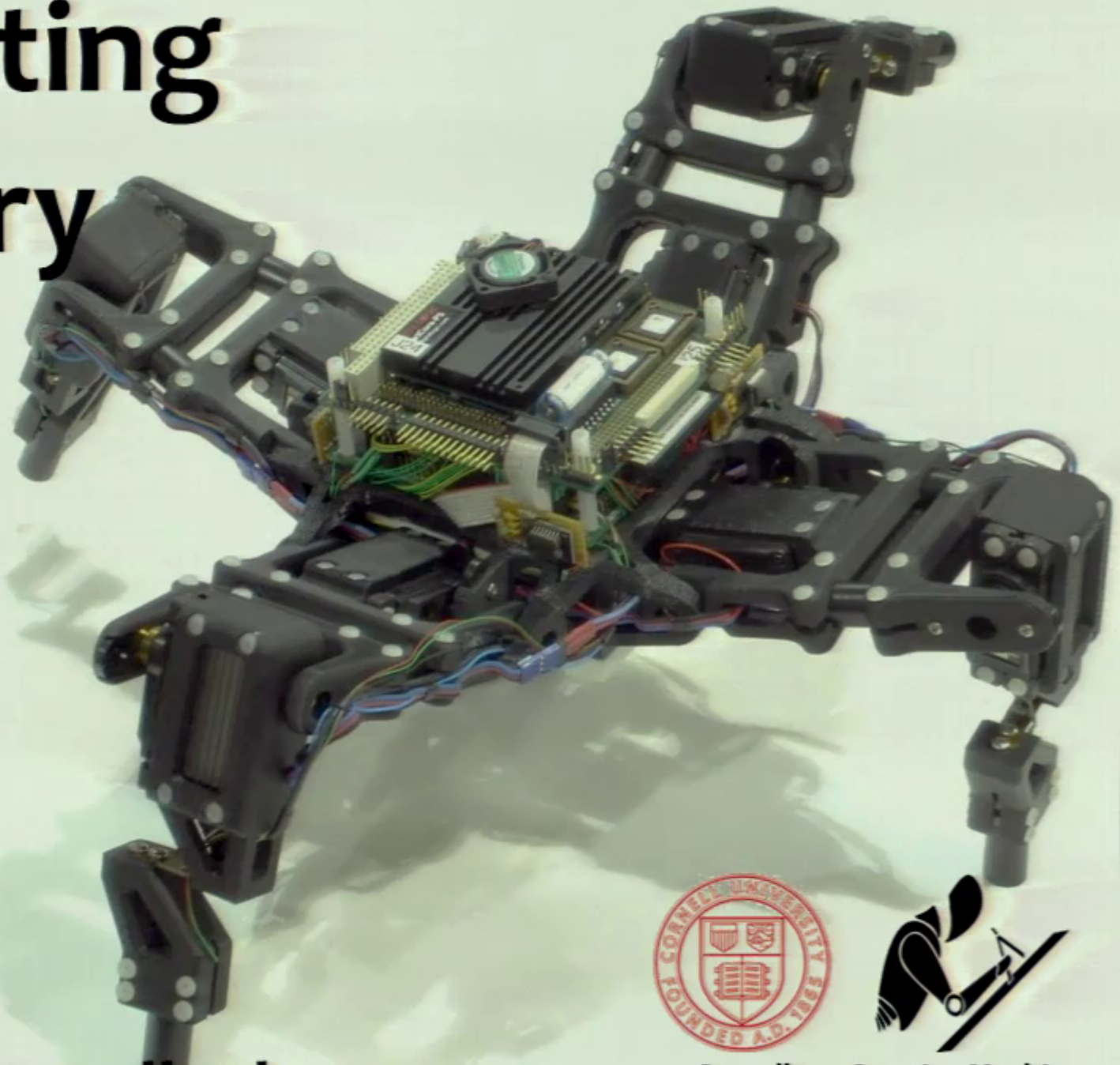


# Automating Discovery



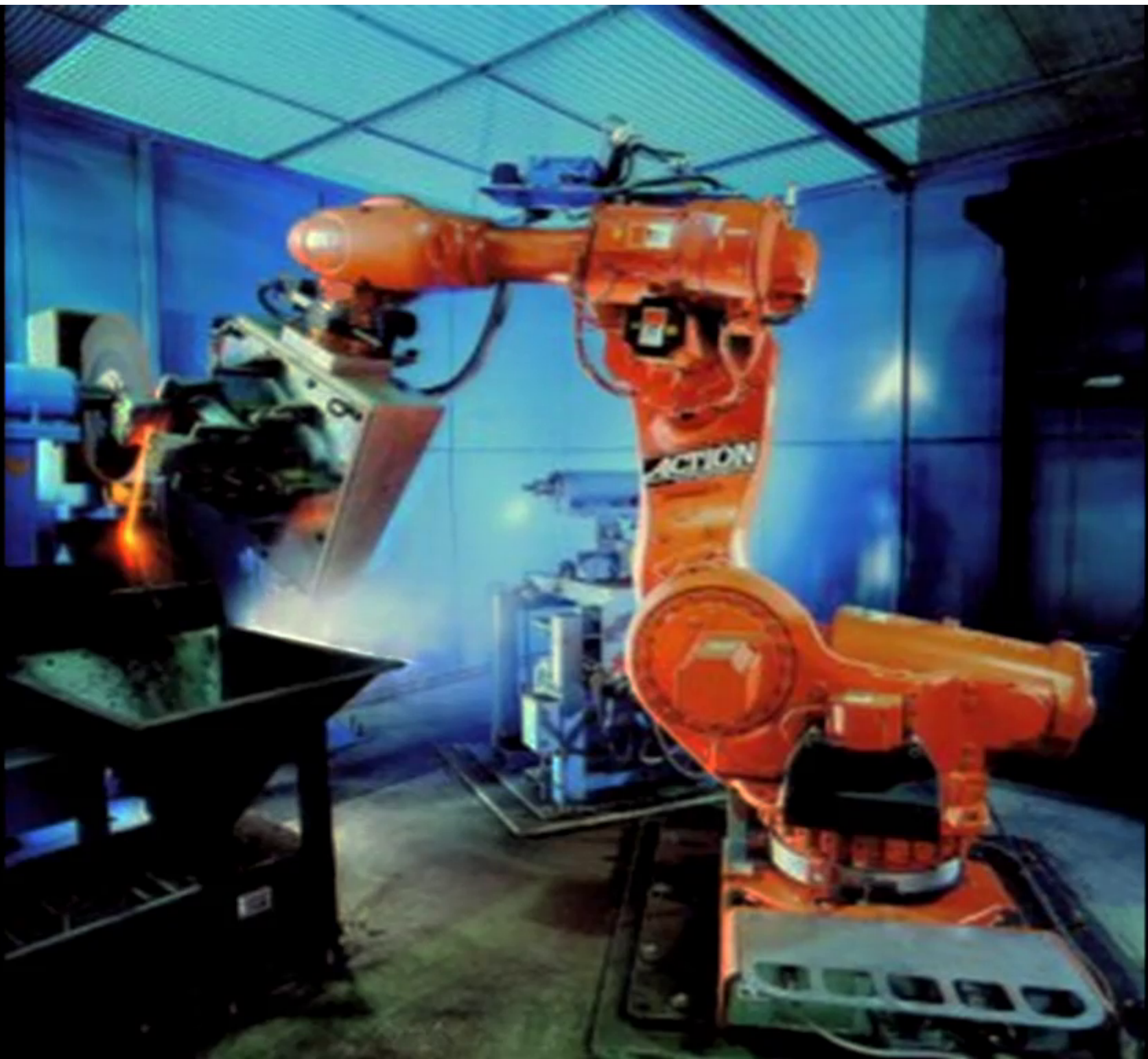
[Hod.Lipson@cornell.edu](mailto:Hod.Lipson@cornell.edu)



