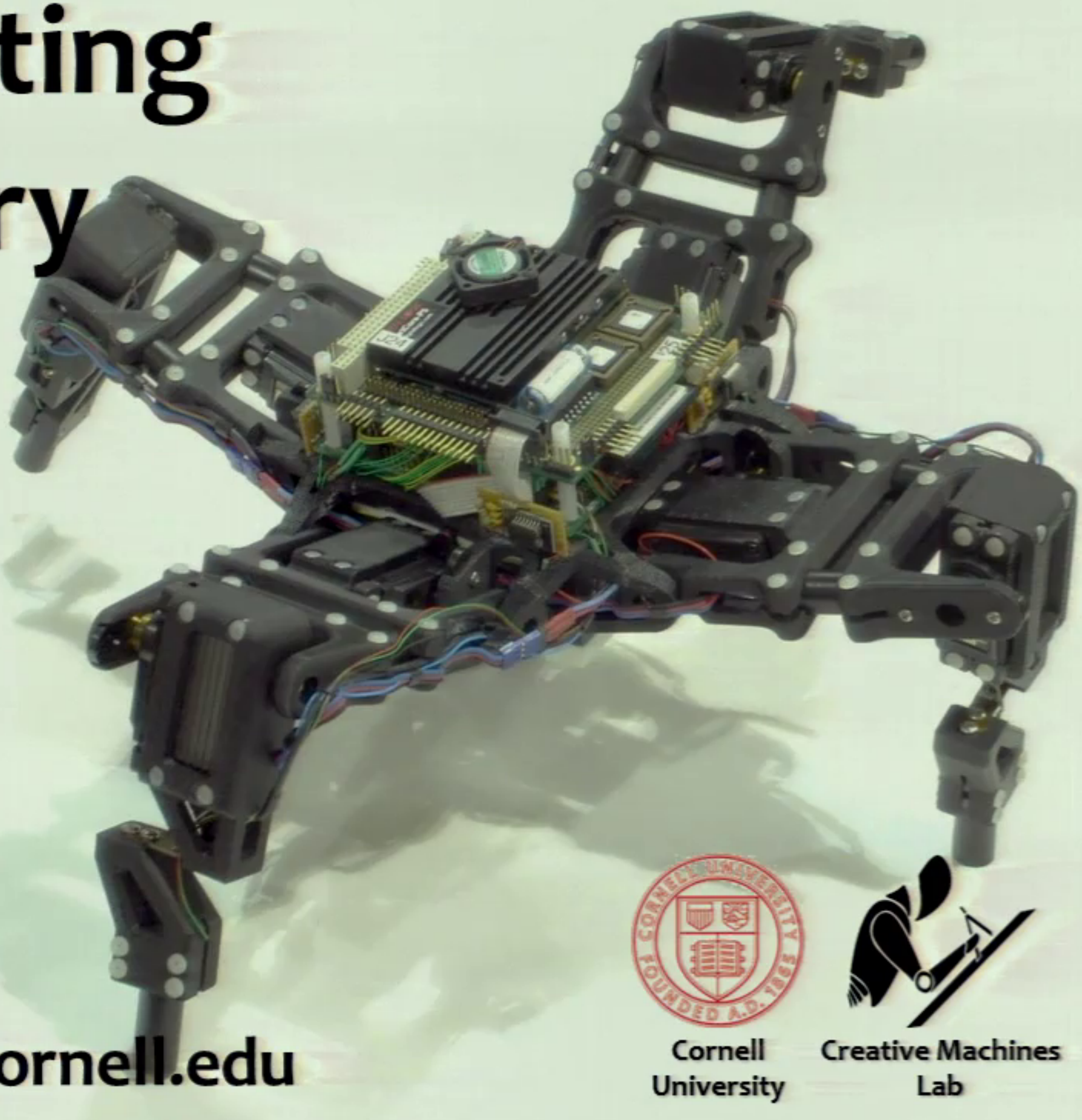
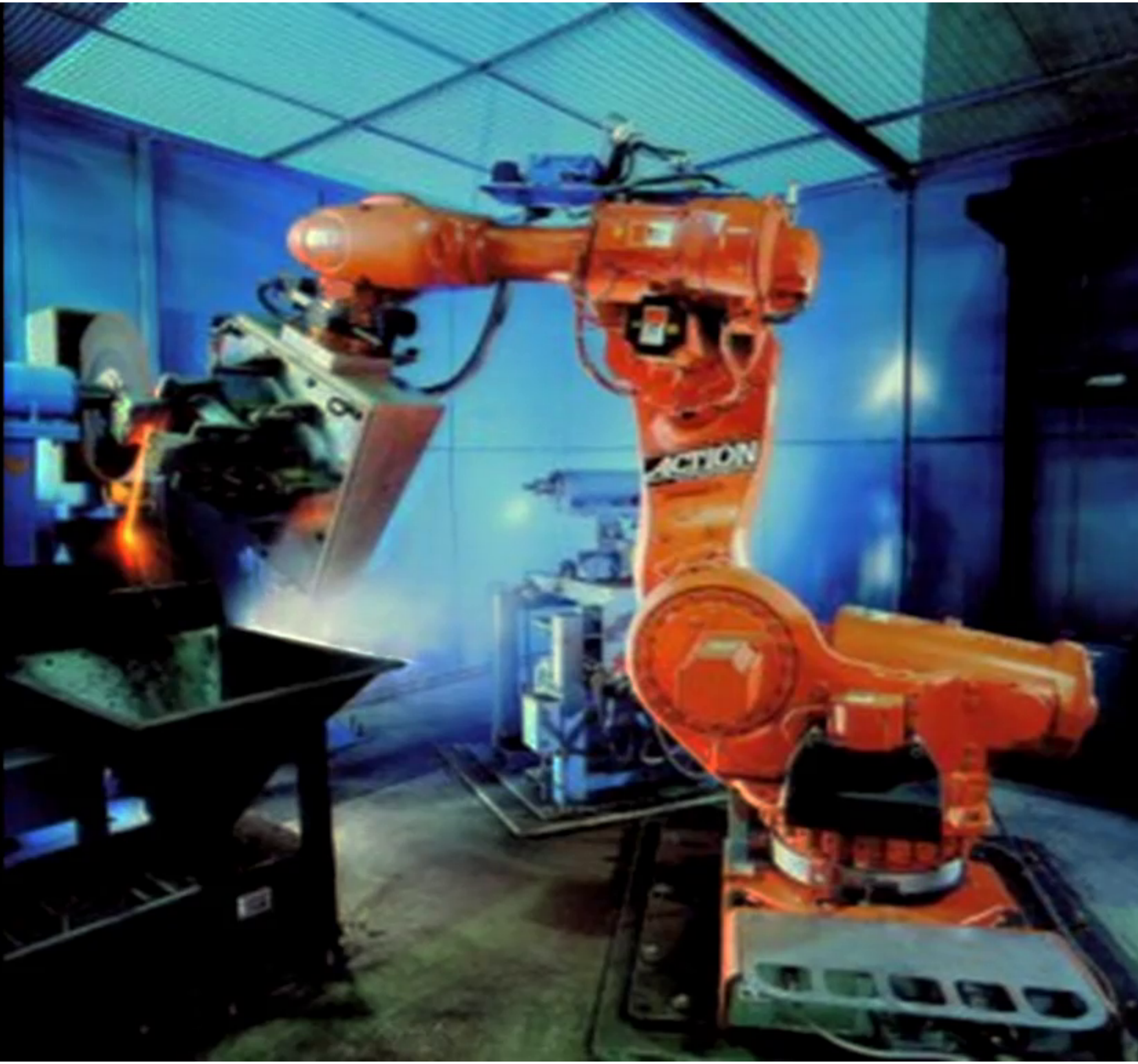
Automating Discovery



Hod.Lipson@cornell.edu

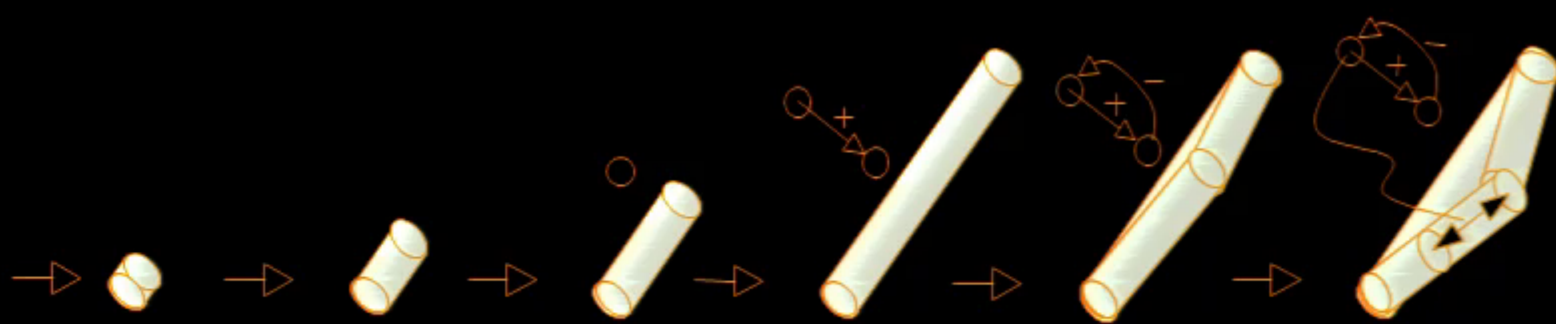
Cornell University

Creative Machines Lab

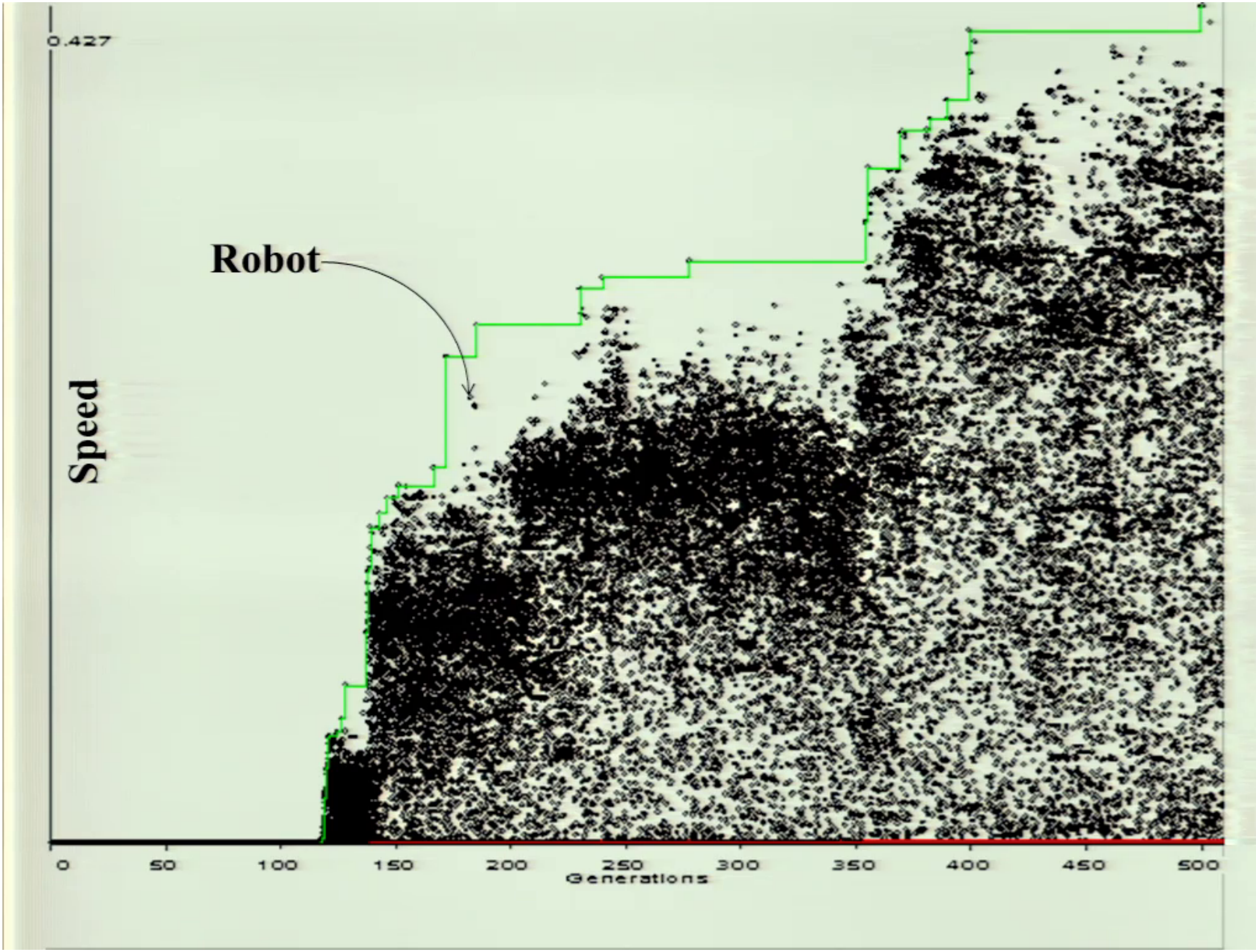

ACTION


Evolution

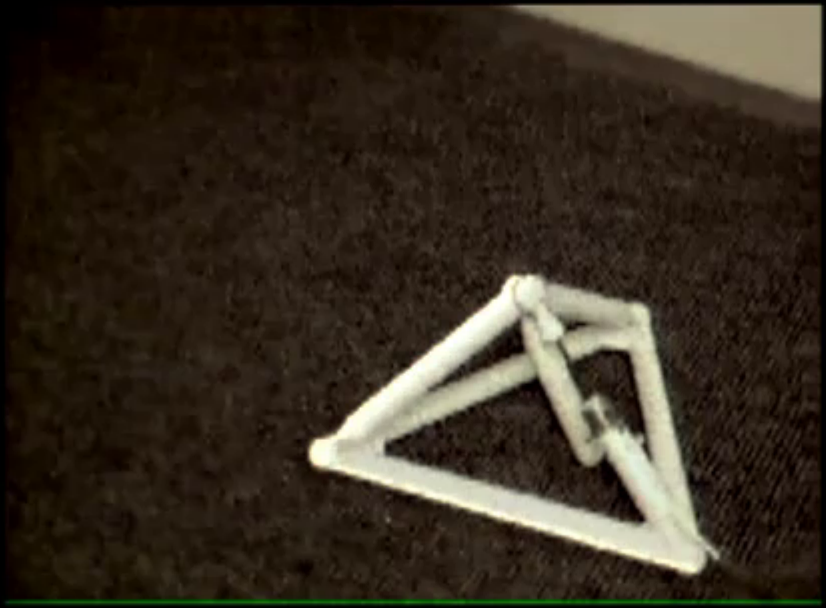
Evolution

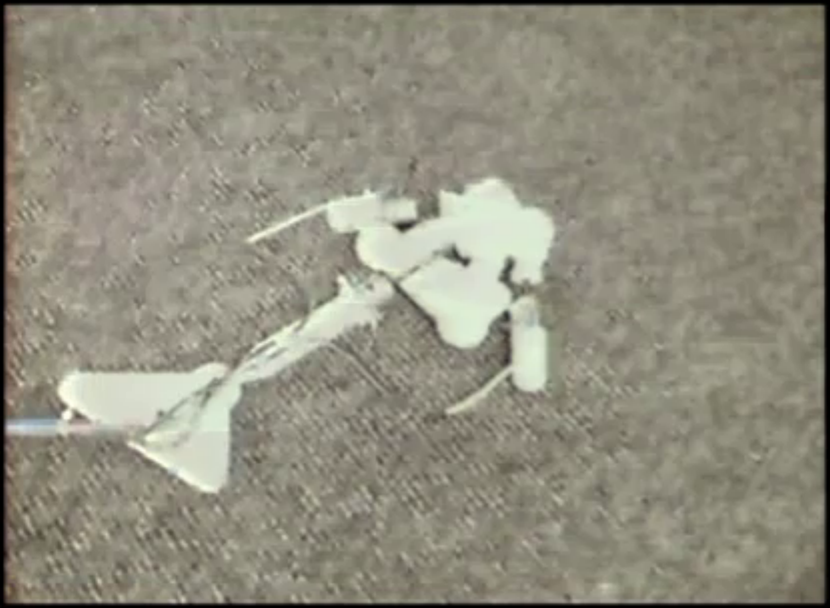



ONYX
SiliconGraphics
Computer Systems

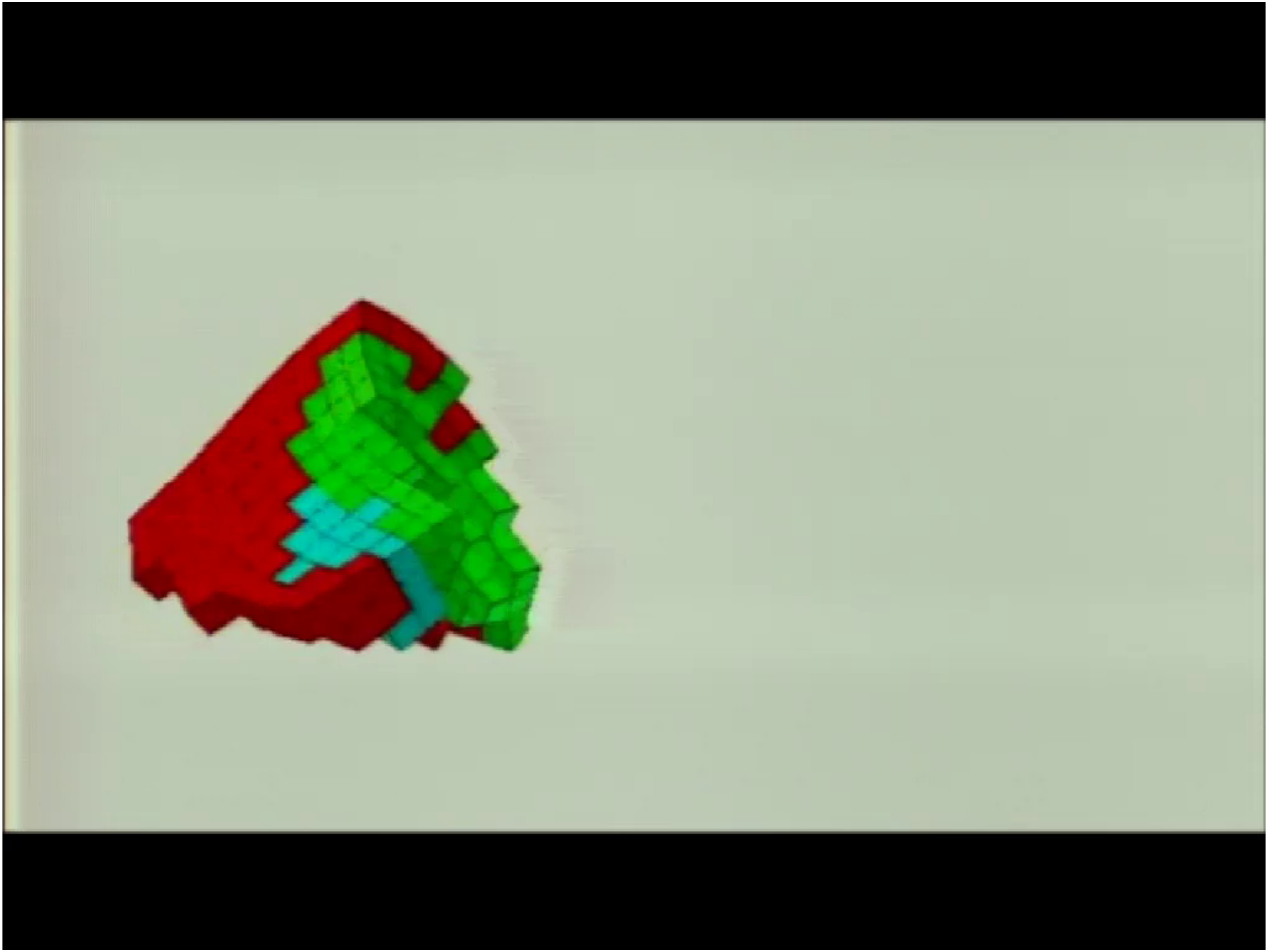


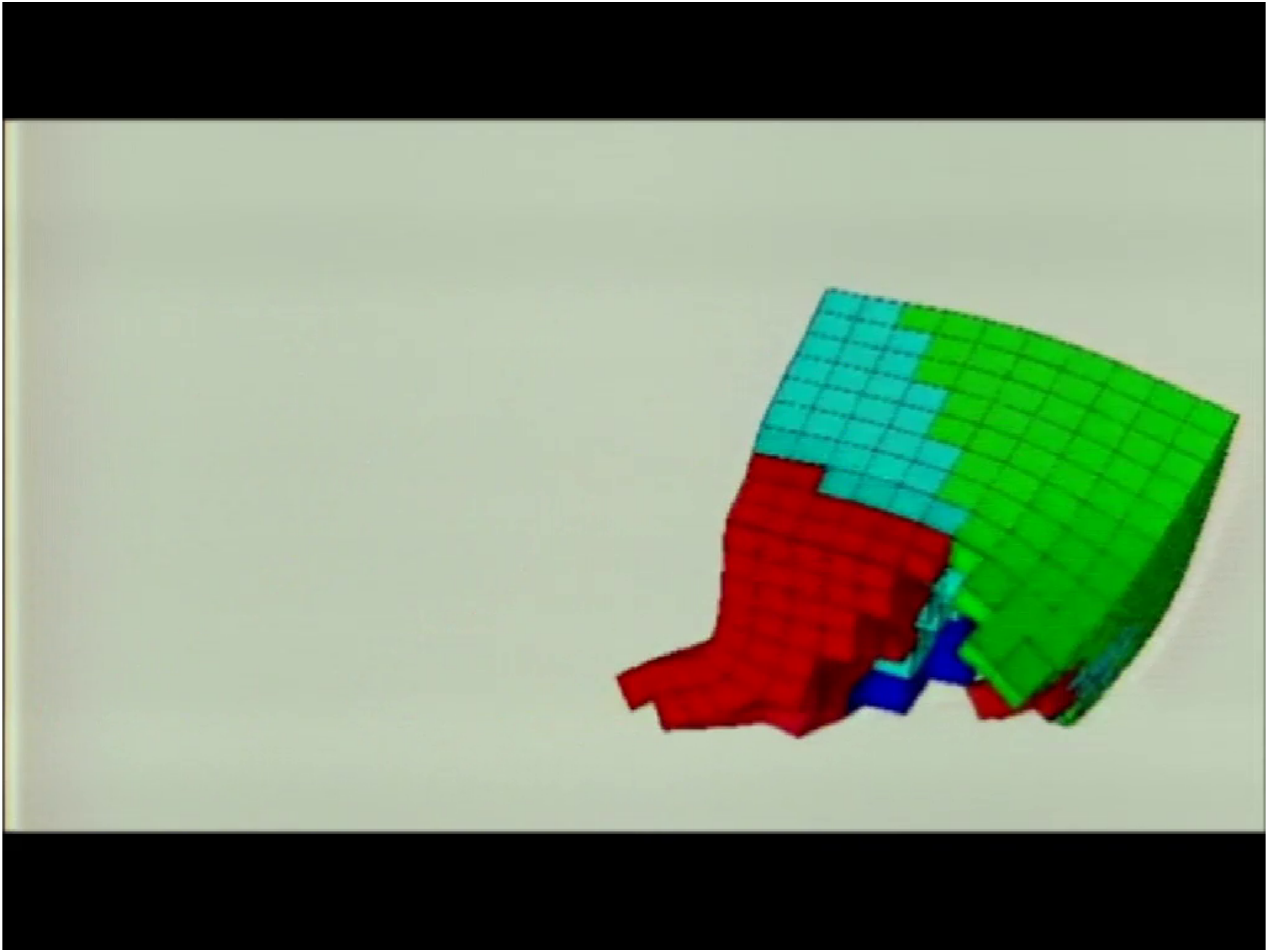
0.427
Robot
Speed
0
50
100
150
200
250
300
350
400
450
500
Generations

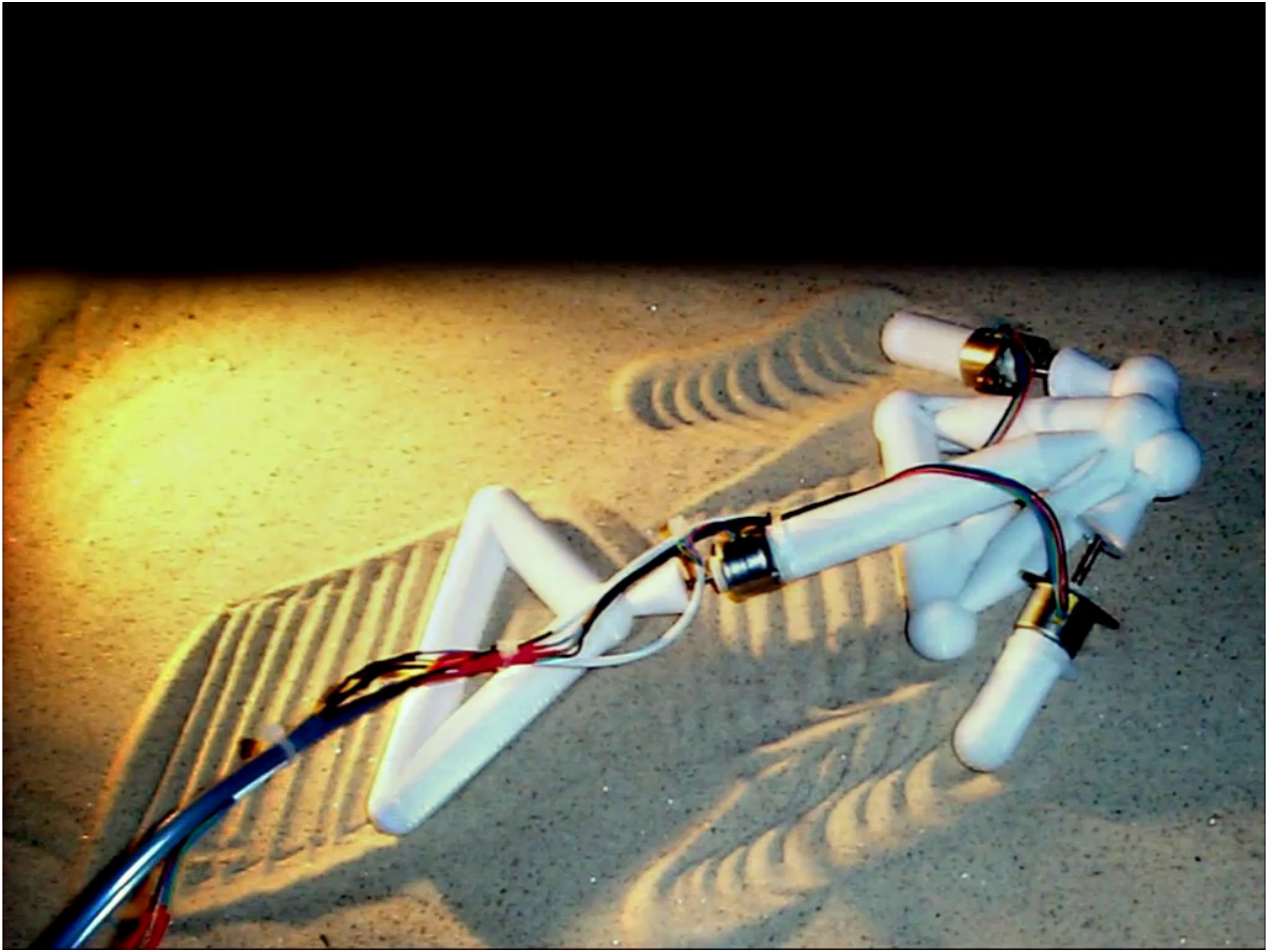




Lipson & Pollack, Nature 406, 2000







The New York Times

THURSDAY, AUGUST 31, 2000

Scientists Report They Have Made Robot That Makes Its Own Robots

By KENNETH CHANG

For the first time, computer scientists have created a robot that designs and builds other robots, almost entirely without human help.

In the short run, this advance could lead to a new industry of inexpensive robots customized for specific tasks. In the long run — decades at least — robots may one day be truly regarded as "artificial life," able to reproduce and evolve, building improved versions of themselves.

Such durable, adaptive robots, astronomers have suggested, could someday be sent into space to explore the galaxy or search for other life.

