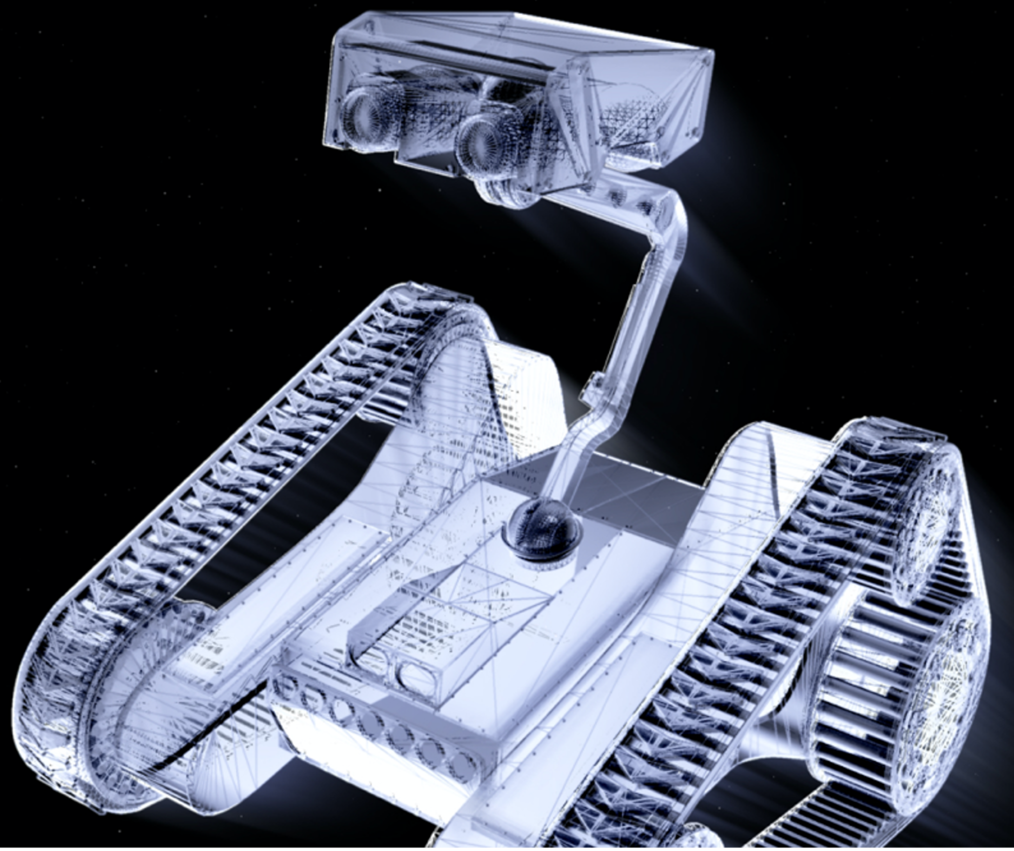


Fuzzing, Feature Deprecation, Causal Models and Metaphors:

How to make AI more Robust from experience on the Vision/Robotics Front.

Gary Bradski

CSO Farm-ng, OpenCV

garybradski@gmail.com

OpenCV.org

Opensource Computer Vision Library Foundation

Mission: Accelerate the Beneficial Uses of Computer Vision in Society

Background Pertinent to this Talk

- Founded the most popular **Open**source **C**omputer **V**ision library **OpenCV**
- Founded, run, involved with many computer vision and robotics startups
  - **SOLD:** Video Surf *(Video Search – Sold to Microsoft),* Industrial *Perception (Robotics in Logistics – Sold to Google)*, Arraiy *(Camera Tracking in Broadcast and Movies – Sold to MatterPort);*
  - **ONGOING***:* Gauss Surgical *(Surgery Monitoring),* OpenCV.ai *(Computer Vision Contracting)*, Farm-ng *(Robotic Tractor, Visual Navigation)*
- Led Vision Team on Stanley, Winner of the DARPA Grand Challenge
  - Now in the Smithsonian – it kickstarted the Autonomous Driving Industry

What is OpenCV?

Open, Free for Commercial or Research Use, Computer Vision and AI Library

Soureforge Downloads:
22,757,138

Github:
~14,980 Unique Clones/Week
~63,560 Unique Visitors/Week

OPENCV LIBRARY

54K — 54K STARS ON GITHUB

Extremely popular Github repo

github.com/opencv/opencv

61.4K stars: opencv + opencv_contrib

1-2M — INSTALLS PER WEEK

Rivals Tensorflow in Python installs using pip

pypistats.org/packages/opencv-python

89% — EMBEDDED VISION ENGINEERS

The most popular Computer Vision library among embedded vision engineers

embedded-vision.com/academy/ToolsAndProcessorsForCV_Jan2019_R1.pdf

2500 — ALGORITHMS

The largest collection of Computer Vision algorithms in a single library

github.com/opencv/opencv

OpenCV.org:

- Maintains the code
- Produces courseware
- Partners to produce hardware
- Sponsors contests
- Runs workshops
- Sponsors Interns

OpenCV at a glance

HighGUI
I/O, Interface

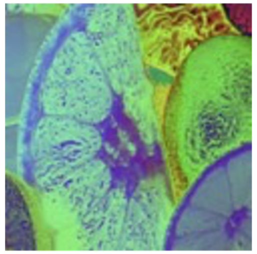
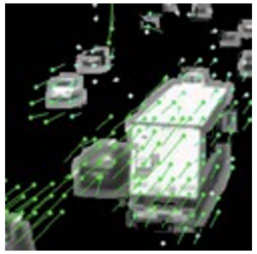