Cornell  
University



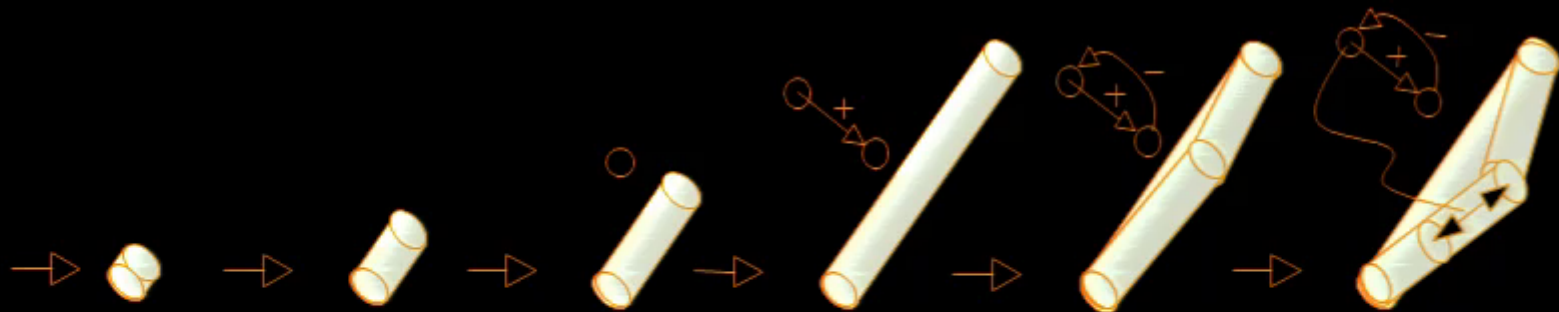
Creative Machines  
Lab



# Evolution



# Evolution



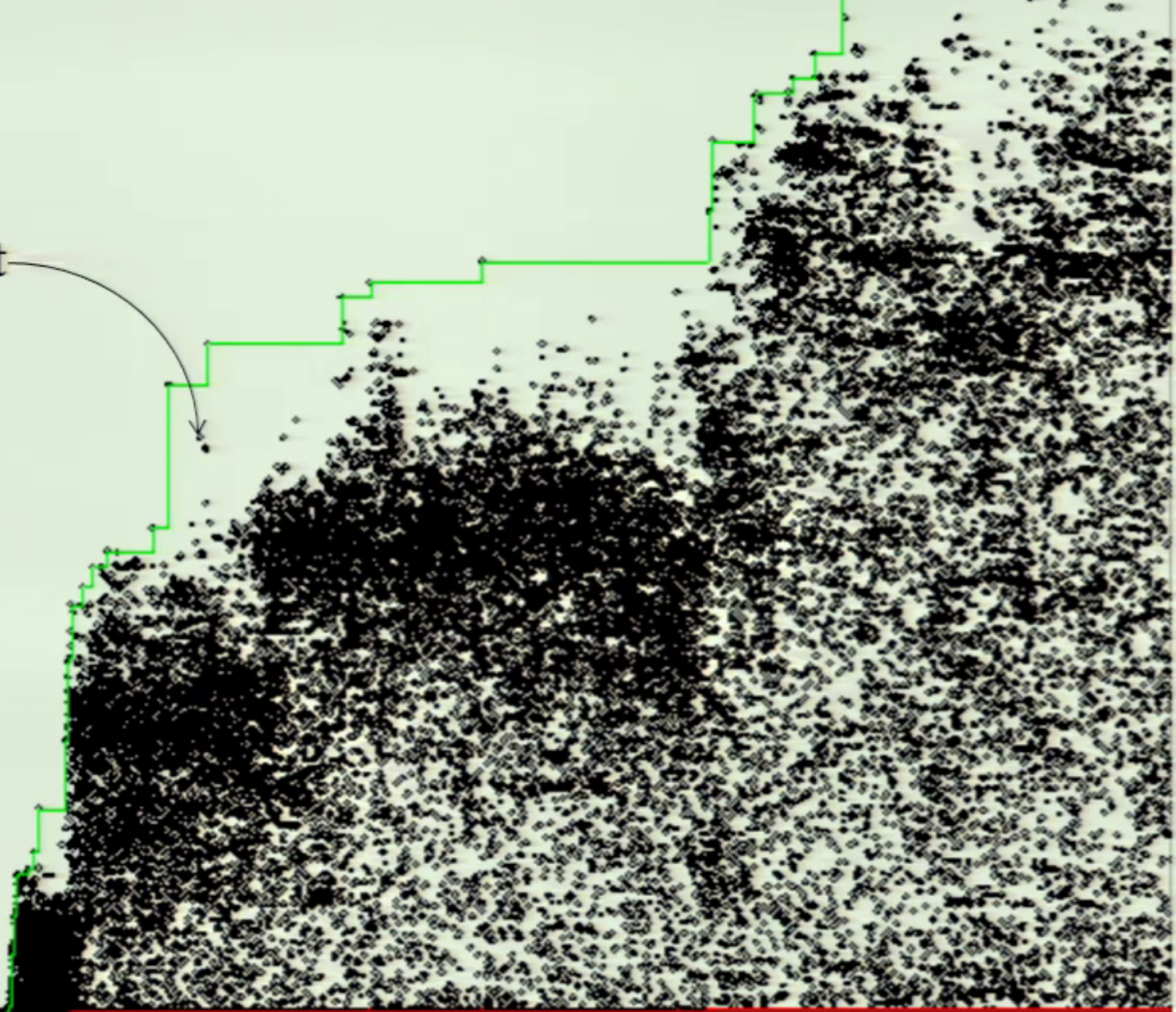
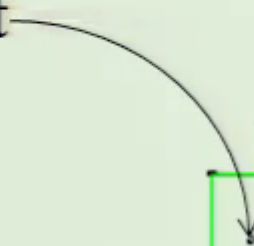


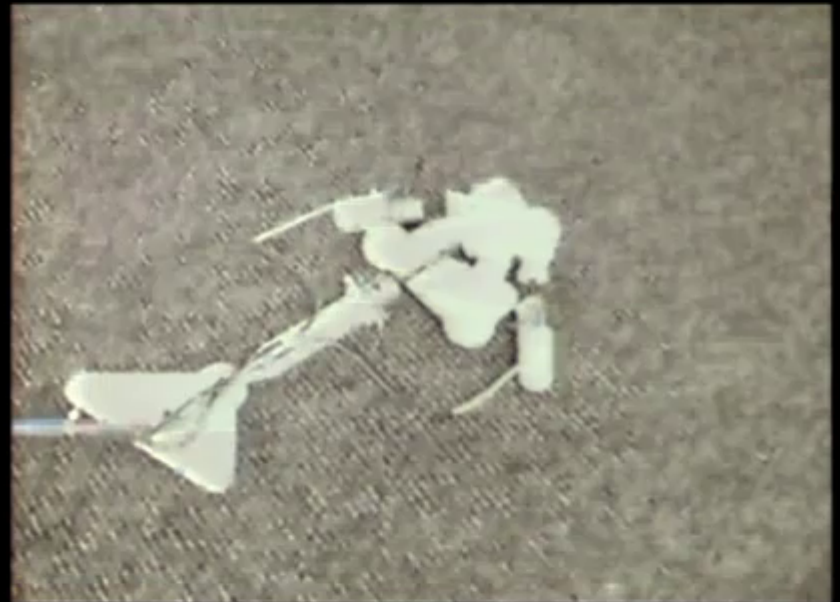
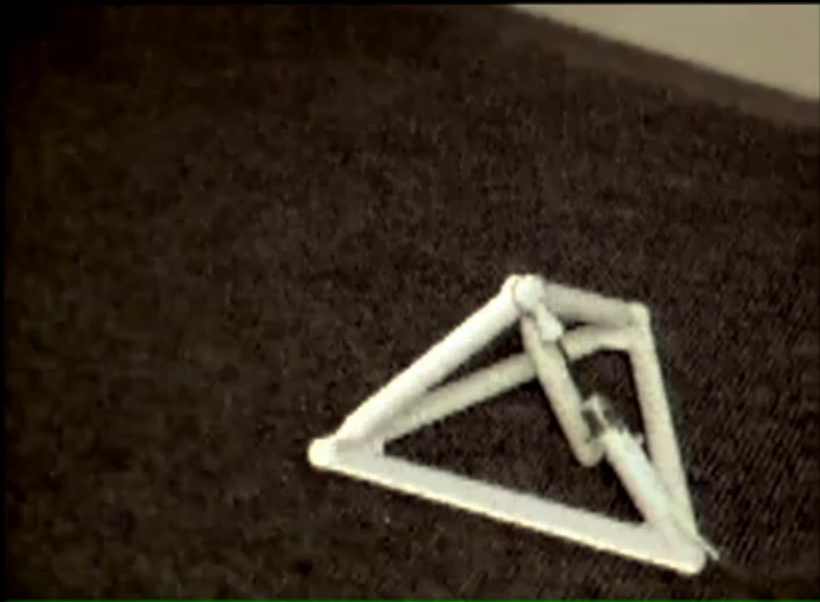
0.427

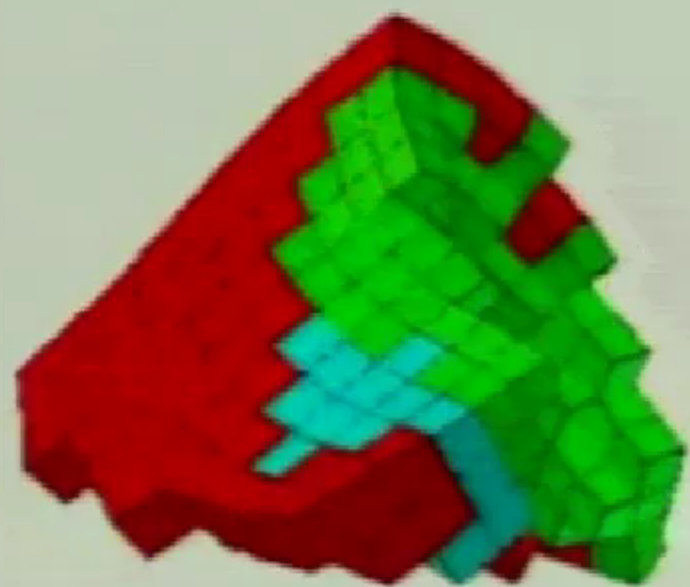
Robot

Speed

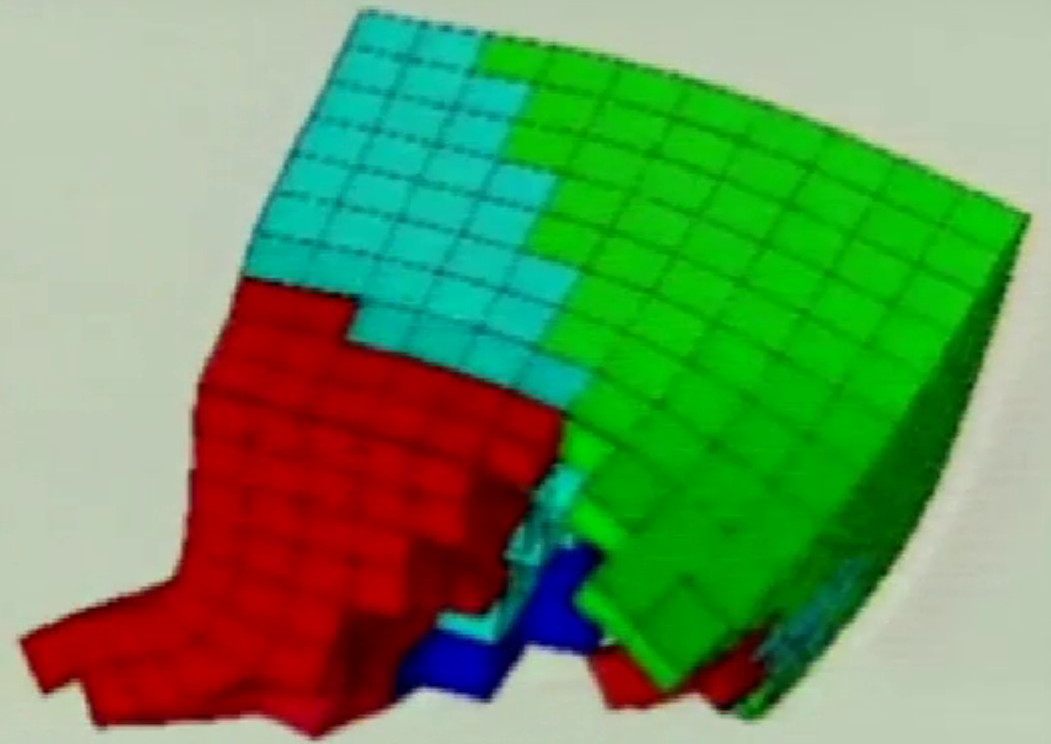
0 50 100 150 200 250 300 350 400 450 500  
Generations