But the quest to create artificial

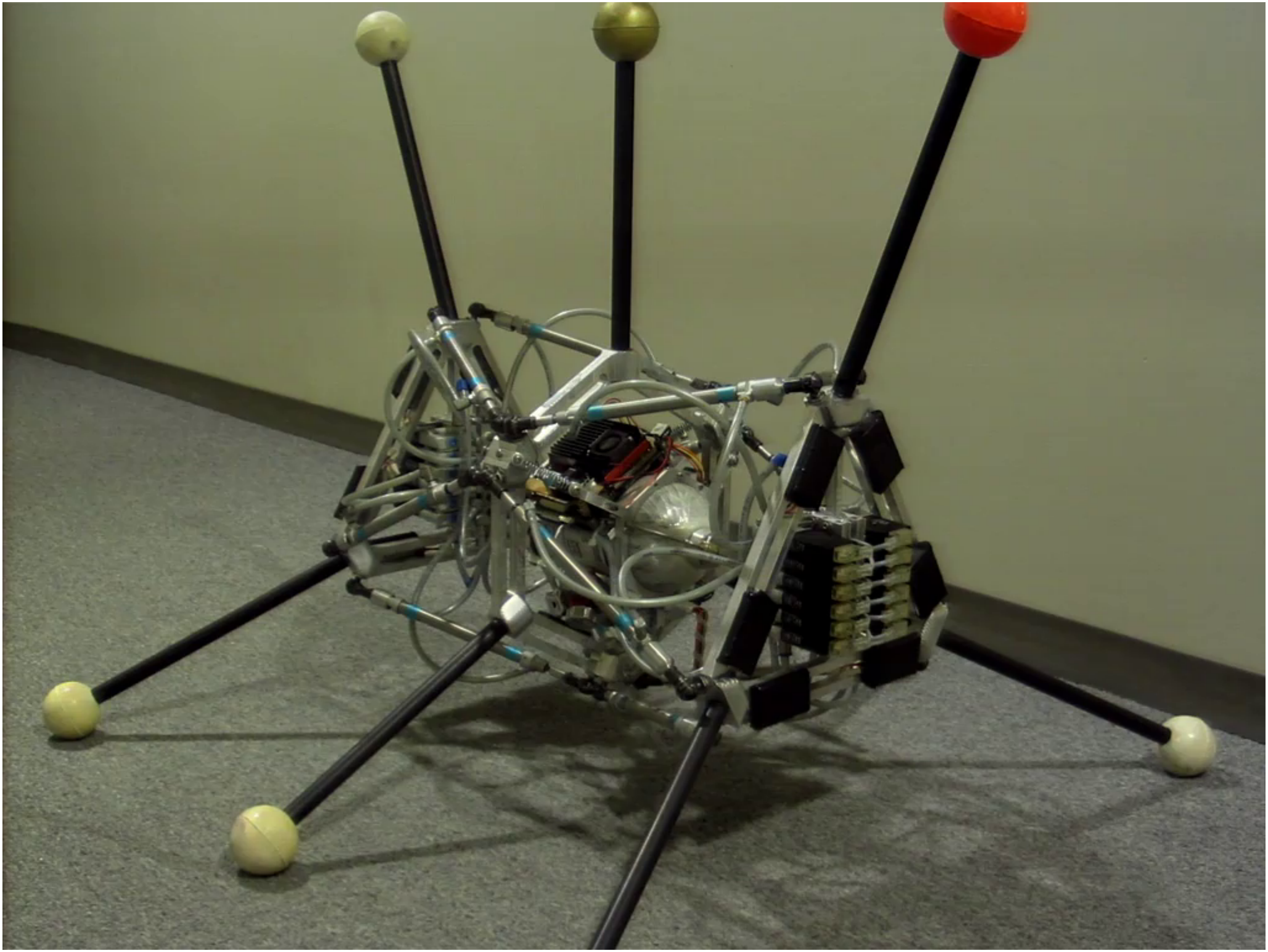
Brandeis University

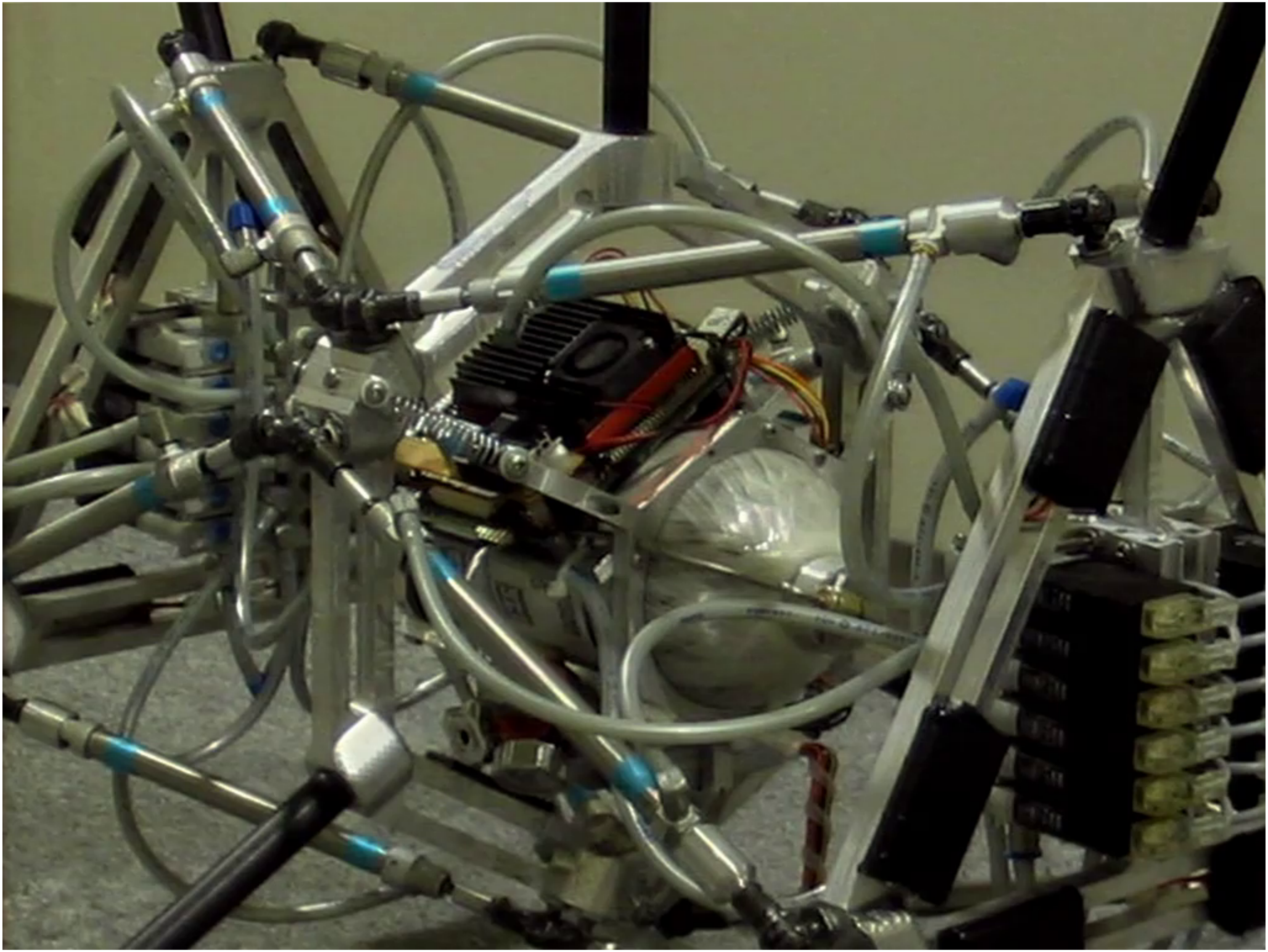
The "Arrow" left a trail as it crawled across a bed of sand.

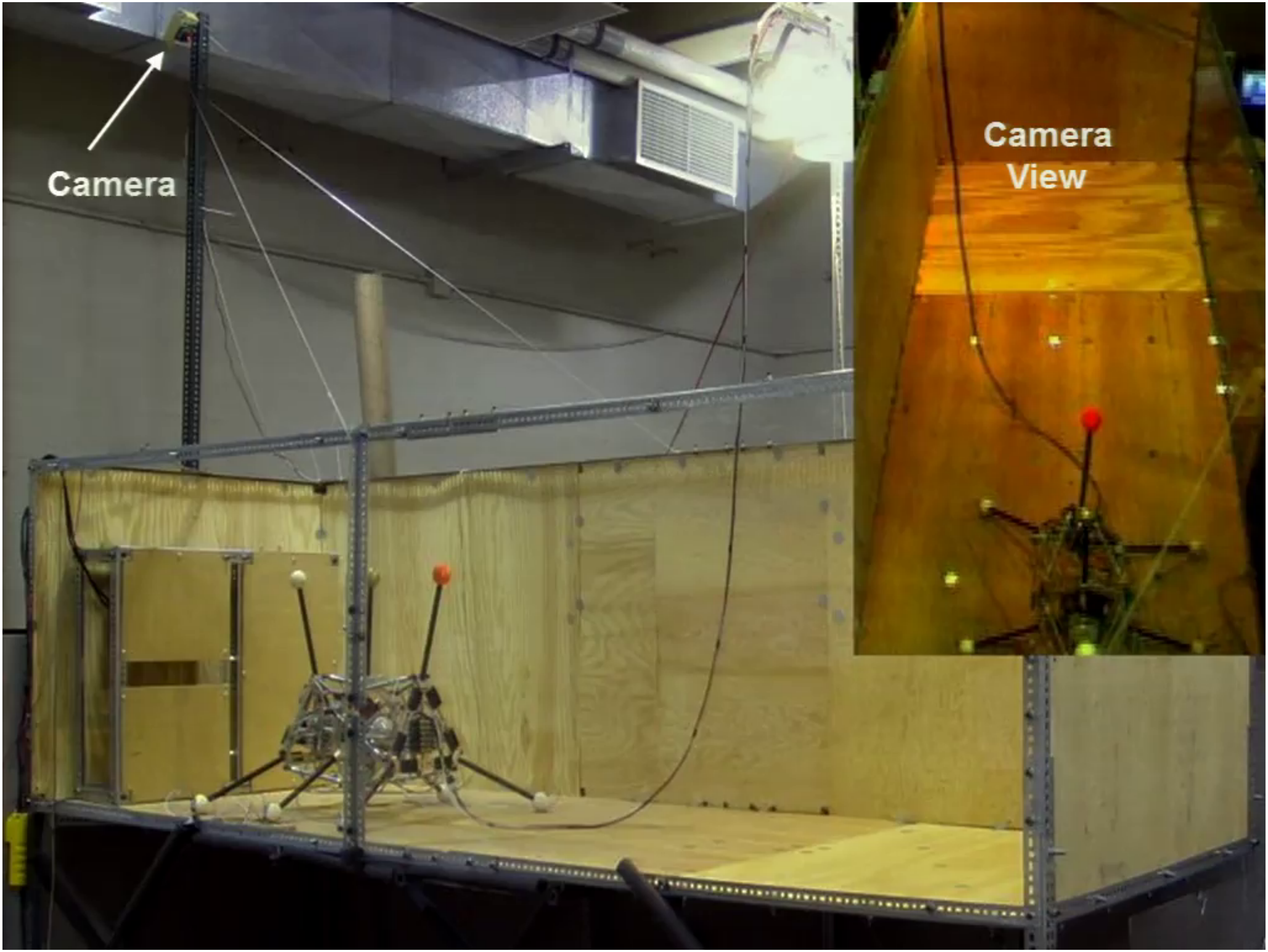
were not manufactured by humans."

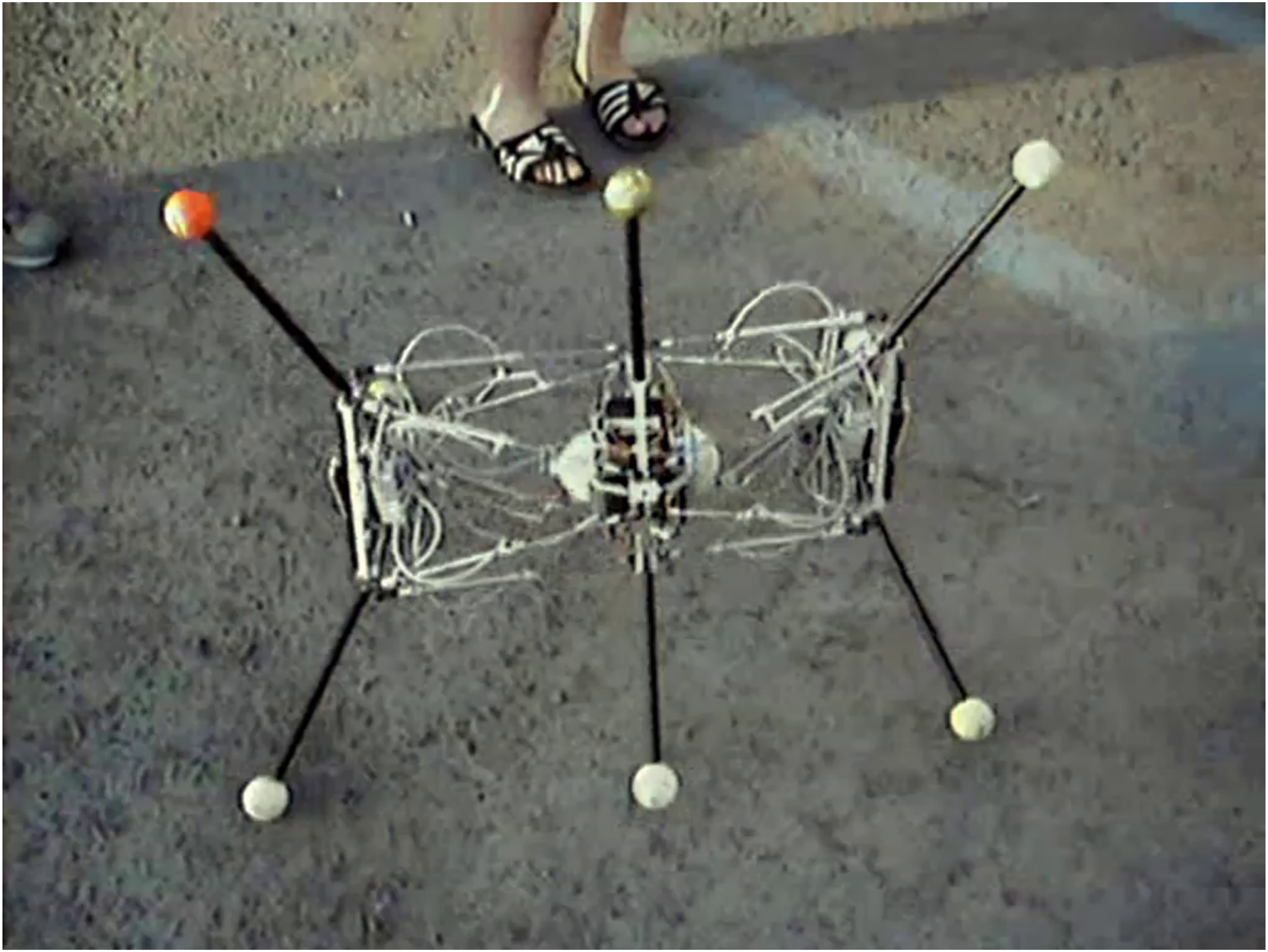
Dr. Pollack and Dr. Lipson, a research scientist, report their results in today's issue of the journal Nature.