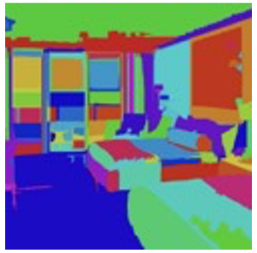
Image
Processing

Transforms

Fitting

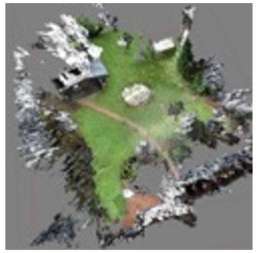


Optical Flow
Tracking



Segmentation

Calibration
Geometry
Color



Features
VSLAM

RGBD
Depth, Pose,
Normals, Planes

Deep Learning,
Machine Learning

Computational
Photography

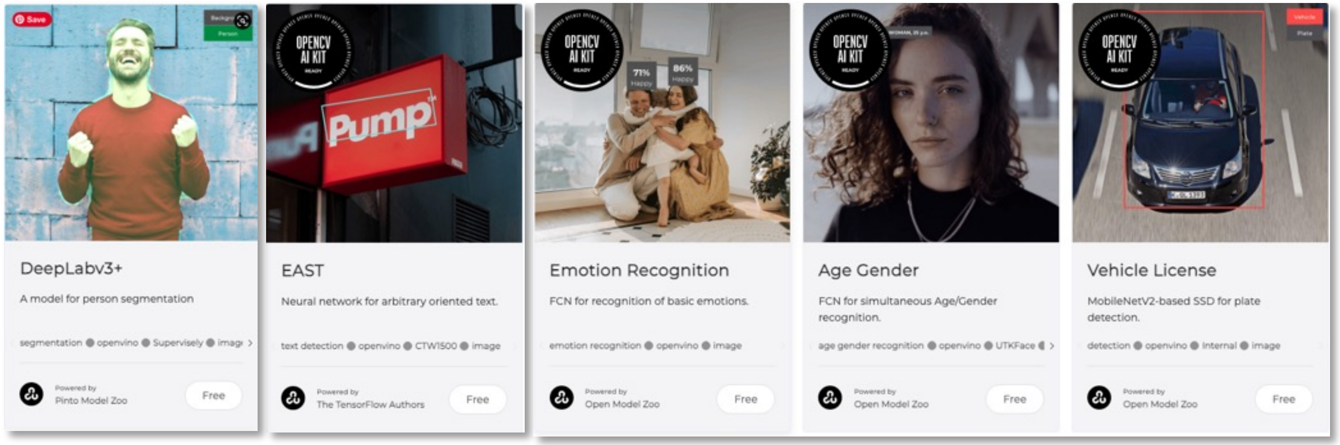
Core
Data structures, Matrix math, Exeptions etc

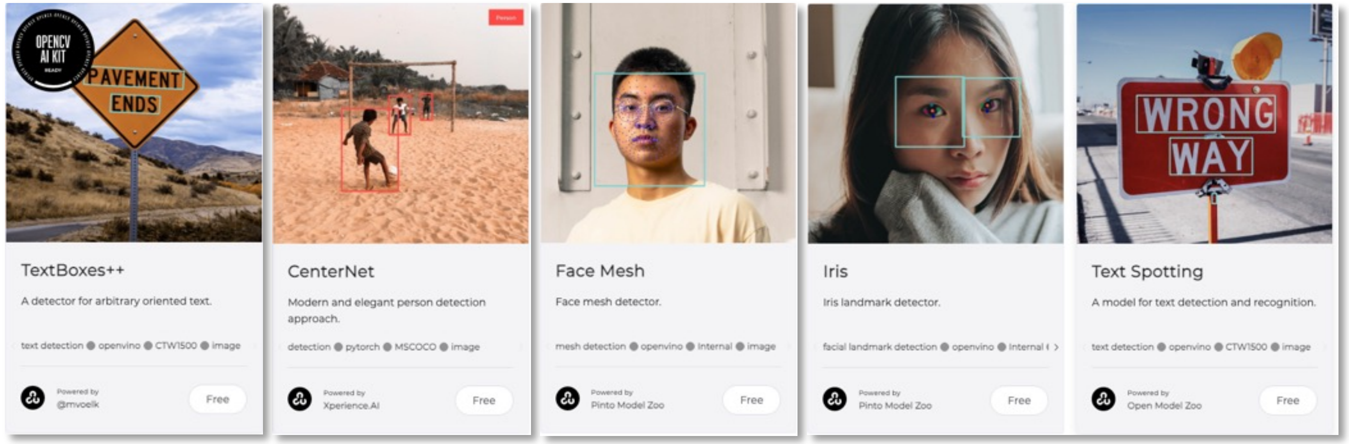
Recent Additions

Recent: Modelplace.AI

https://modelplace.ai/

- A model zoo for OpenCV, Community and Commercial models:
  - Optimized, trainable, tunable, memory efficient models
  - Drop image on browser to test:





Recent: HARDWARE -- OPENCV AI KIT (OAK)

OpenCV AI Kit with Depth (OAK-D) is OpenCV's spatial AI camera based on Intel® Myriad™ X.