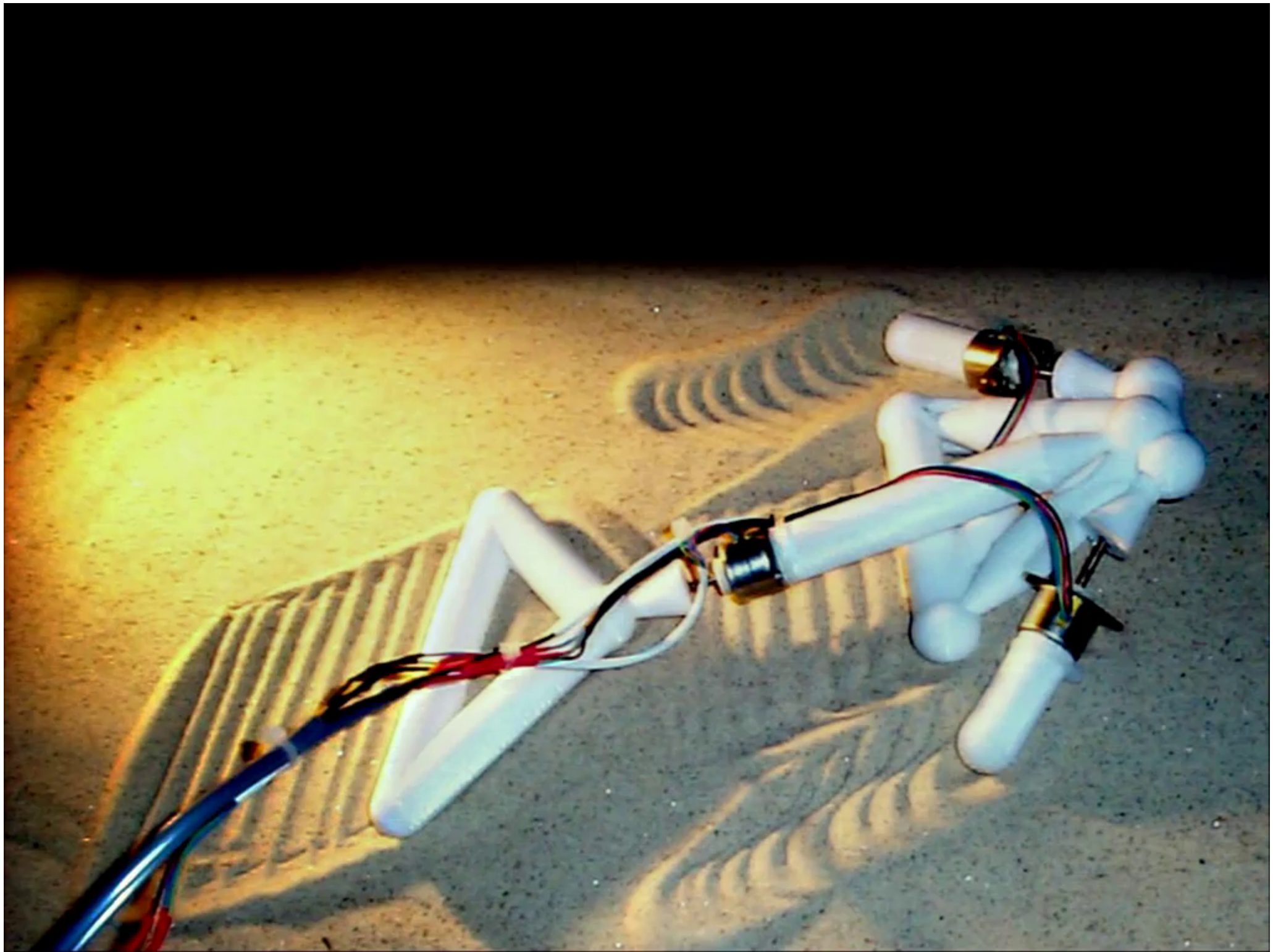












# The New York Times

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



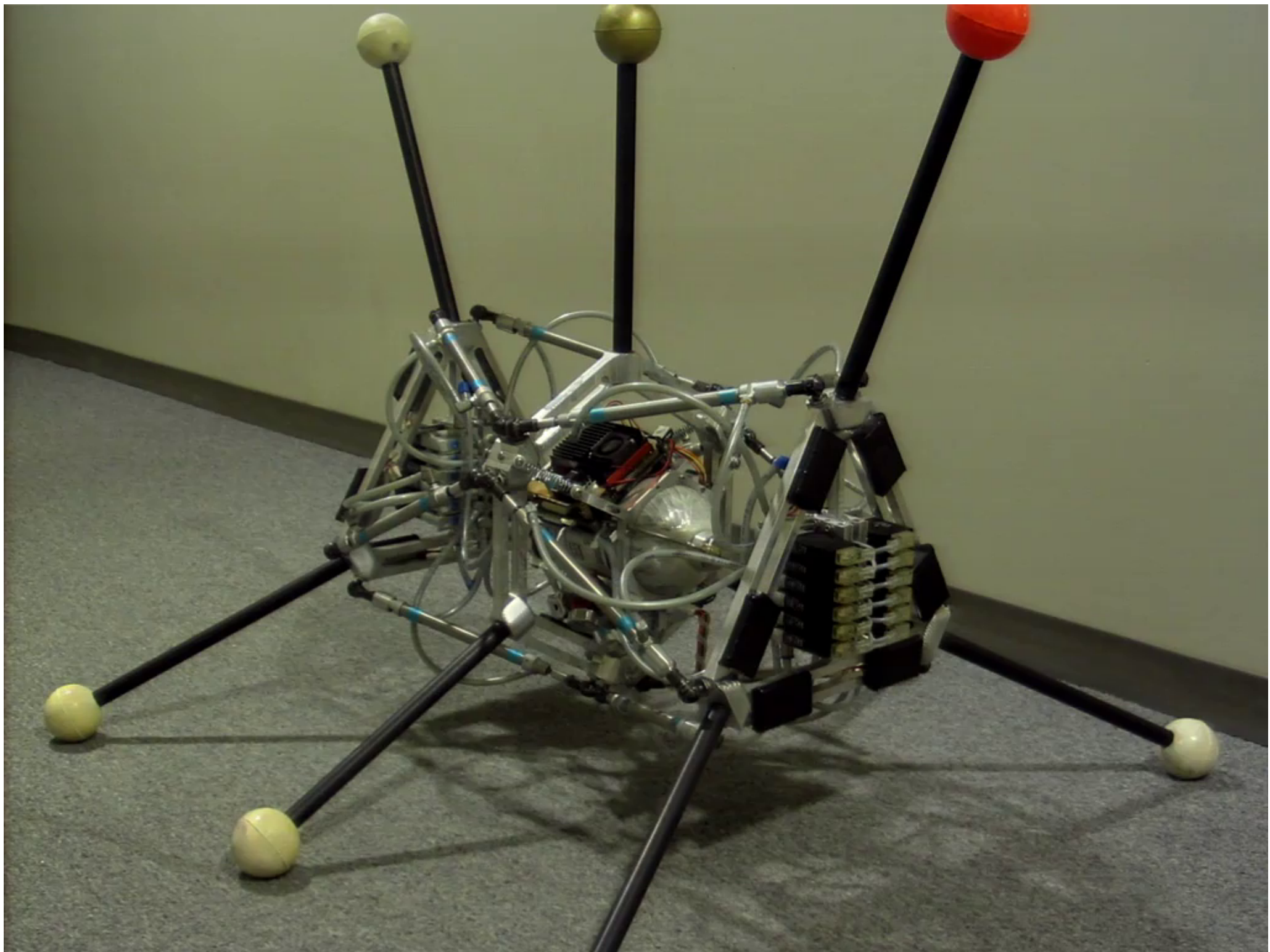
Brandeis University

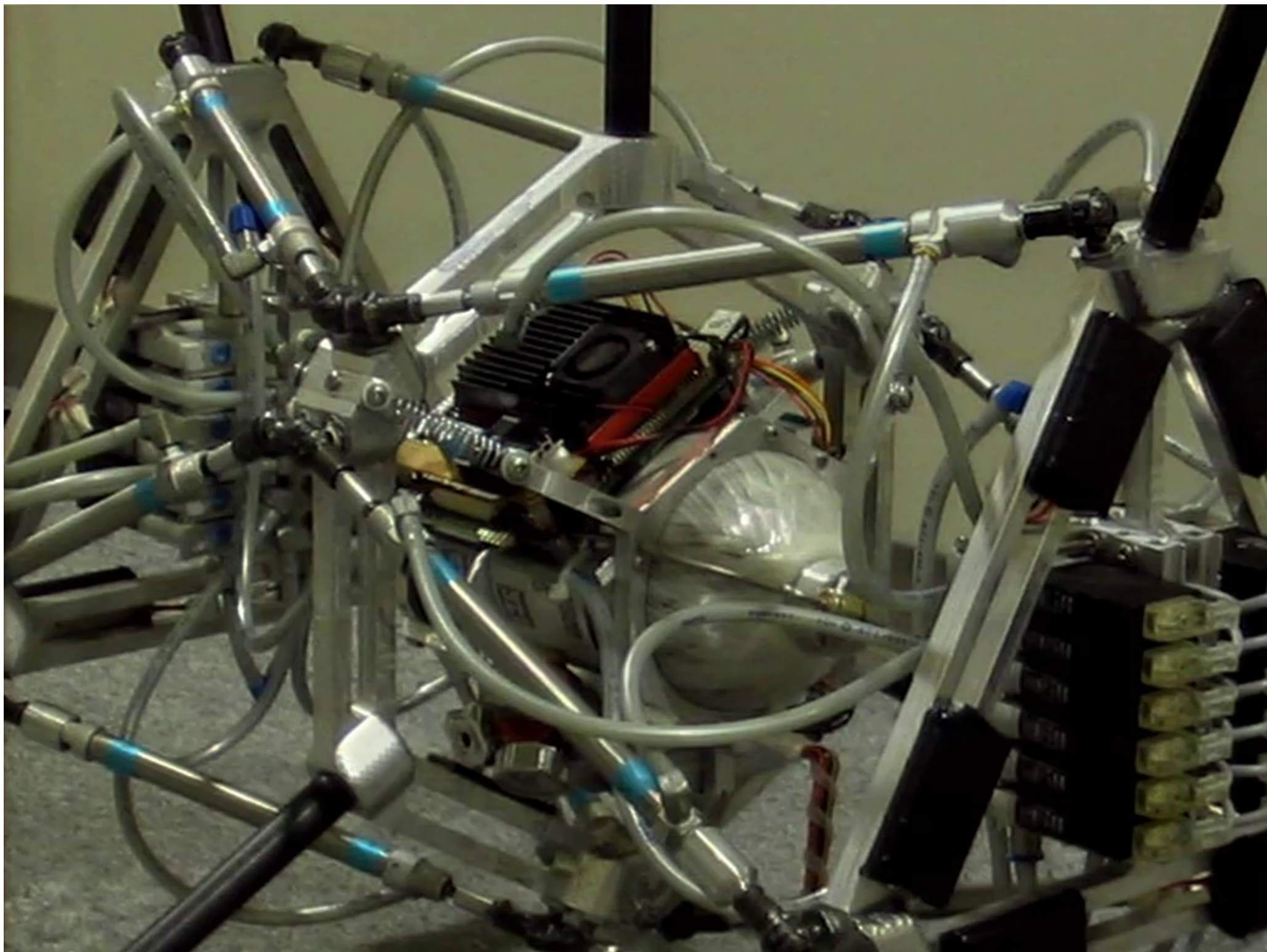
The “Arrow” left a trail as it crawled across a bed of sand.