"This is the first example of pretty

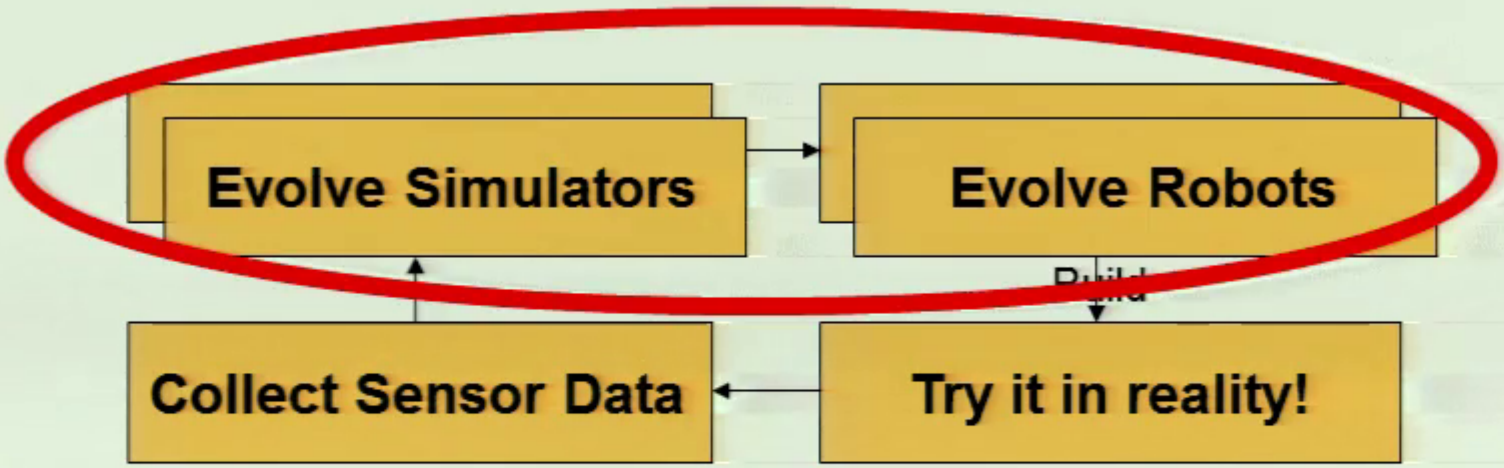


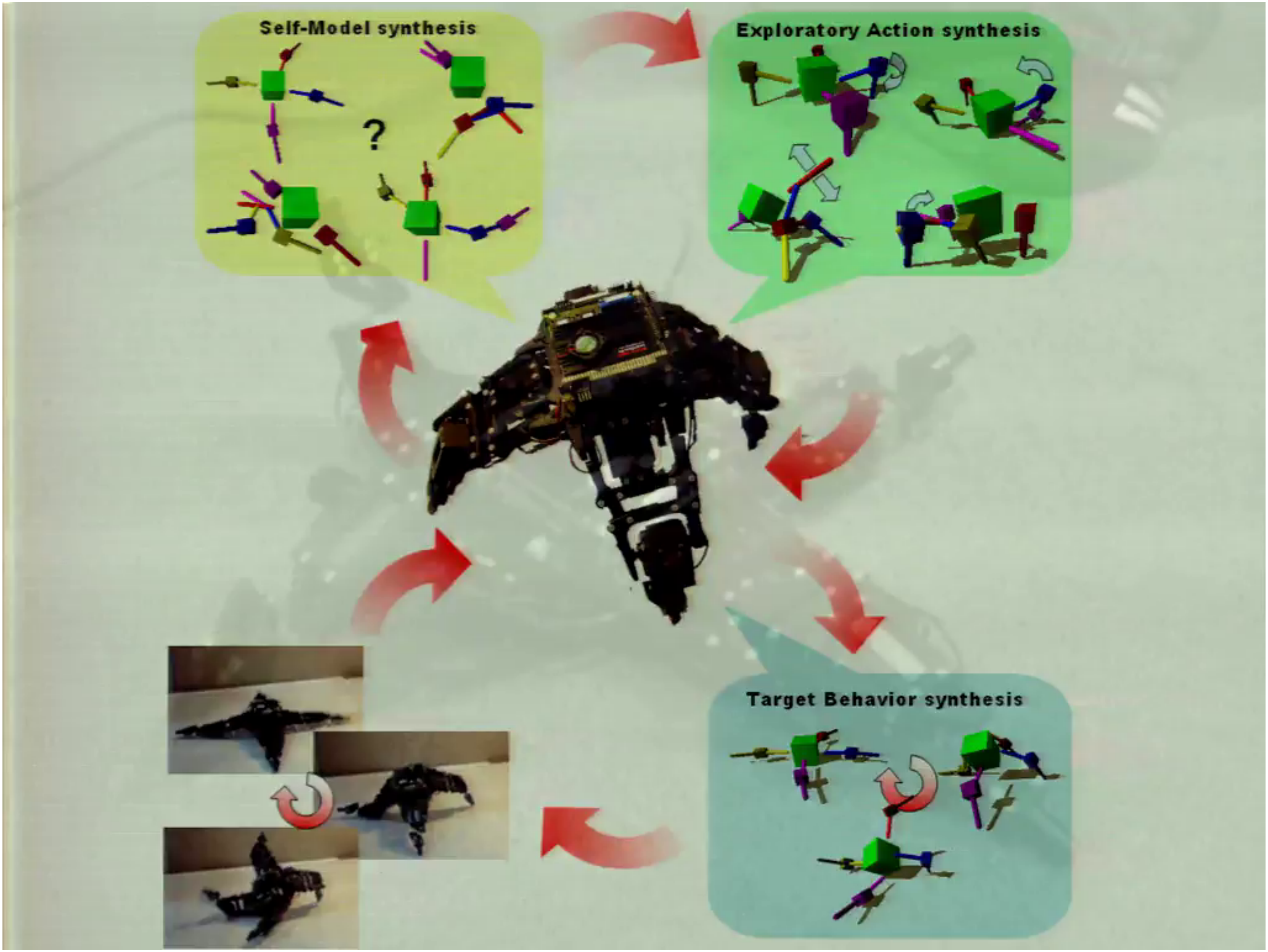


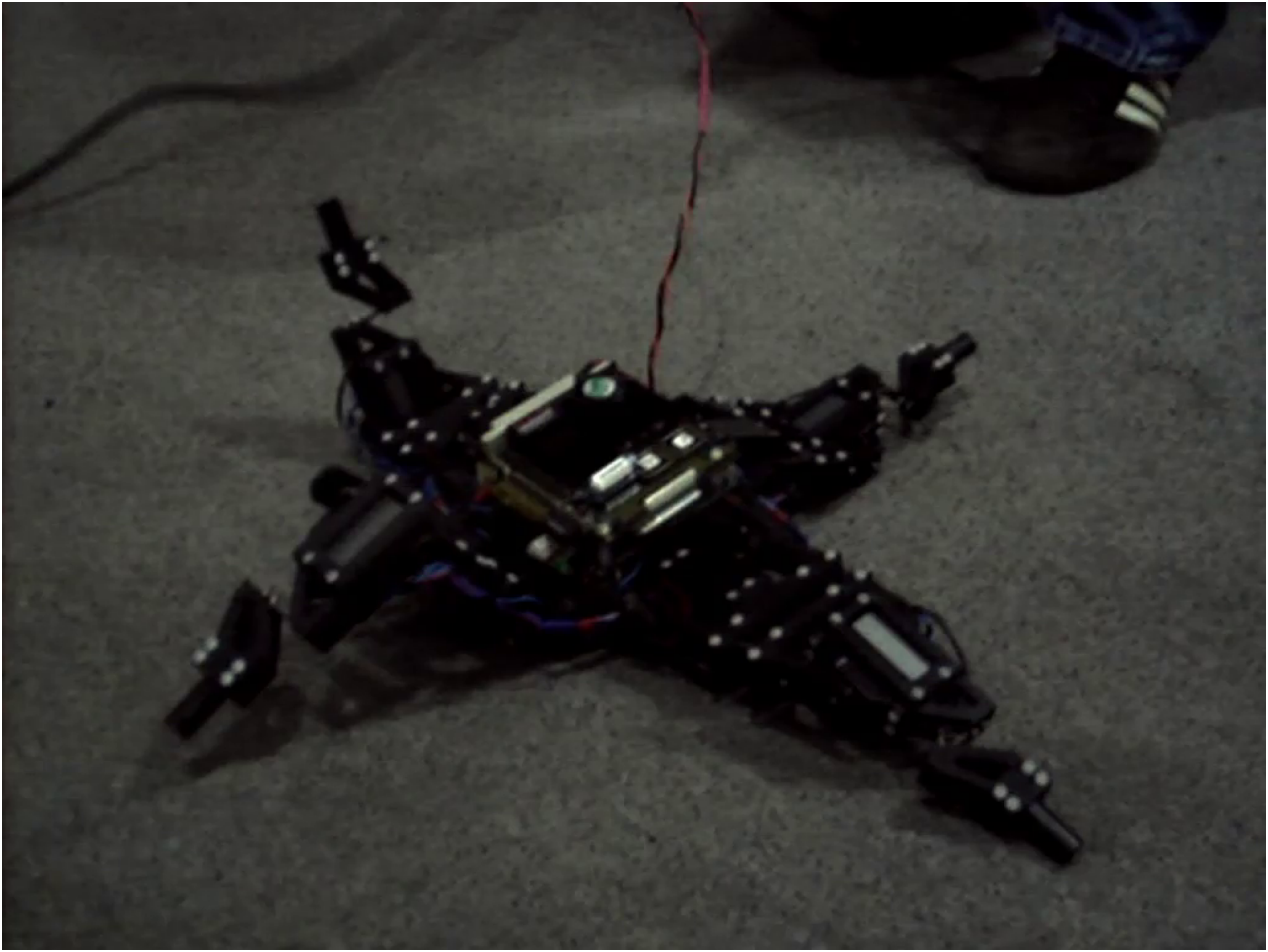

Camera
Camera
View



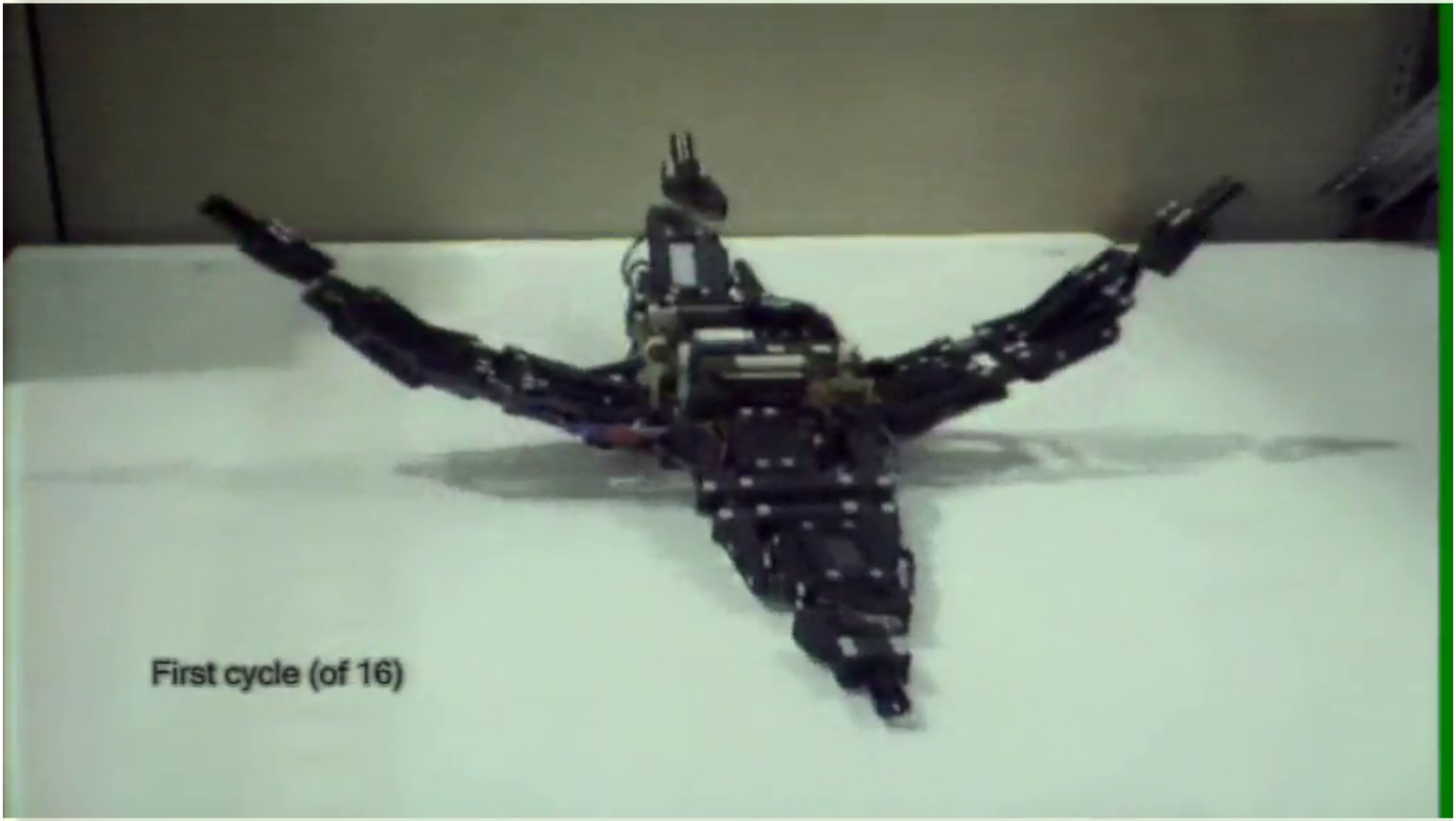
Simulation & Reality




Self-Model synthesis
?
Exploratory Action synthesis
Target Behavior synthesis

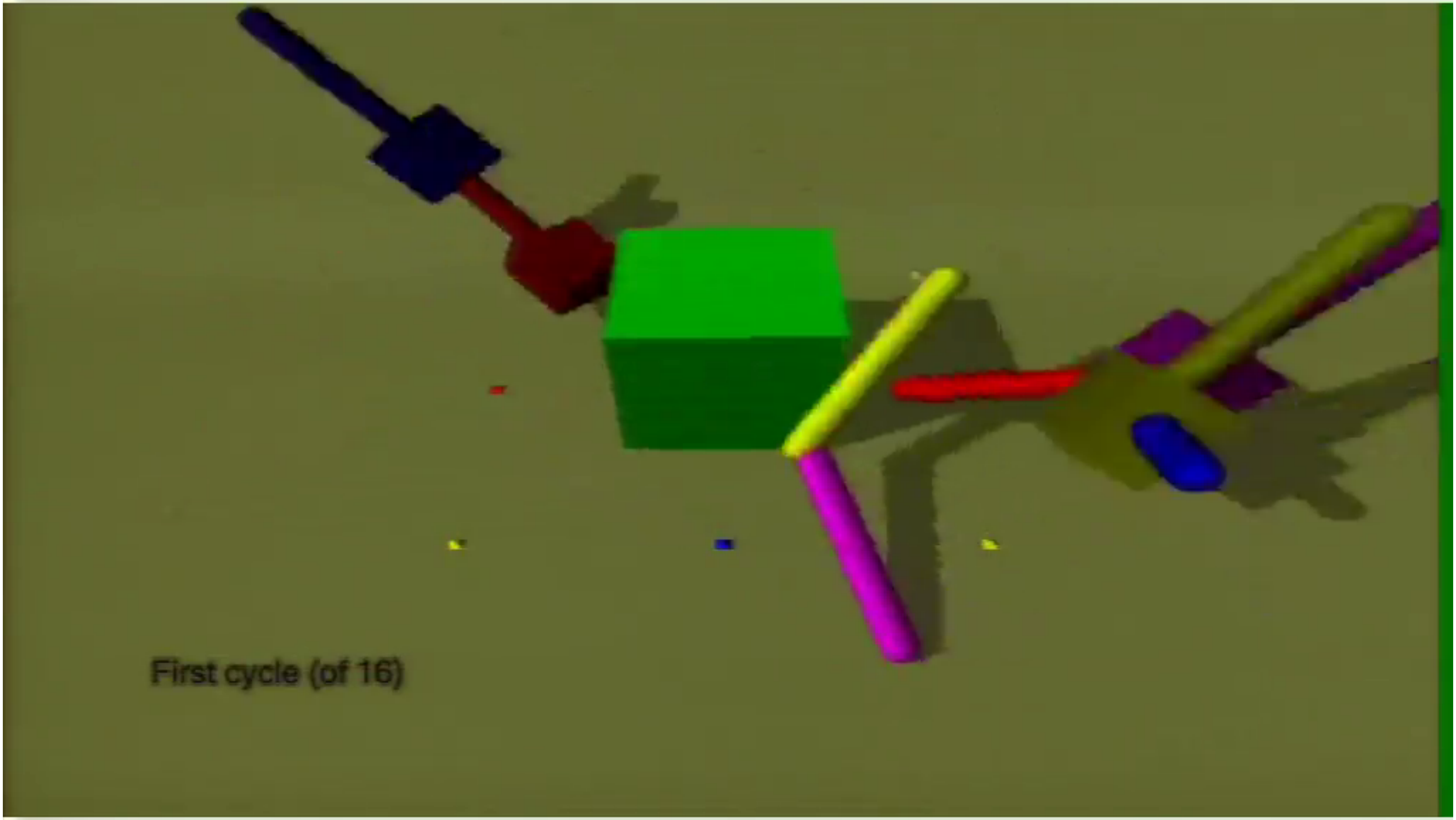


Emergent Self-Model



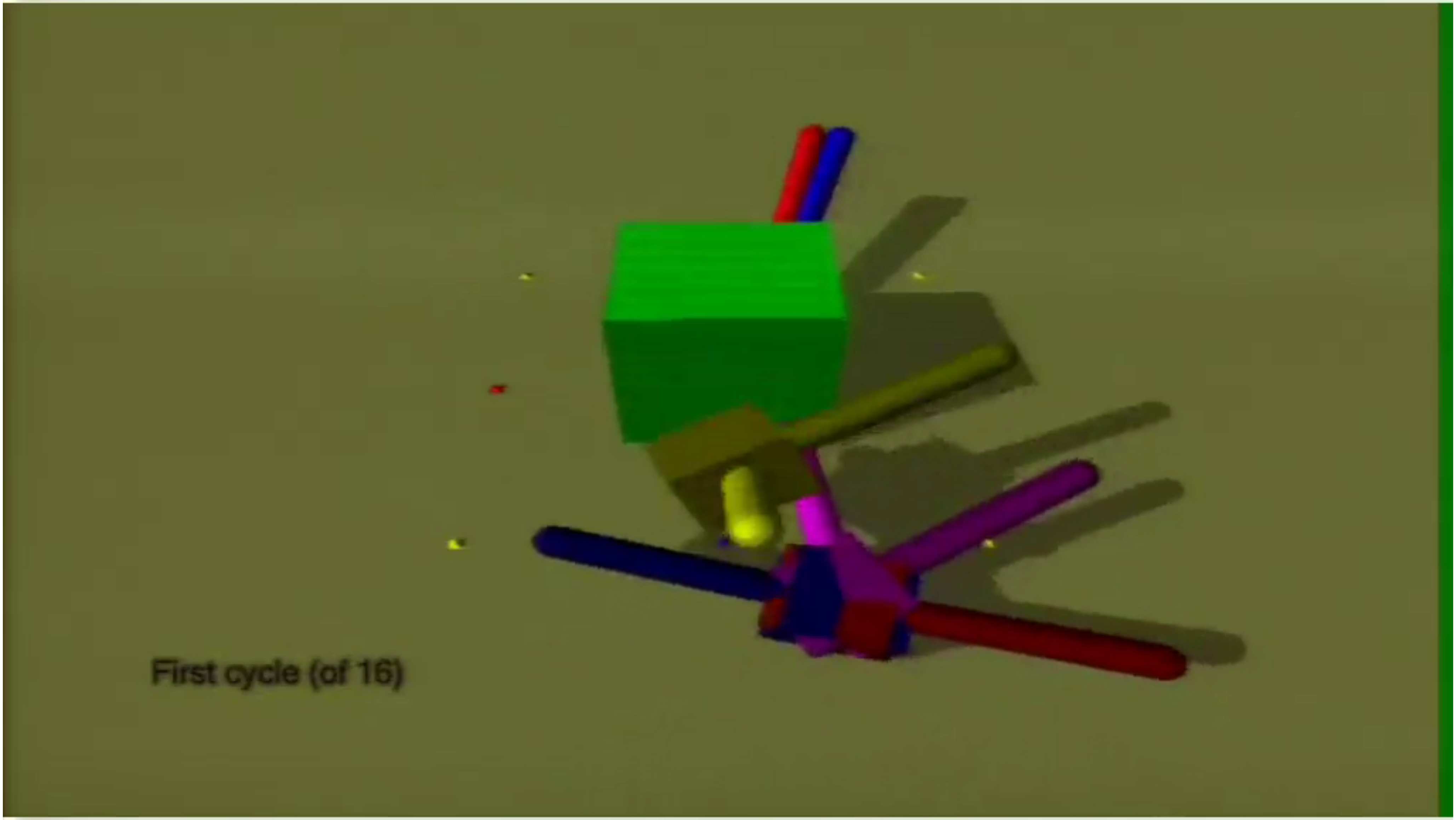
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



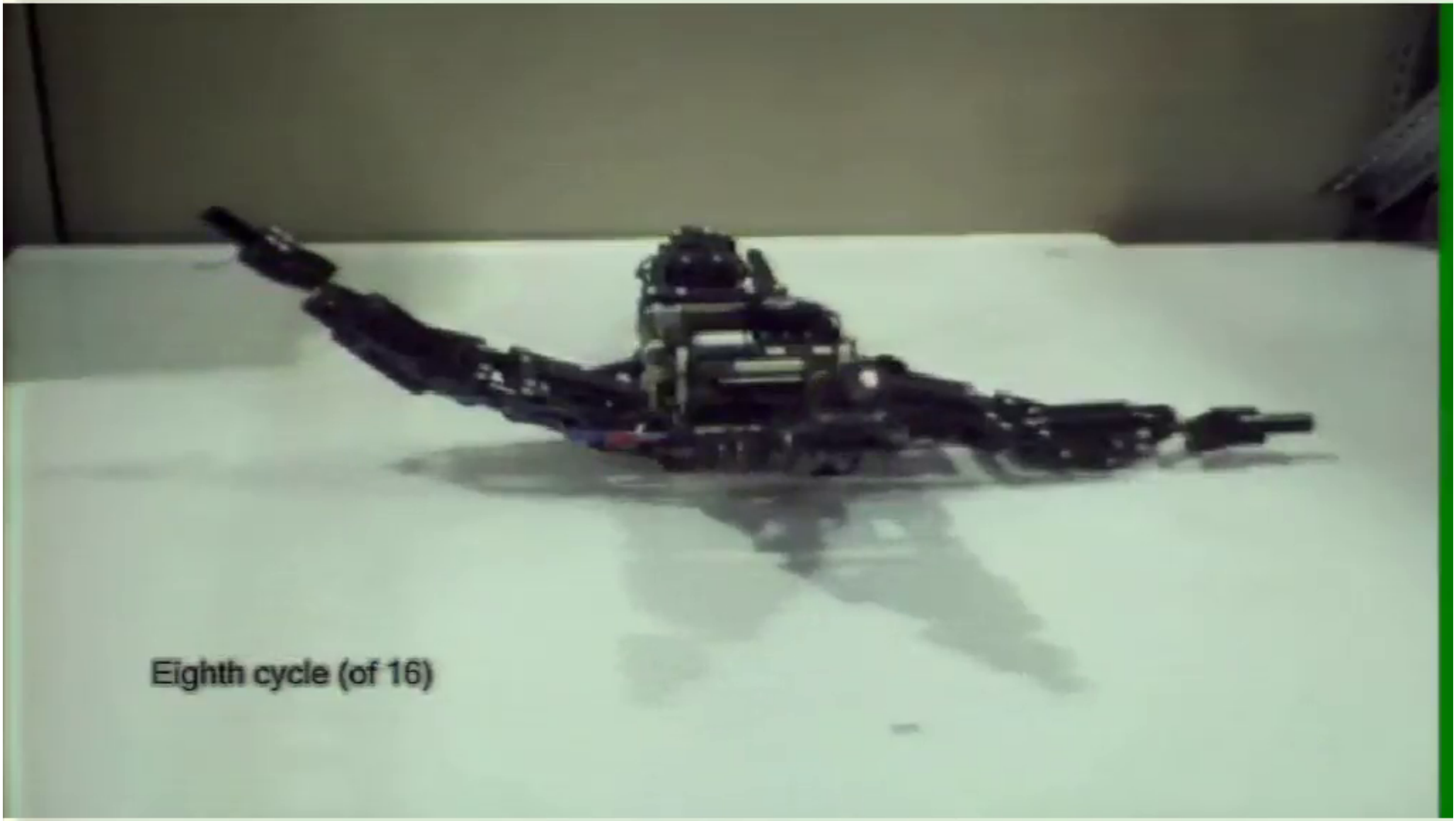
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



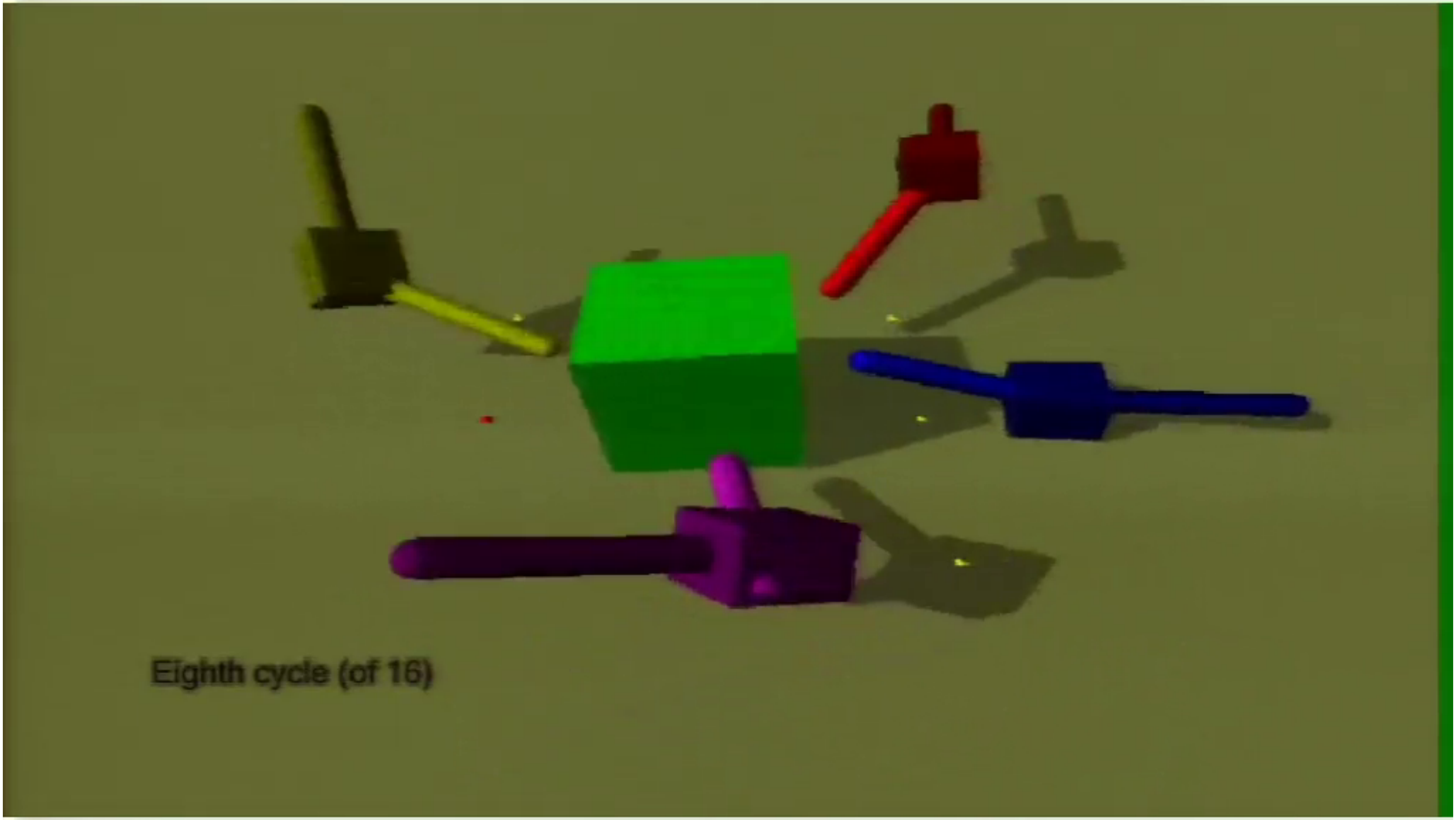
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



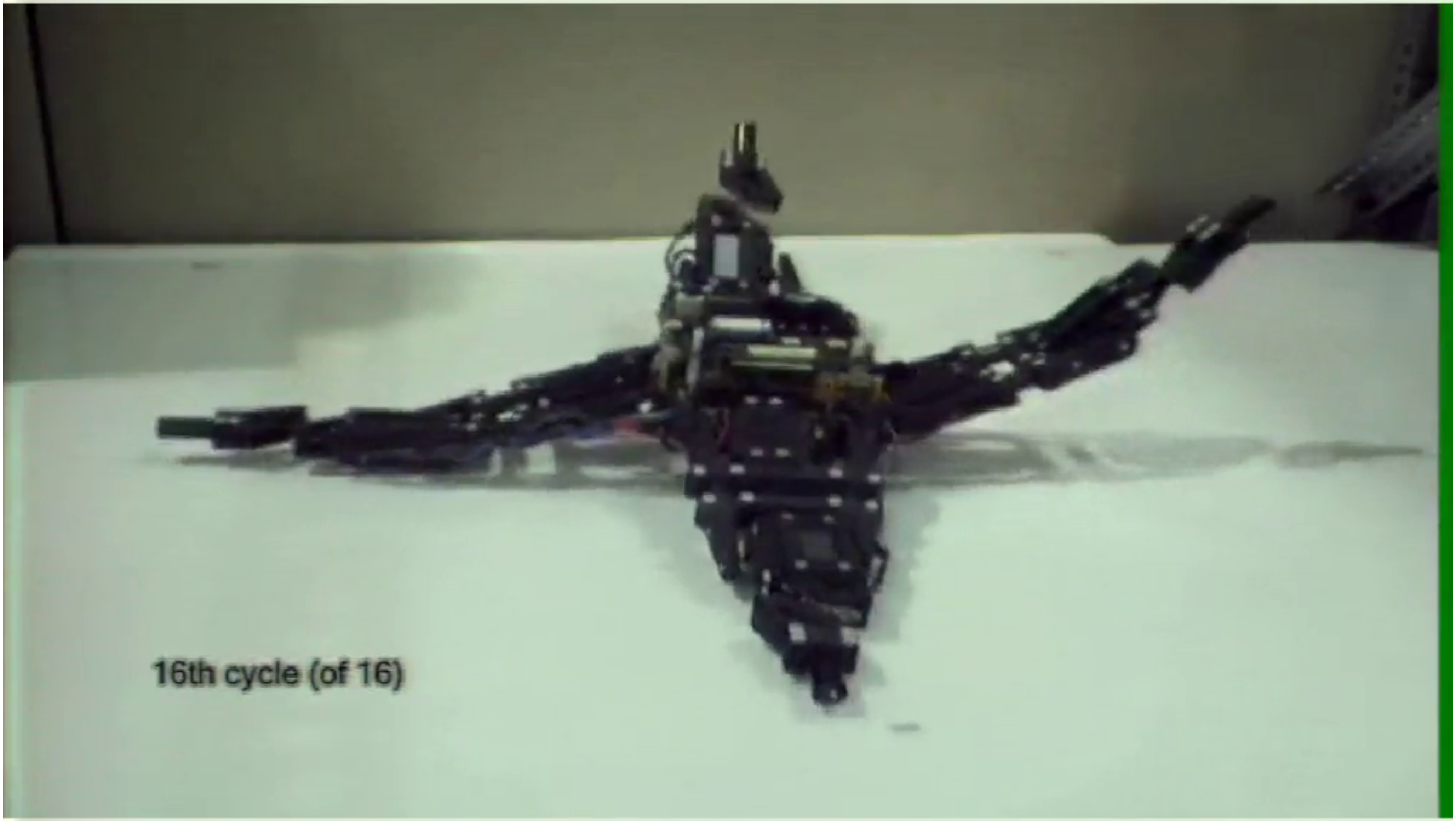
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



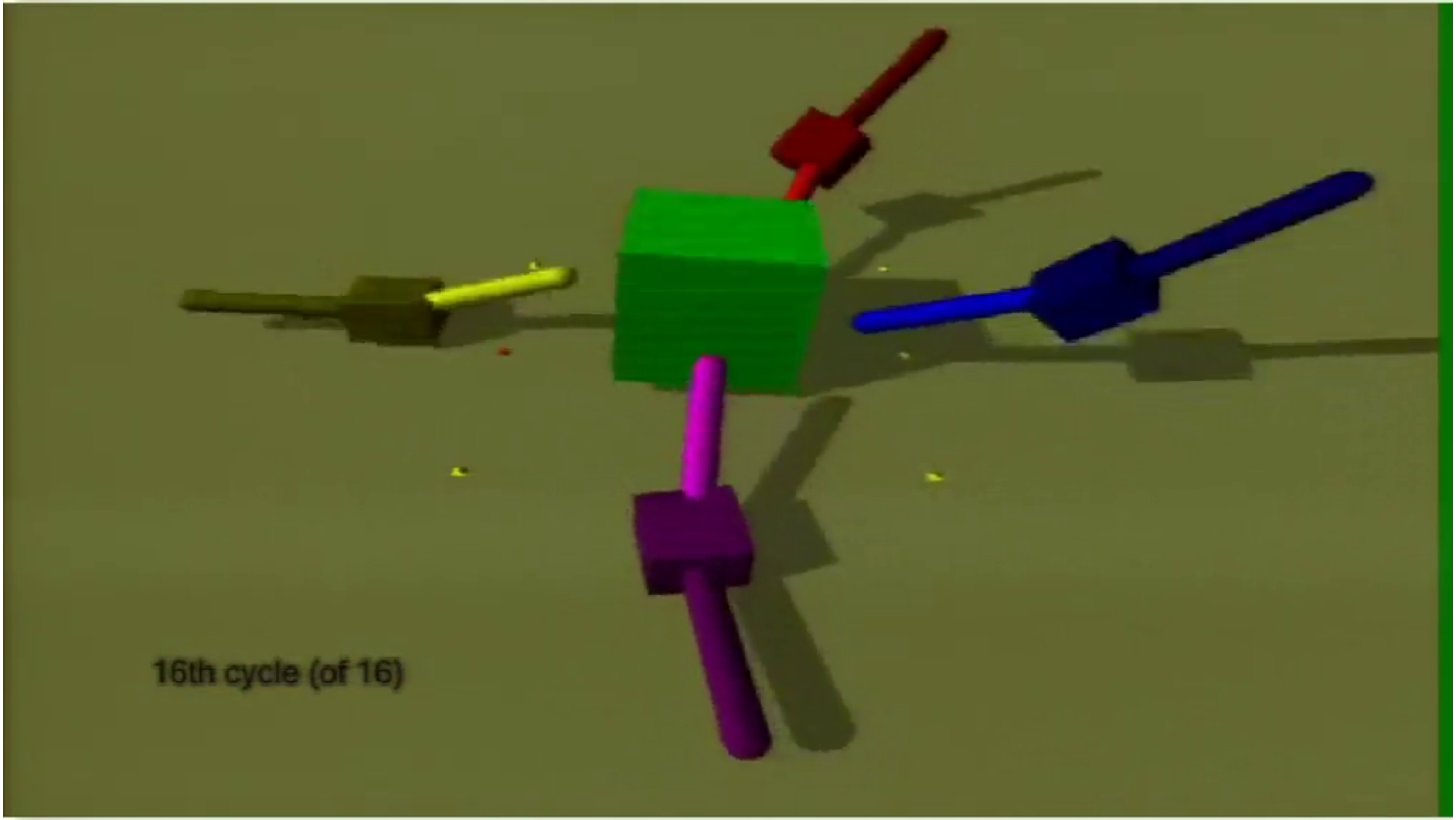
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



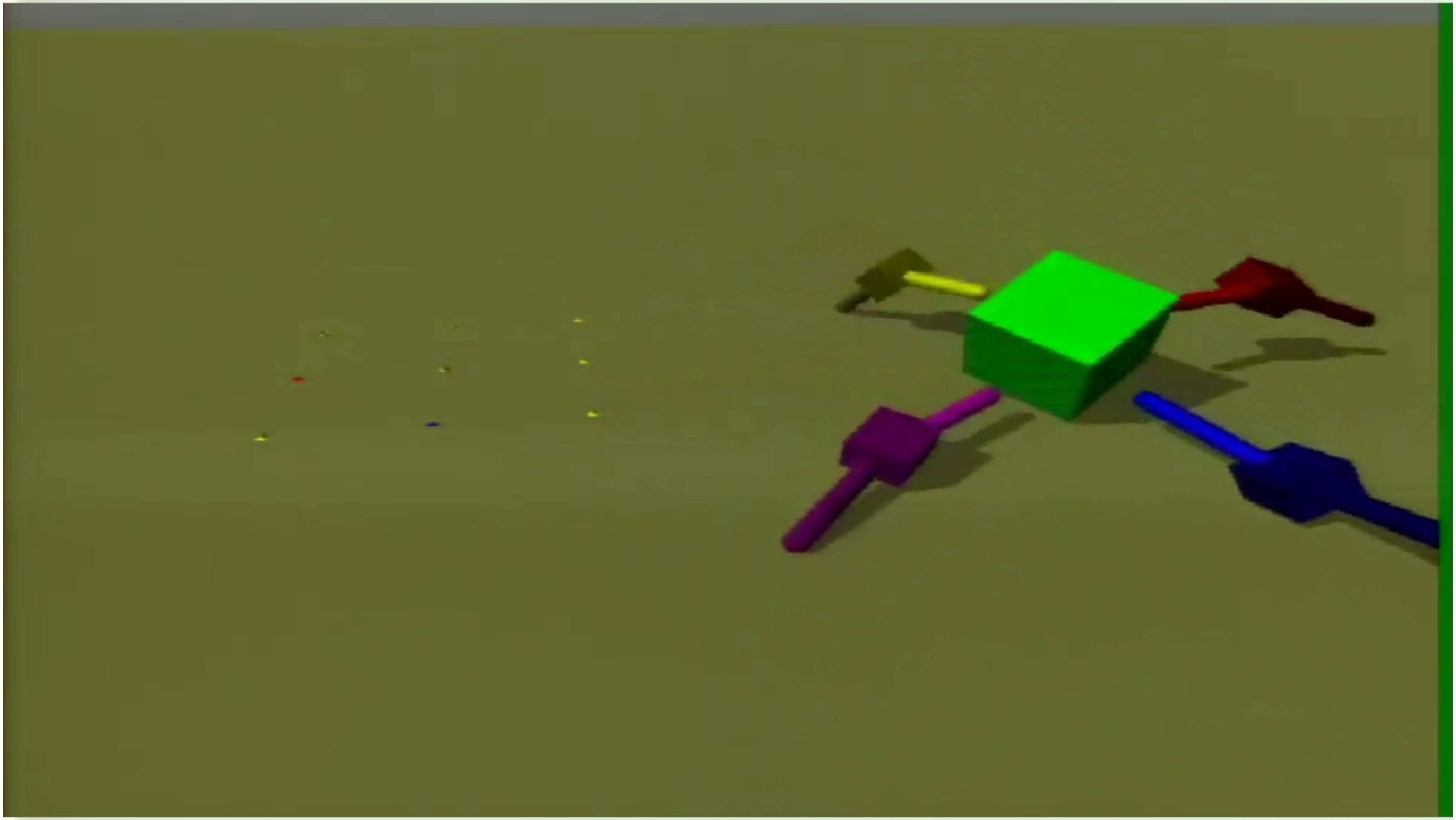
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model

Stage II:
Using the model to generate a gait

With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



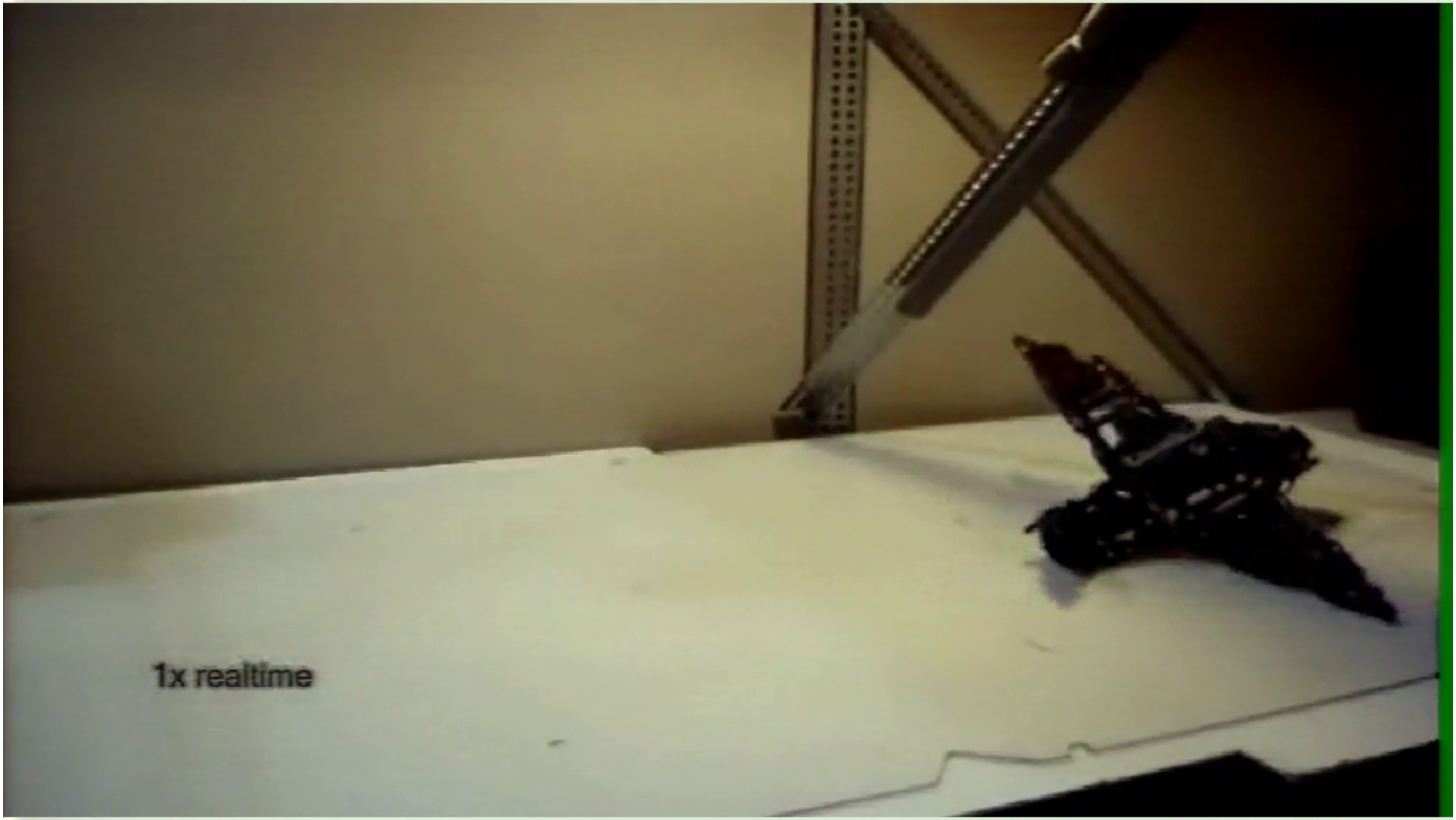
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model