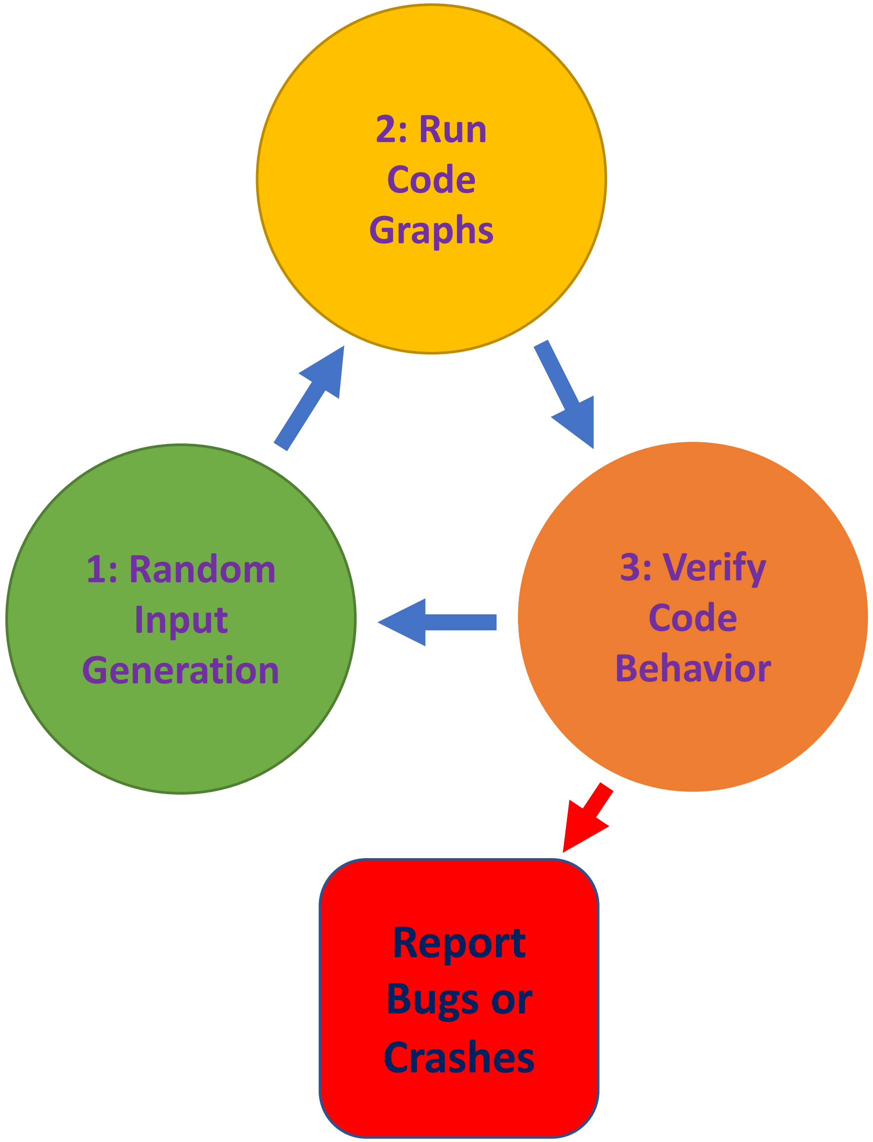
4 Tera Ops/second

It can **run neural networks** for tasks like image classification, object detection, segmentation, pose estimation etc. in real time.

It comes with a stereo pair for **depth perception** in real time.

Problem #1:
Secure Code

Solution 1: Fuzzing



Google runs OpenCV through 1000s of machines every day

Deep Nets

- Need traditional Fuzzing
- And “Structural fuzzing”:
  - In vision, this might be random shapes, noise, lighting, glare, 3D perspective changes, object orientations etc
  - Check that it doesn’t destabilize results

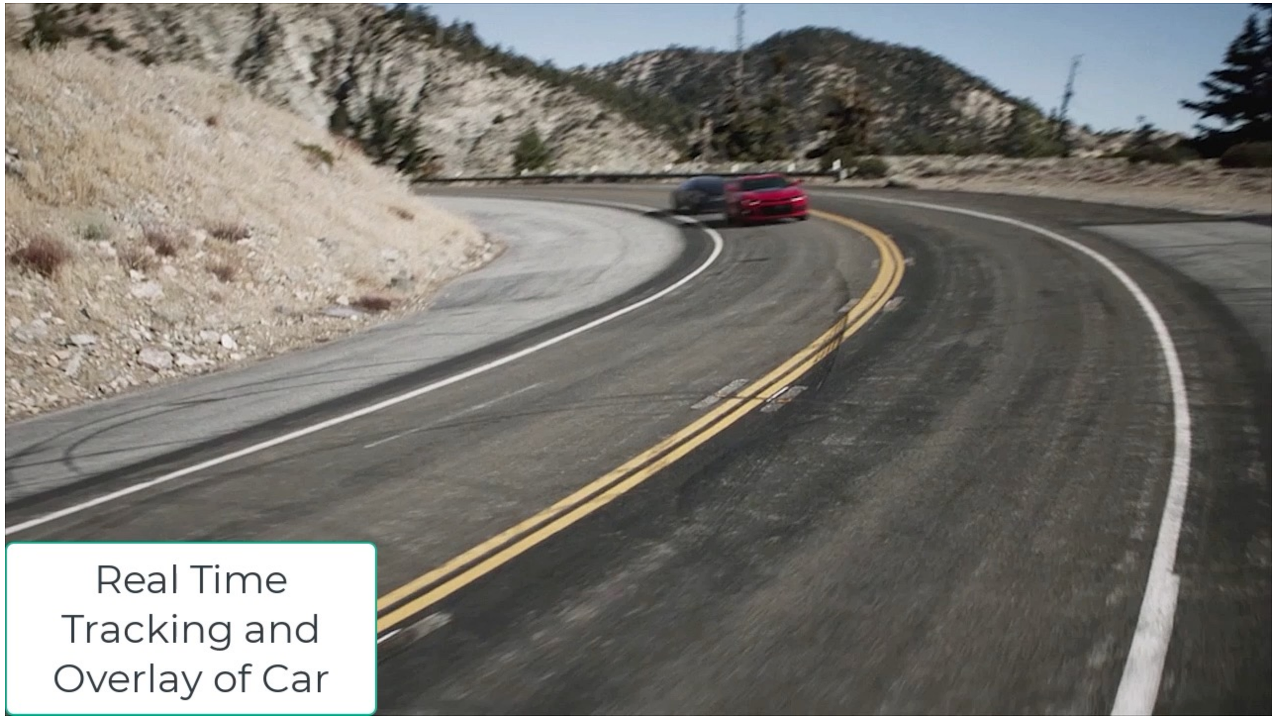
Company: Arraiy.com

Problem #2: Robustness in the field

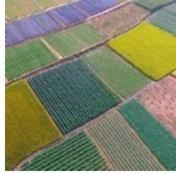
- We must <u>not</u> fail at tracking camera location on set
- Regardless of objects or people in the way
- Lighting changes
- Changes to the set

Solution #2: Feature Deprecation

- Do not let the network get over reliant on any feature (here, geometric structures in the scene)
  - Randomly occlude parts of the scene during learning
  - If a feature starts growing strong, hide it (blur it out)
- We thus cause the network to rely on robust overall context in a scene