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 progre

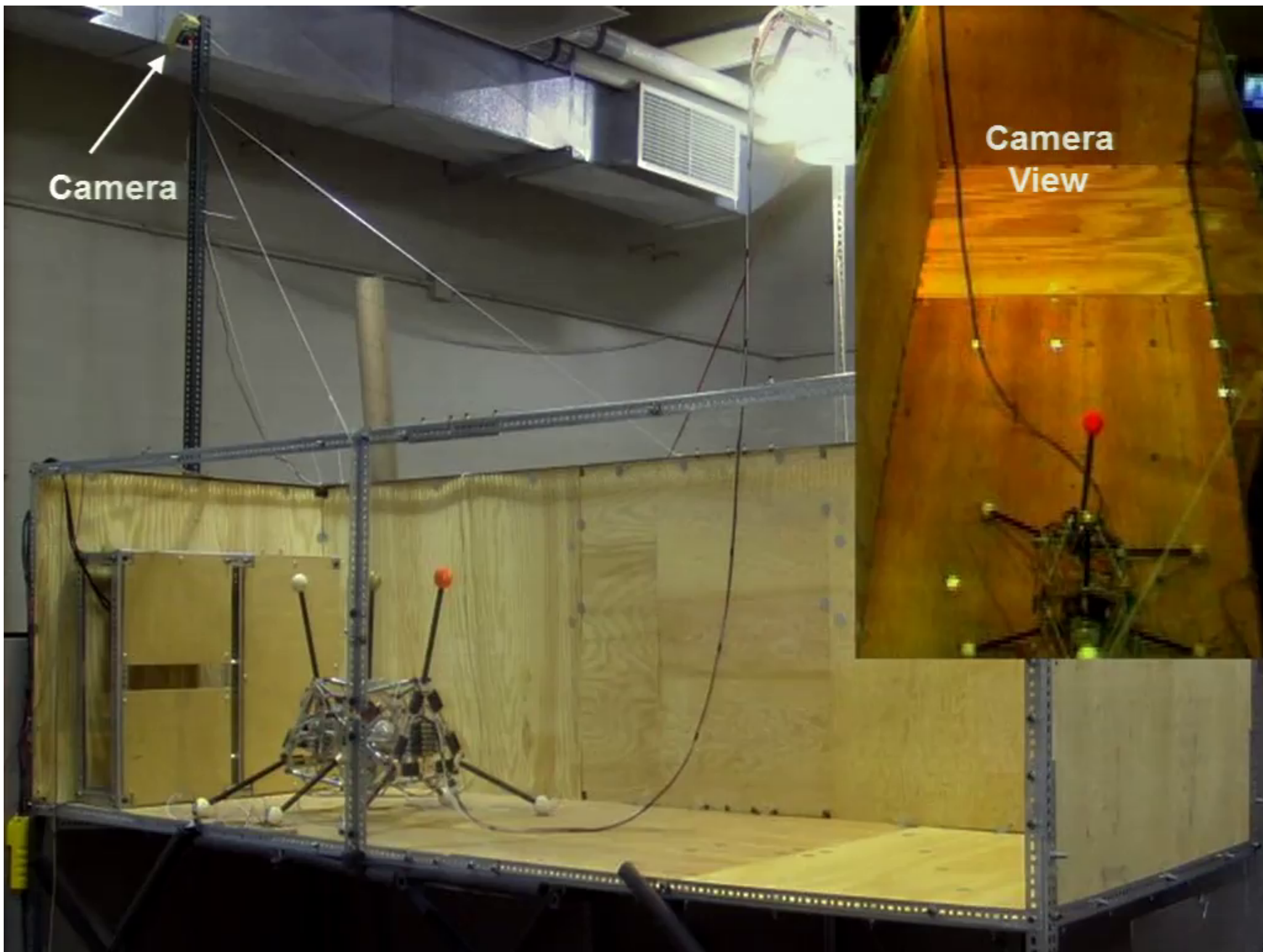


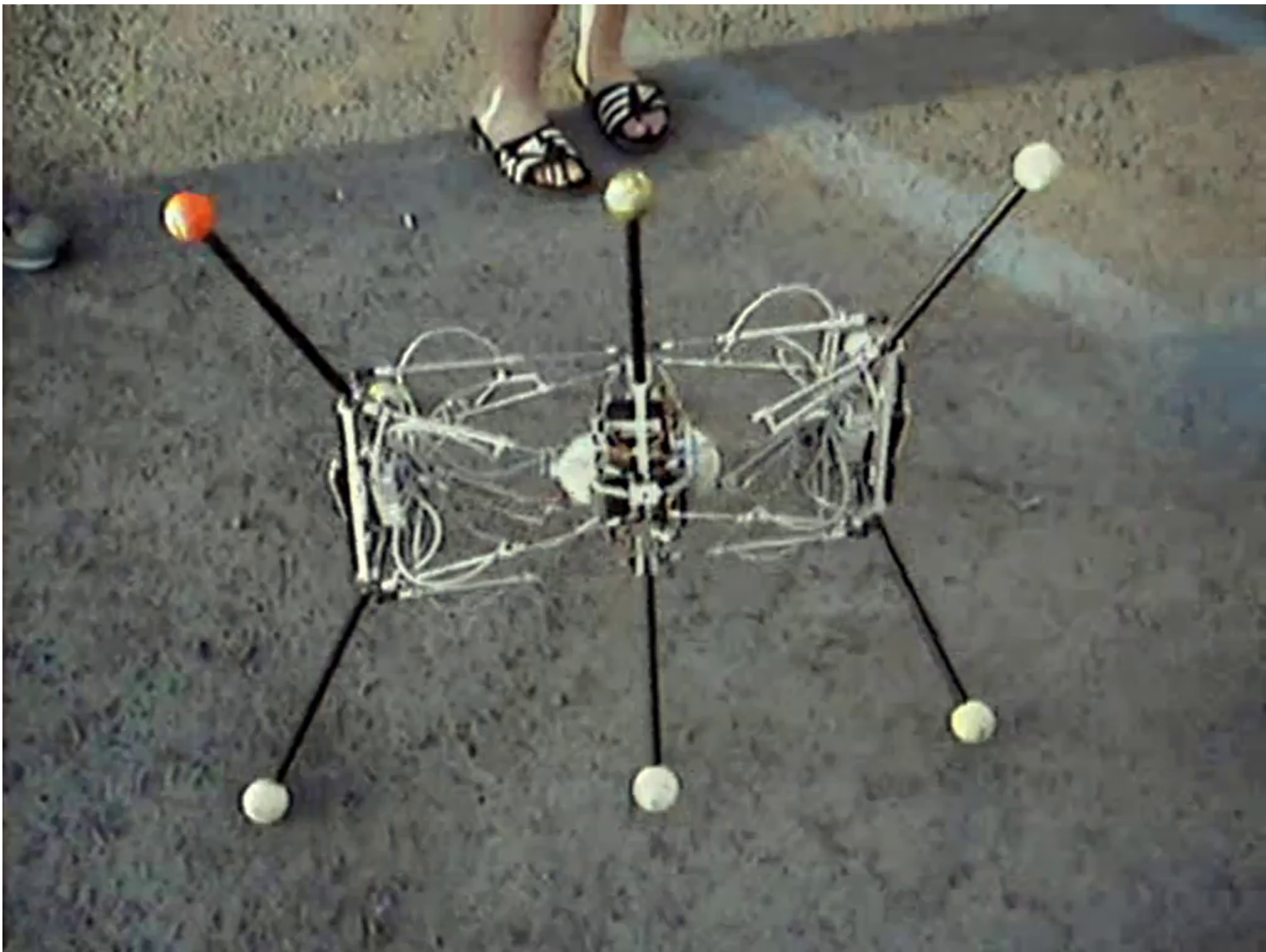




Camera

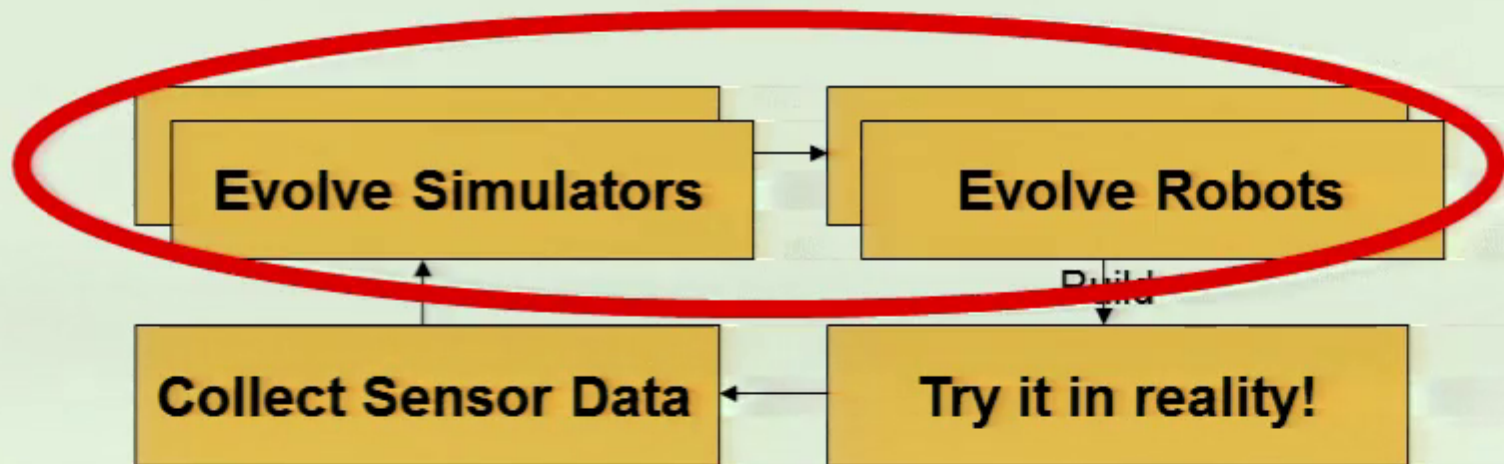
Camera  
View



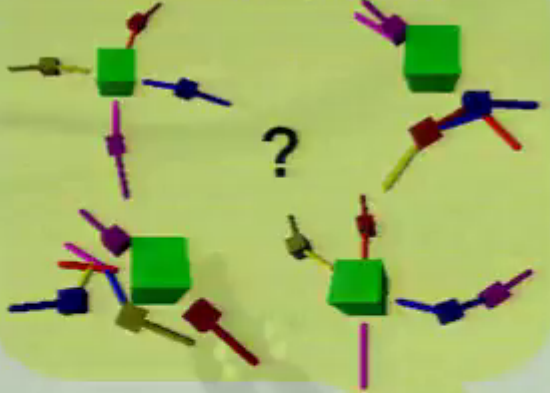




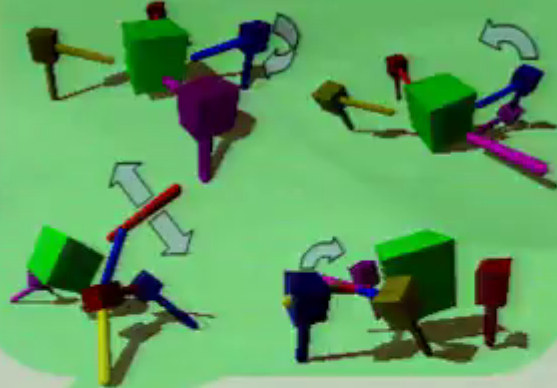