Stage III:
Trying gait in reality

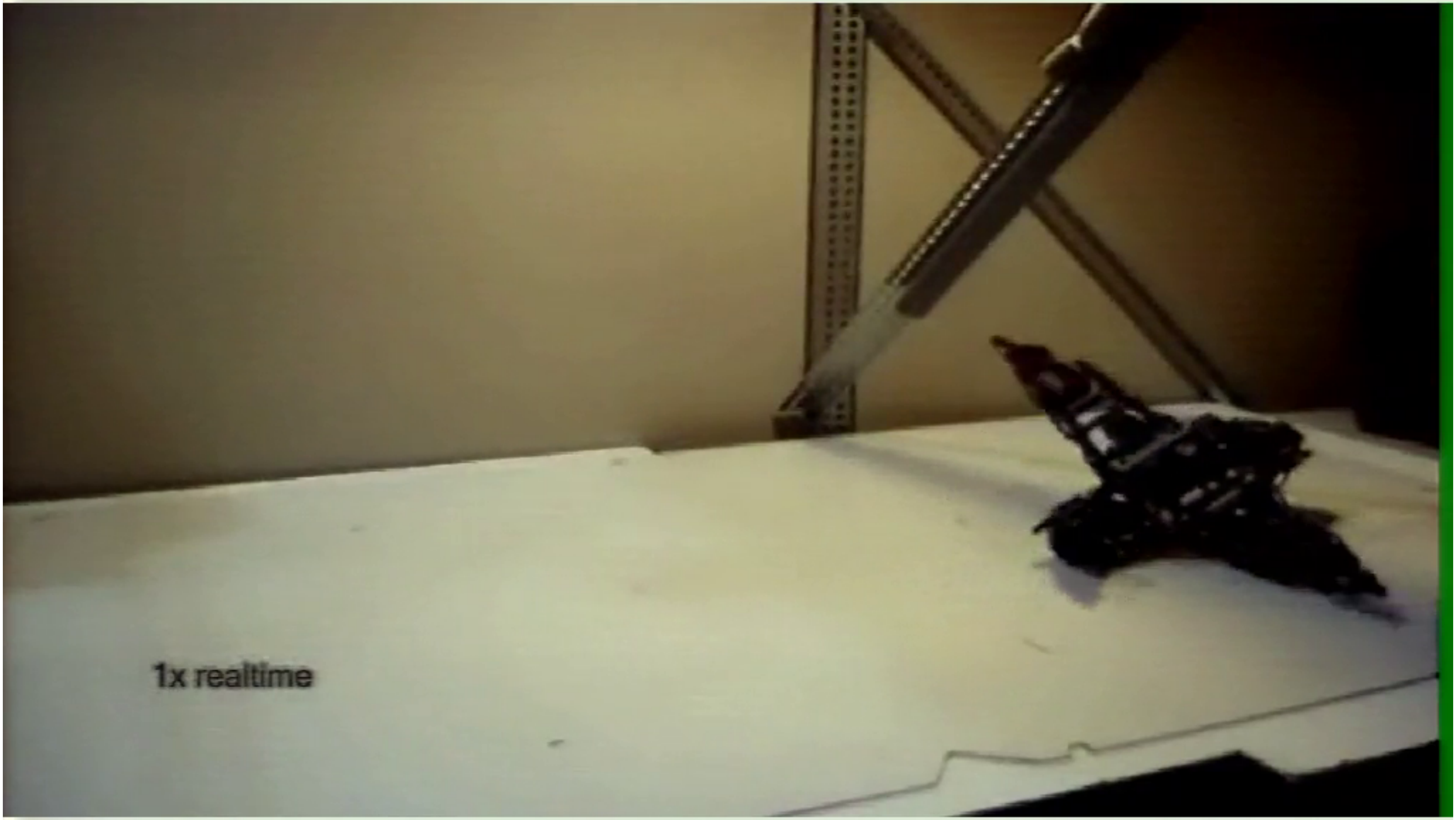
With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model

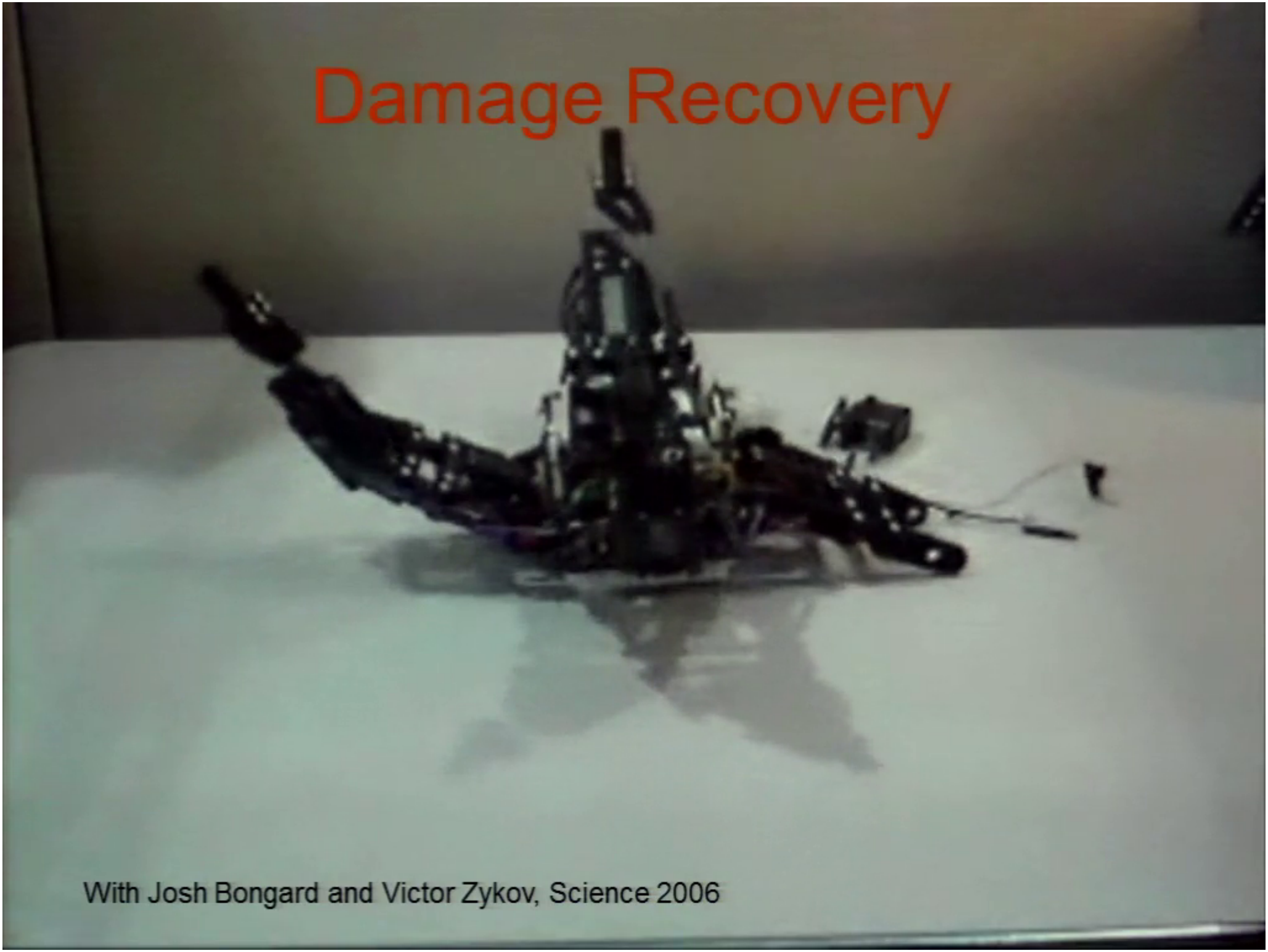


With Josh Bongard and Victor Zykov, Science 2006

Emergent Self-Model



With Josh Bongard and Victor Zykov, Science 2006


Damage Recovery
With Josh Bongard and Victor Zykov, Science 2006


Damage Recovery
With Josh Bongard and Victor Zykov, Science 2006


Damage Recovery
With Josh Bongard and Victor Zykov, Science 2006


Damage Recovery
With Josh Bongard and Victor Zykov, Science 2006

Damage Recovery

Other gaits generated for damaged morphology

With Josh Bongard and Victor Zykov, Science 2006

Damage Recovery

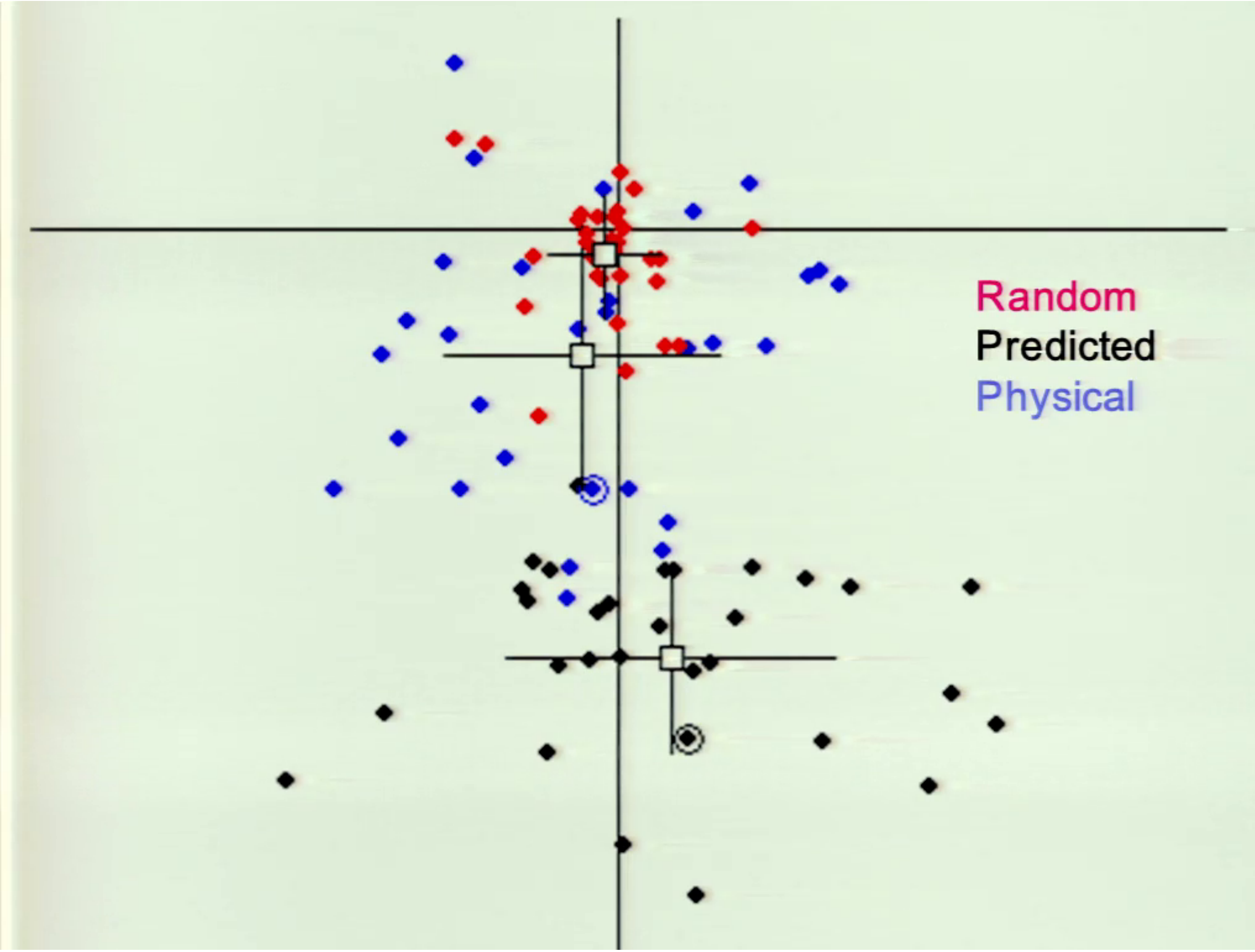
Other gaits generated for damaged morphology

With Josh Bongard and Victor Zykov, Science 2006

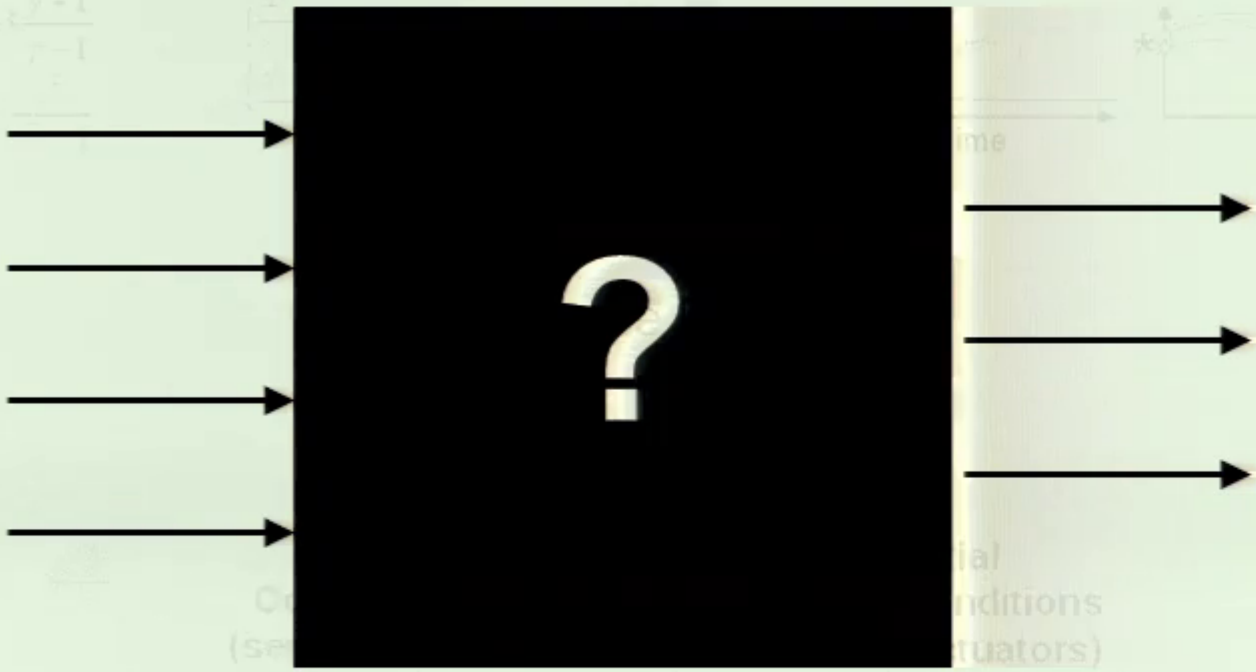
Damage Recovery

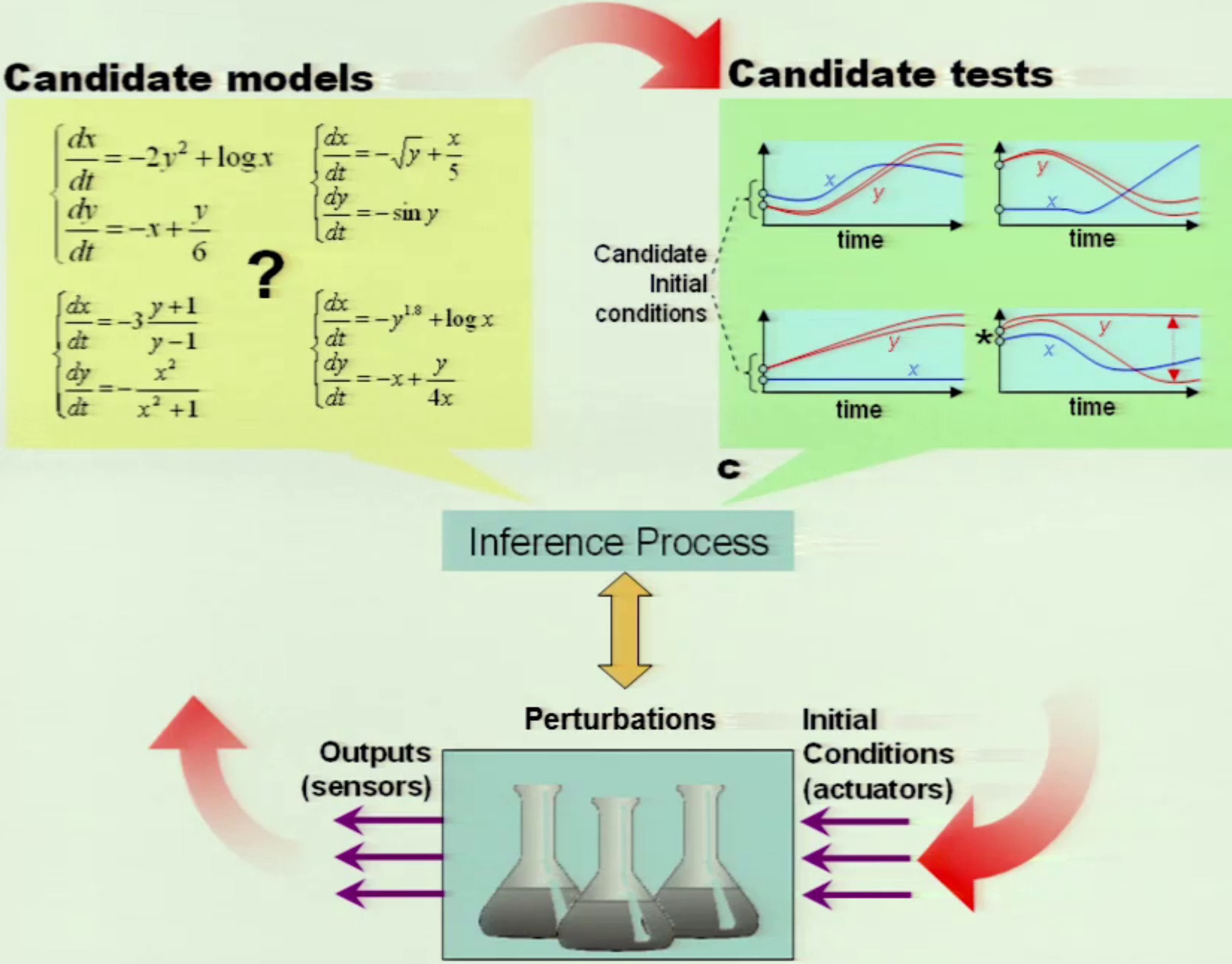
With Josh Bongard and Victor Zykov, Science 2006

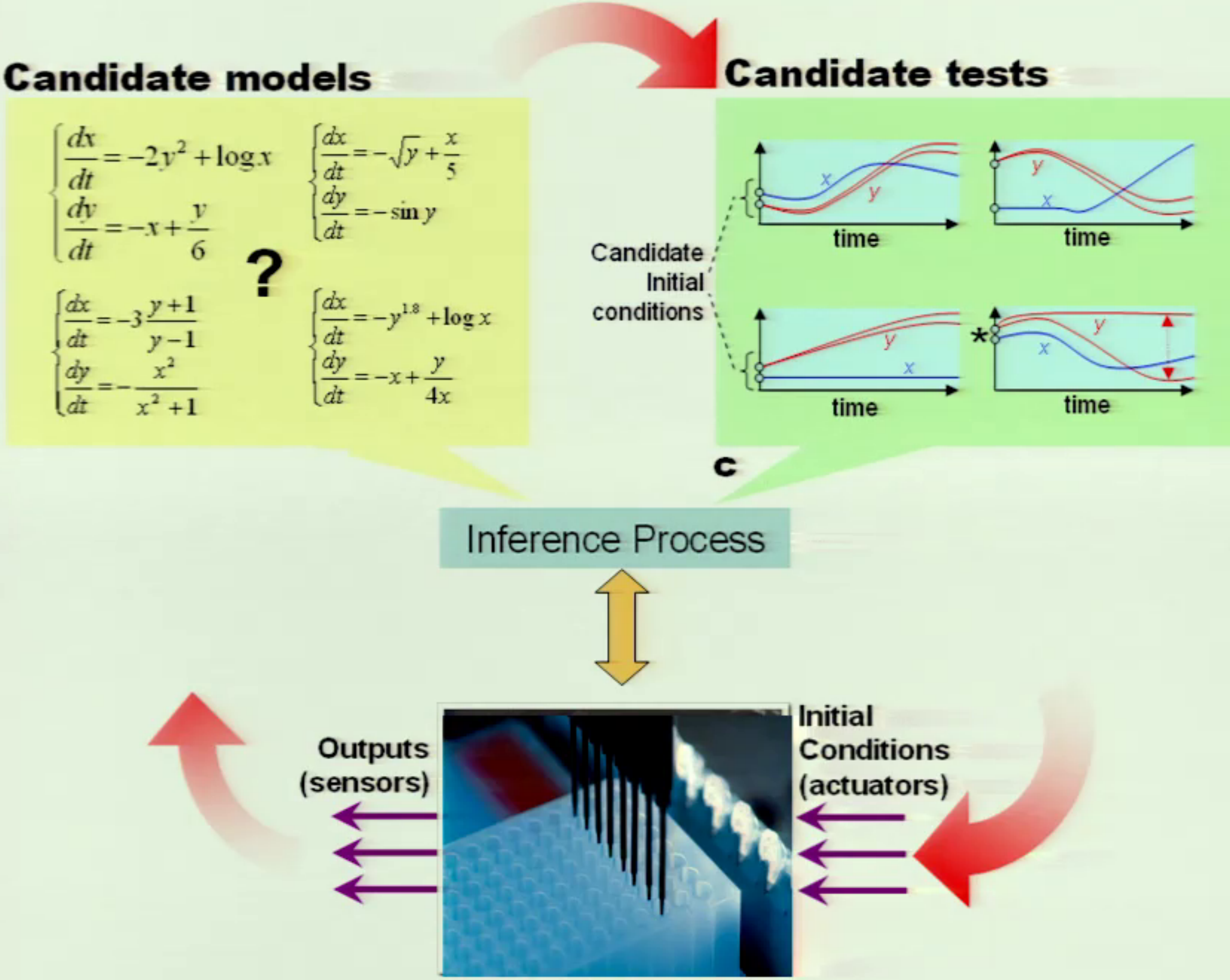

Damage Recovery
With Josh Bongard and Victor Zykov, Science 2006


Random
Predicted
Physical

System Identification




Candidate models
$\frac{dx}{dt} = -2y^2 + \log x$, $\frac{dy}{dt} = -x + \frac{y}{6}$
$\frac{dx}{dt} = -\sqrt{y} + \frac{x}{5}$, $\frac{dy}{dt} = -\sin y$
?
$\frac{dx}{dt} = -3\frac{y+1}{y-1}$, $\frac{dy}{dt} = -\frac{x^2}{x^2+1}$
$\frac{dx}{dt} = -y^{1.8} + \log x$, $\frac{dy}{dt} = -x + \frac{y}{4x}$
Candidate tests
Candidate Initial conditions
x
y
time
*
c
Inference Process
Perturbations
Outputs (sensors)
Initial Conditions (actuators)


Candidate models
$\frac{dx}{dt} = -2y^2 + \log x$, $\frac{dy}{dt} = -x + \frac{y}{6}$
$\frac{dx}{dt} = -\sqrt{y} + \frac{x}{5}$, $\frac{dy}{dt} = -\sin y$
?
$\frac{dx}{dt} = -3\frac{y+1}{y-1}$, $\frac{dy}{dt} = -\frac{x^2}{x^2+1}$
$\frac{dx}{dt} = -y^{1.8} + \log x$, $\frac{dy}{dt} = -x + \frac{y}{4x}$
Candidate tests
Candidate Initial conditions
x
y
time
*
c
Inference Process
Outputs (sensors)
Initial Conditions (actuators)

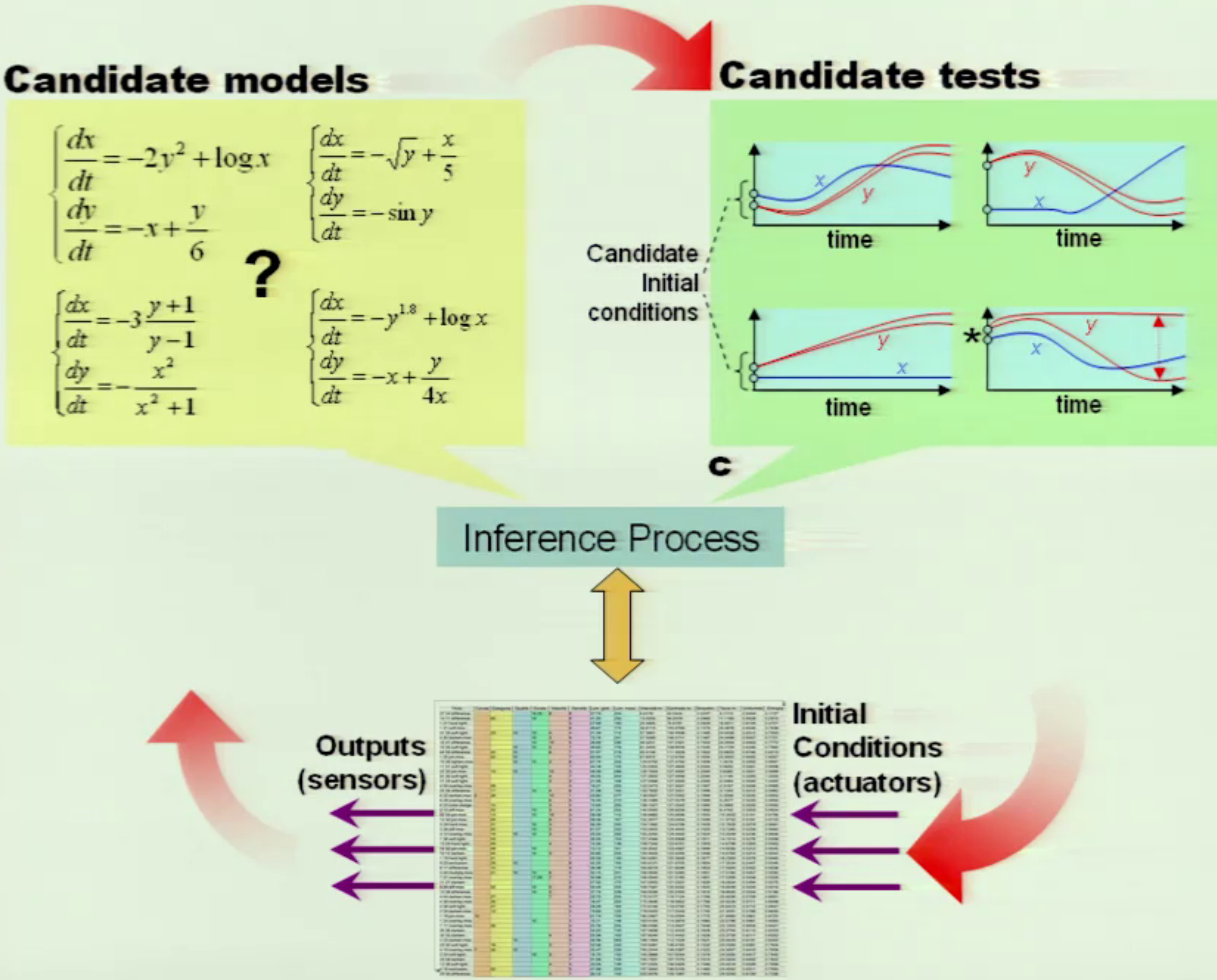
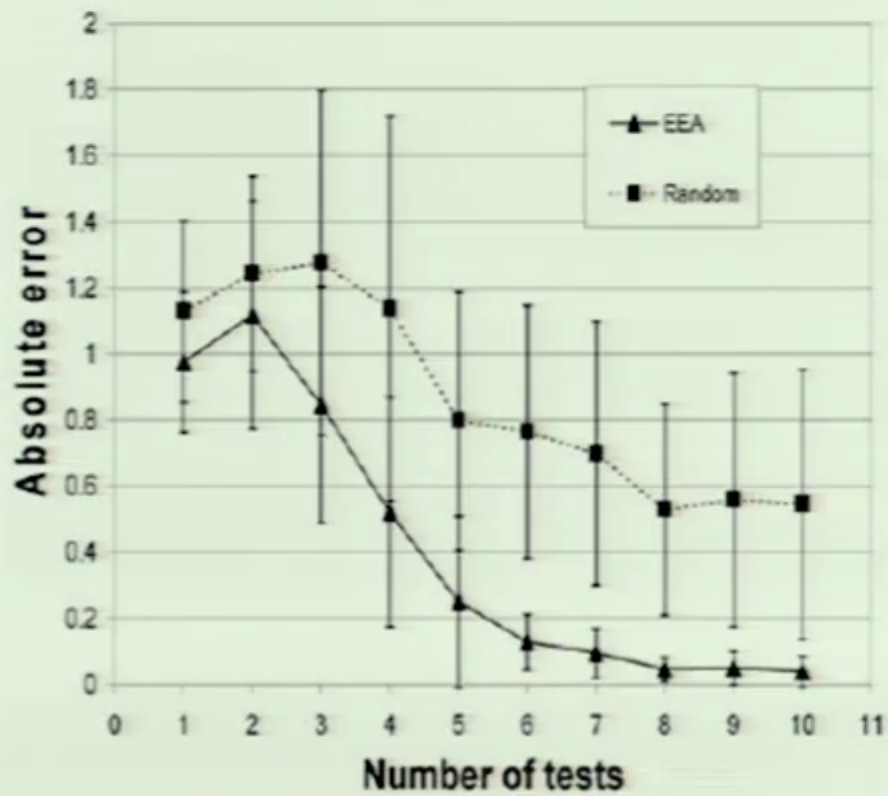
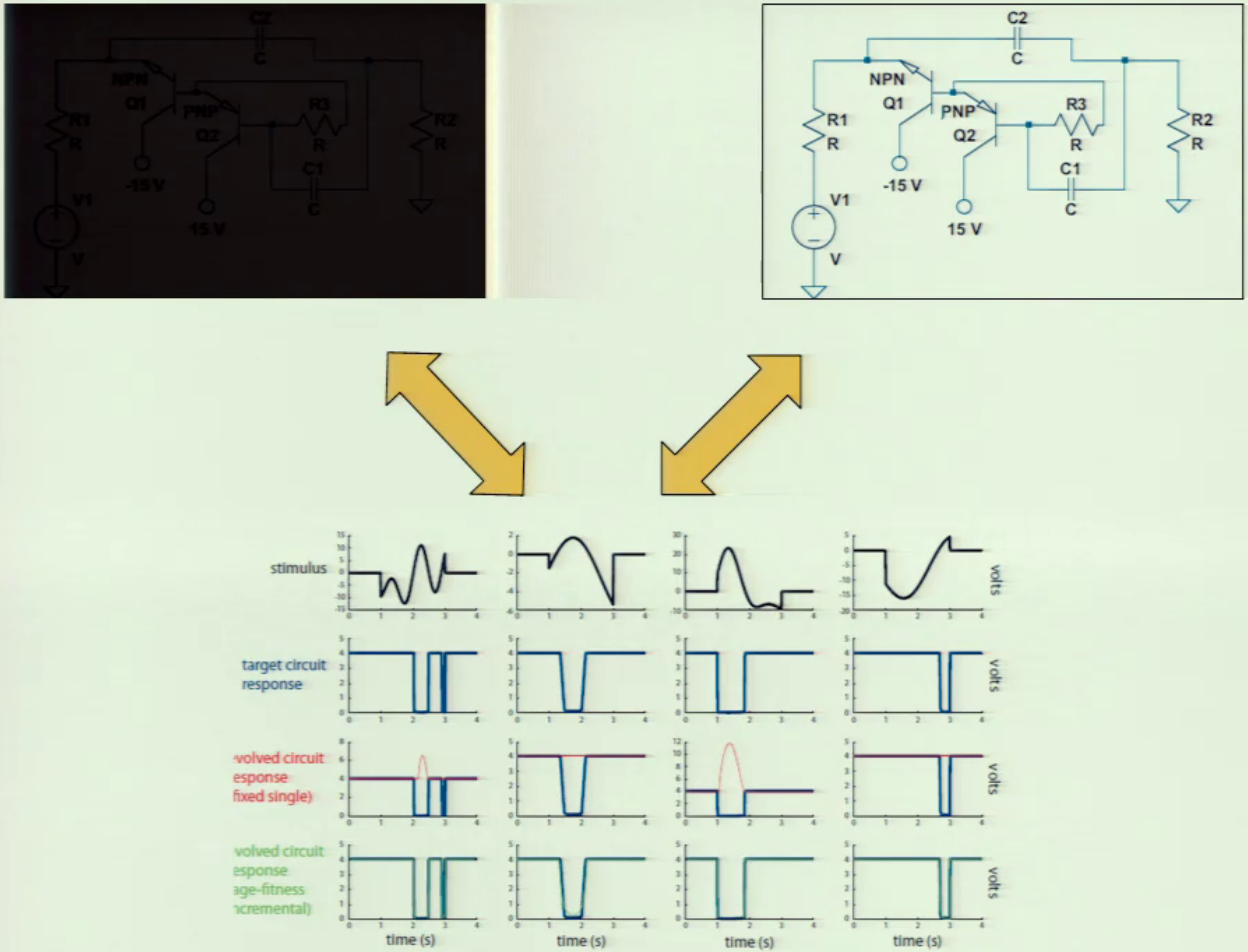

