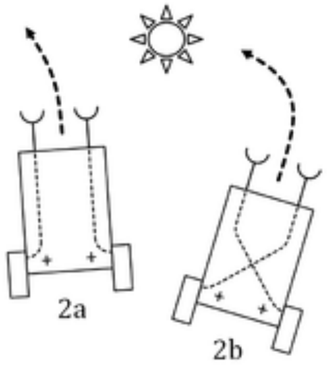
—> If agents solve the same task with different info, how do these solutions relate?

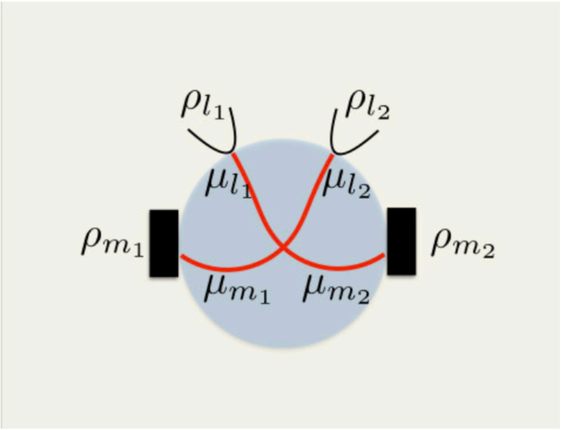
Can we describe the relation between interpretations of agents performing tasks?

Is it a lattice (ideal) or some weaker kind of order?



A vehicle's beliefs





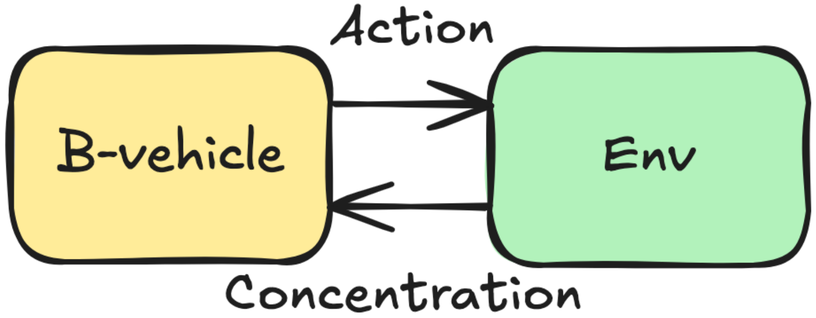
Taxis in terms of POMDPs and their possible compressions (not covered here)

What vehicles can "know":

- Stimulus concentrations (observations)
- Motor output (actions)

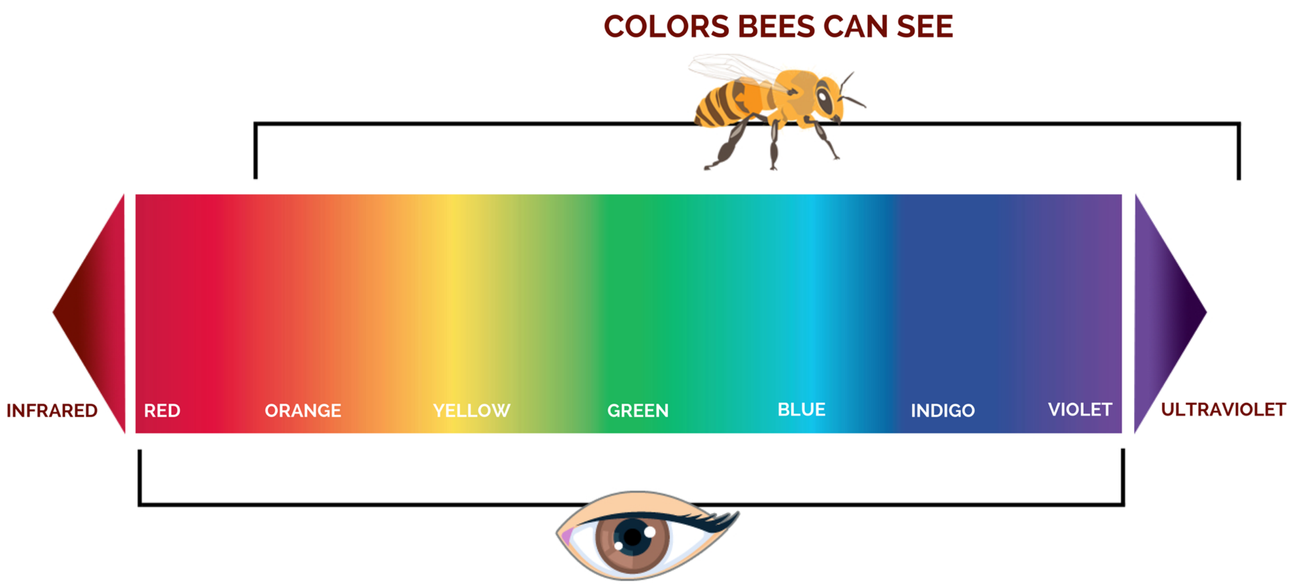
Structure of the problem/environment:

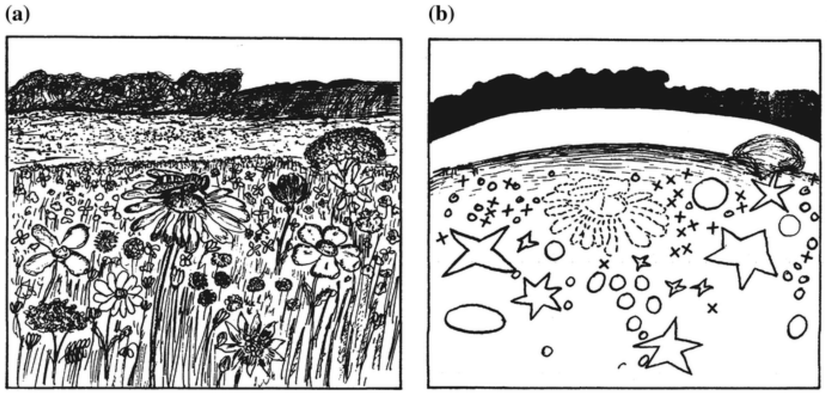
- Reward/Observations: chemical/ light/... concentration
- Transitions: navigation in space