# Simulation & Reality



### Self-Model synthesis

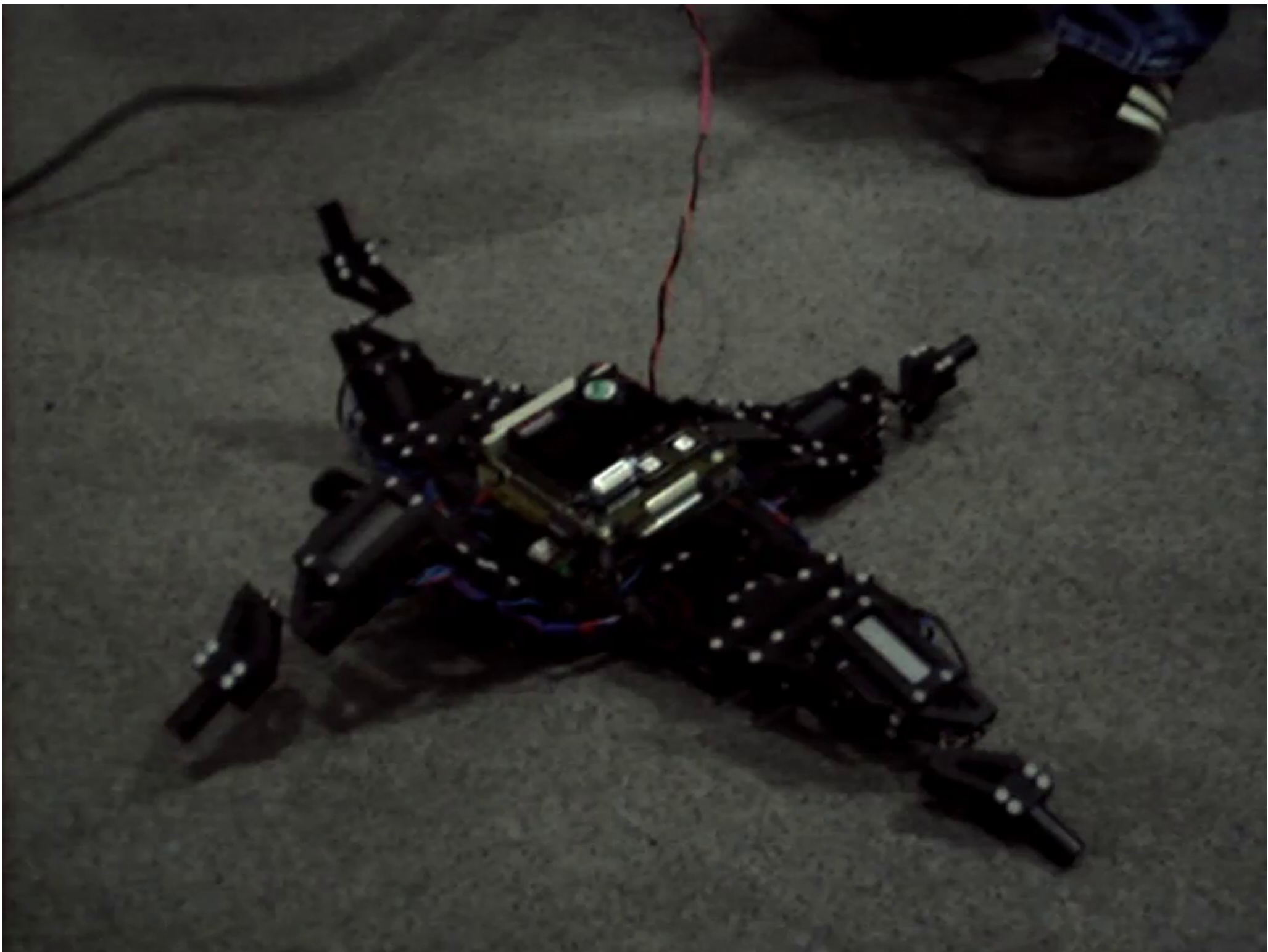


### Exploratory Action synthesis

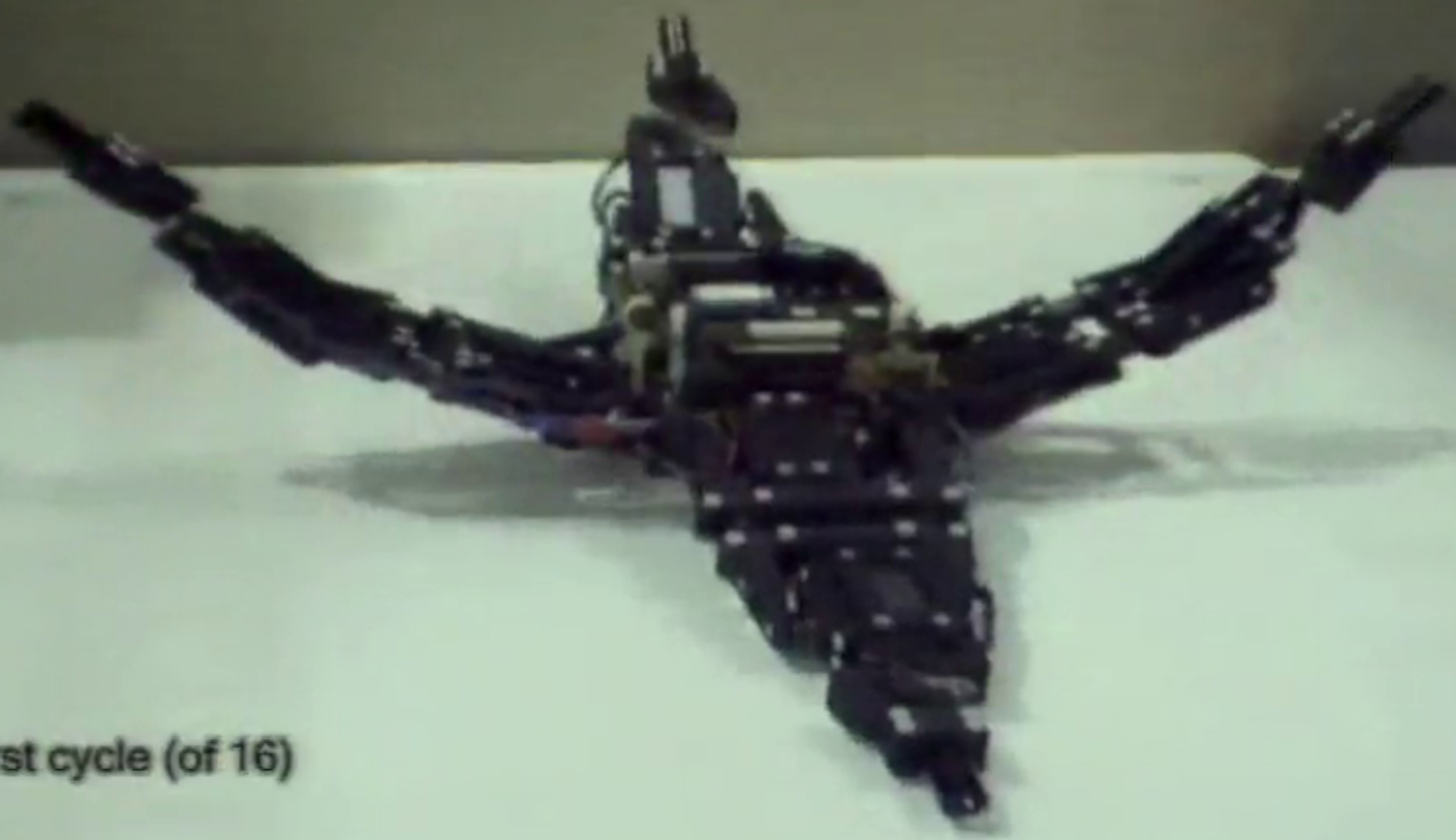


### Target Behavior synthesis





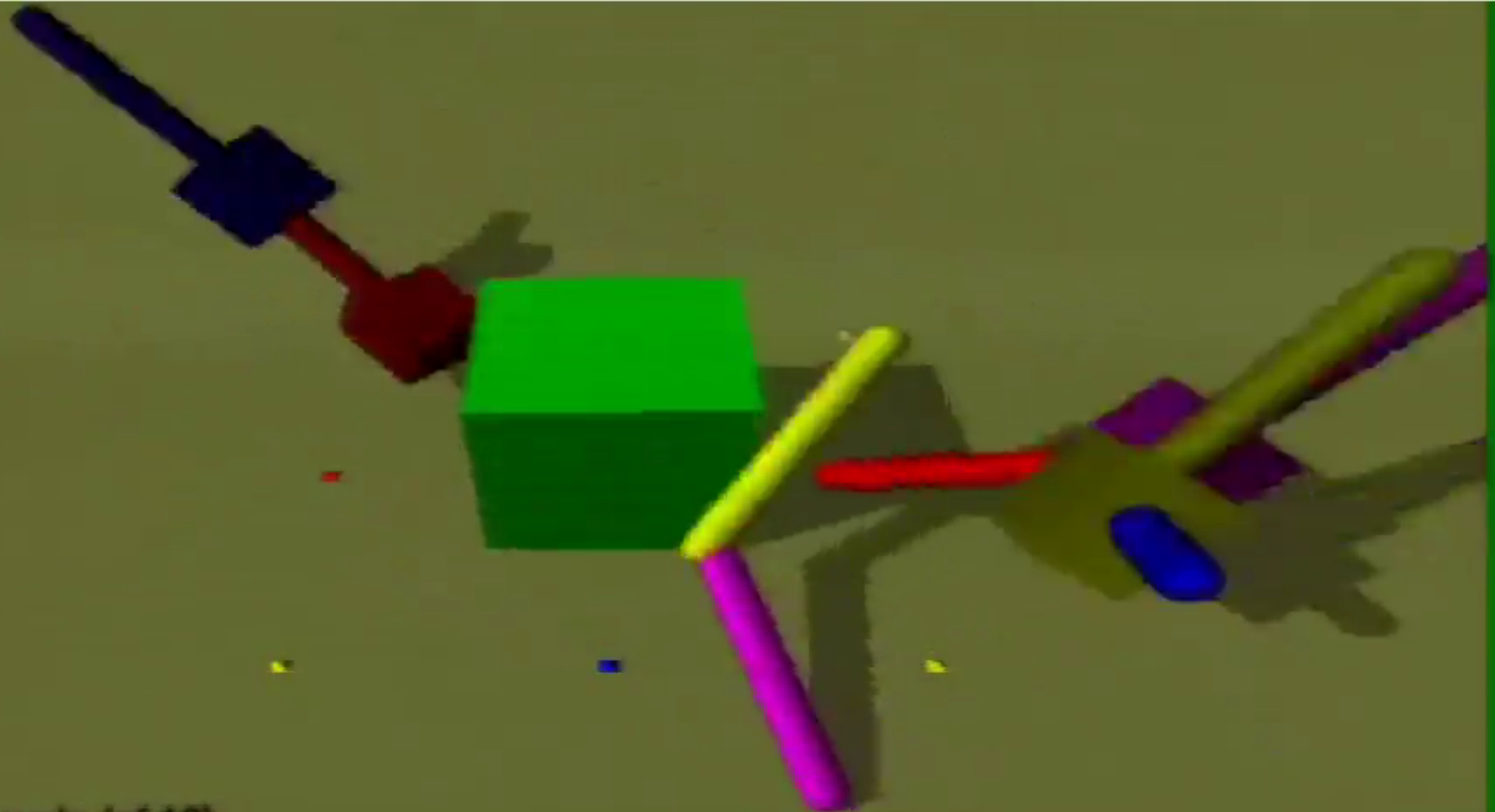
# Emergent Self-Model



First cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

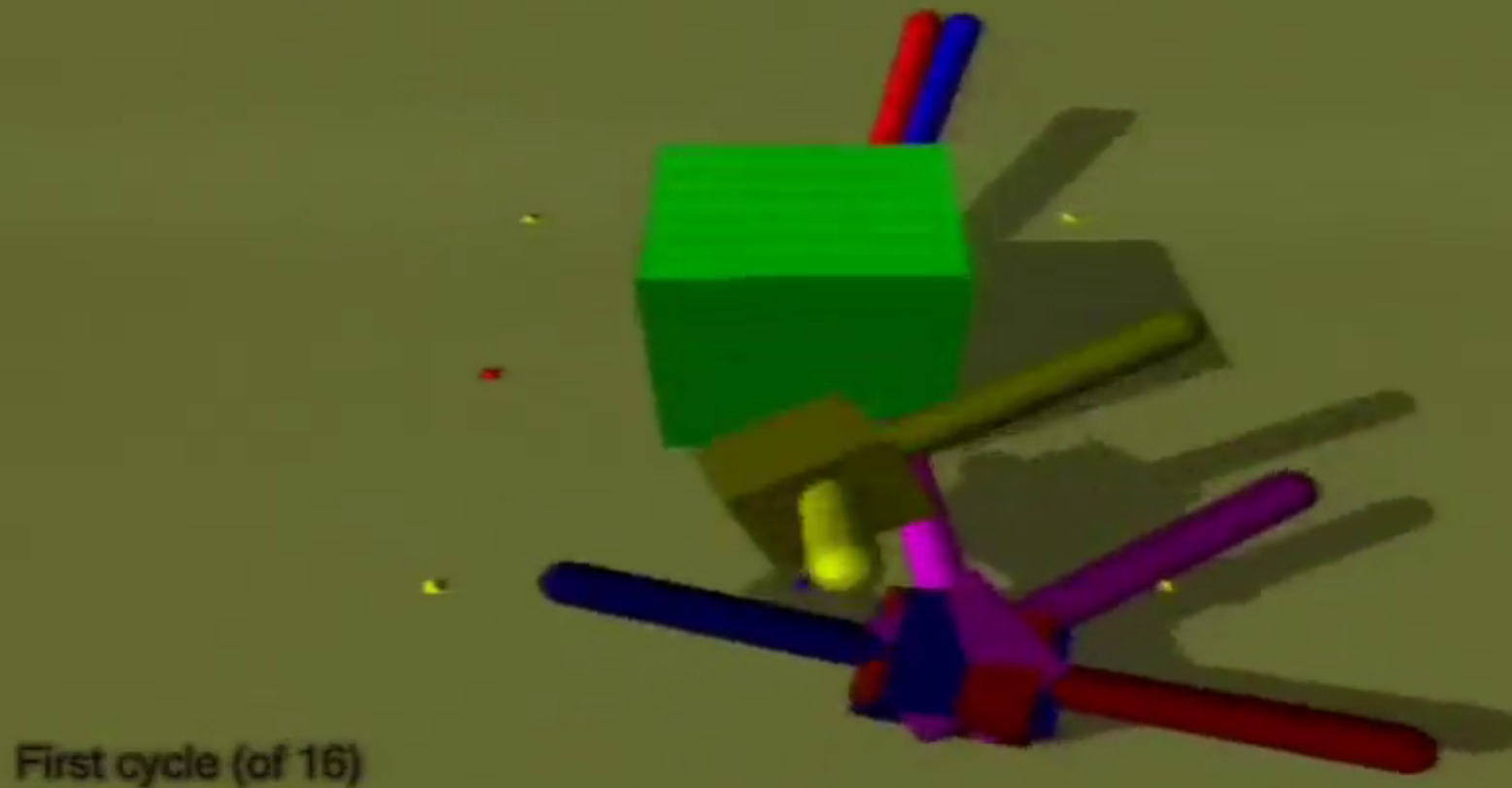