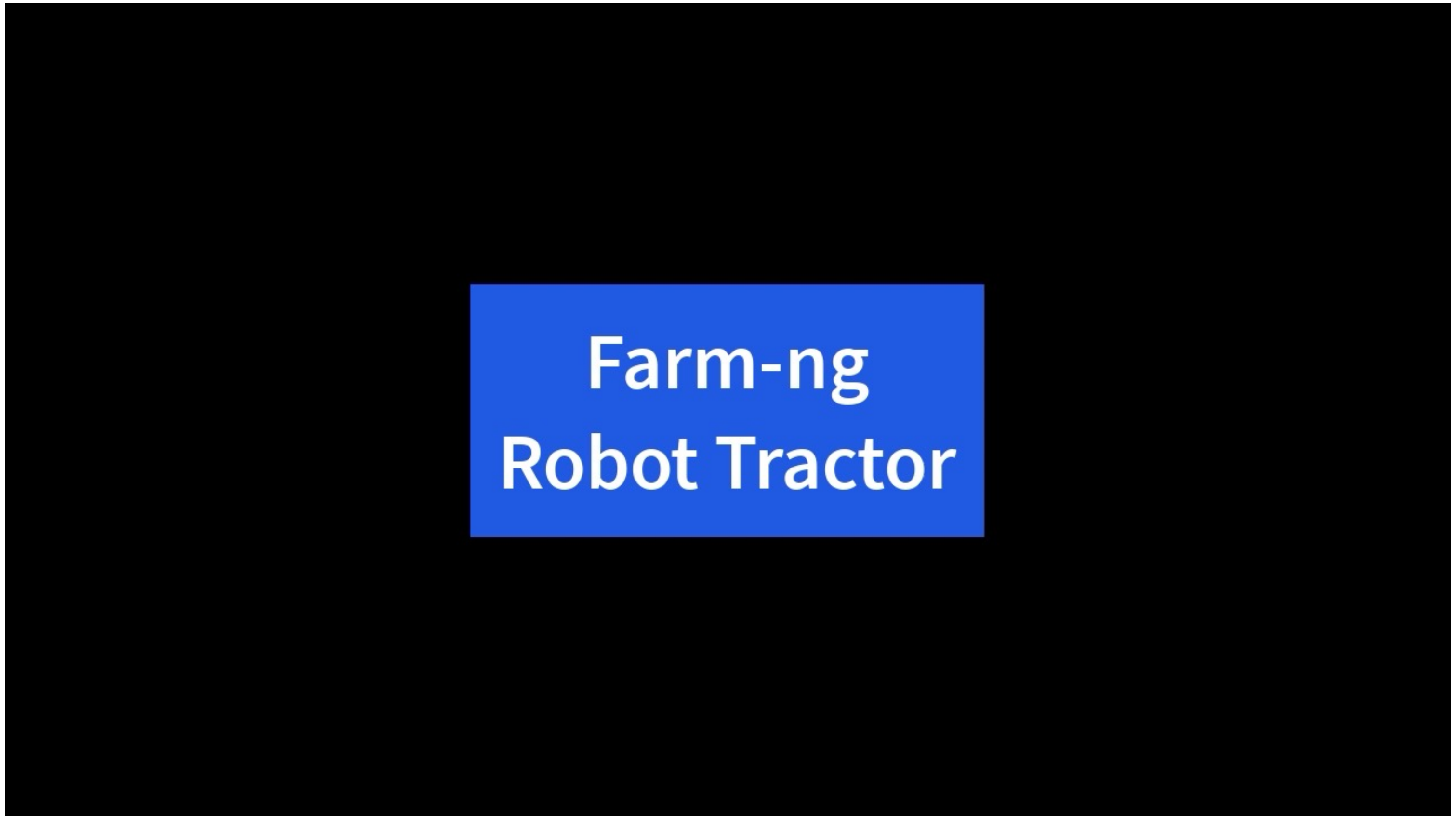


Farm-ng.com



These film techniques are now generalized (licensed) to:

- Robot navigation in agricultural fields and in logistics
- Solving GPS denial in Defense
  - Drones, Planes, Rovers, Boats
- NASA moon lander



Gap #3:

How to use Programmatic Operators in Deep Nets

Solution #3: (for Vision)

Use PyTorch to create differentiable vision operators that can work w/in and around networks

kornia → OpenCV-X

Open Source Differentiable Computer Vision Library for PyTorch

Edgar Riba; Dmytro Mishkin; Daniel Ponsa; Ethan Rublee; Gary Bradski, “Kornia: an Open Source Differentiable Computer Vision Library for PyTorch”, WACV 2020

E. Riba, D. Mishkin, J. Shi, D. Ponsa, F. Moreno-Noguer and G. Bradski, “A survey on Kornia: an Open Source Differentiable Computer Vision Library for PyTorch”, https://arxiv.org/abs/2009.10521, 2020

E. Riba, M. Fathollahi, W. Chaney, E. Rublee and G. Bradski, “Torchgeometry: when PyTorch meets geometry”, PyTorch Developer Conference, 2018, https://drive.google.com/file/d/1xiao1Xj9WzjJ08YY_nYwsthE-wxfyfhG/view?usp=sharing

OpenCV-X

core features

Computer Vision

Differentiable

Transparent API

Parallel Programming

Distributed

Production

Problem #4:

Common sense

and

Reasoning

Solution #4:

Causal Models

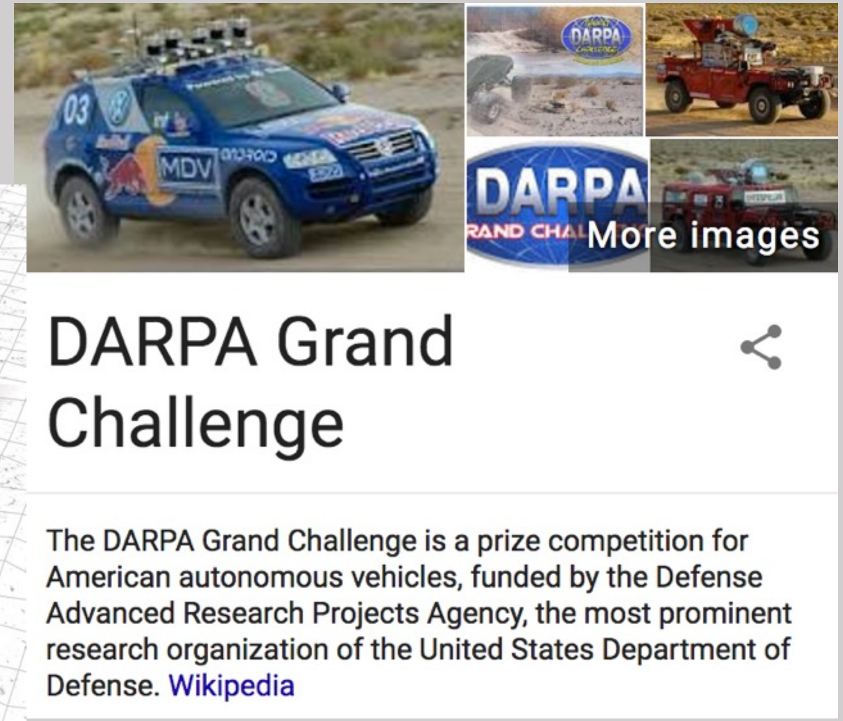
and

Metaphors

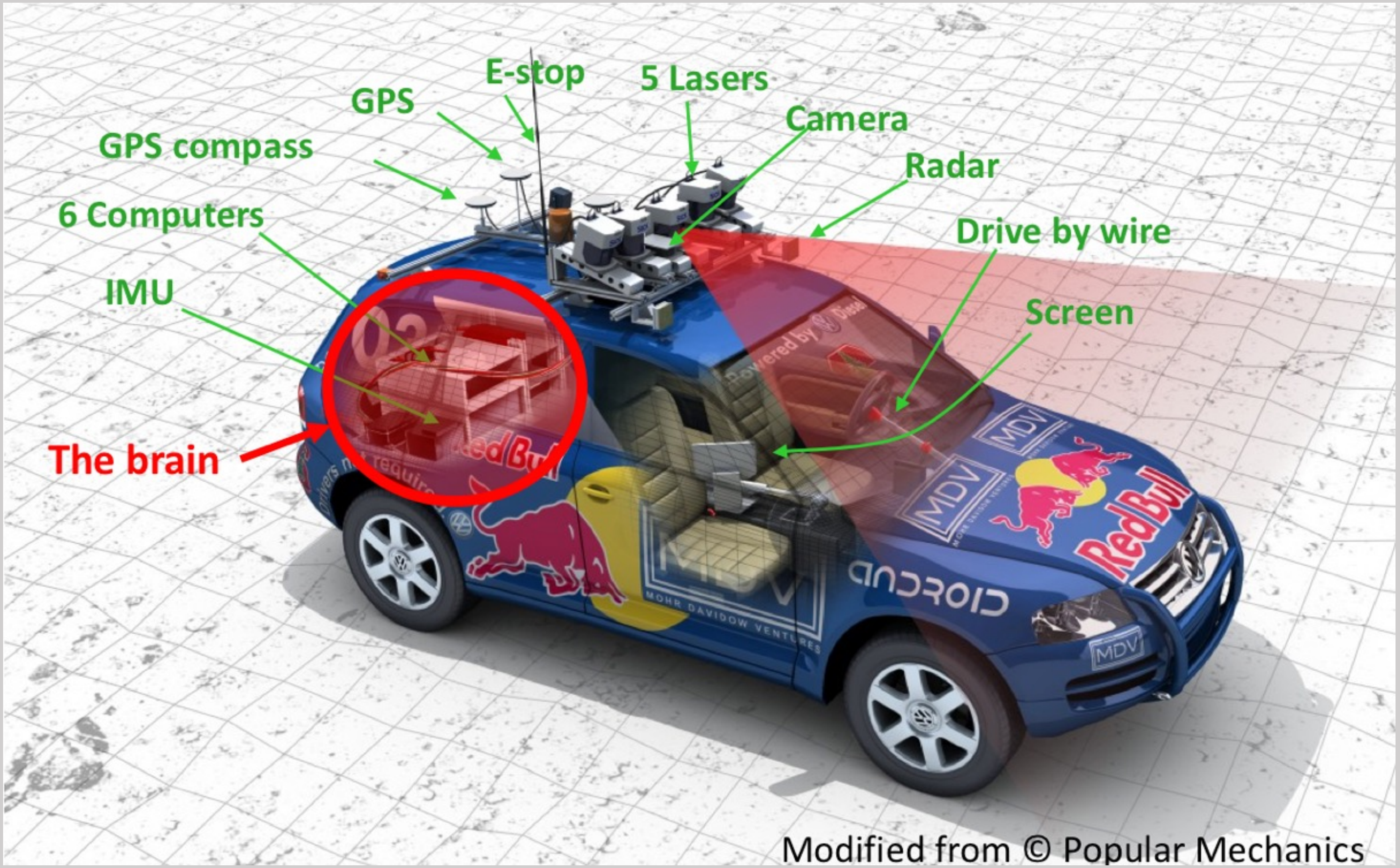
The DARPA Grand Challenge

- $2M robot car race across the desert
- Kicked off the autonomous driving industry
- Stanford's Stanley won
- Now in the Smithsonian