Example: how vision shapes our reality

Things we can see and things we can't





Human eye view

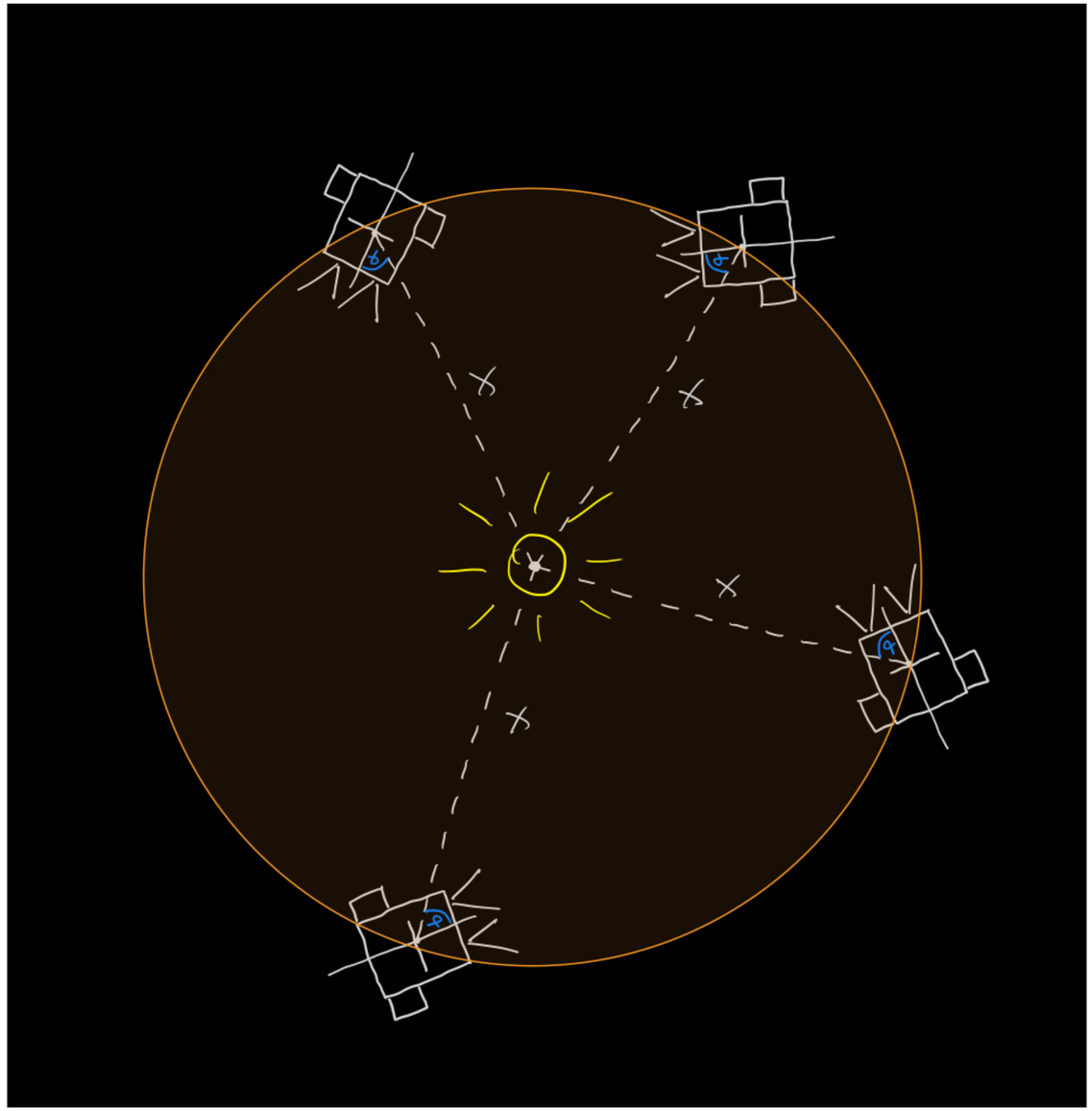
Ant's eye view with 650 ommatidia

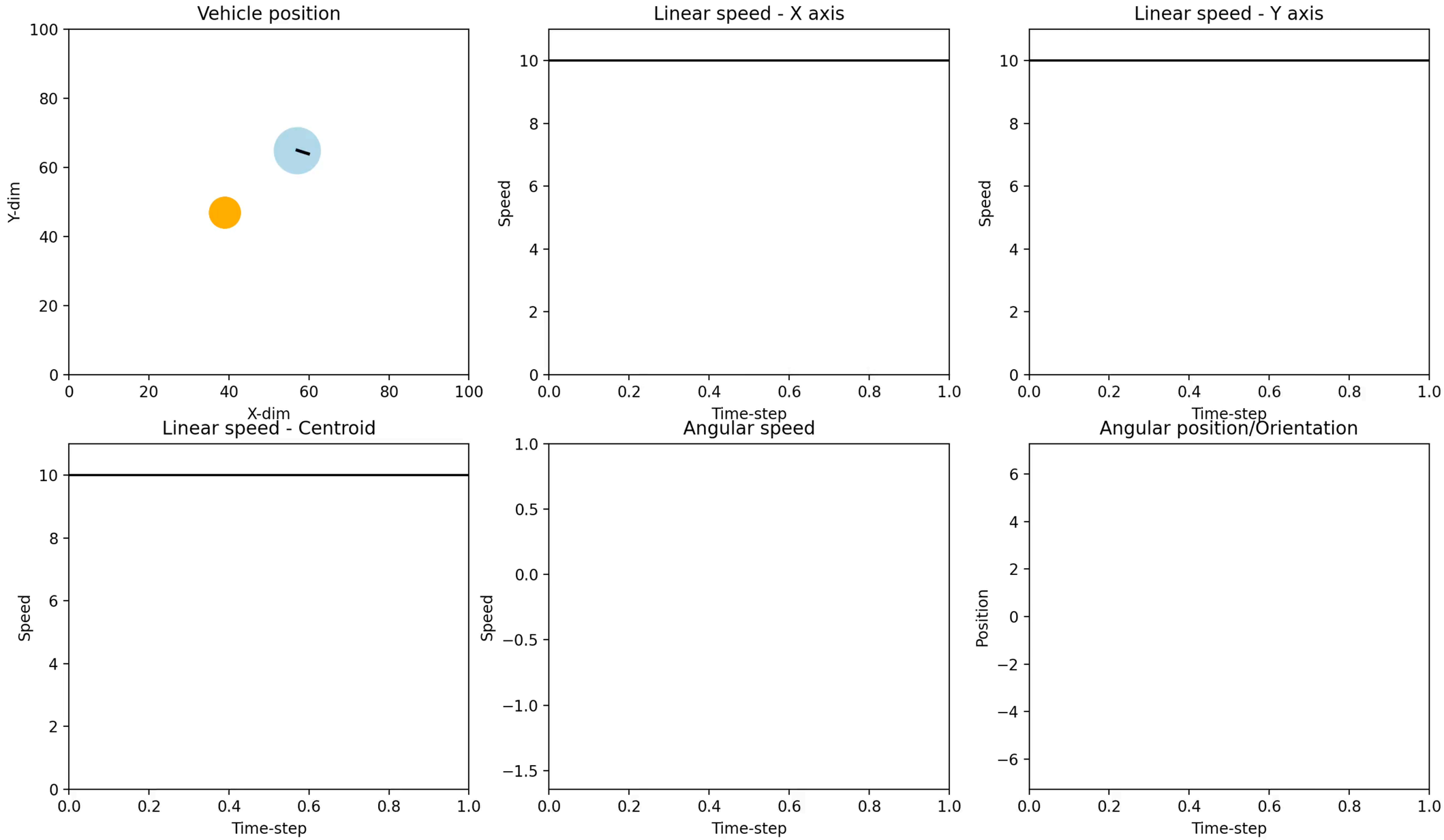
Ant's eye view with 150 ommatidia