Candidate models
$\frac{dx}{dt} = -2y^2 + \log x$, $\frac{dy}{dt} = -x + \frac{y}{6}$
$\frac{dx}{dt} = -\sqrt{y} + \frac{x}{5}$, $\frac{dy}{dt} = -\sin y$
?
$\frac{dx}{dt} = -3\frac{y+1}{y-1}$, $\frac{dy}{dt} = -\frac{x^2}{x^2+1}$
$\frac{dx}{dt} = -y^{1.8} + \log x$, $\frac{dy}{dt} = -x + \frac{y}{4x}$
Candidate tests
Candidate Initial conditions
x
y
time
*
c
Inference Process
Outputs (sensors)
Initial Conditions (actuators)


Photo: Floris van Breugel

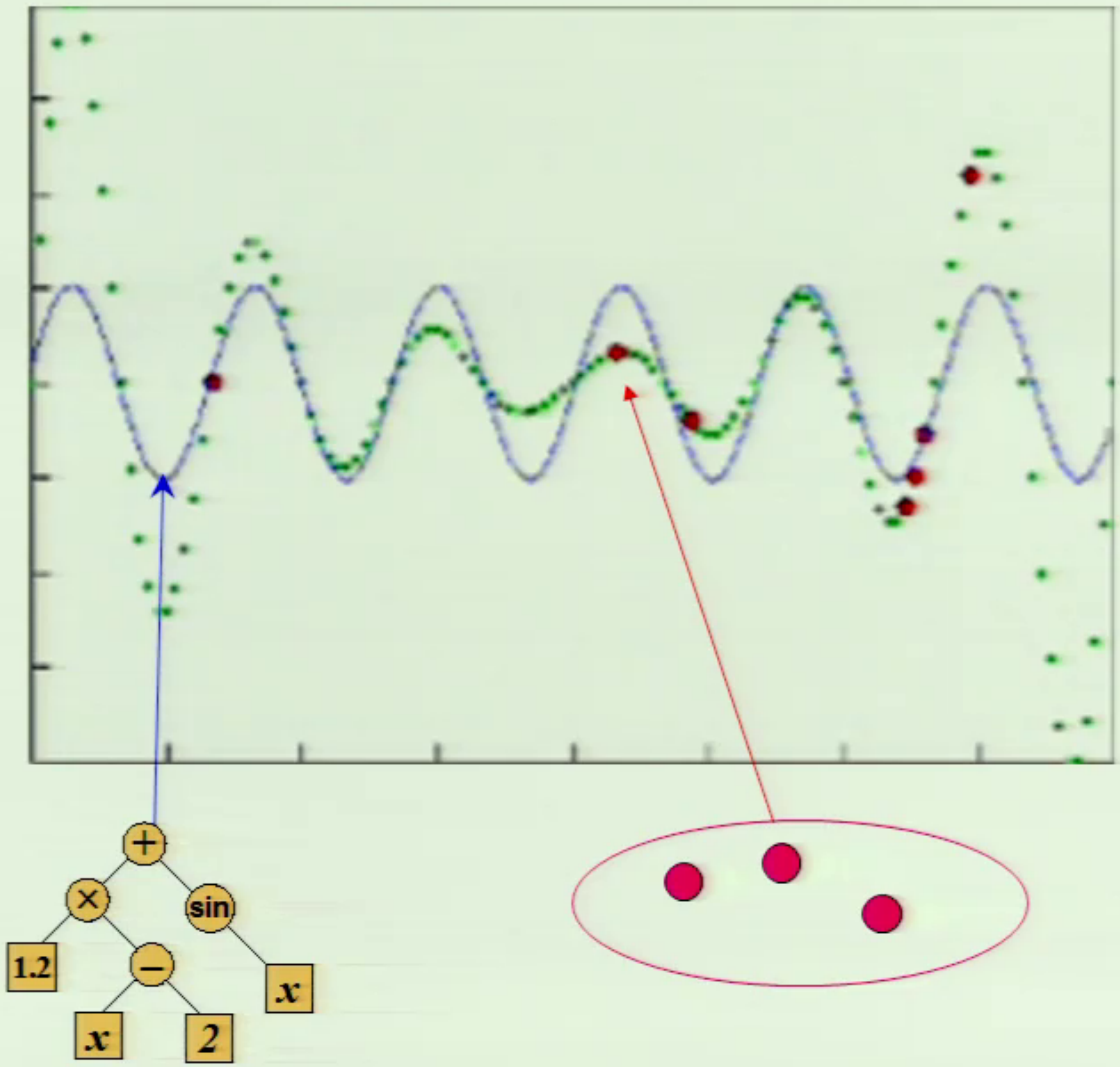
Structural Damage Diagnosis



With Wilkins Aquino



Ted Cornforth

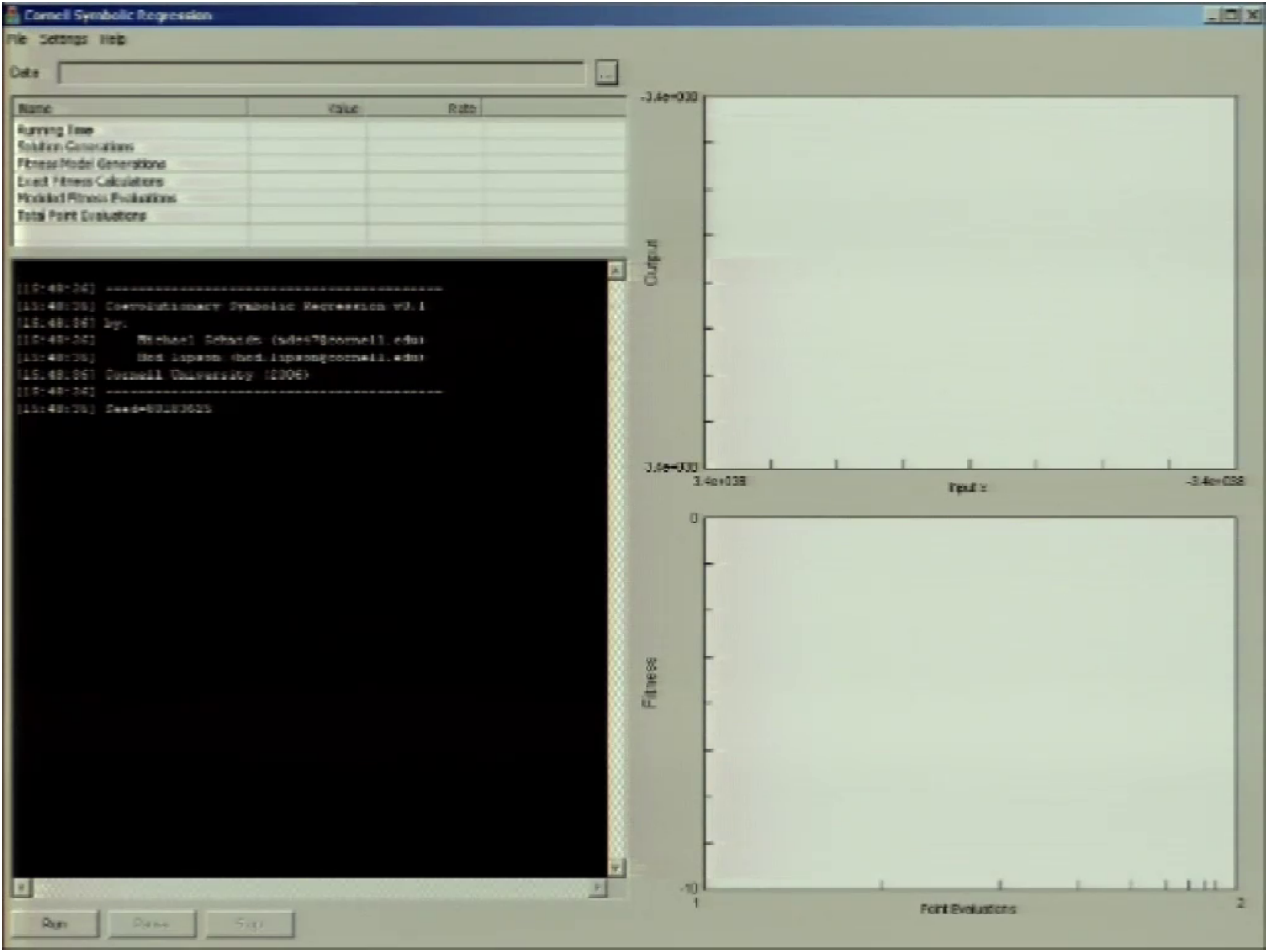


Models: Expression trees
Subject to mutation and selection

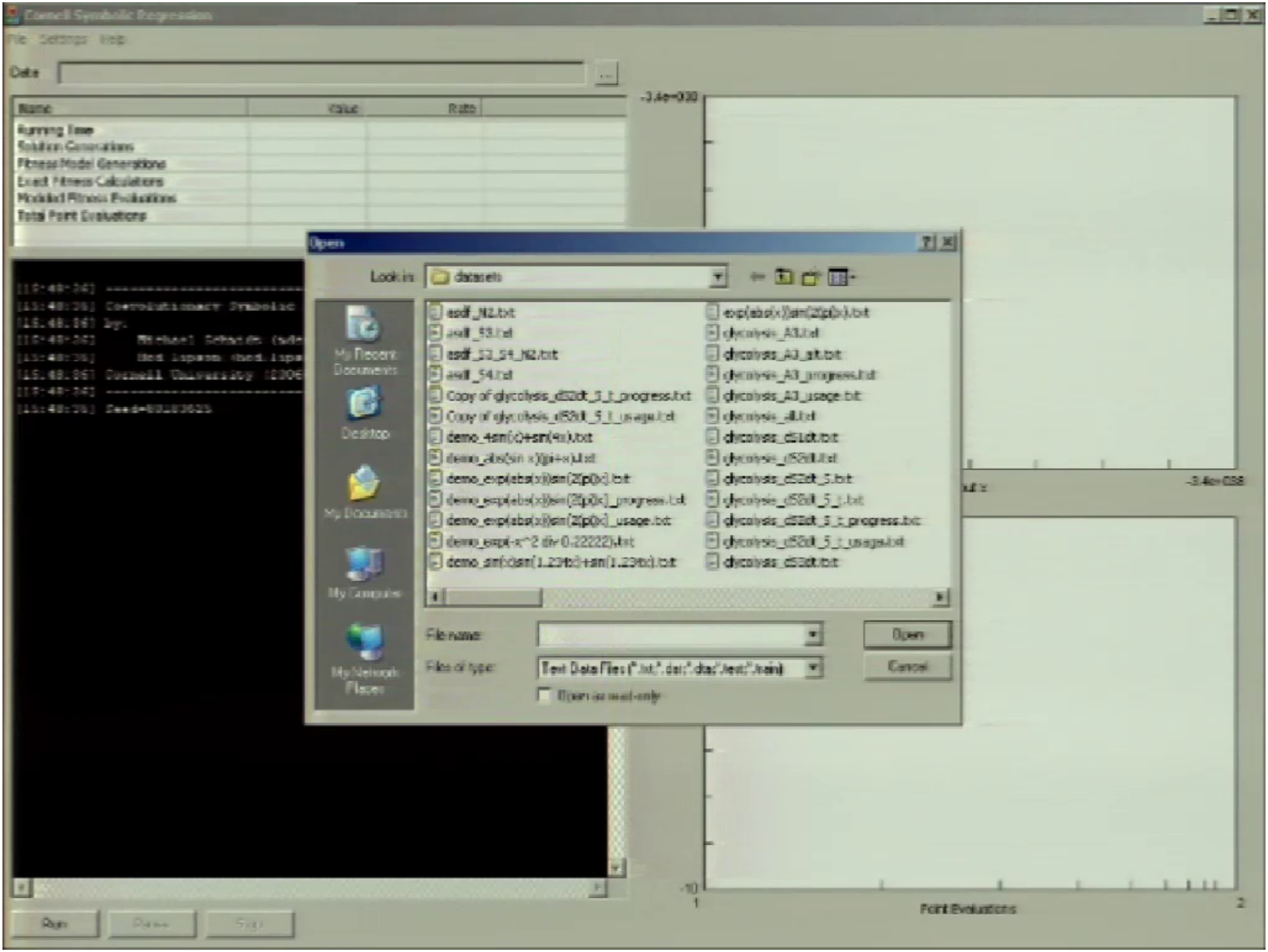
Experiments: Data-points
Subject to mutation and selection

`{const,+,-,*,/,sin,cos,exp,log,abs}`

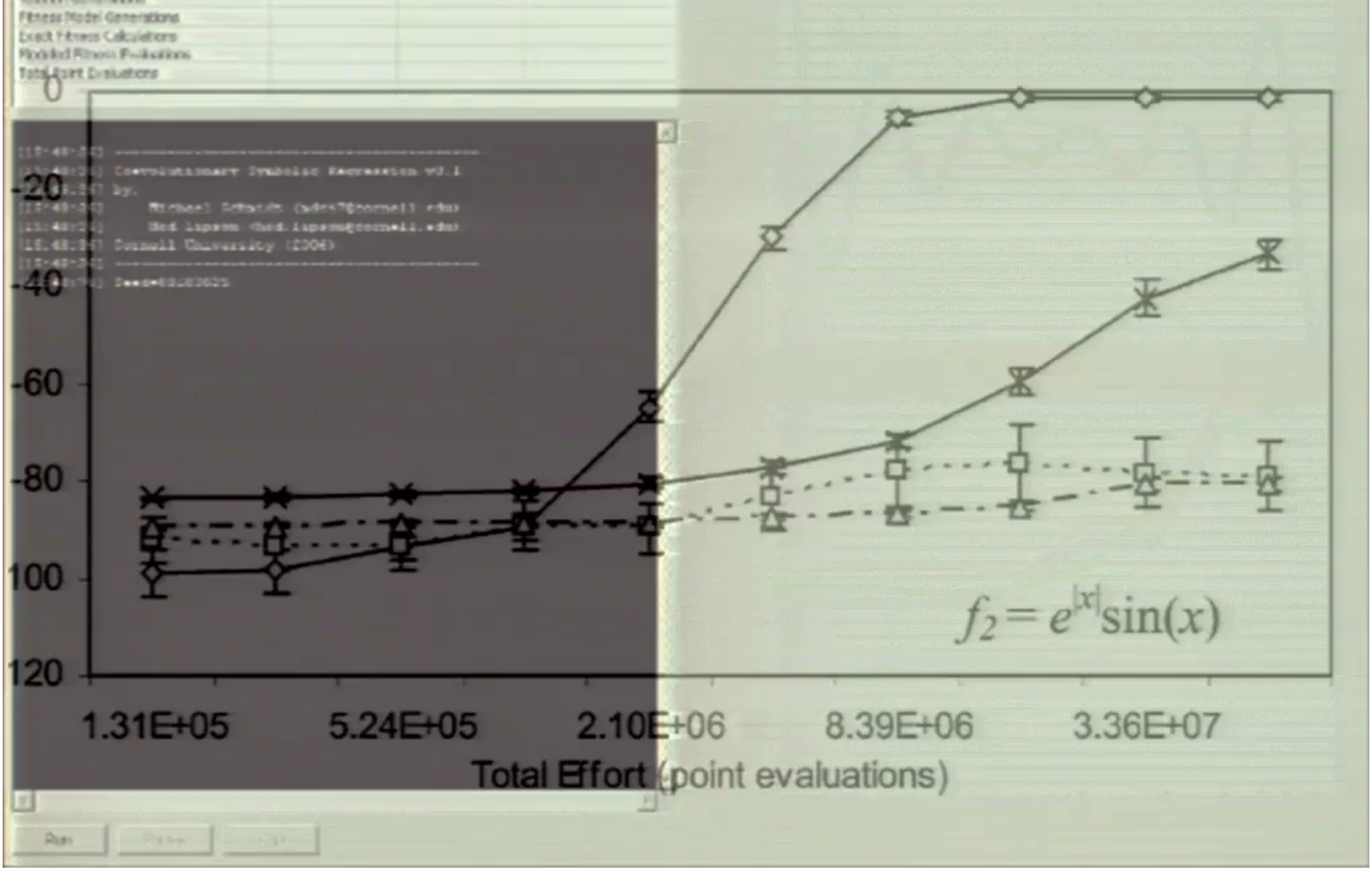
Michael D. Schmidt, Hod Lipson (2006)



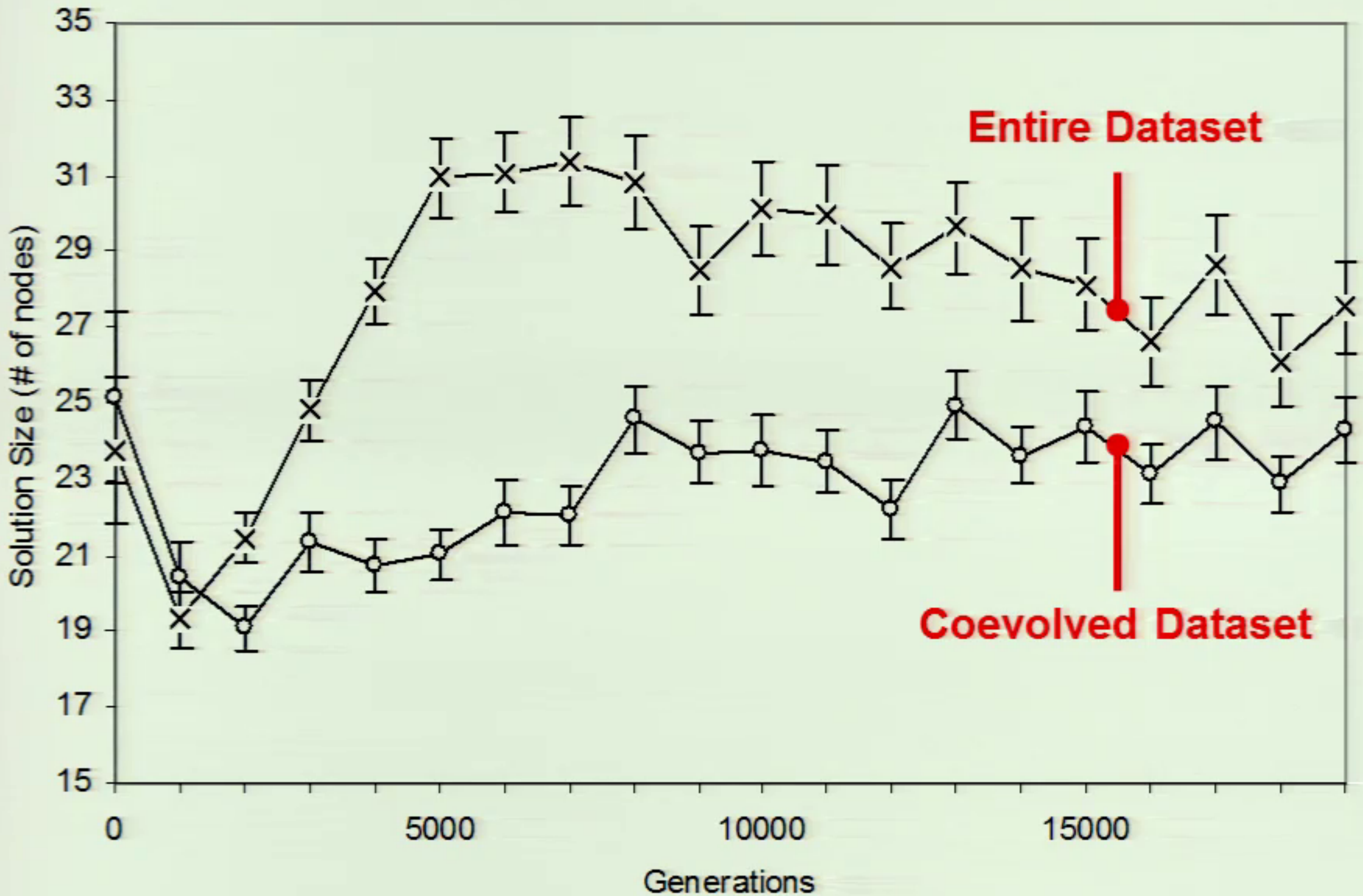
Cornell Symbolic Regression
File Settings Help
Data
Name
Value
Ratio
Running Time
Solution Generations
Fitness Model Generations
Exact Fitness Calculations
Modeled Fitness Evaluations
Total Point Evaluations
Output
Input x
Fitness
Point Evaluations
Run


Cornell Symbolic Regression
File Settings Help
Data
Name
Value
Rate
Running Time
Solution Generations
Fitness Model Generations
Exact Fitness Calculations
Modeled Fitness Evaluations
Total Point Evaluations
Open
Look in:
datasets
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
asdf_N2.txt
asdf_S3.txt
asdf_S4.txt
File name:
Files of type:
Open
Cancel
Run
Point Evaluations

Solution Accuracy



Solution Complexity



Inferring Biological Networks

$$\frac{dS_1}{dt} = \underbrace{2.5}_{J_1} \underbrace{-100\left(\frac{S_1 * A_3}{1+13.6769 * A_3^4}\right)}_{-v_1}$$

$$\frac{dS_2}{dt} = \underbrace{200\left(\frac{S_1 * A_3}{1+13.6769 * A_3^4}\right)}_{2*v_1} \underbrace{-6 * S_2 * N_1}_{-v_2} \underbrace{-12 * S_2 * N_1}_{-v_6}$$

$$\frac{dS_3}{dt} = \underbrace{6 * S_2 * N_1}_{v_2} + \underbrace{16 * S_3 * A_2}_{v_3}$$

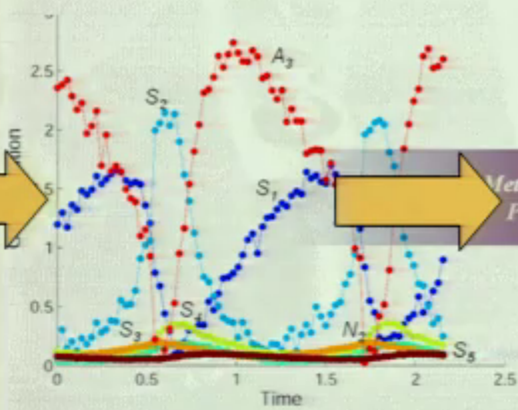
$$\frac{dS_4}{dt} = \underbrace{16 * A_2 * S_3}_{v_3} \underbrace{-100 * N_2 * S_4}_{-v_4}$$

$$\frac{dN_2}{dt} = \underbrace{6 * S_2 * N_1}_{v_2} \underbrace{-100 * N_2 * S_4}_{-v_4}$$

$$\frac{dA_3}{dt} = \underbrace{-200\left(\frac{S_1 * A_3}{1+13.6769 A_3^4}\right)}_{-2*v_1} + \underbrace{32 * A_2 * S_3}_{2v_3} \underbrace{-1.28 * A_3}_{-v_5}$$

$$\frac{dS_5}{dt} = \underbrace{-1.3 * S_5}_{\varphi J}$$

Original Equations



$$\frac{dS_1}{dt} = \underbrace{2.42114}_{J_0} \underbrace{-99.2721\left(\frac{S_1 * A_3}{1+13.5956 * A_3^4}\right)}_{-v_1}$$

$$\frac{dS_2}{dt} = \underbrace{199.935\left(\frac{S_1 * A_3}{1+13.6734 * A_3^4}\right)}_{2*v_1} \underbrace{-5.99475 * S_2 * N_1}_{-v_2} \underbrace{-11.9895 * S_2 * N_1}_{-v_6}$$

$$\frac{dS_3}{dt} = \underbrace{5.99857 * S_2 * N_1}_{v_2} + \underbrace{15.99606 * S_3 * A_2}_{v_3} \underbrace{-0.01286 * S_3}_{extraneous}$$

$$\frac{dS_4}{dt} = \underbrace{15.997 * A_2 * S_3}_{v_3} \underbrace{-100.015 * N_2 * S_4}_{-v_4}$$

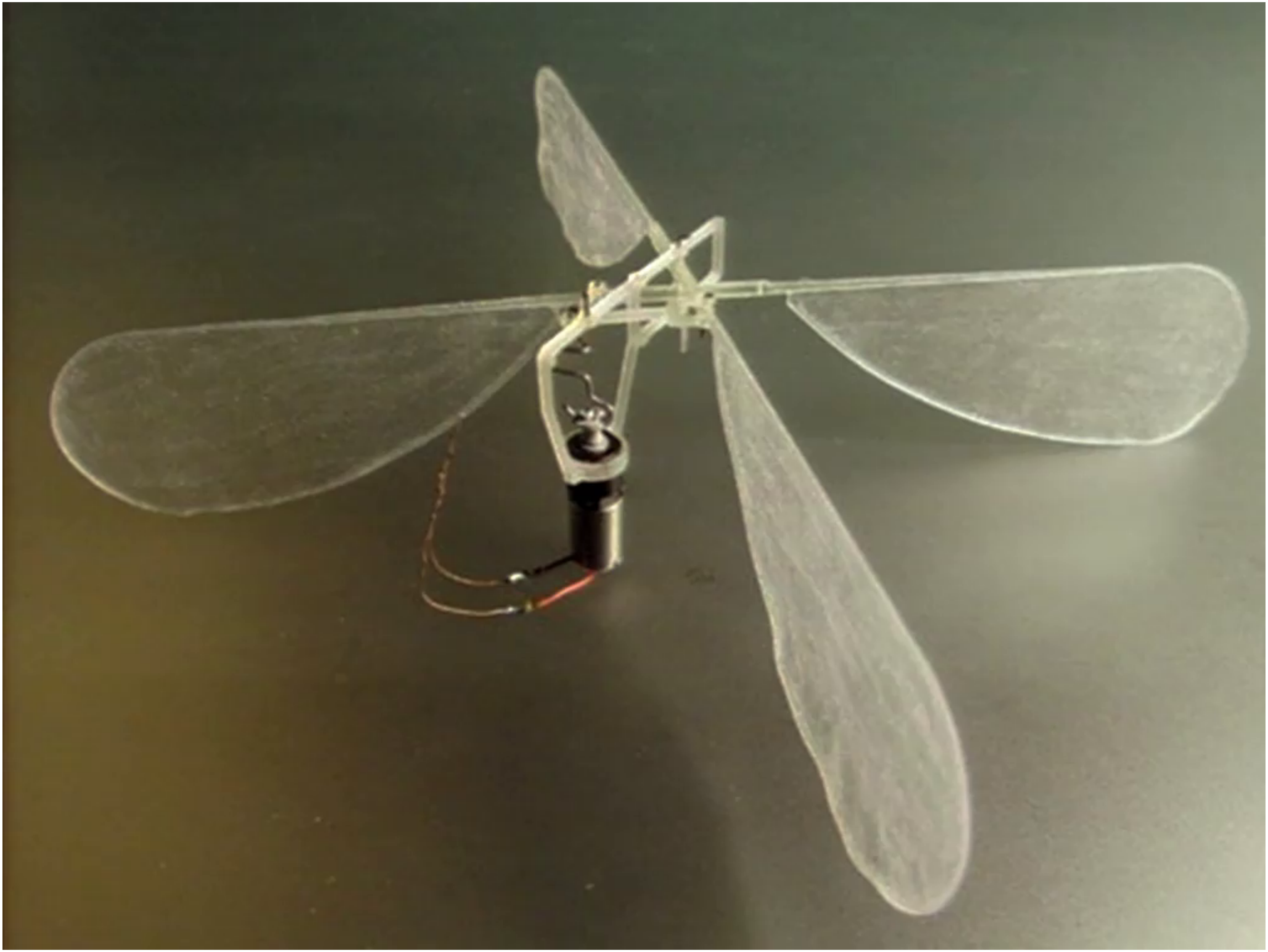
$$\frac{dN_2}{dt} = \underbrace{5.99857 * S_2 * N_1}_{v_2} \underbrace{-99.9963 * N_2 * S_4}_{-v_4}$$