© Popular Mechanics



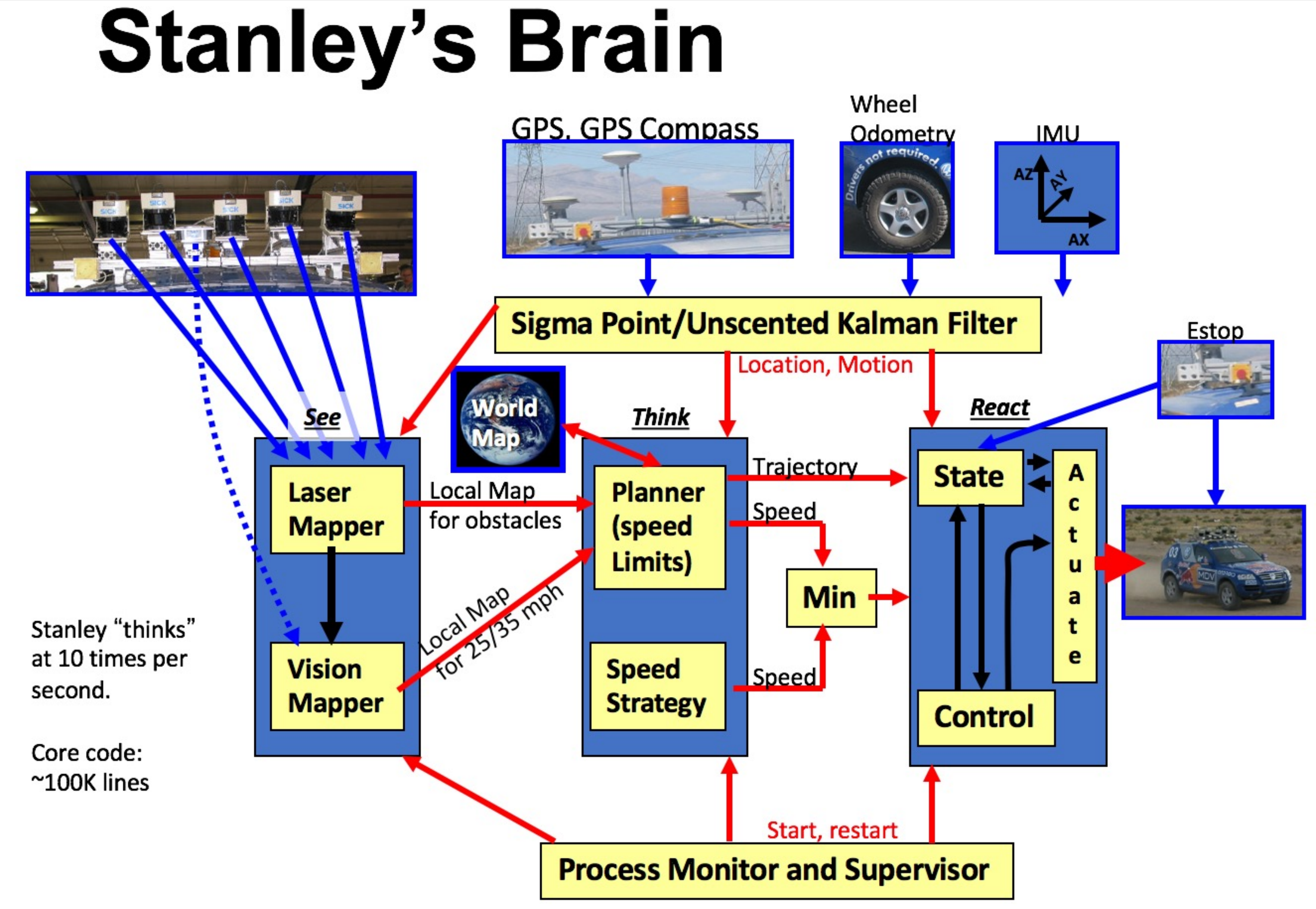
Stanley's Sensing



Picture 2

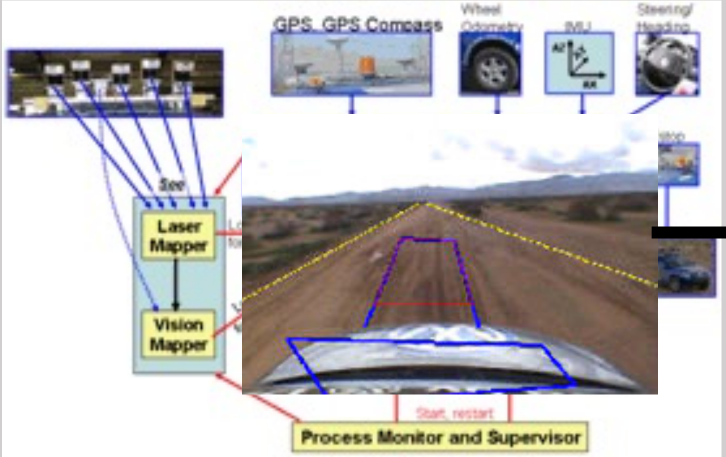
Stanley's Brain

- Many sensors being fused in a planning map
- Sensing is reformatted for the mind



Stanley's "Mind"

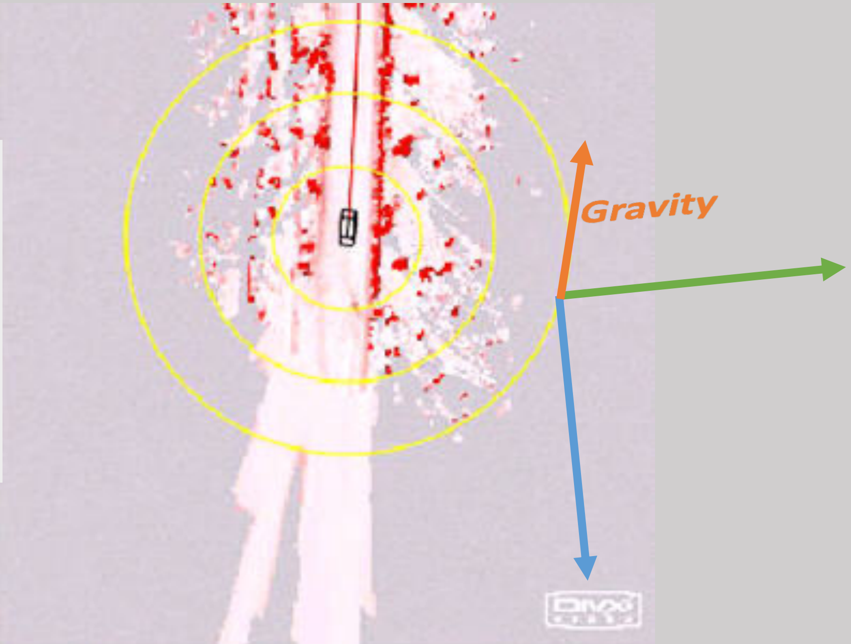
Sensors are fused into a map: Seeing => Perception



Fused into bird's eye world model

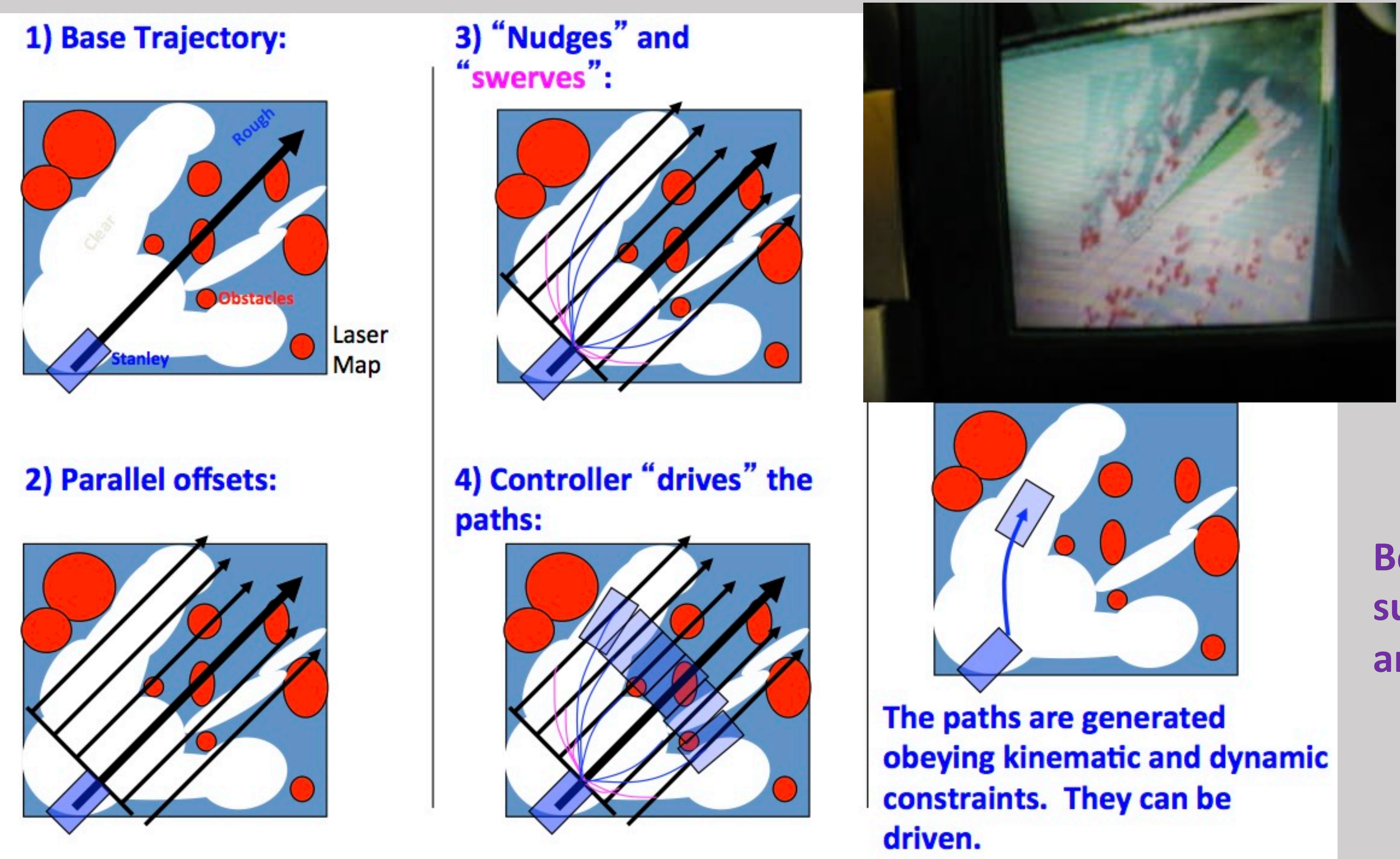
The map has only: red => bad; gray => don't know; **white => drivable.**
The map also has **tilt** and **direction of gravity**.