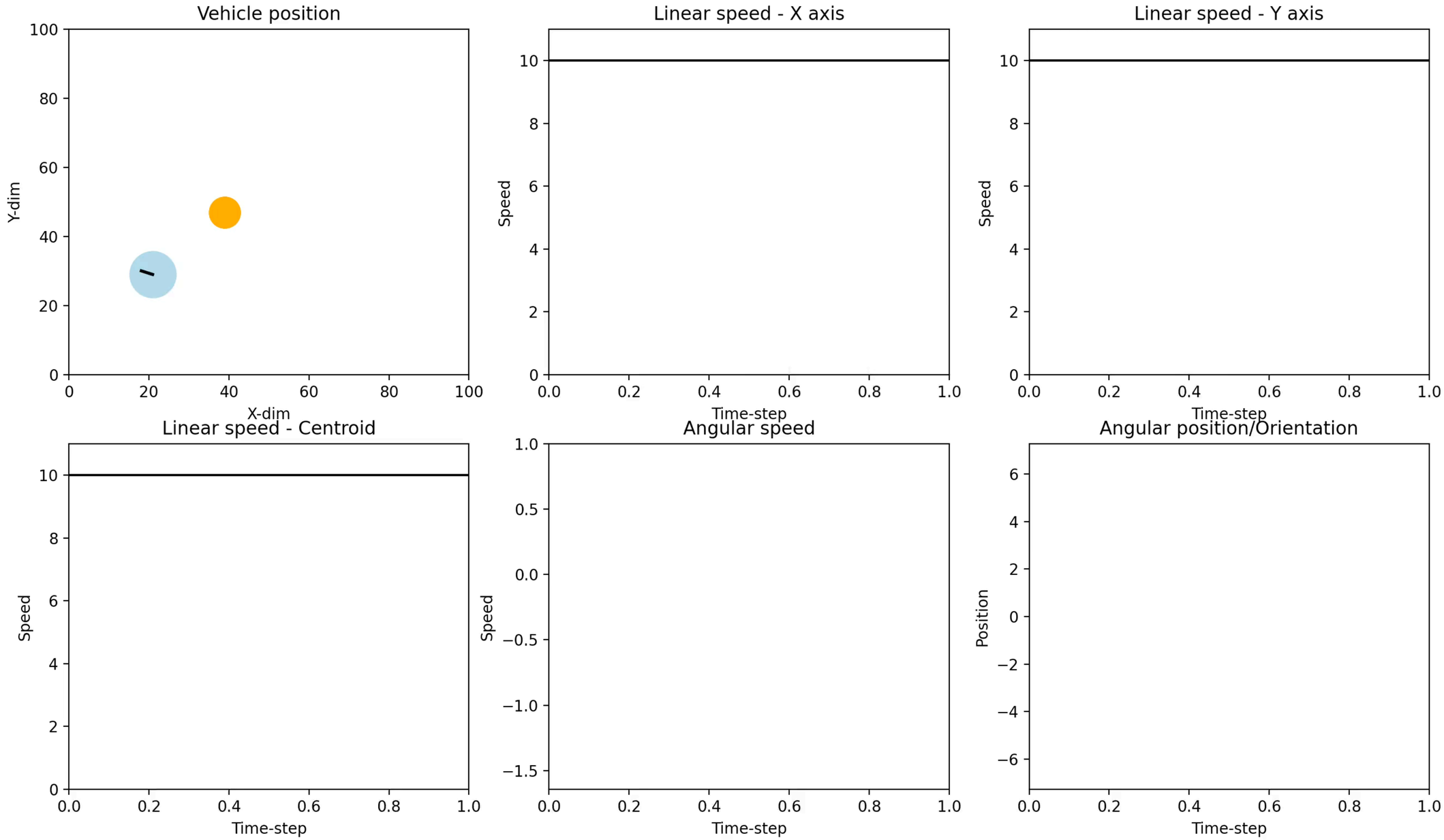
Standard vehicles

Properties

- Given same distance to source
- ...and same angle between front-facing direction and direction to source
- ...there is an invariance to rotations around sources