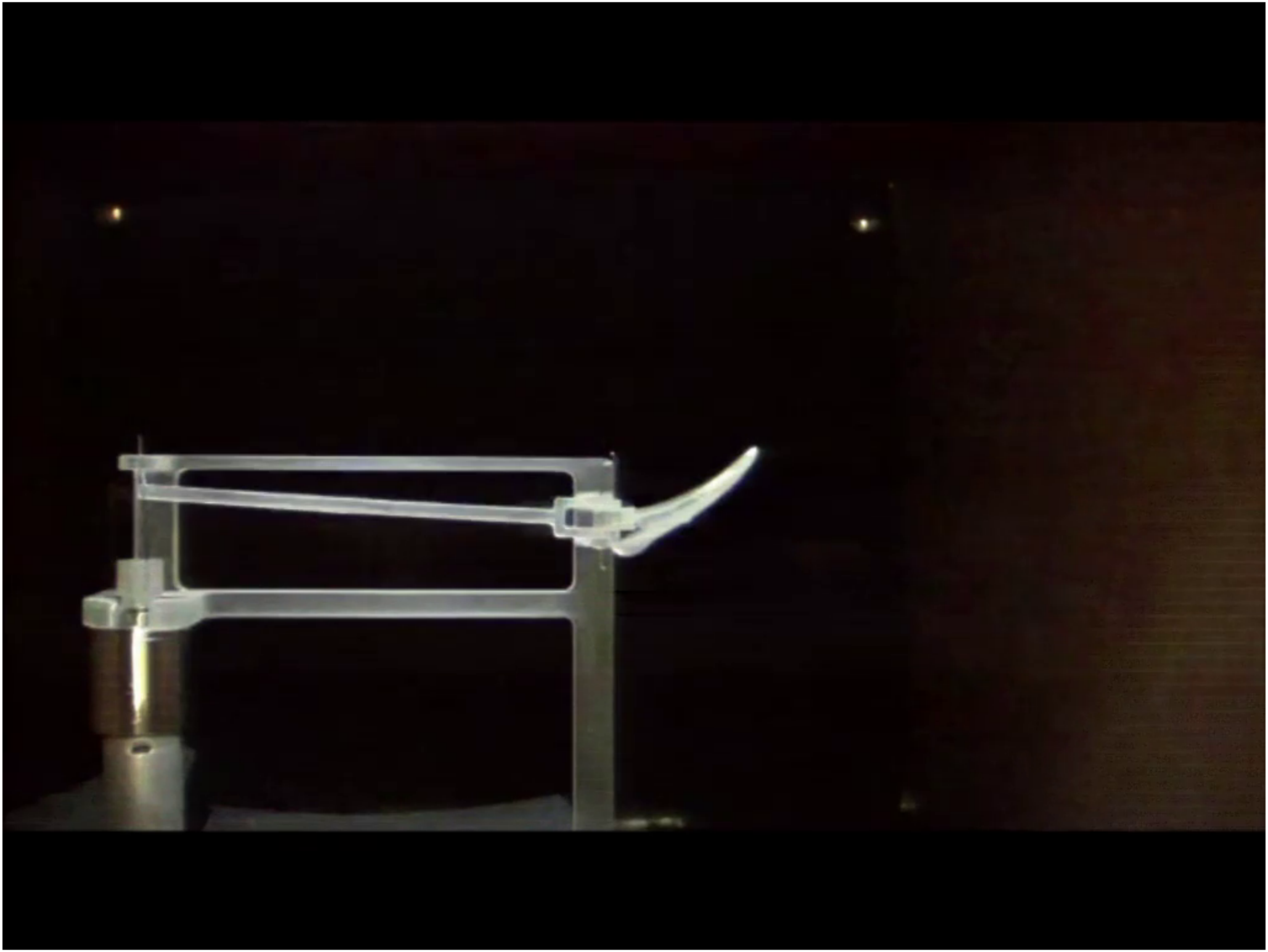
$$\frac{dA_3}{dt} = \underbrace{-197.781\left(\frac{S_1 * A_3}{1+13.2633 A_3^4}\right)}_{-2*v_1} + \underbrace{31.9682 * A_2 * S_3}_{2v_3} \underbrace{-1.29659 * A_3}_{-v_5}$$

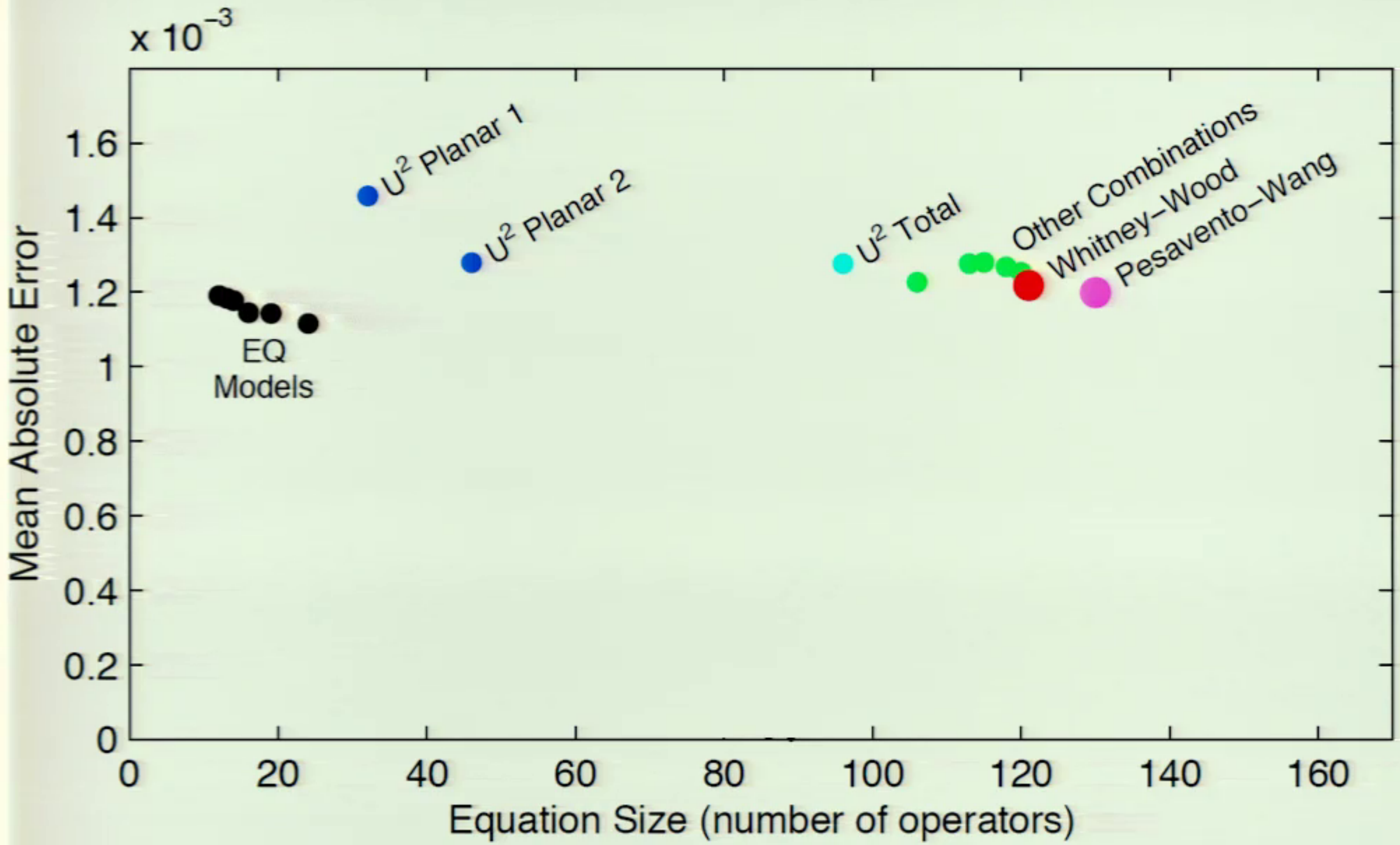
$$\frac{dS_5}{dt} = \underbrace{-1.29626 * S_5}_{\varphi J}$$

Inferred Equations

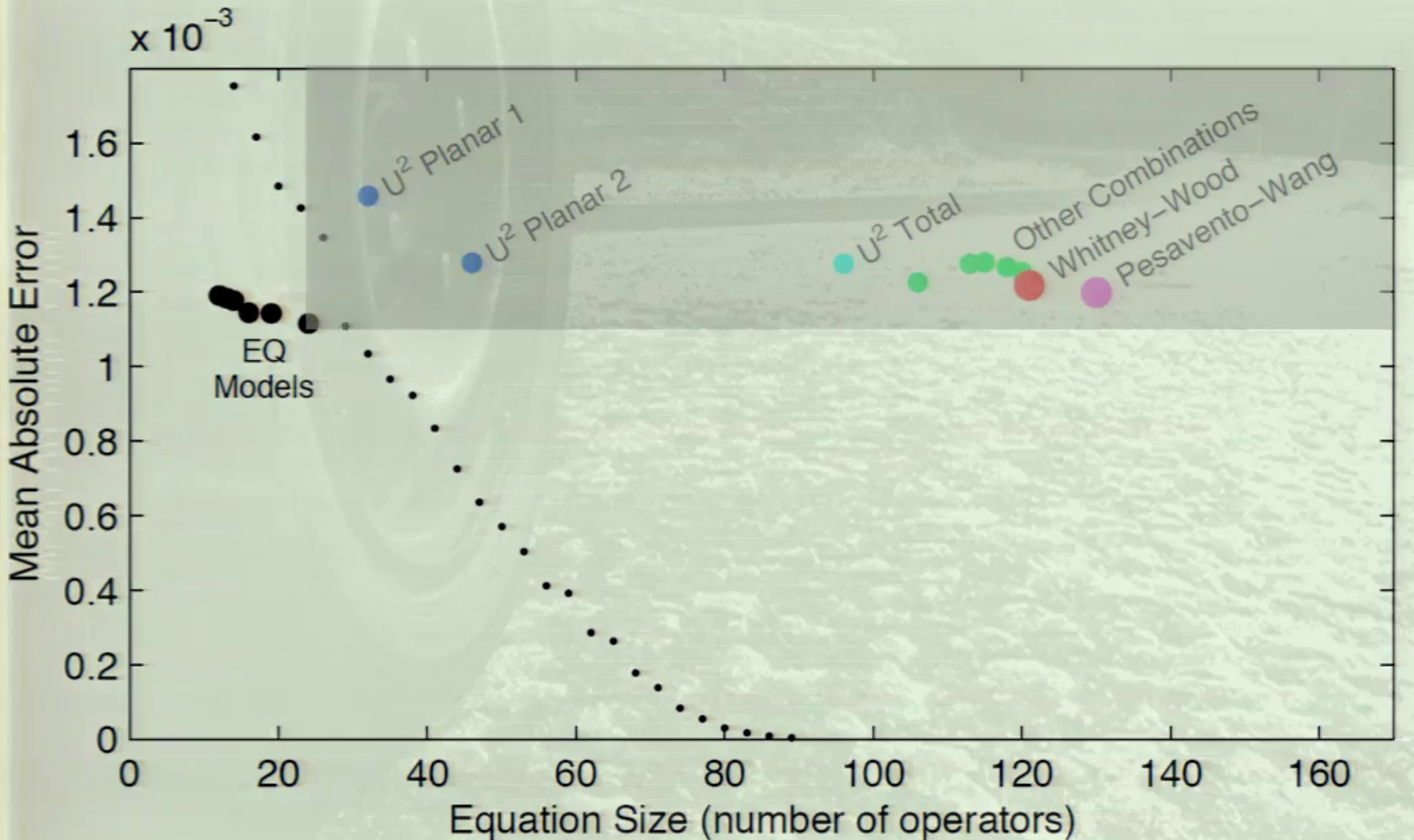
With Michael Schmidt, John Wikswo (Vanderbilt), Jerry Jenkins (CFDRC)

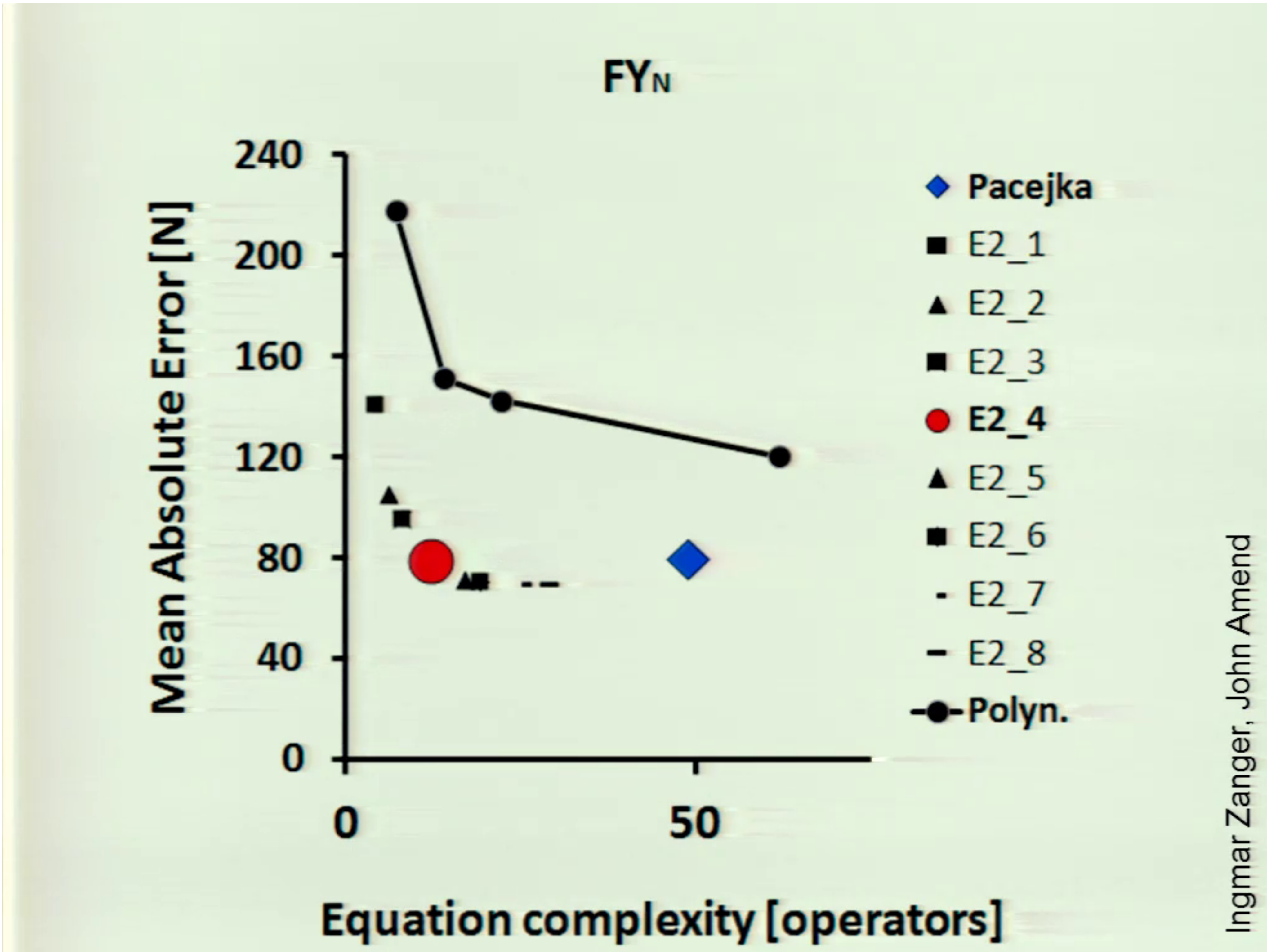






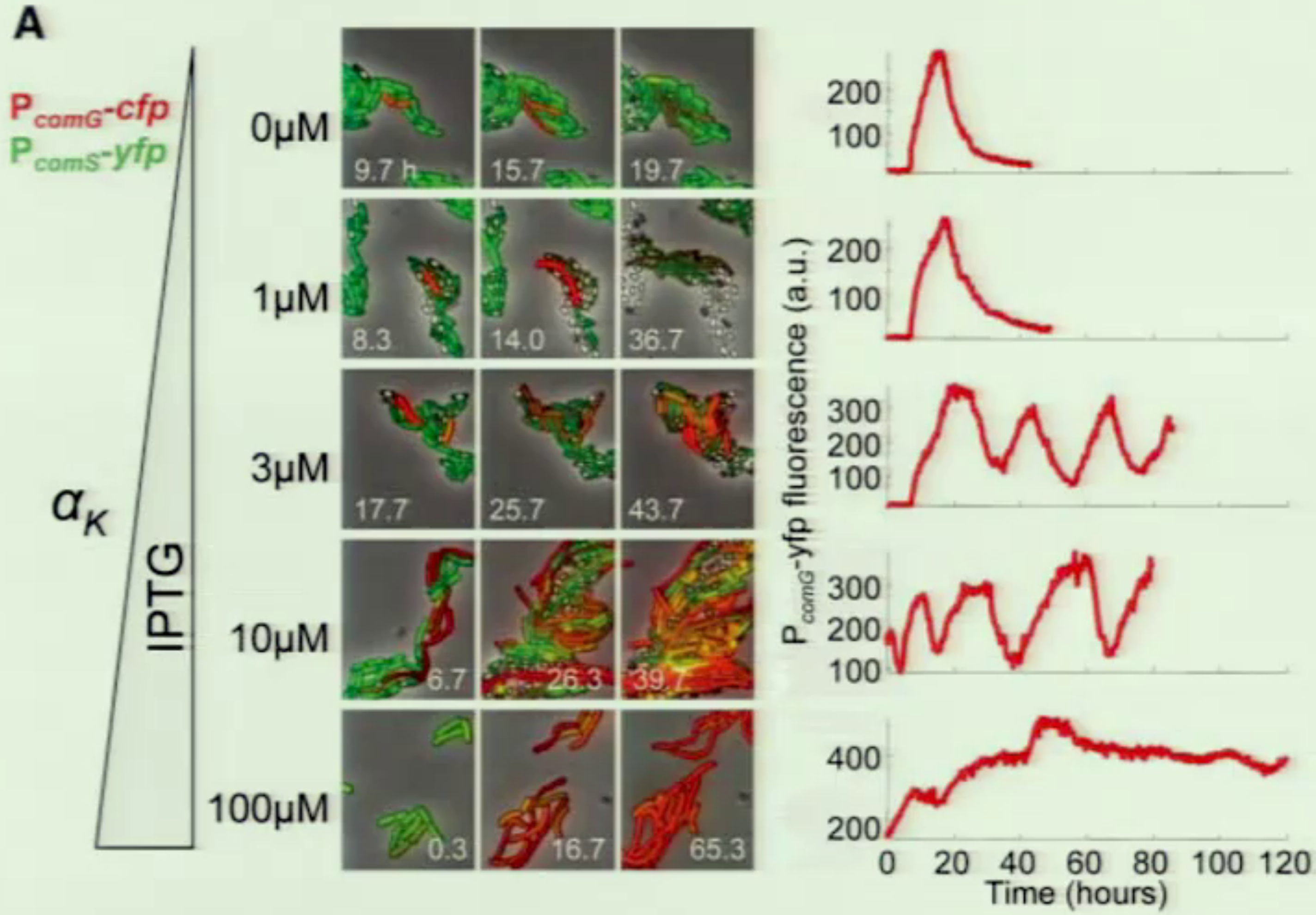
Charles Richter


x 10^-3
Mean Absolute Error
Equation Size (number of operators)
U^2 Planar 1
U^2 Planar 2
U^2 Total
Other Combinations
Whitney–Wood
Pesavento–Wang
EQ Models
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
0
20
40
60
80
100
120
140
160
Charles Richter



FY_N
Mean Absolute Error [N]
240
200
160
120
80
40
0
0
50
Equation complexity [operators]
Pacejka
E2_1
E2_2
E2_3
E2_4
E2_5
E2_6
E2_7
E2_8
Polyn.
Ingmar Zanger, John Amend

Wet Data, Unknown System



With Michael Schmidt (Cornell) and Gurol Suel (UT Southwestern)

Symbolic Regression Inferred *Time-Delay* Model:

$$\frac{dK}{dt} = a_K + \frac{b_K + c_K S}{K}$$

$$\frac{dS}{dt} = a_S + \frac{b_S + c_S K}{S}$$

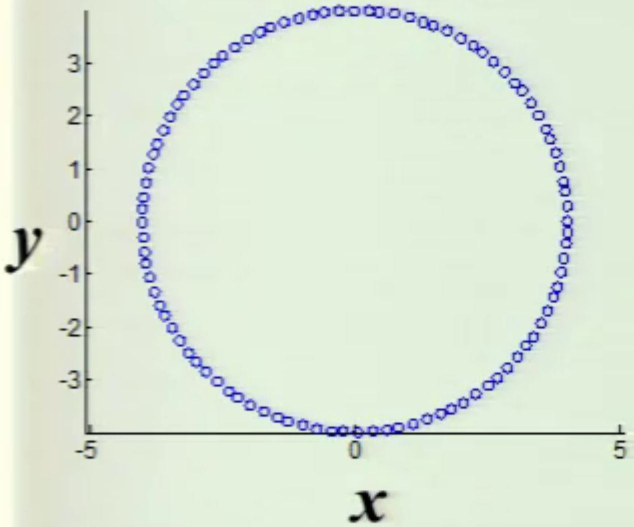
Searching for meaning


Searching for meaning

Correlations

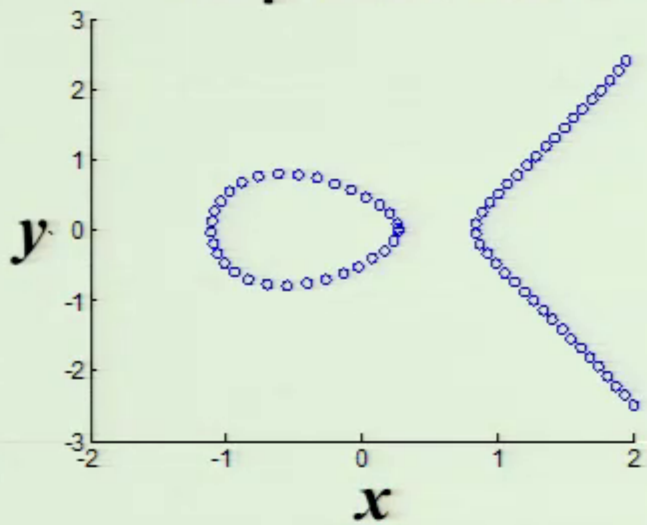
Homework

Circle



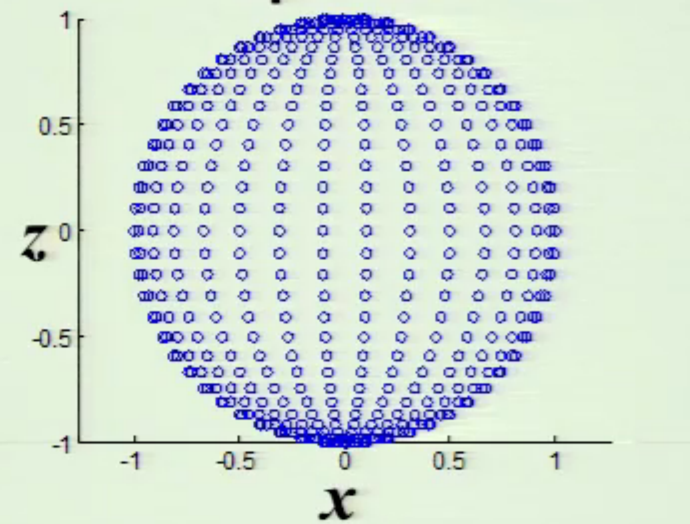
$$x^2 + y^2 - 16 = 0$$

Elliptic Curve

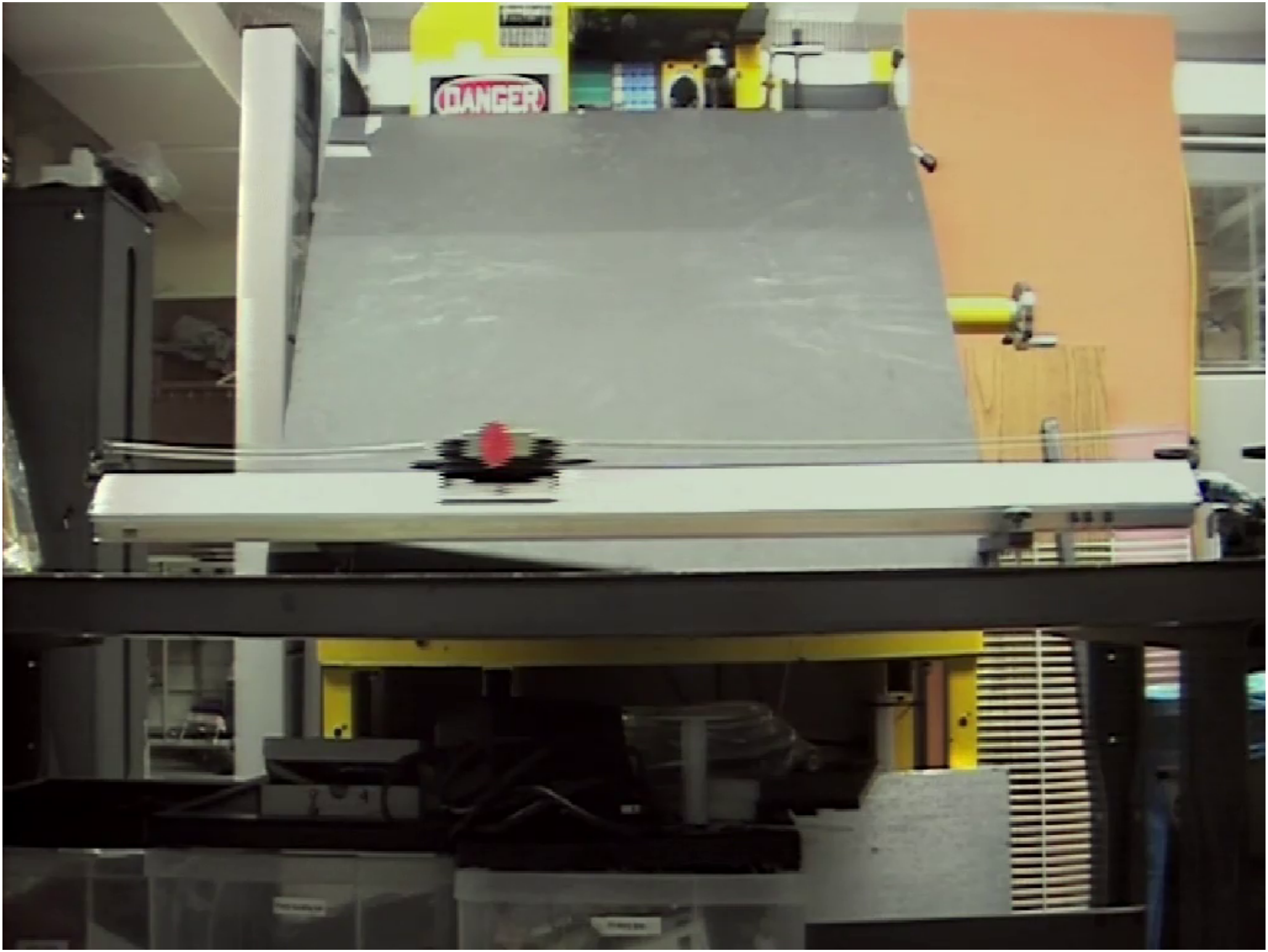


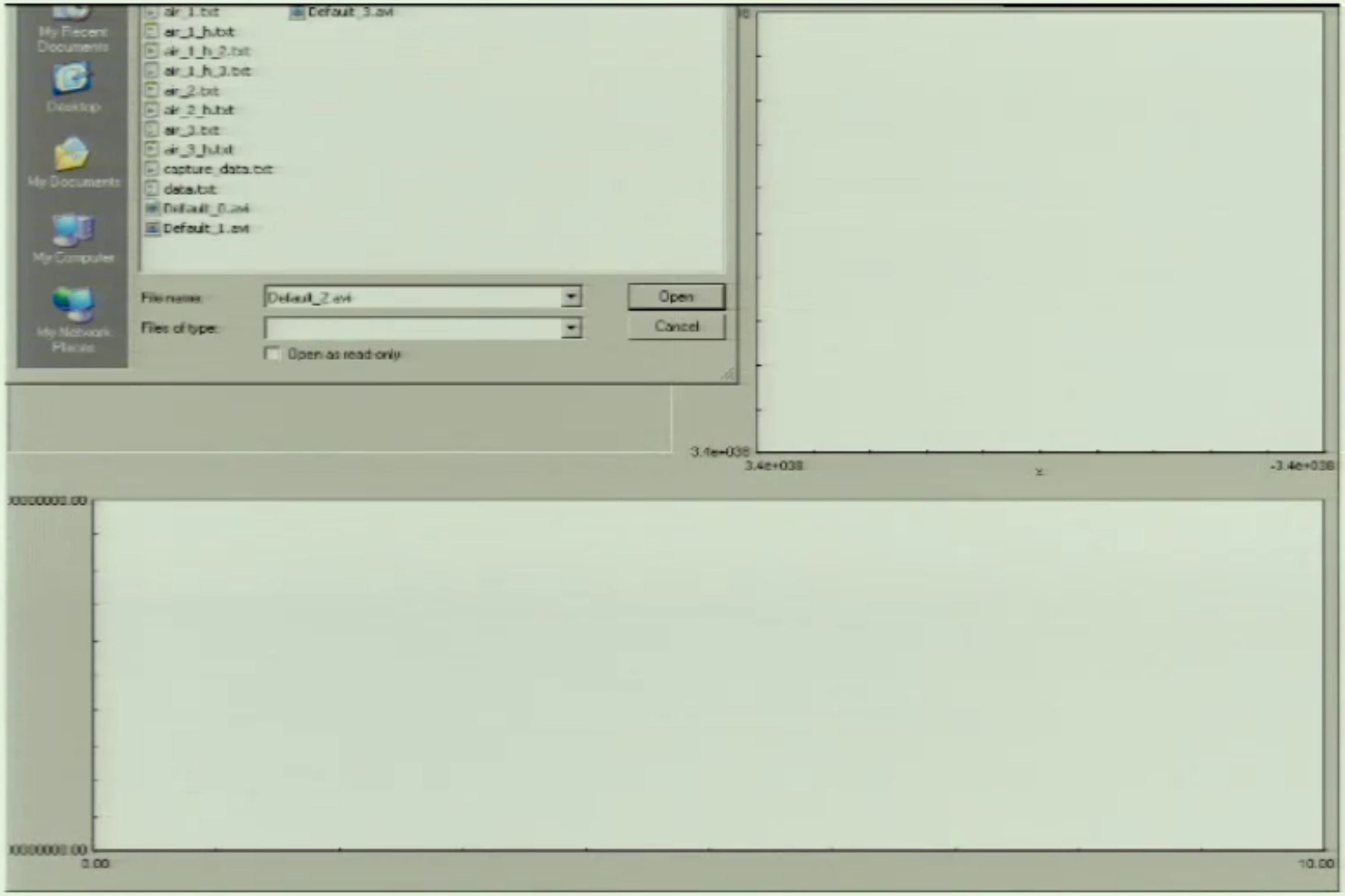
$$x^3 + x - y^2 - 1.5 = 0$$

Sphere

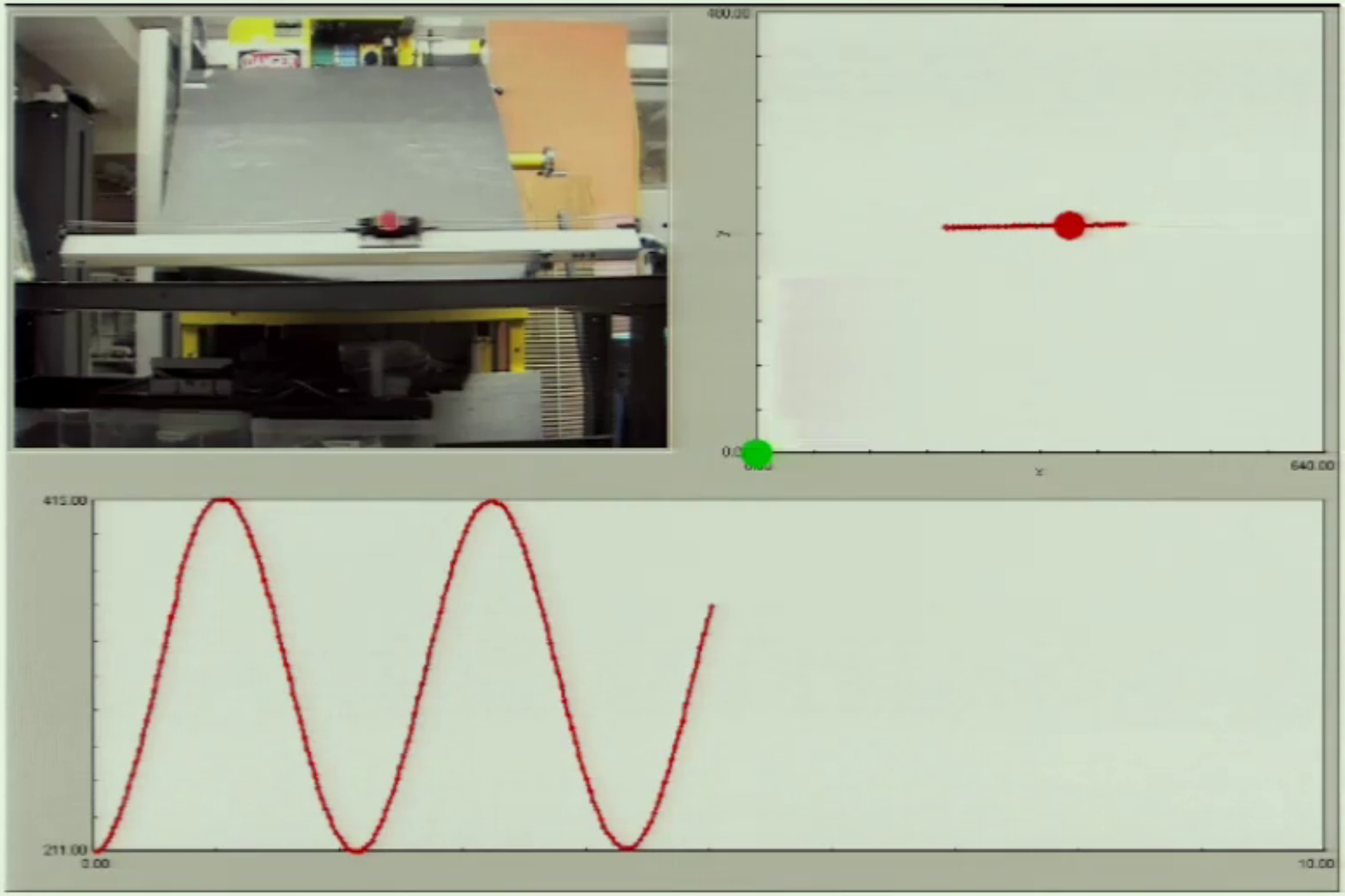


$$x^2 + y^2 + z^2 - 1 = 0$$


DANGER



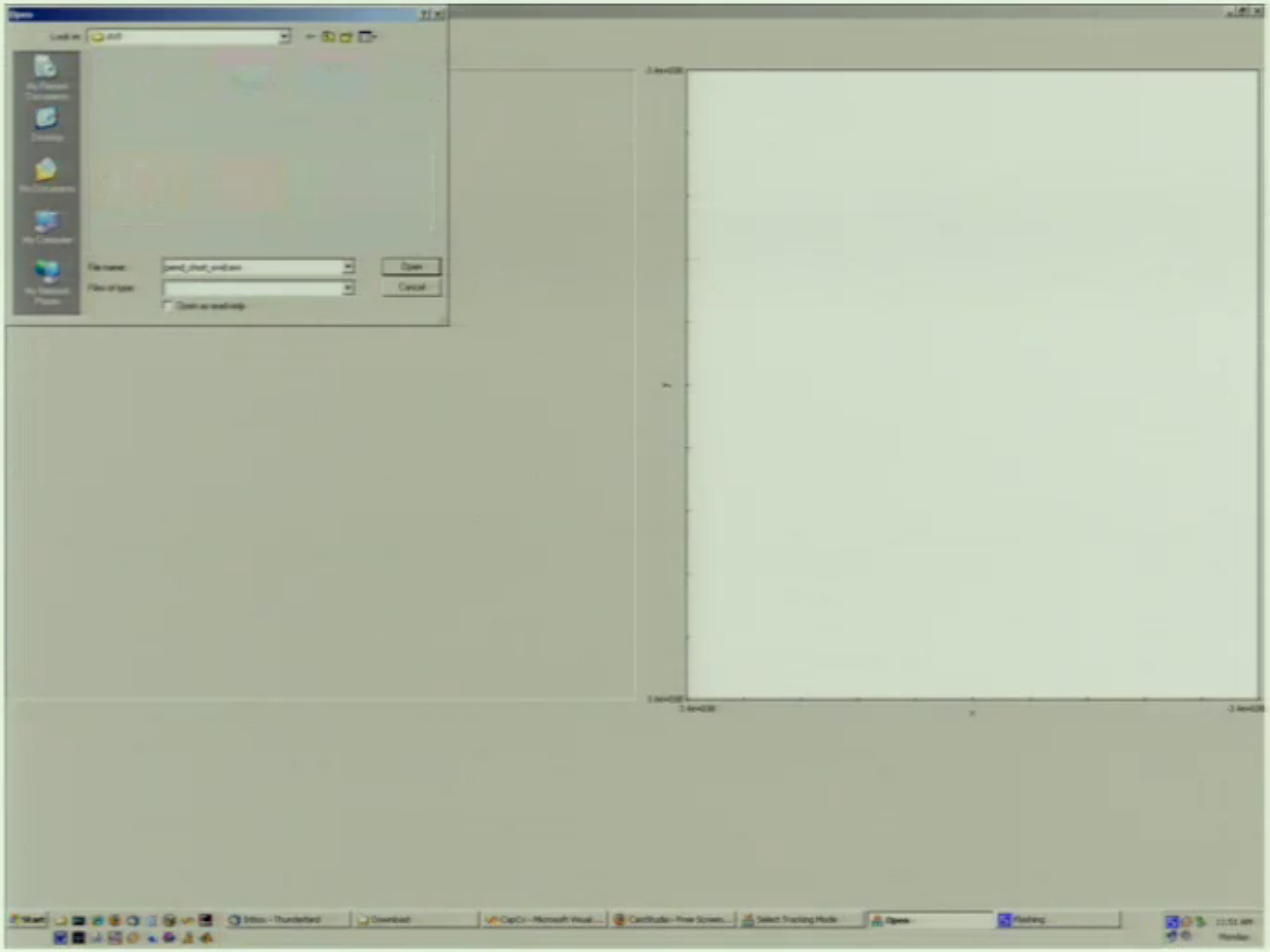
air_1.txt
Default_3.avi
air_1_h.txt
air_1_h_2.txt
air_1_h_3.txt
air_2.txt
air_2_h.txt
air_3.txt
air_3_h.txt
capture_data.txt
data.txt
Default_0.avi
Default_1.avi
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
File name:
Default_2.avi
Files of type:
Open as read only
Open
Cancel
3.4e+038
x
-3.4e+038
0.00
10.00



$$H = 114.28 * \left(\frac{dx}{dt}\right)^2 + 692.322 * x^2$$

$$L = 61.591 * \left(\frac{dx}{dt}\right)^2 - 369.495 * x^2$$

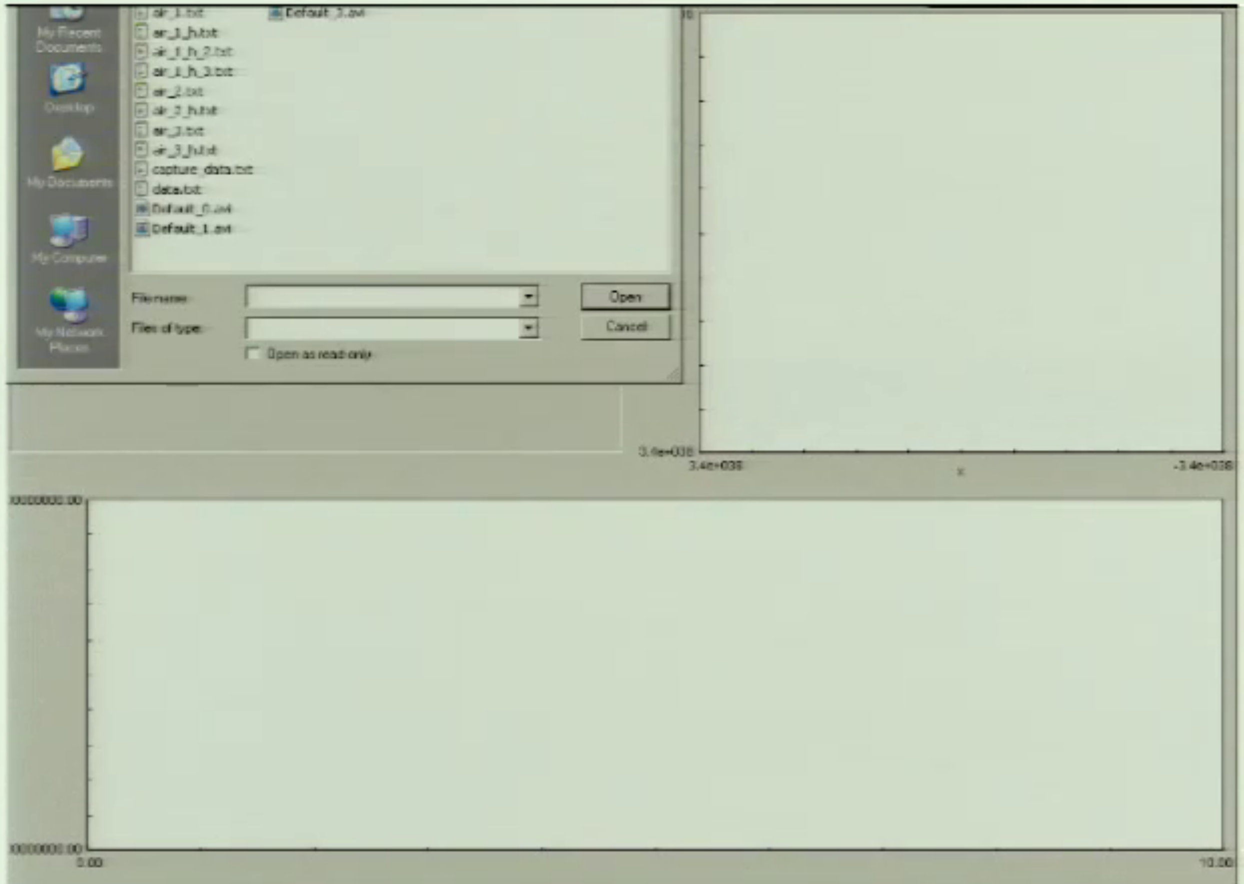
- Coefficients may have different scales and offsets each run



$$\mathbf{H} = \left(\frac{d\theta}{dt}\right)^2 + 2.42847 * \cos(\theta)$$

$$\mathbf{L} = 3.52768 * \left(\frac{d\theta}{dt}\right)^2 - 9.43429 * \cos(\theta)$$

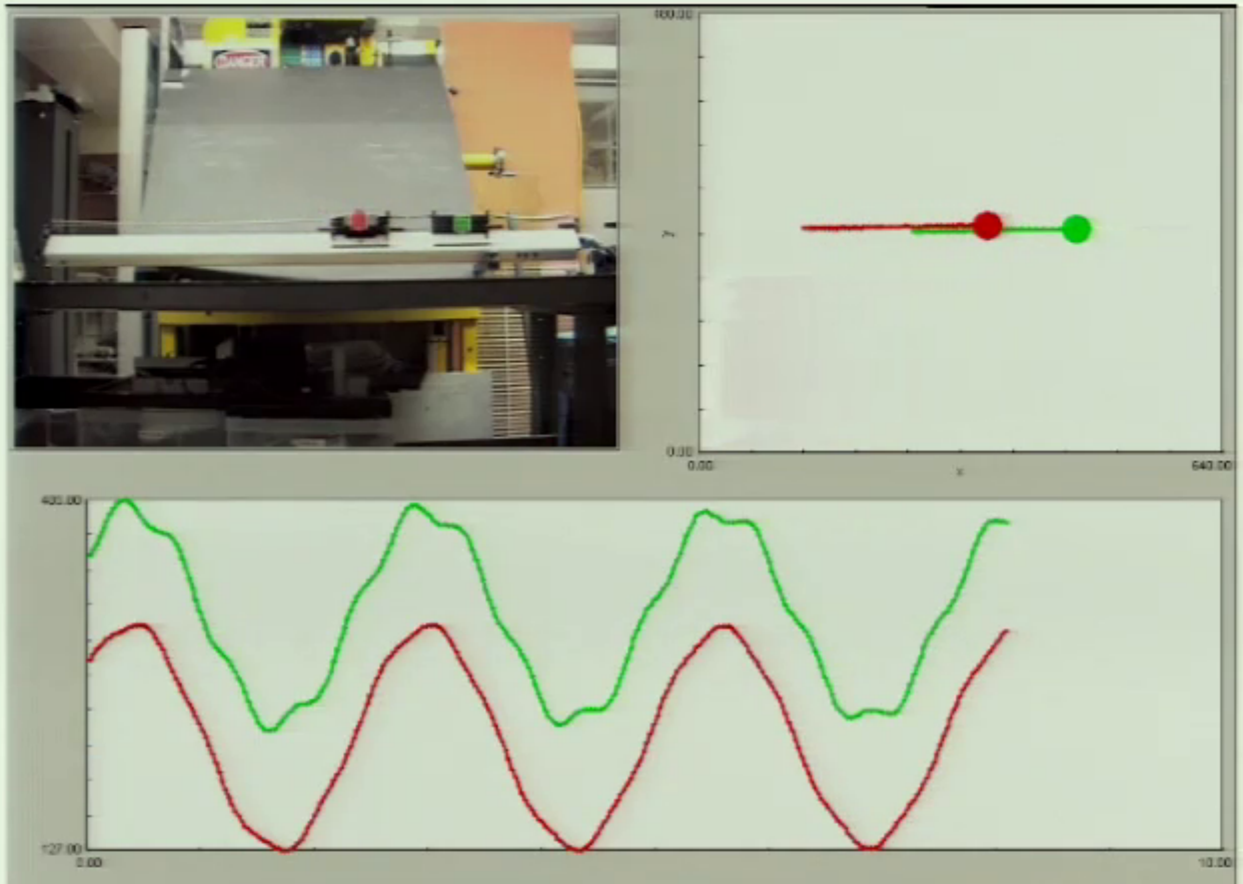
Double Linear Oscillator



$$H = -14.691 * x_1^2 - 15.551 * x_2^2 - 21.676 * x_1 x_2 + 8.3808 * \left(\frac{dx_2}{dt}\right)^2 + 2.6046 * \left(\frac{dx_1}{dt}\right)^2$$

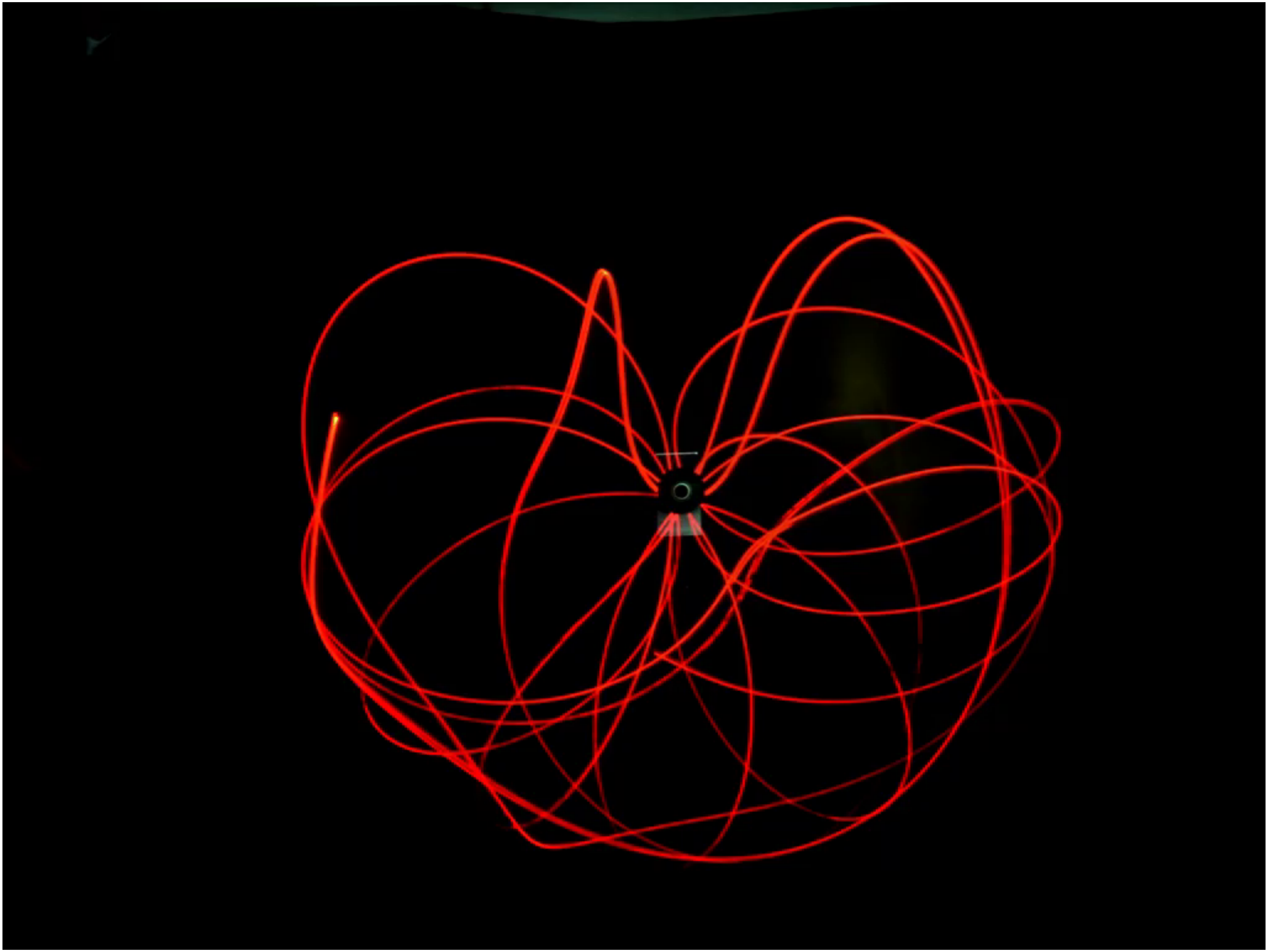
would be plus for Lagrangian

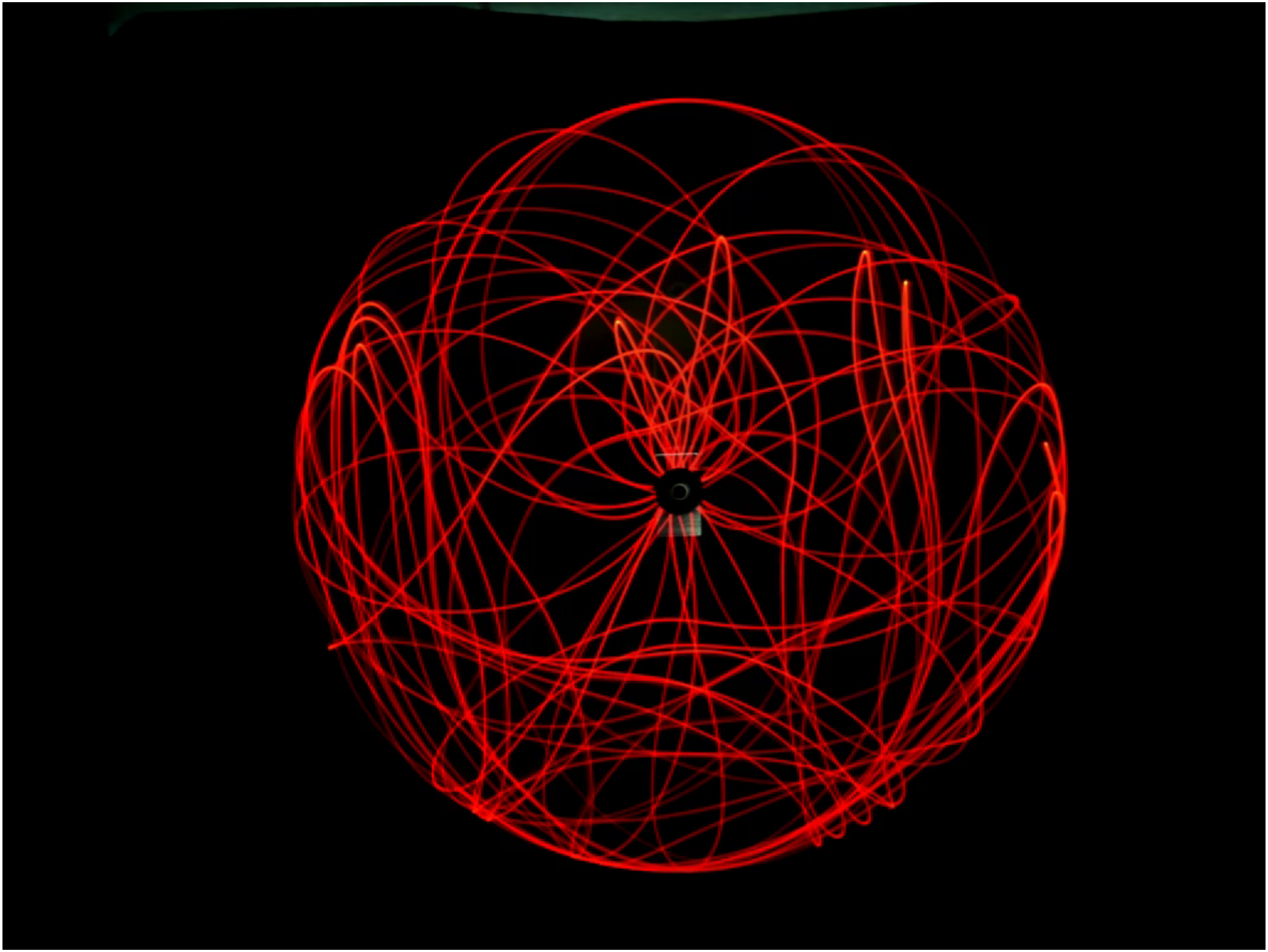
Double Linear Oscillator



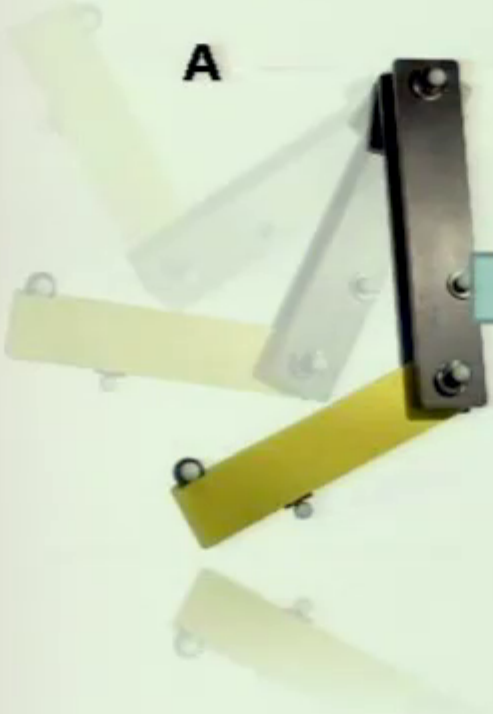
$$H = -14.691 * x_1^2 - 15.551 * x_2^2 - 21.676 * x_1 x_2 + 8.3808 * \left(\frac{dx_2}{dt}\right)^2 + 2.6046 * \left(\frac{dx_1}{dt}\right)^2$$

would be plus for Lagrangian

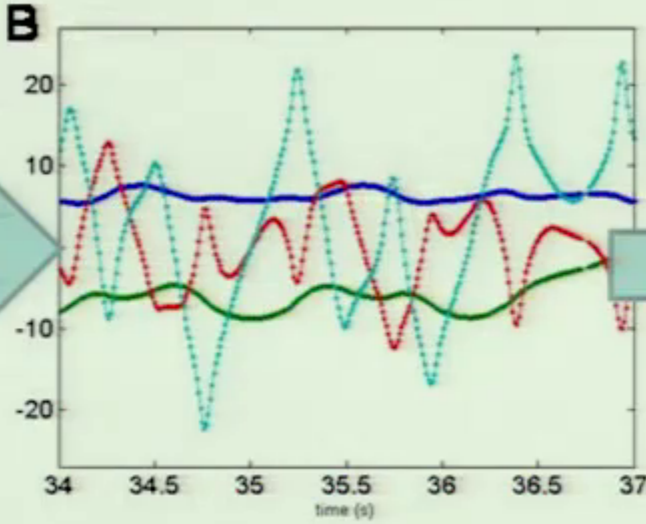




A



B



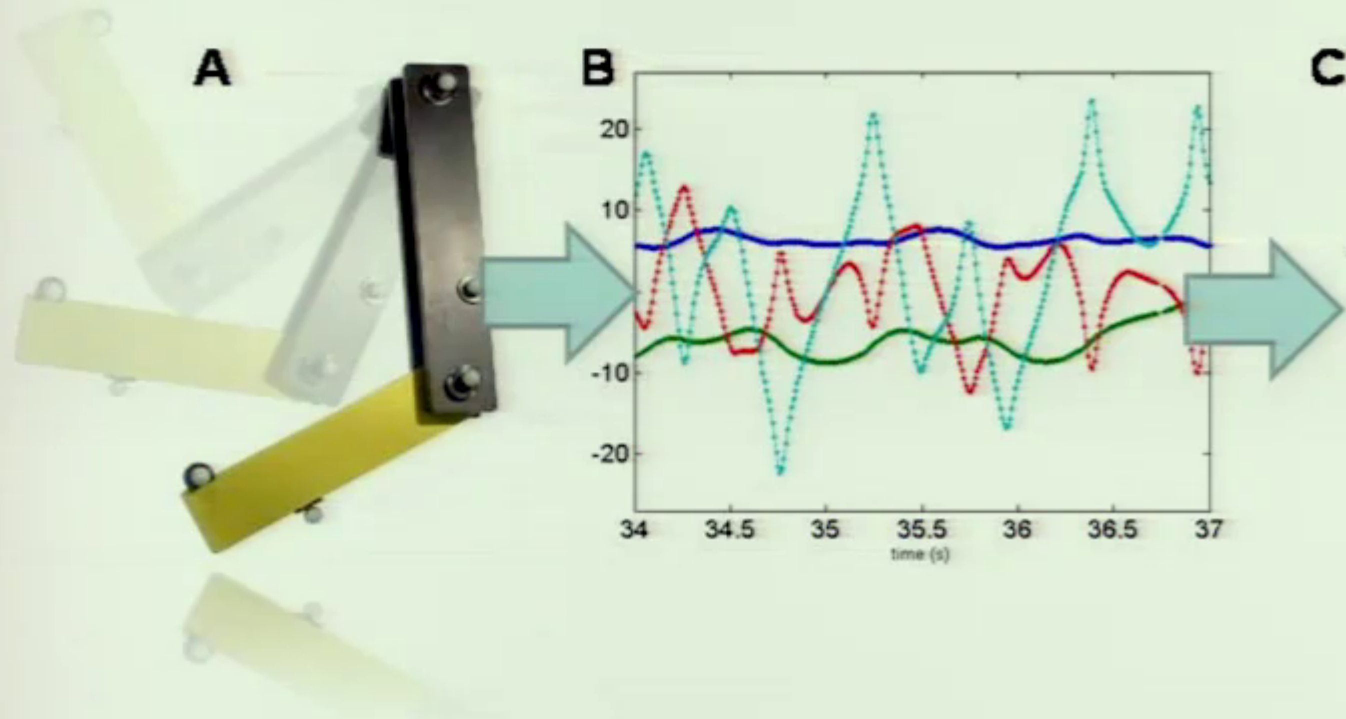
C

Detected Invariance:

$$L_1^2(m_1+m_2)\omega_1^2 + m_2L_2^2\omega_2^2 + m_2L_1L_2\omega_1\omega_2\cos(\theta_1-\theta_2) - 19.6L_1(m_1+m_2)\cos\theta_1 - 19.6m_2L_2\cos\theta_2$$