# Emergent Self-Model



First cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

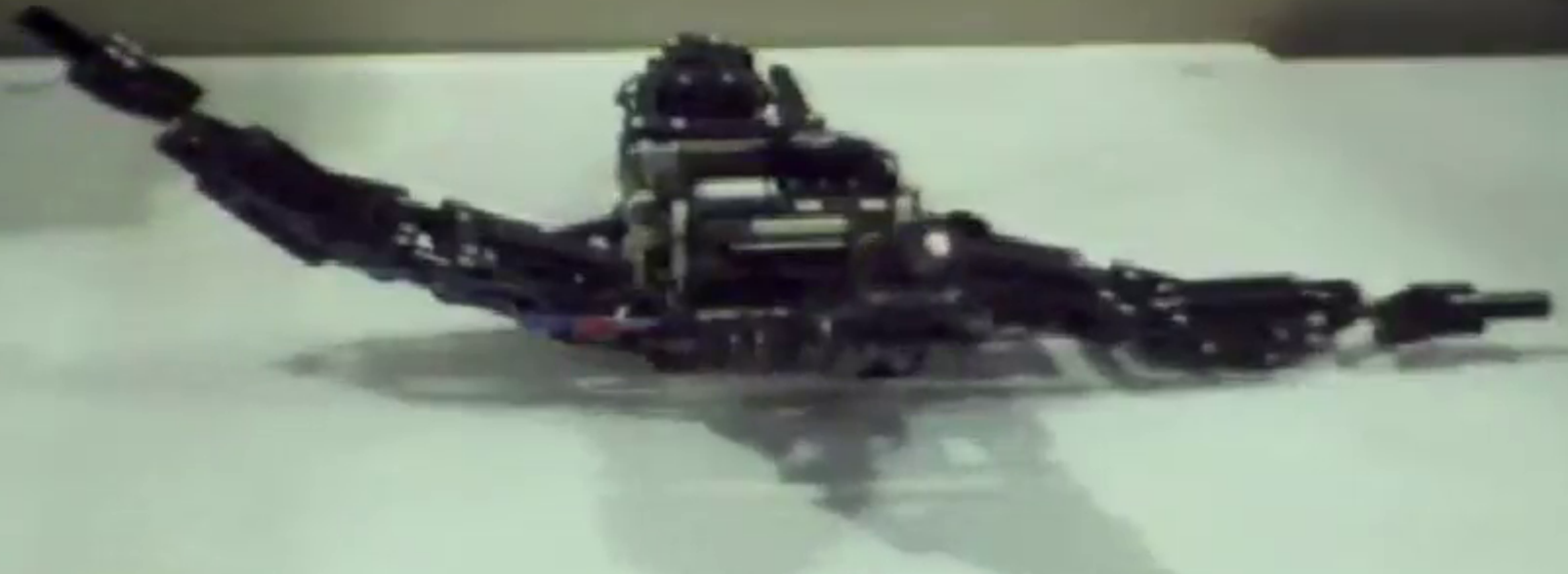
# Emergent Self-Model



First cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

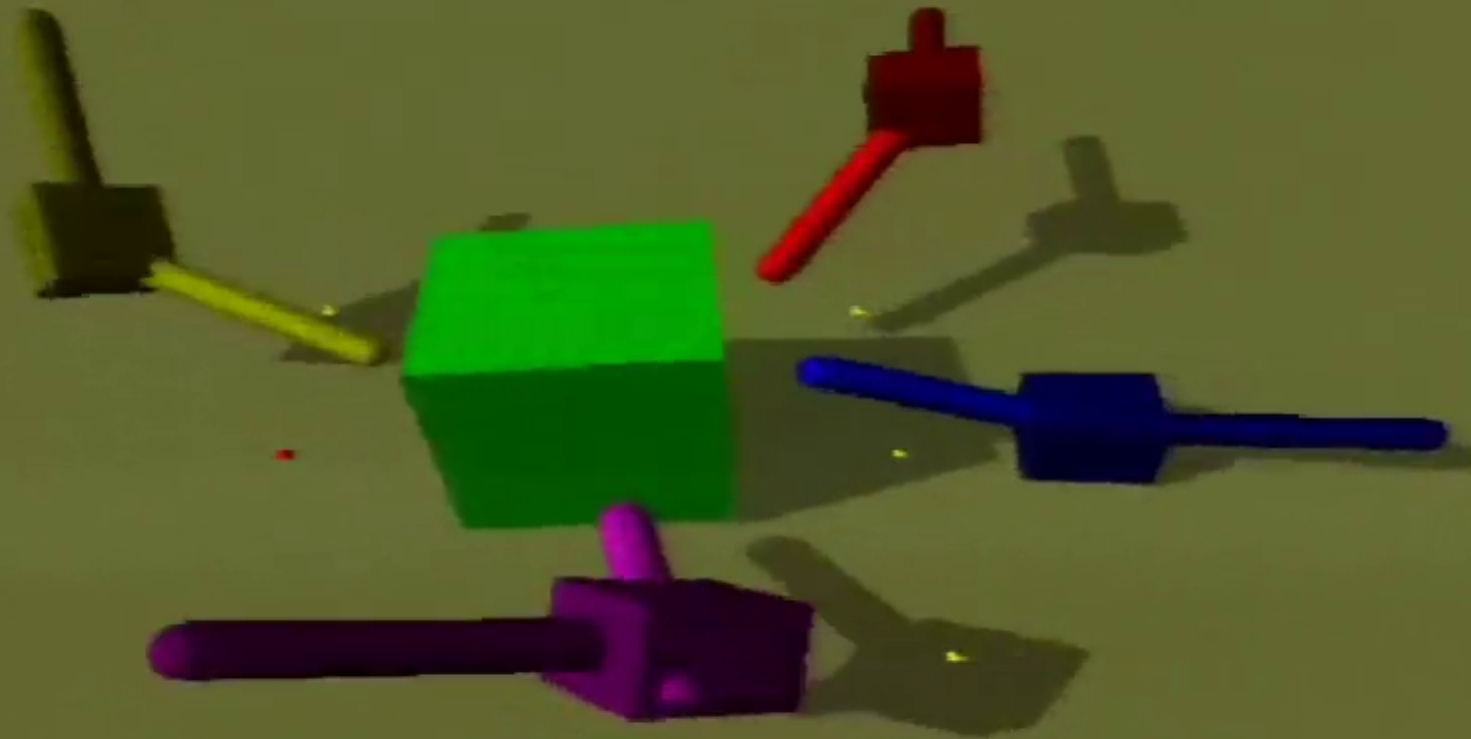
# Emergent Self-Model



Eighth cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