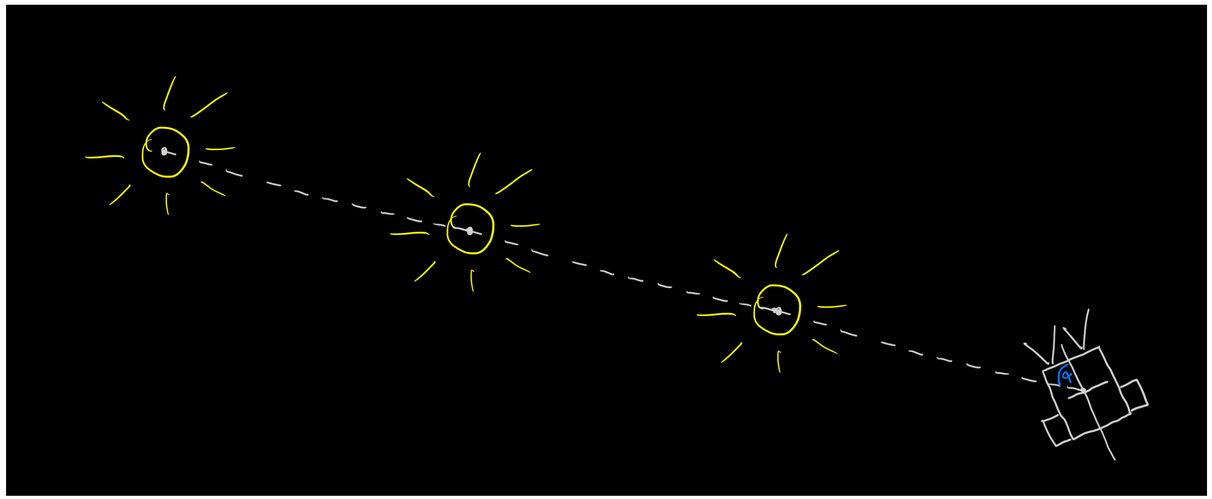

Frame 0
Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step

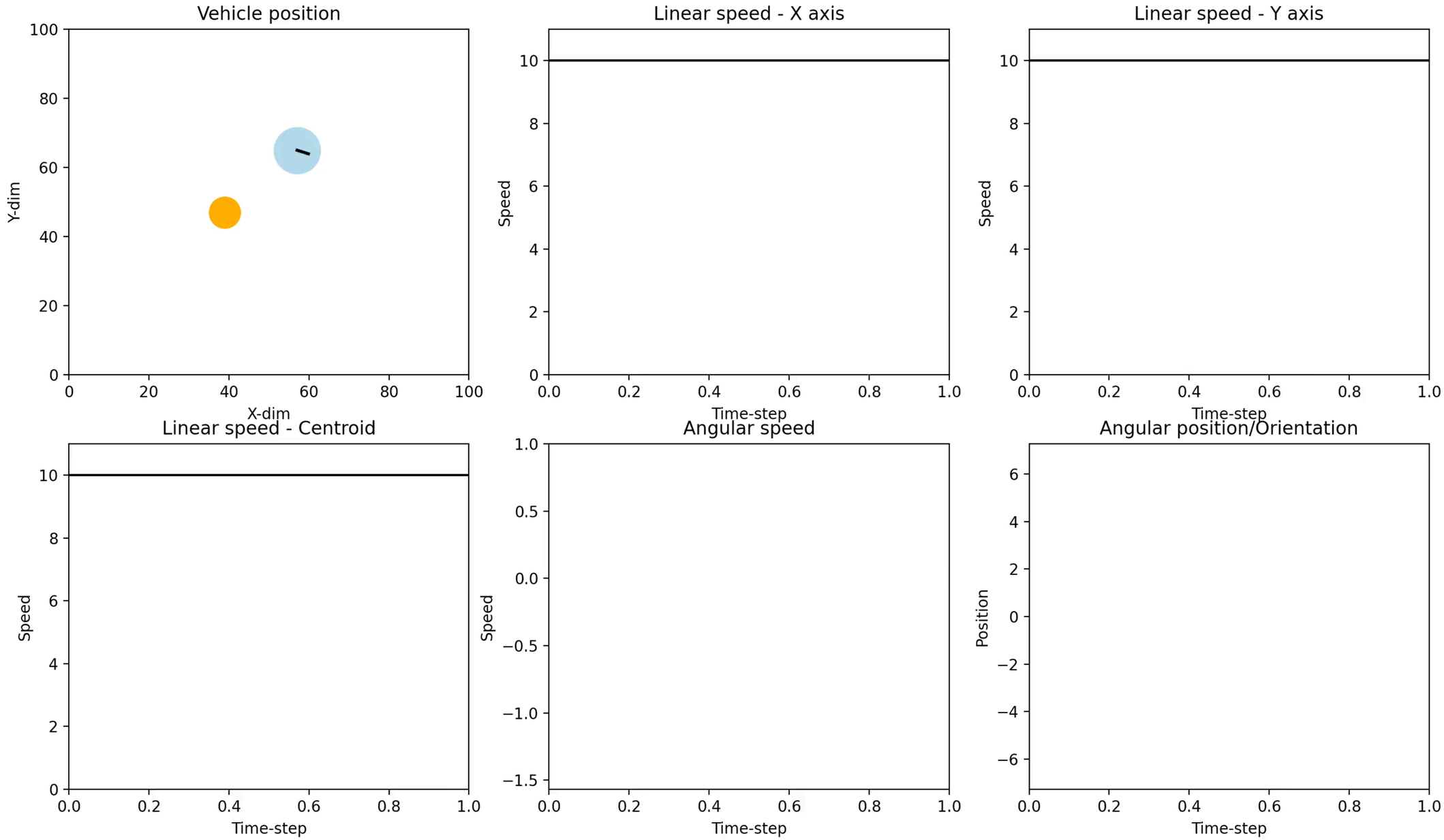

Frame 0
Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step

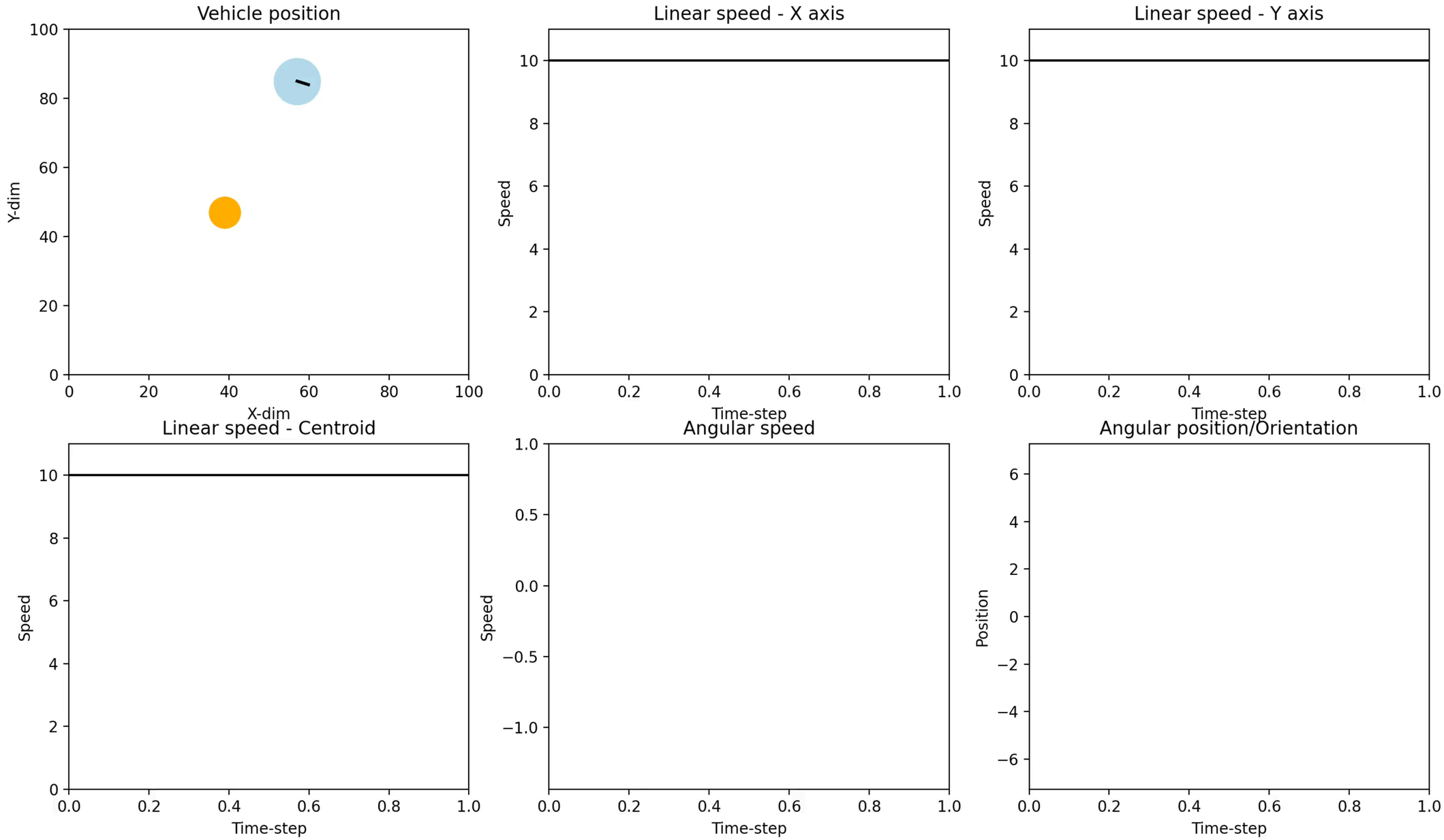
Simplified vehicle 1

Properties

- ~~Only emits one bit of motor information per motor (full speed/no move)~~
- Normalise sensory values in interval (e.g., [0, 1]), meaning sensors are only sensitive to concentration change
- Roughly the same as “spatial sensing” in bacteria
- Cannot distinguish distance to source