Bird's eye world model
+
Gravity direction
+
Physics



Stanley's Mind
Recognition + Simulation/Plan => Action

From the map and its tilt, the robot and it's driving plans/goals are (physics) simulated. THIS SIMULATOR IS A CAUSAL MODEL that is matched against the world



Best action chosen subject to external and learned rules

And a final decision is made.

Solution #4: Matching Learned Models

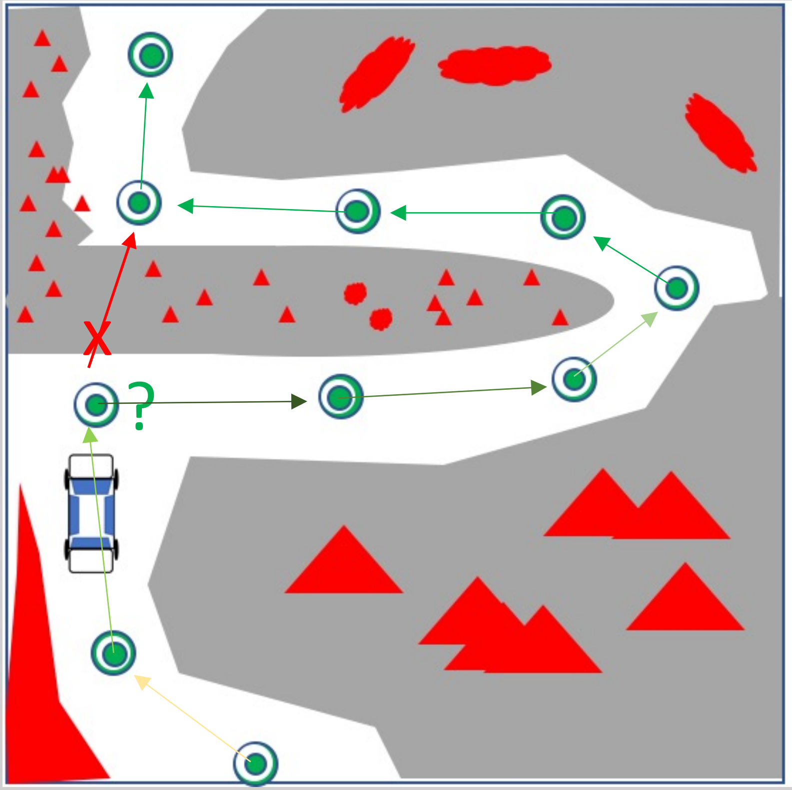
Key takeaways for why our perception is more robust than deep nets:

- **Deep Nets use**
  - matching (against loss functions) to learn
  - Inference to predict
- **Robots/Humans use**
  - Correlation to learn
  - Matching against a causal model to recognize

Note on: Stanley's Programming:

- Emotions program "WHAT to do"
- Mind finds "HOW to do it"

Stanley "wanted" to follow GPS way points quickly and in sequential order

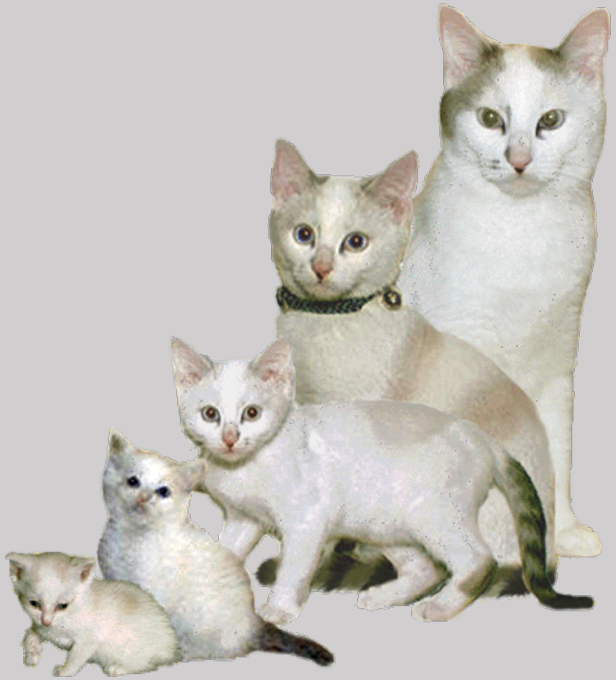


The "Mind" is "How"

The "Emotions" are "What"

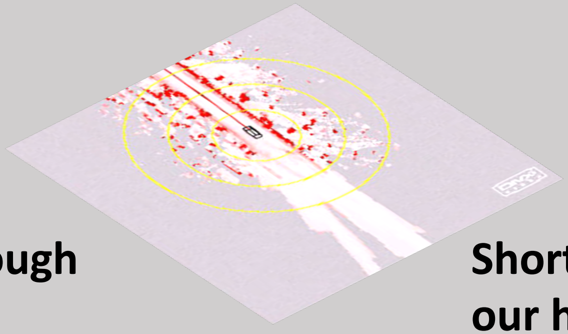
Generalizing Learning:
Reason with the metaphors you've got

- **Explain Shakespeare to a cat:**
- Even if your cat had listened to Shakespeare since it was a kitten.
  - Would you expect it to understand Shakespeare?
  - Perhaps a morality tale of a kitten that is tragically too bold?