# Emergent Self-Model

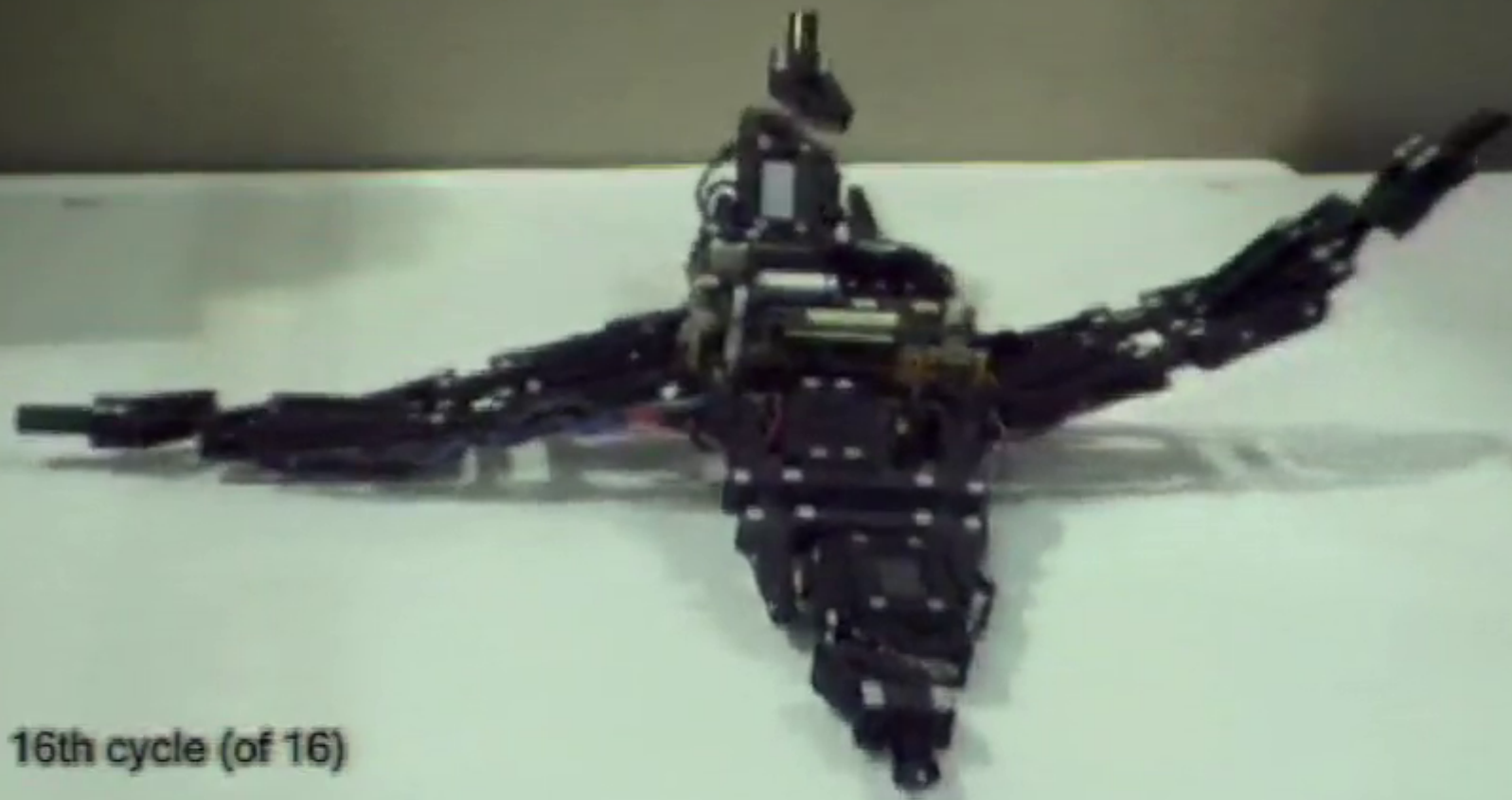


Eighth cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006



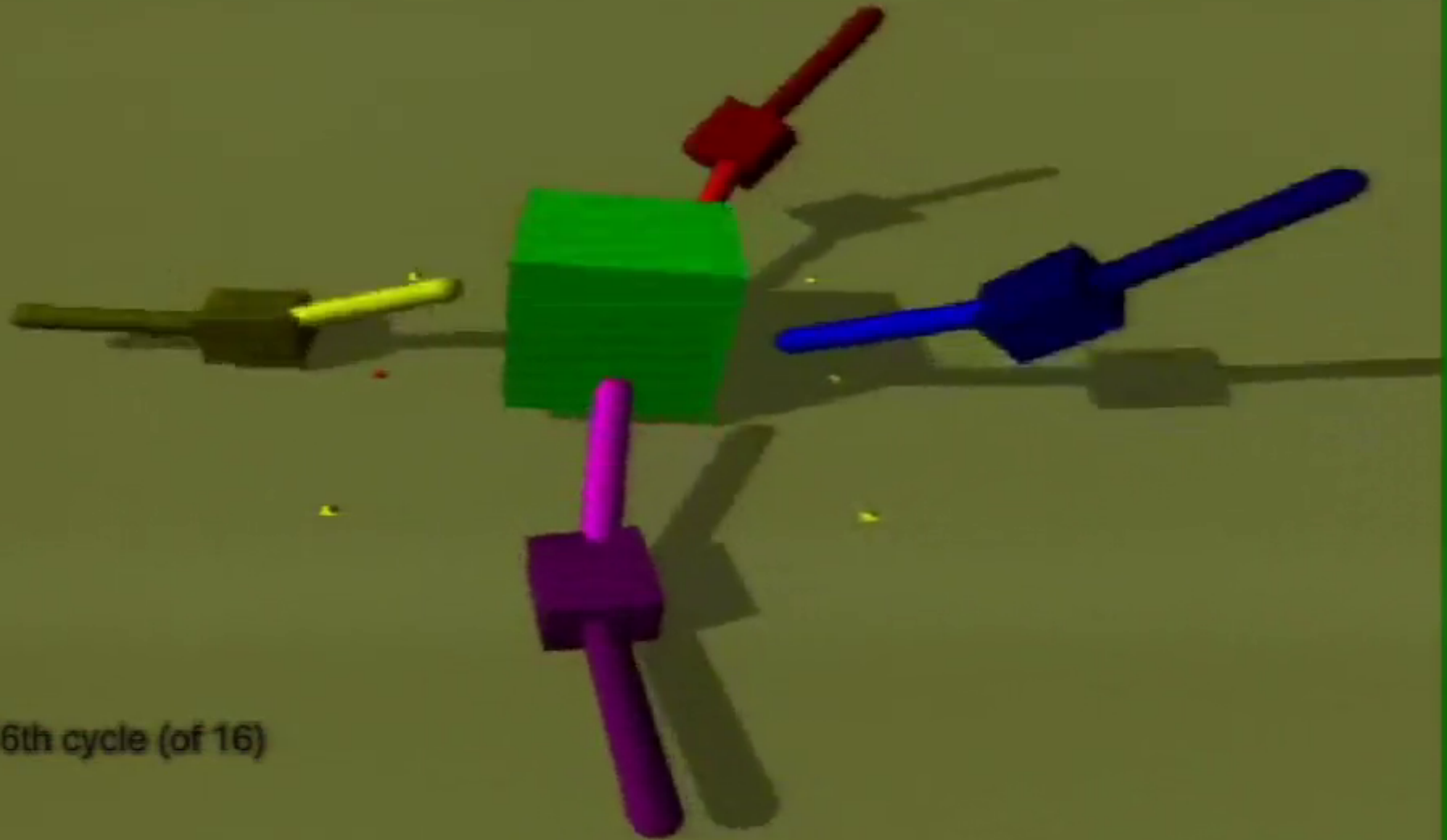
# Emergent Self-Model



16th cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

# Emergent Self-Model



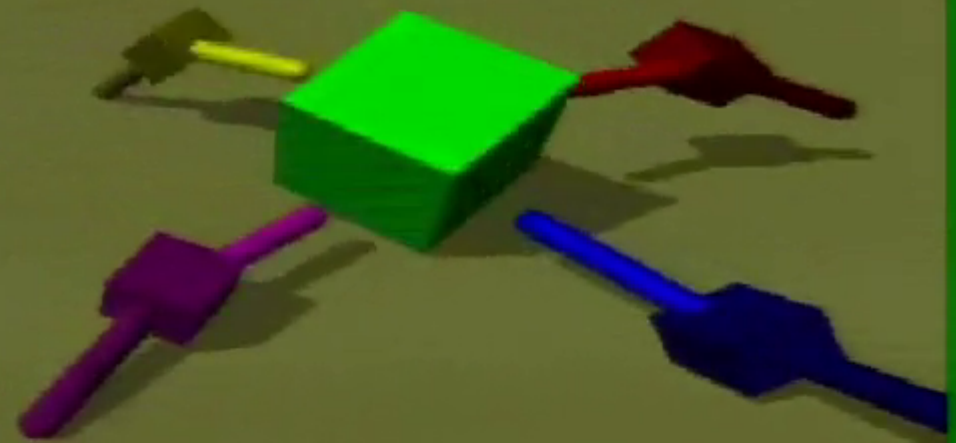
16th cycle (of 16)