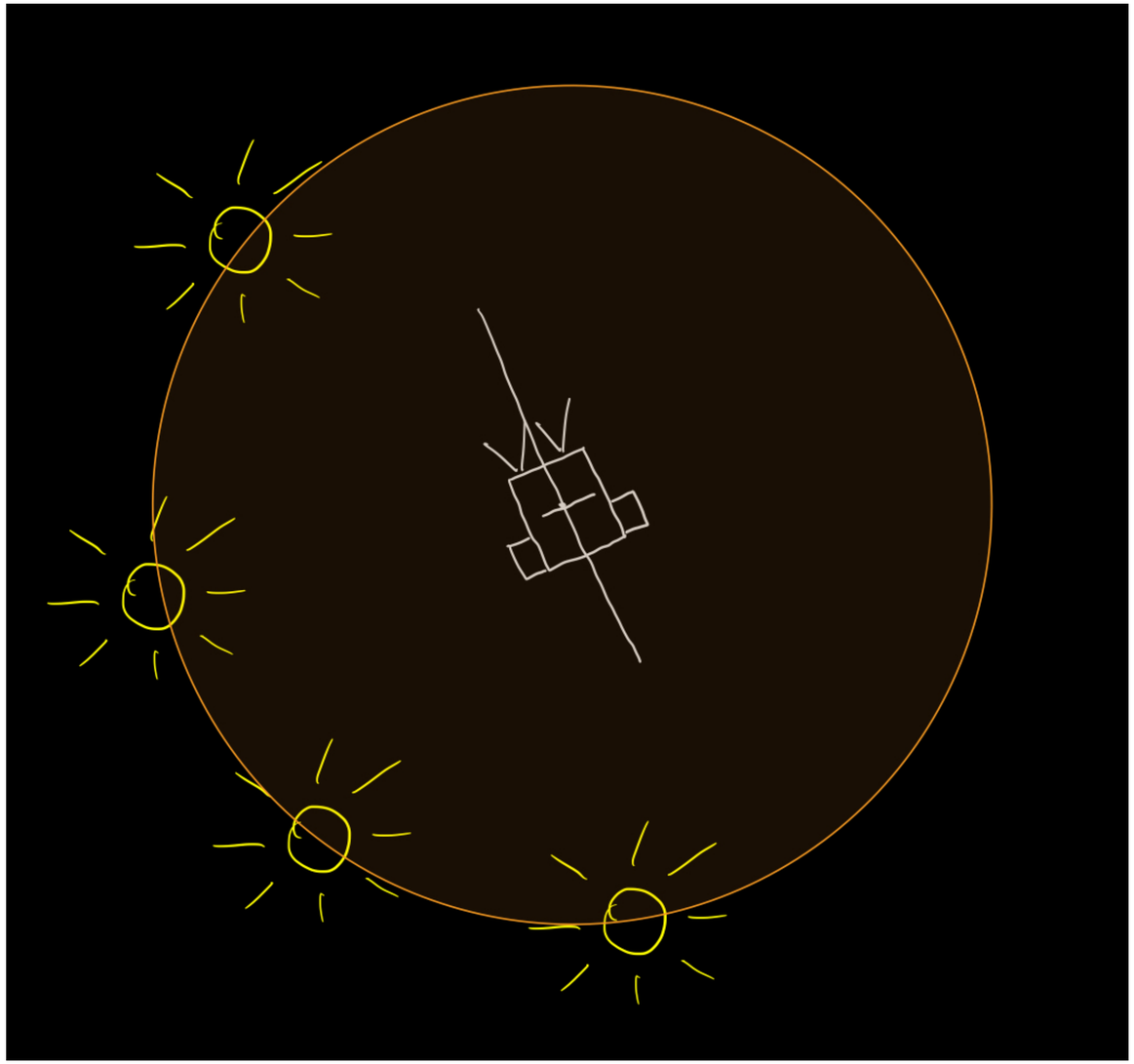

Frame 0
Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step


Frame 0
Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step

Simplified vehicles 2

Properties

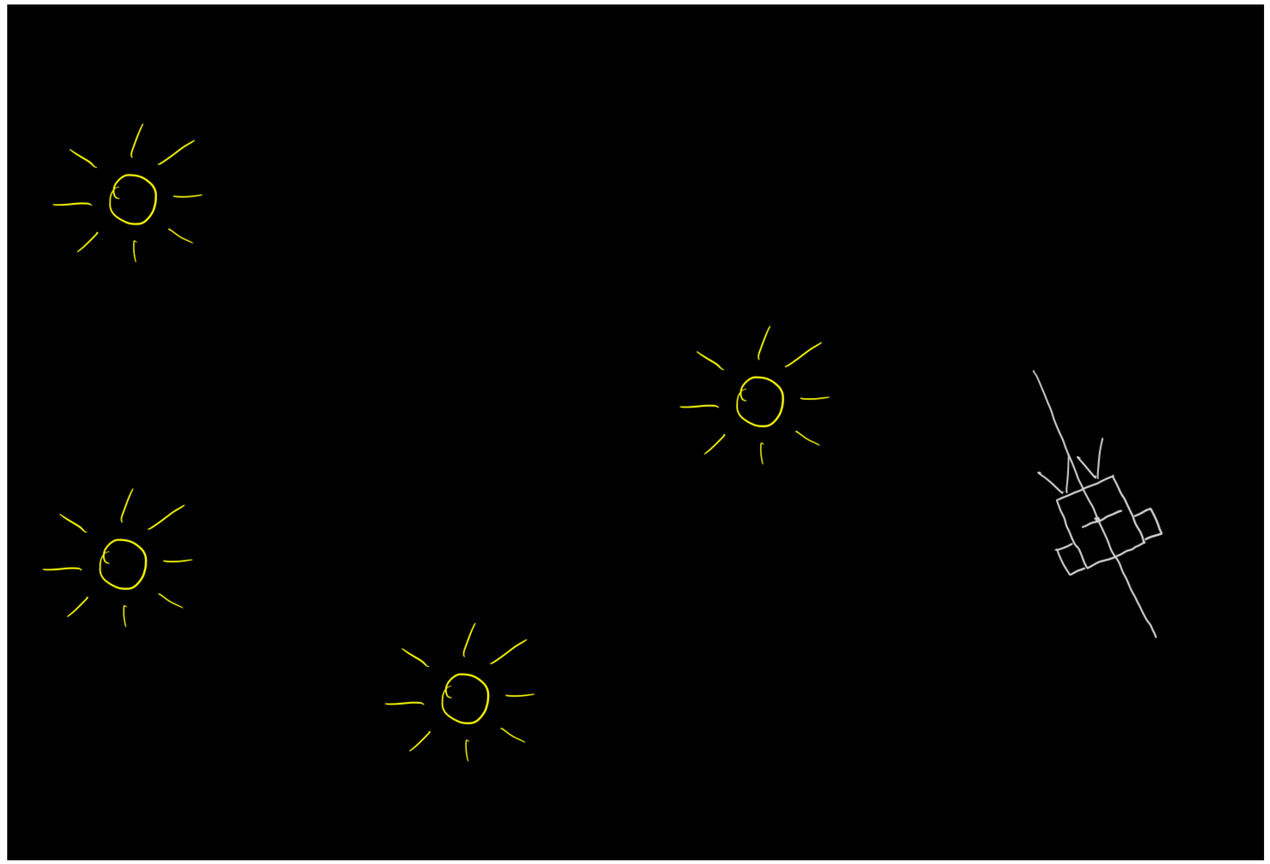
- Only gets one bit of sensory information per sensor (light/no light, chemical/no chemical)
- Cannot distinguish angle to source
- Broke the code, don’t remember how I got it to (maybe) works