A

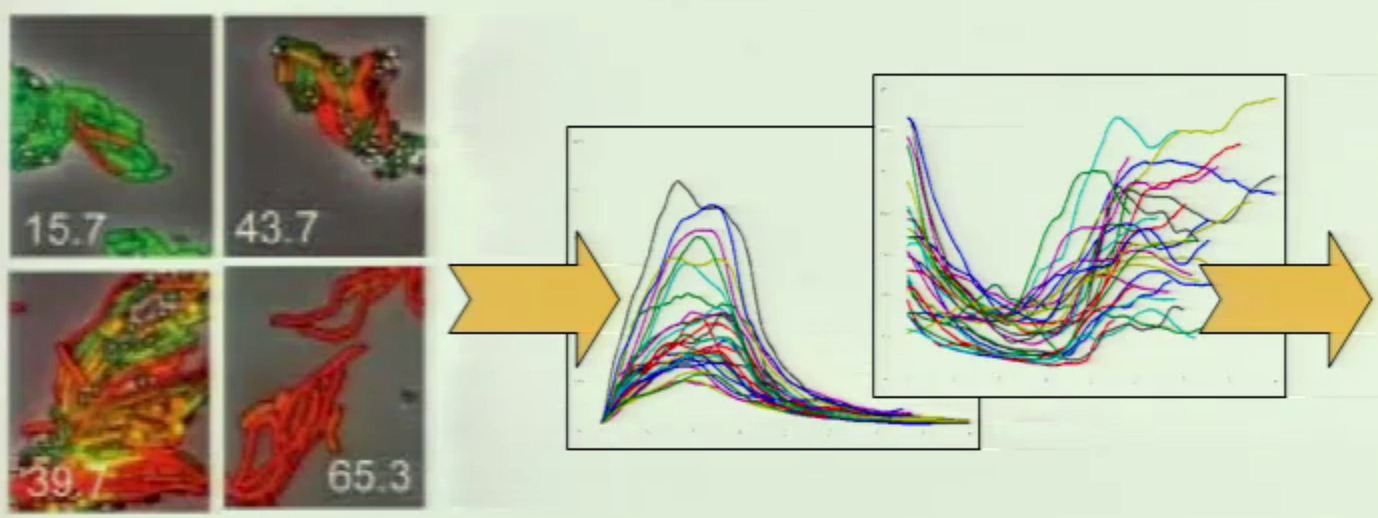
B



C

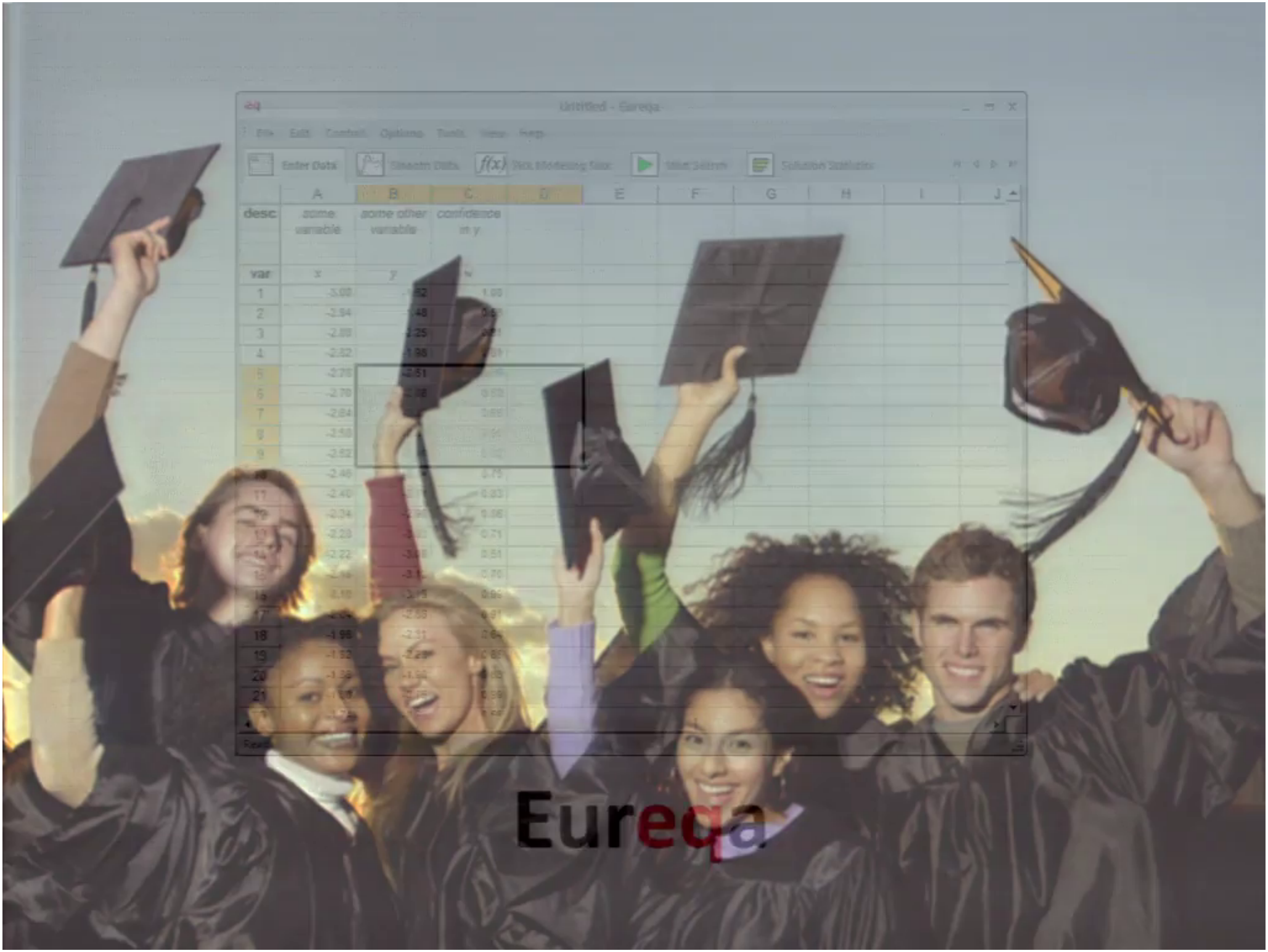
Detected Invariance:

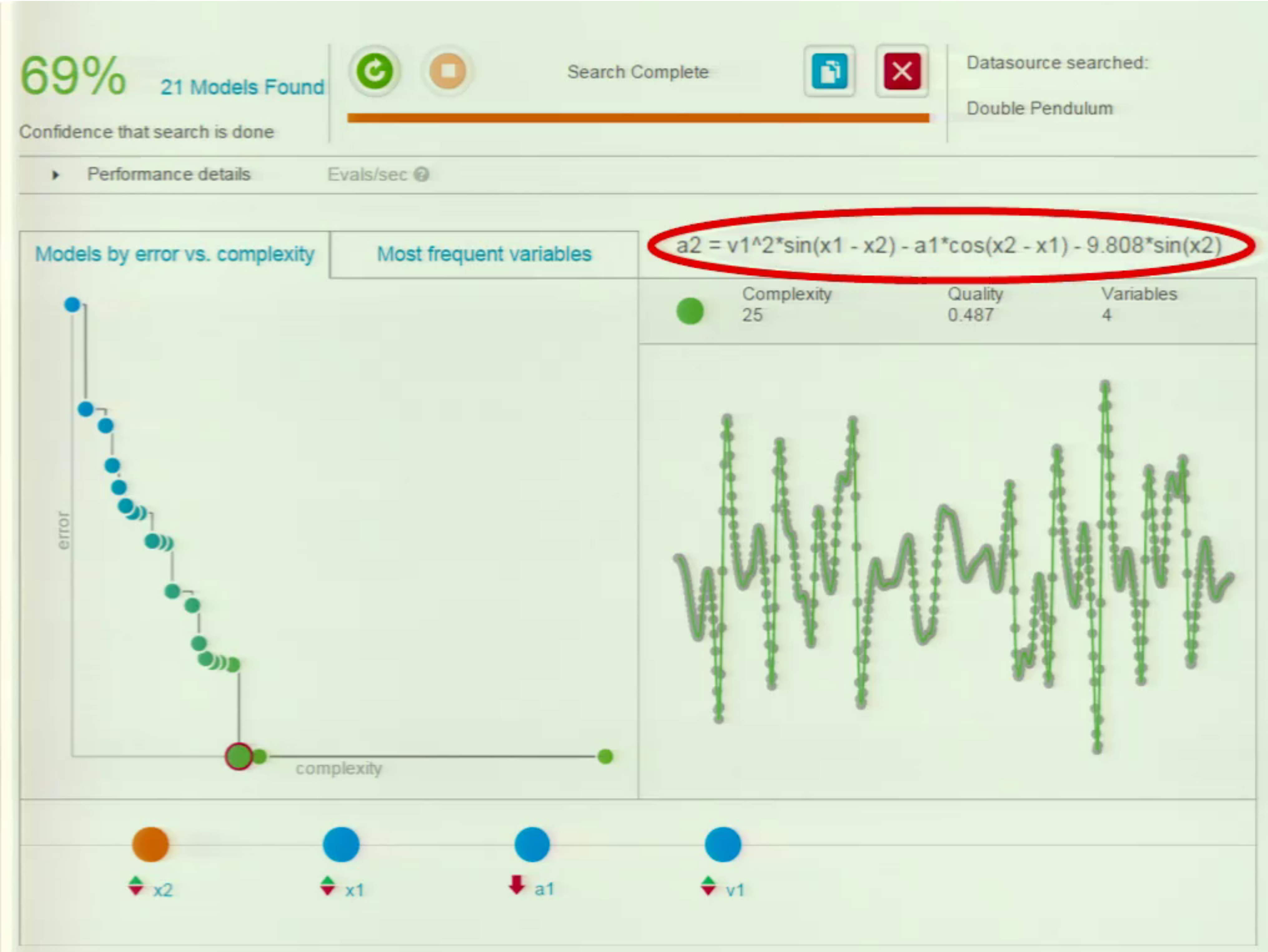
$$L_1^2(m_1+m_2)\omega_1^2 + m_2L_2^2\omega_2^2 + m_2L_1L_2\omega_1\omega_2\cos(\theta_1-\theta_2) - 19.6L_1(m_1+m_2)\cos\theta_1 - 19.6m_2L_2\cos\theta_2$$



$$\frac{dK}{dt} = a_K + \frac{b_K + c_K S_{t-t_1}}{K_{t-t_2}}$$

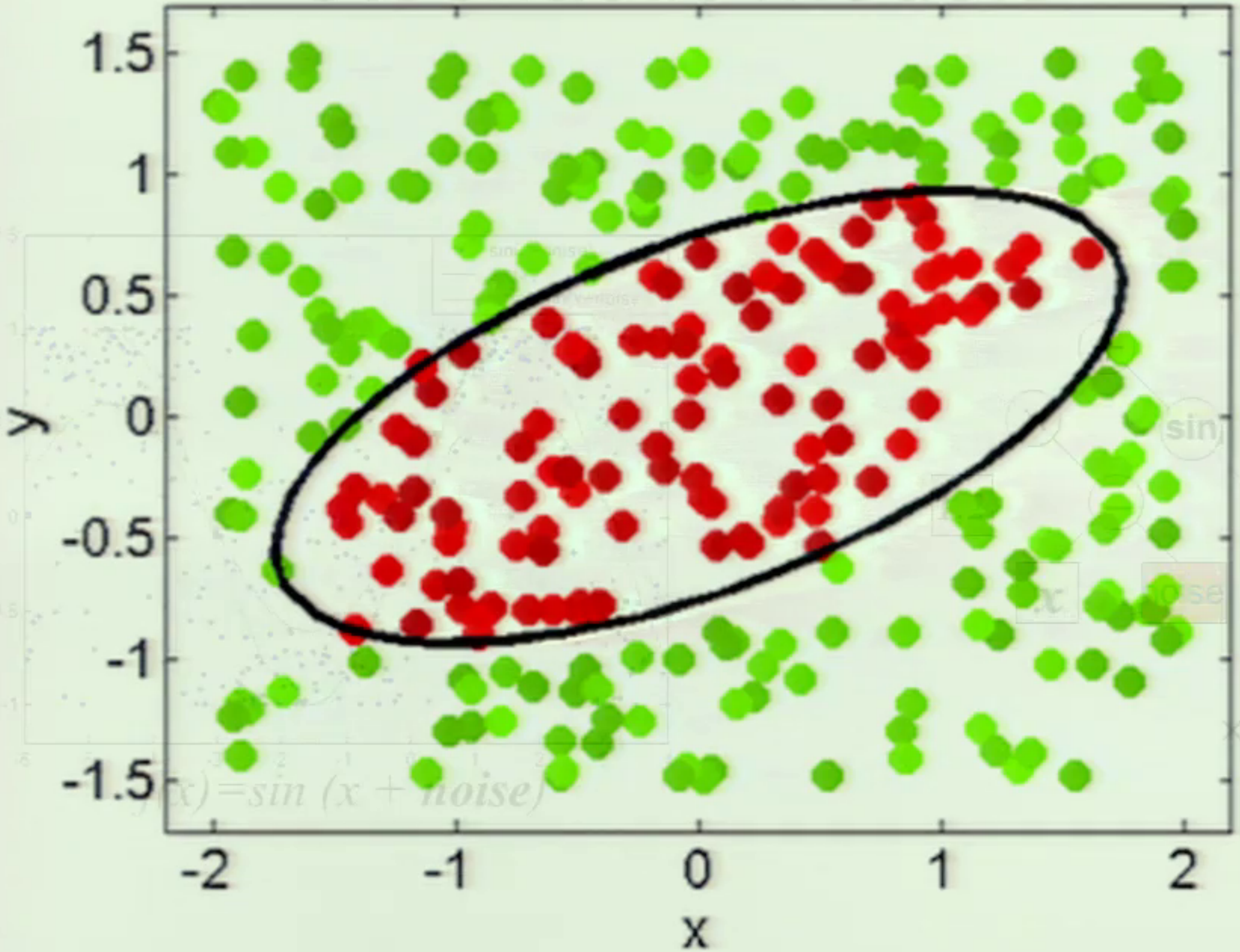
$$\frac{dS}{dt} = a_S + \frac{b_S + c_S K_{t-t_3}}{S_{t-t_4}}$$


Eureqa

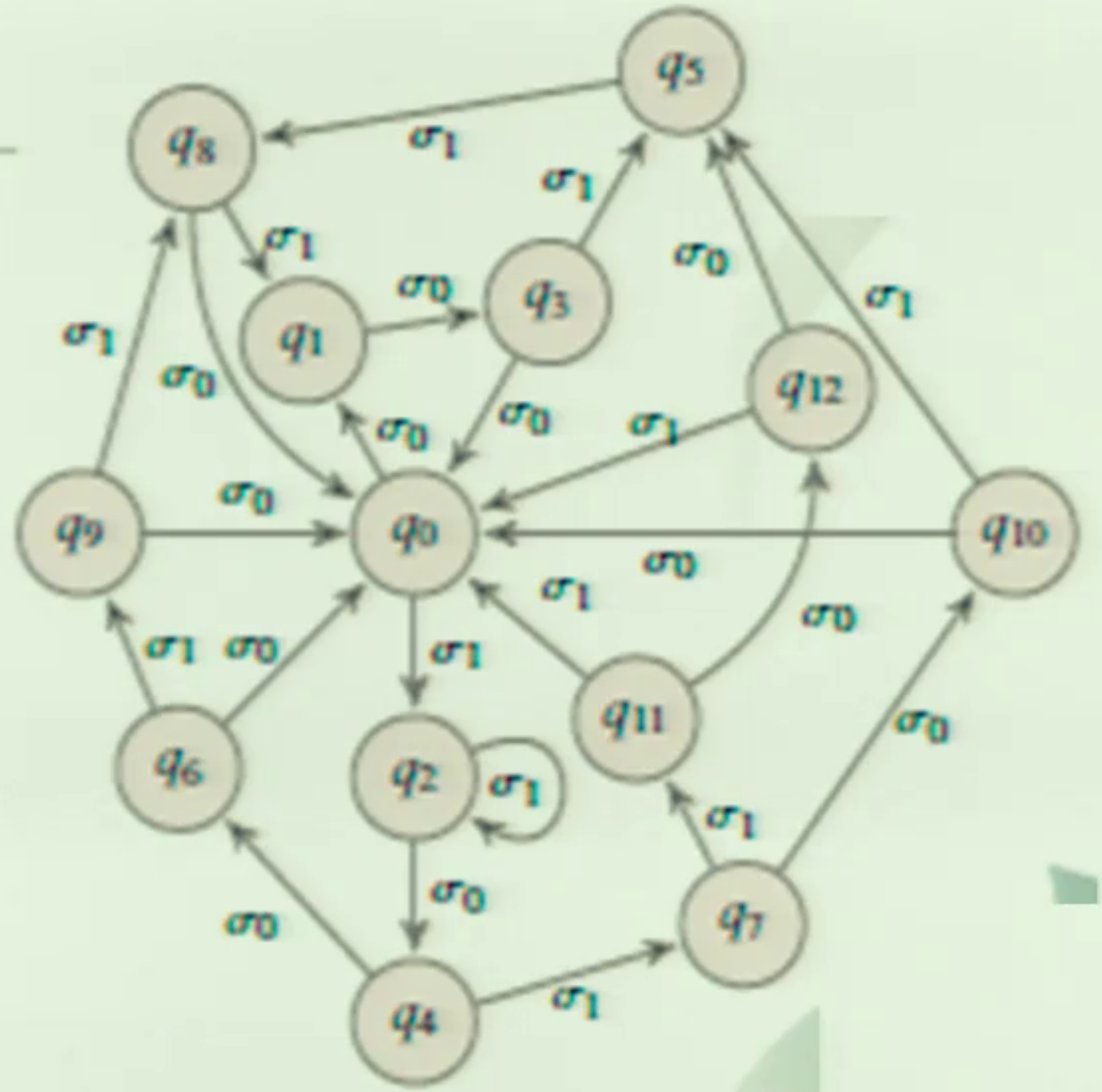

69%
21 Models Found
Confidence that search is done
Search Complete
Datasource searched:
Double Pendulum
Performance details
Evals/sec
Models by error vs. complexity
Most frequent variables
a2 = v1^2*sin(x1 - x2) - a1*cos(x2 - x1) - 9.808*sin(x2)
Complexity
25
Quality
0.487
Variables
4
error
complexity
x2
x1
a1
v1


RADON
DETECTOR
R.P.II. Tel: 353 1 269 77 66
RADON
DETECTOR
R.P.II. Tel: 353 1 269 77 66

Stochastic Models

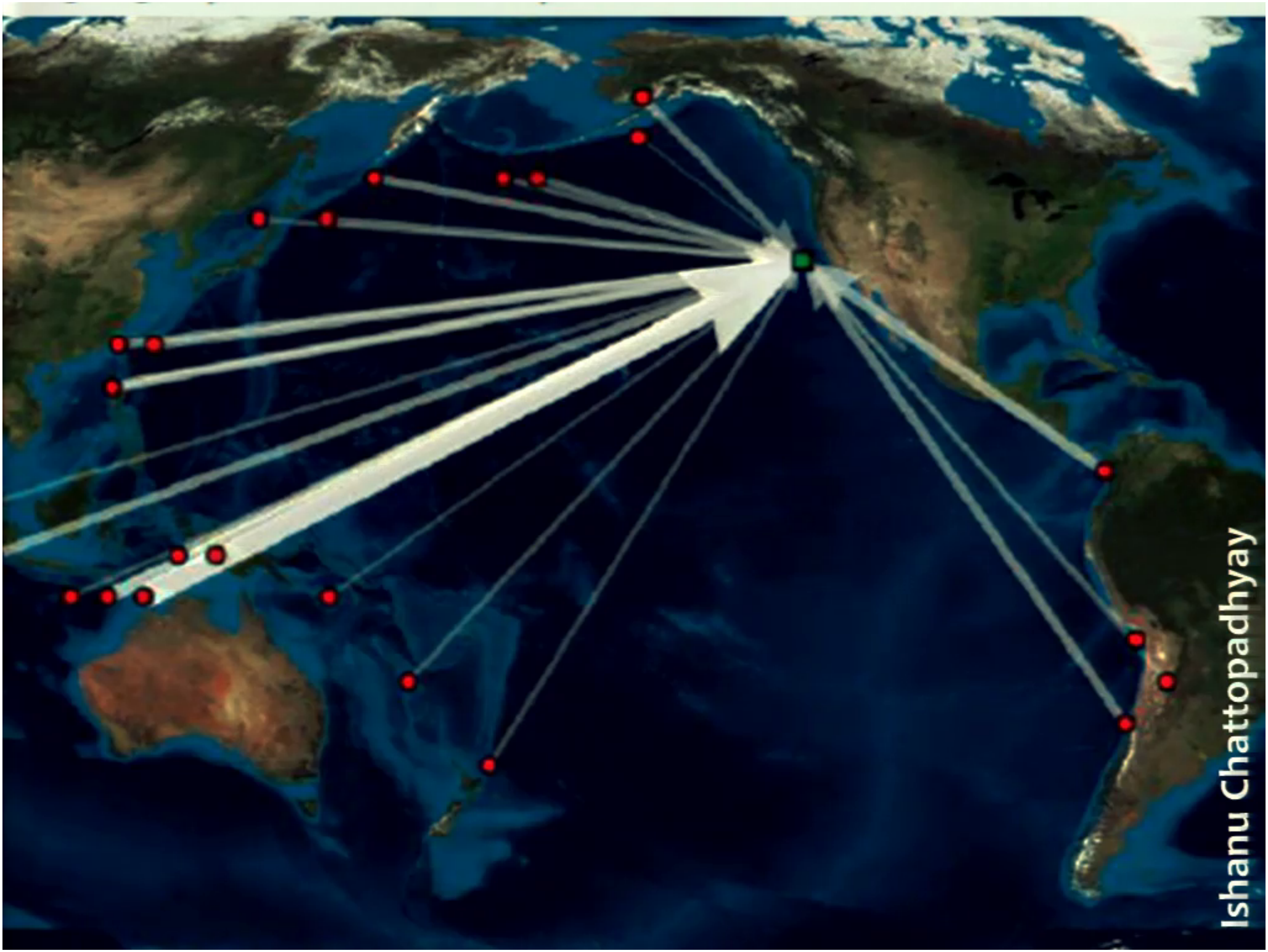


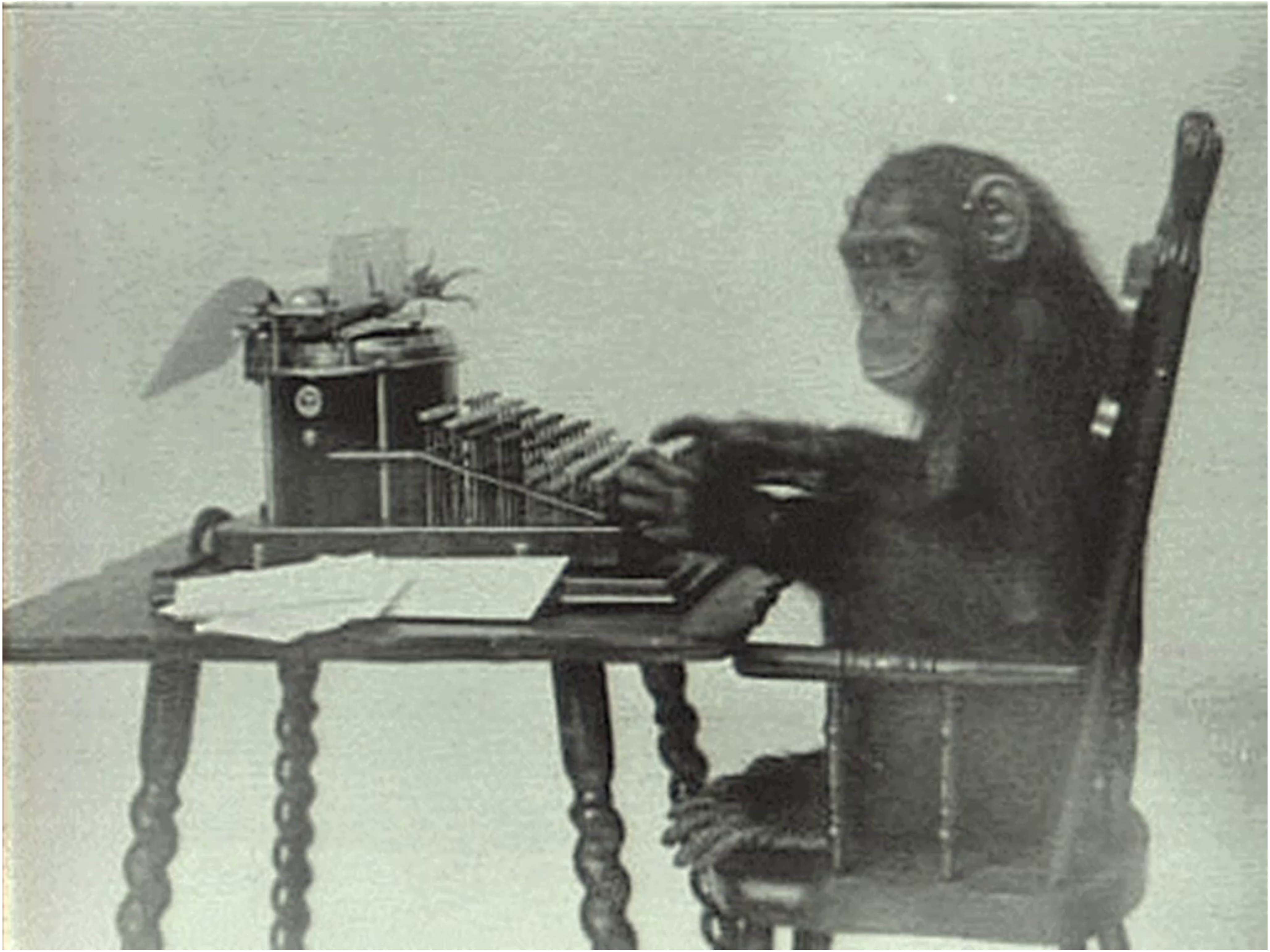
Probabilistic Finite State Machines



$\sigma_1\ \sigma_1\sigma_2\ \sigma_1\sigma_2\ \sigma_2\ \sigma_2\ \sigma_1\sigma_2\ \sigma_1\sigma_2\ \sigma_1\ \sigma_1\sigma_2\ \sigma_1\sigma_2\ \sigma_2\ \sigma_2\ \sigma_1\sigma_2$

Ishanu Chattopadhyay


Ishanu Chattopadhyay



The New York Times

Theoretical physicists are not yet obsolete, but scientists have taken steps toward replacing themselves