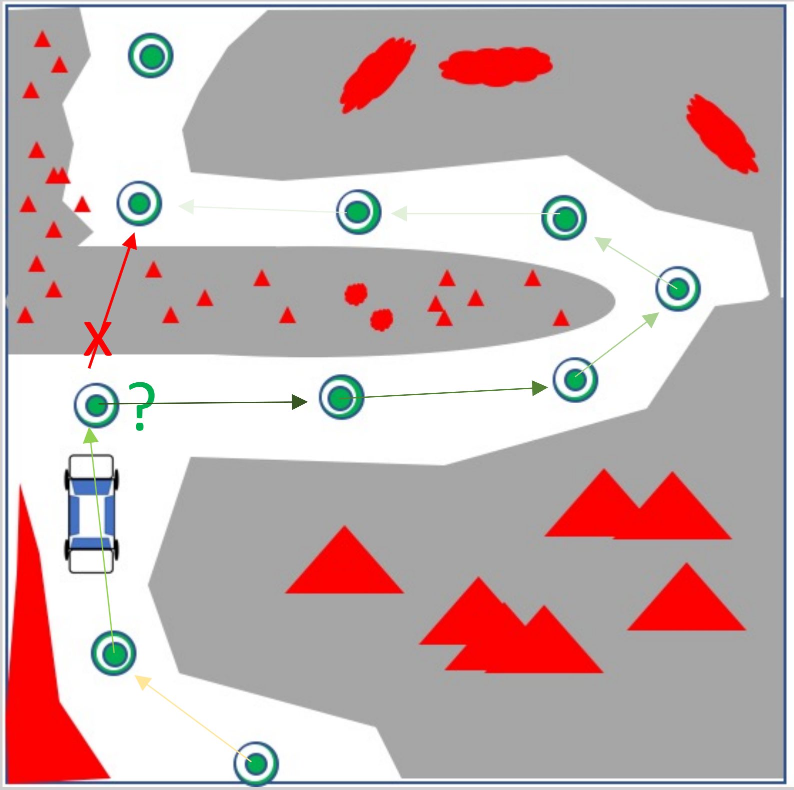


OpenCV

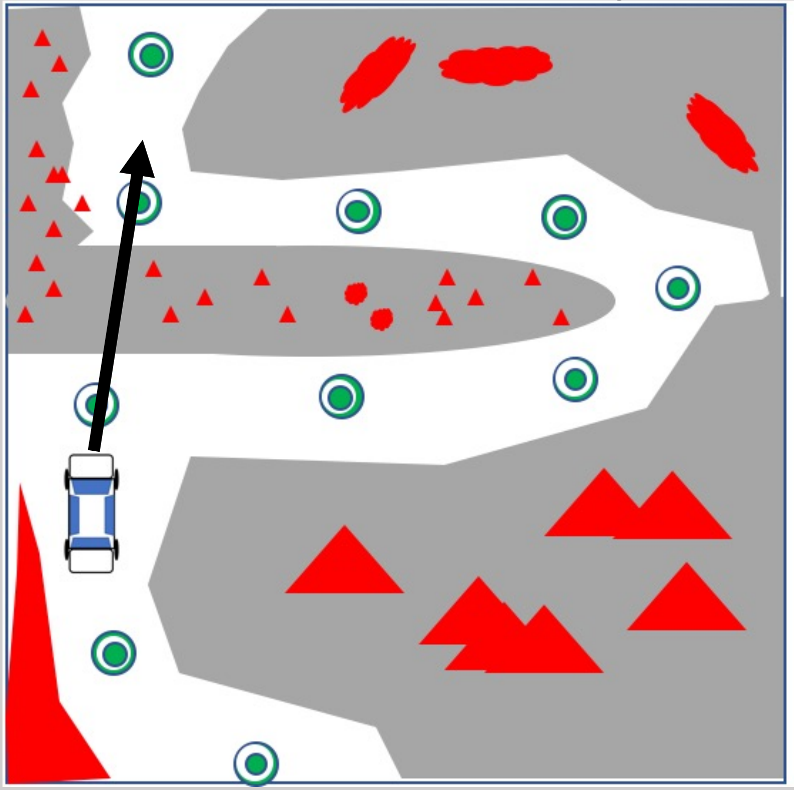
Metaphors: Communicating Shakespeare's **Hamlet** plot to the robot Stanley

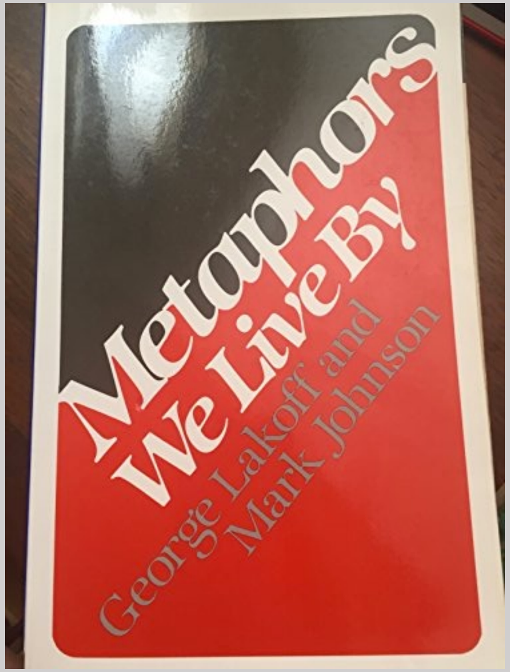


The "hero" needs get through waypoints fast:



Shortsightedness in thinking causes our hero to take a destructive path:





Takeaways (Solution #4, extending learning via reasoning):

(1) Reasoning can extend the mind far beyond just raw "training"
(2) Reason breaks down when the causal primitives aren't rich enough

Key Points about a robust mind:

STABILZED PERCEPTION/PLANNING

- Some causal model priors are tuned into a causal model of the world
- Matching this model against world data provides robust perception
- The causal model is used in simulations of mind's world representation in order to plan
  - The causal grounding of the mind's model in the external world stabilizes perception
  - The simulation is not a "*true representation*", but it is **causally** accurate for that mind

GENERALIZING LEARNING

- The models that the mind learns may be used to generalize learning via metaphor
- This allows "transfer learning reasoning"
- The brain is ultimately limited by the richness of its causal models


DARPA
GRAND CHALLENGE
FINISH
QUESTIONS?
Powered by Diesel
Red Bull
Photo: Gary Bradski