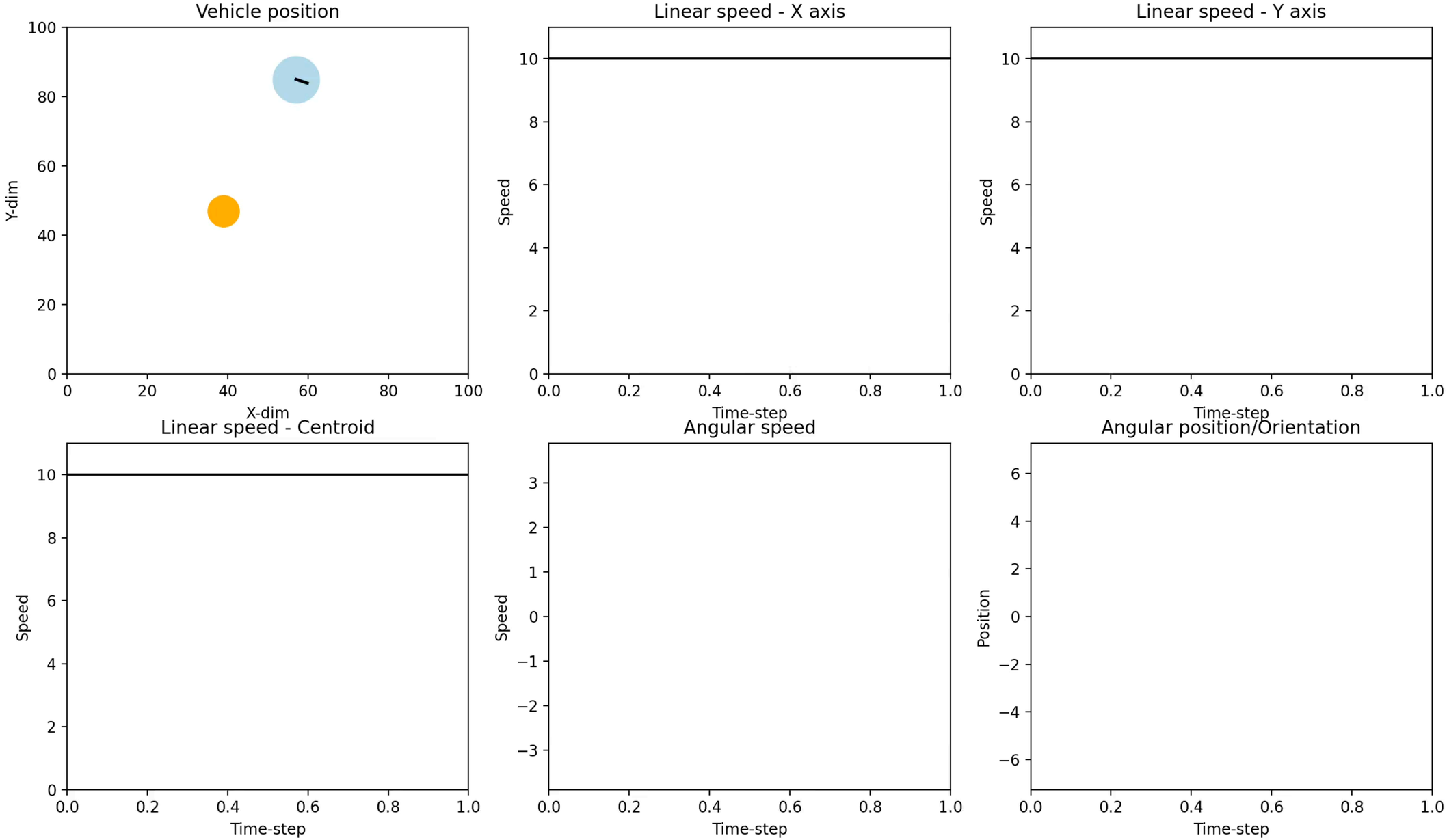
Simplified vehicles 3

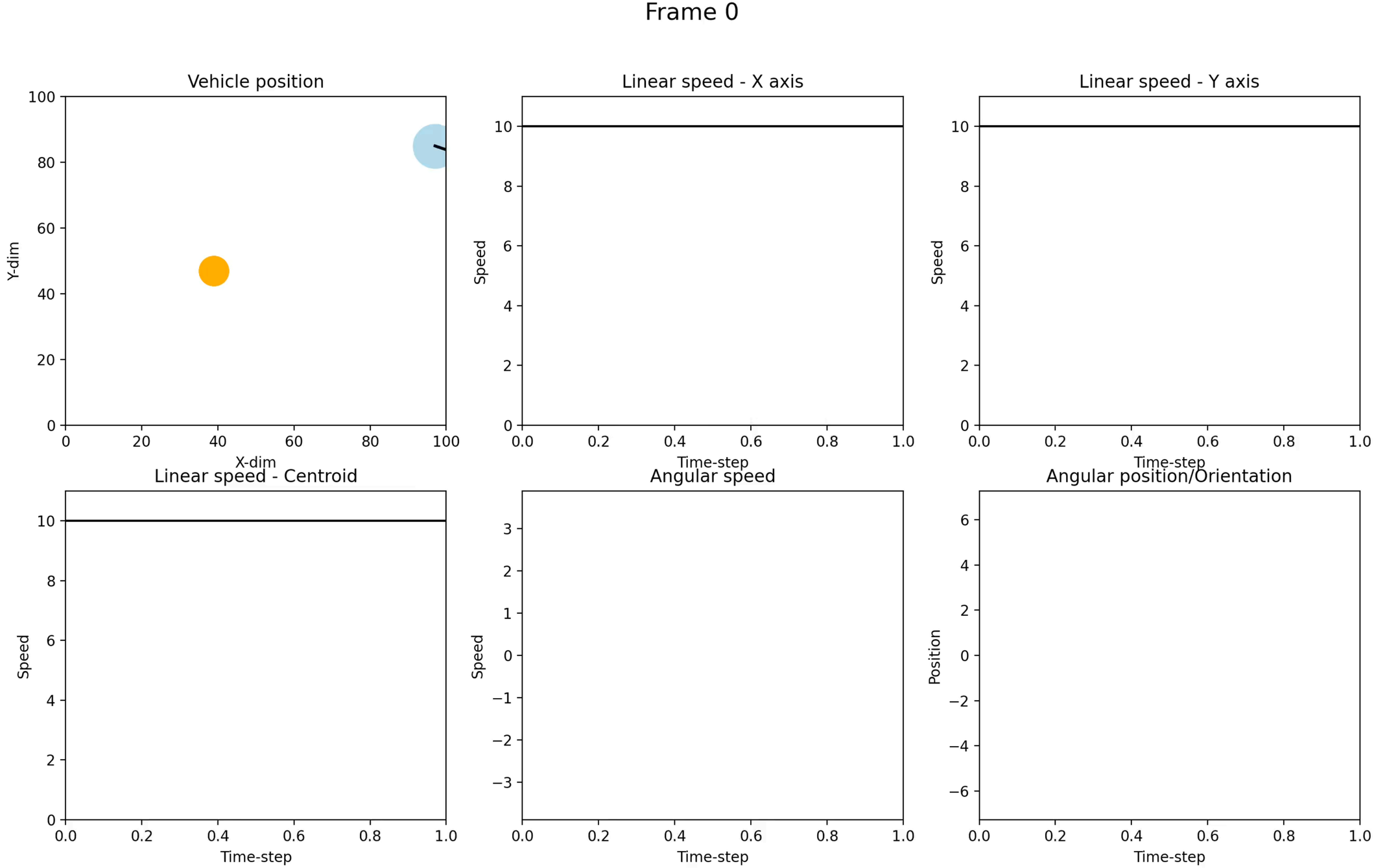
Properties

- Only gets one bit of sensory information per sensor (light/no light, chemical/no chemical)
- Only emits one bit of motor information per motor (full speed/no move)
- Cannot distinguish angle or distance to source



Frame 0

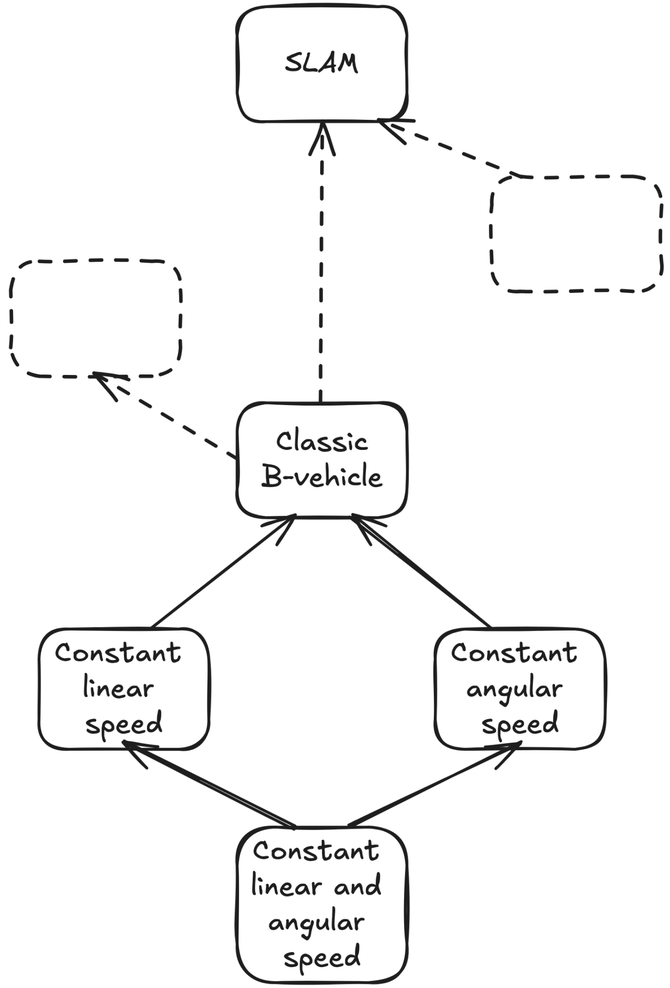

Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step


Frame 0
Vehicle position
Y-dim
X-dim
Linear speed - X axis
Speed
Time-step
Linear speed - Y axis
Speed
Time-step
Linear speed - Centroid
Speed
Time-step
Angular speed
Speed
Time-step
Angular position/Orientation
Position
Time-step

Doing this formally

No maths here, but ask me later if interested!

- Belief MDPs, epsilon machines/transducers, filtering, etc.
- Coarse-grainings through bisimulations of various kinds
- Obtain model order?
- …
- Can also be implemented with learning (through approximations of bisimulations)



Conclusion

- Looking for simple explanations/models + examples of when they work
- Different solutions to same problem are probably related (how to formalise this order?)
- Examples of simplified Braitenberg vehicles to study their beliefs