With Josh Bongard and Victor Zykov, Science 2006

# Emergent Self-Model

Stage II:  
Using the model to generate a gait

# Emergent Self-Model



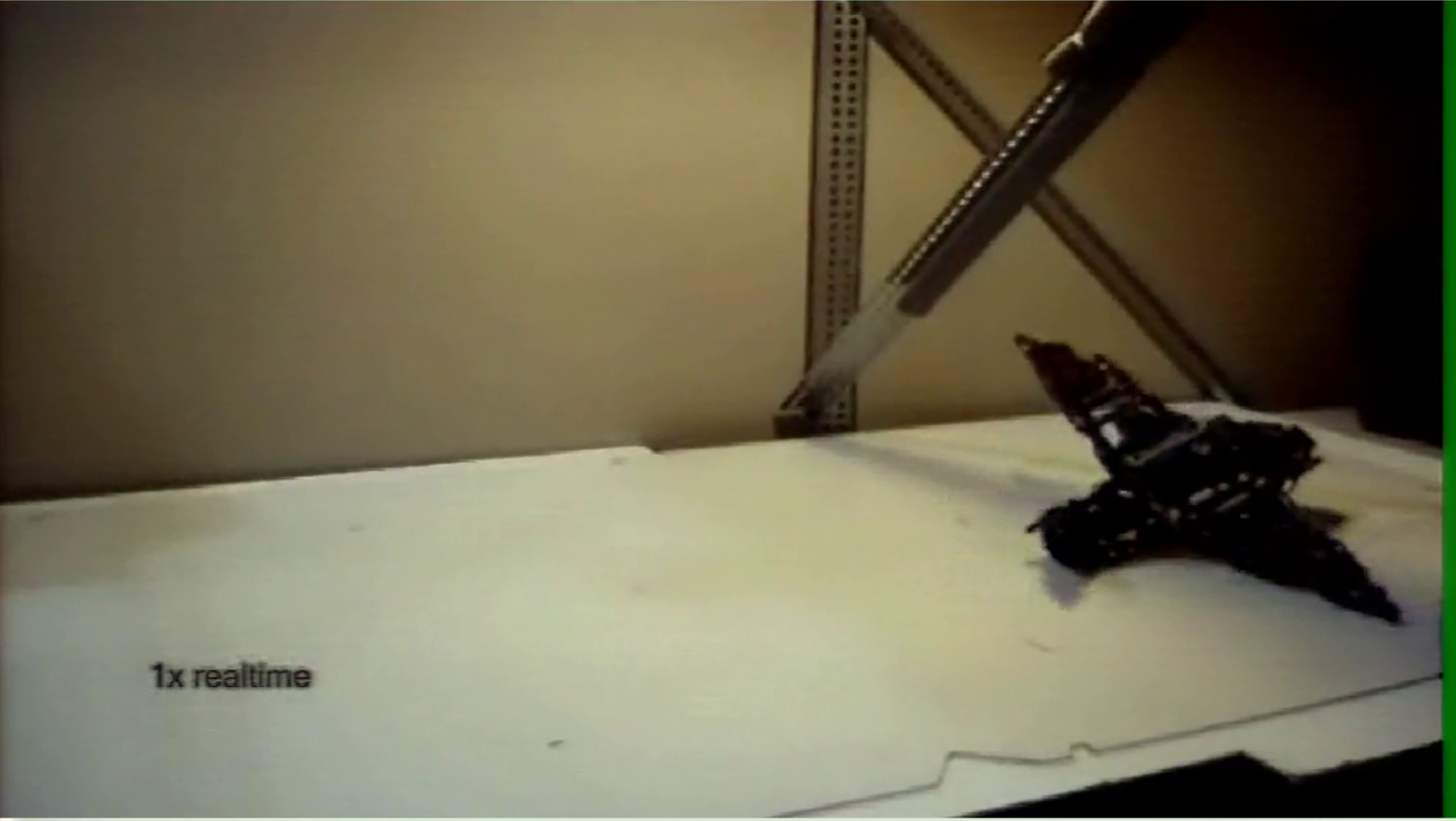
With Josh Bongard and Victor Zykov, Science 2006

# Emergent Self-Model

Stage III:  
Trying gait in reality

With Josh Bongard and Victor Zykov, Science 2006

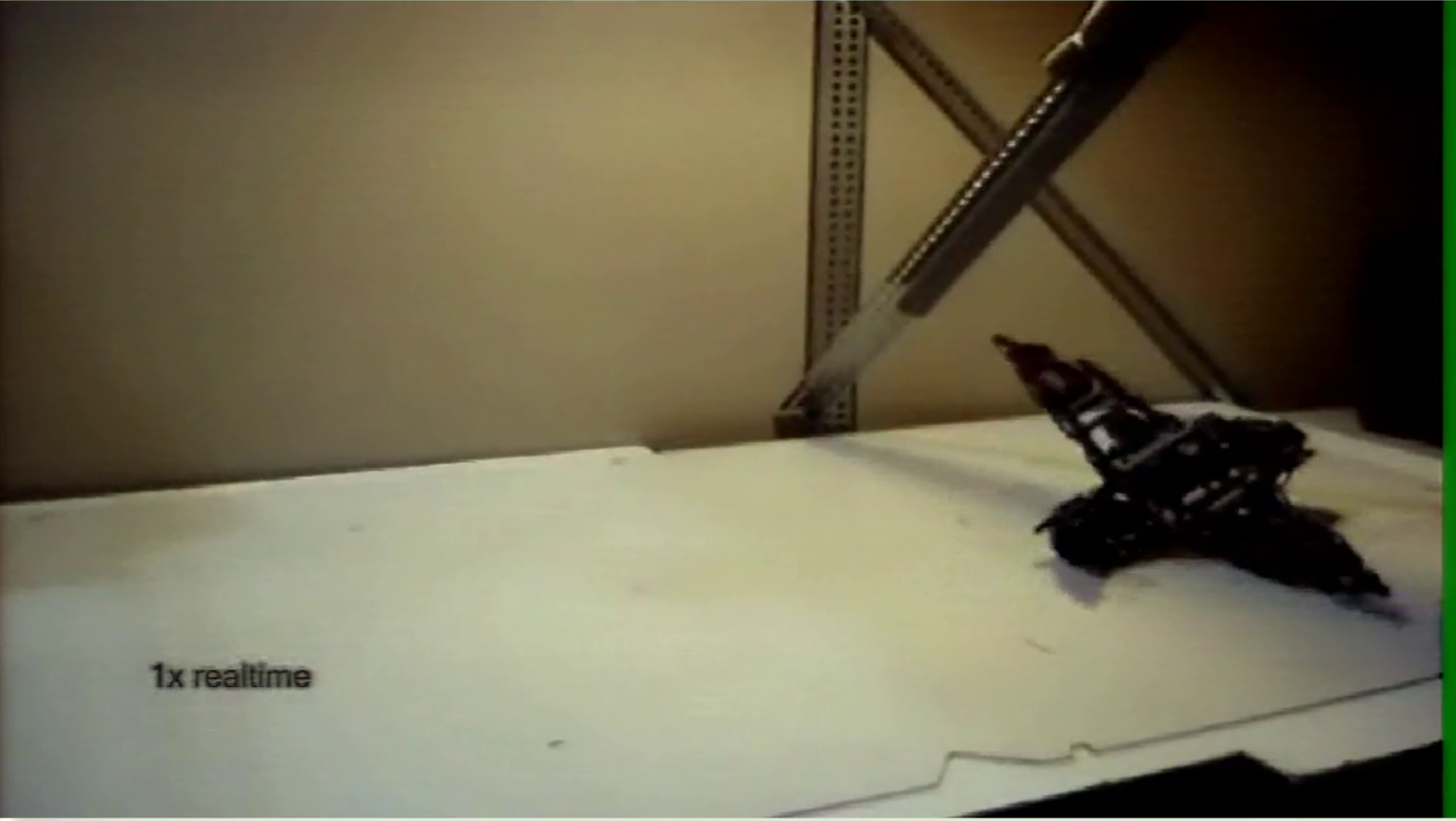
# Emergent Self-Model



1x realtime

With Josh Bongard and Victor Zykov, Science 2006

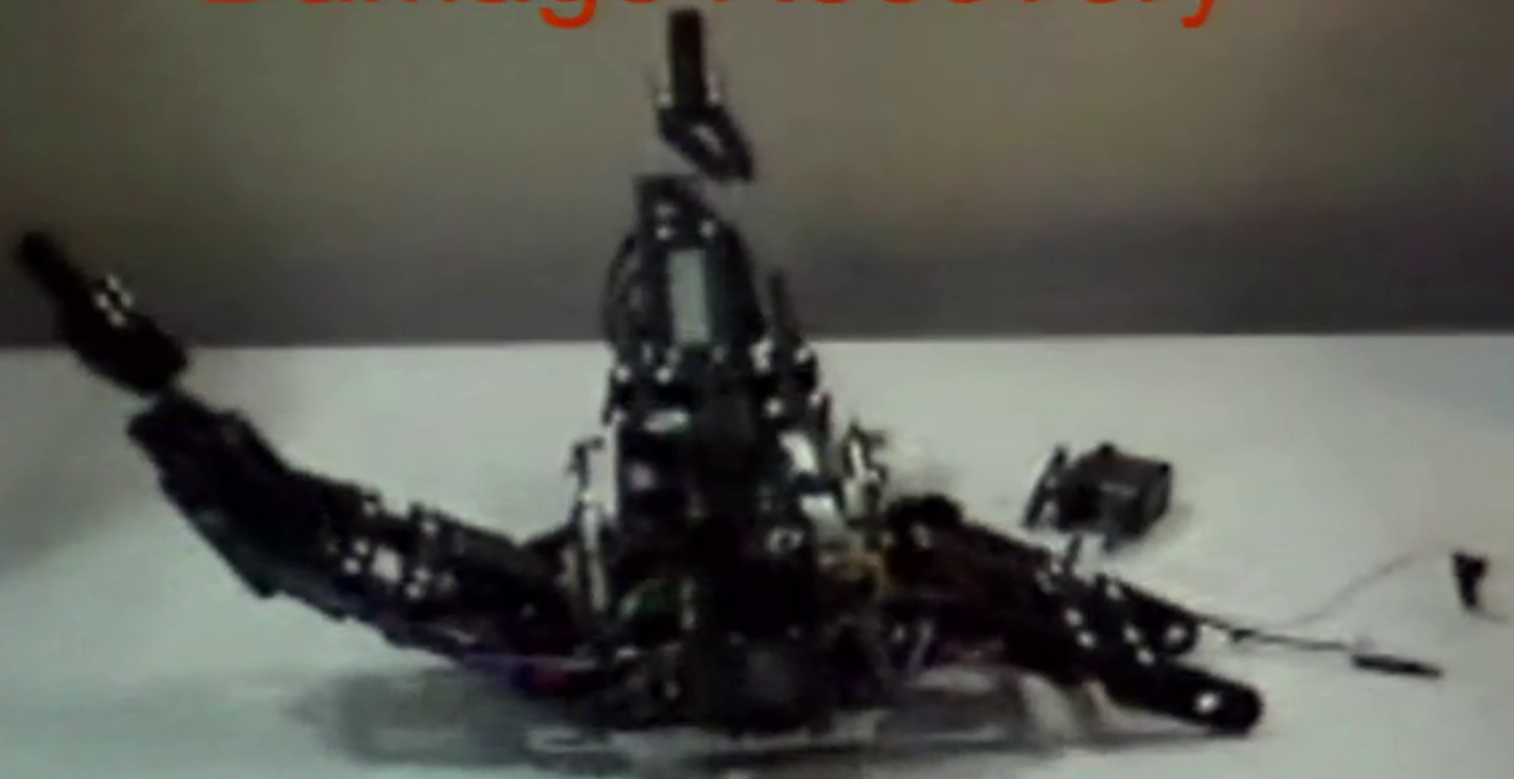
# Emergent Self-Model



1x realtime

With Josh Bongard and Victor Zykov, Science 2006

# Damage Recovery



With Josh Bongard and Victor Zykov, Science 2006

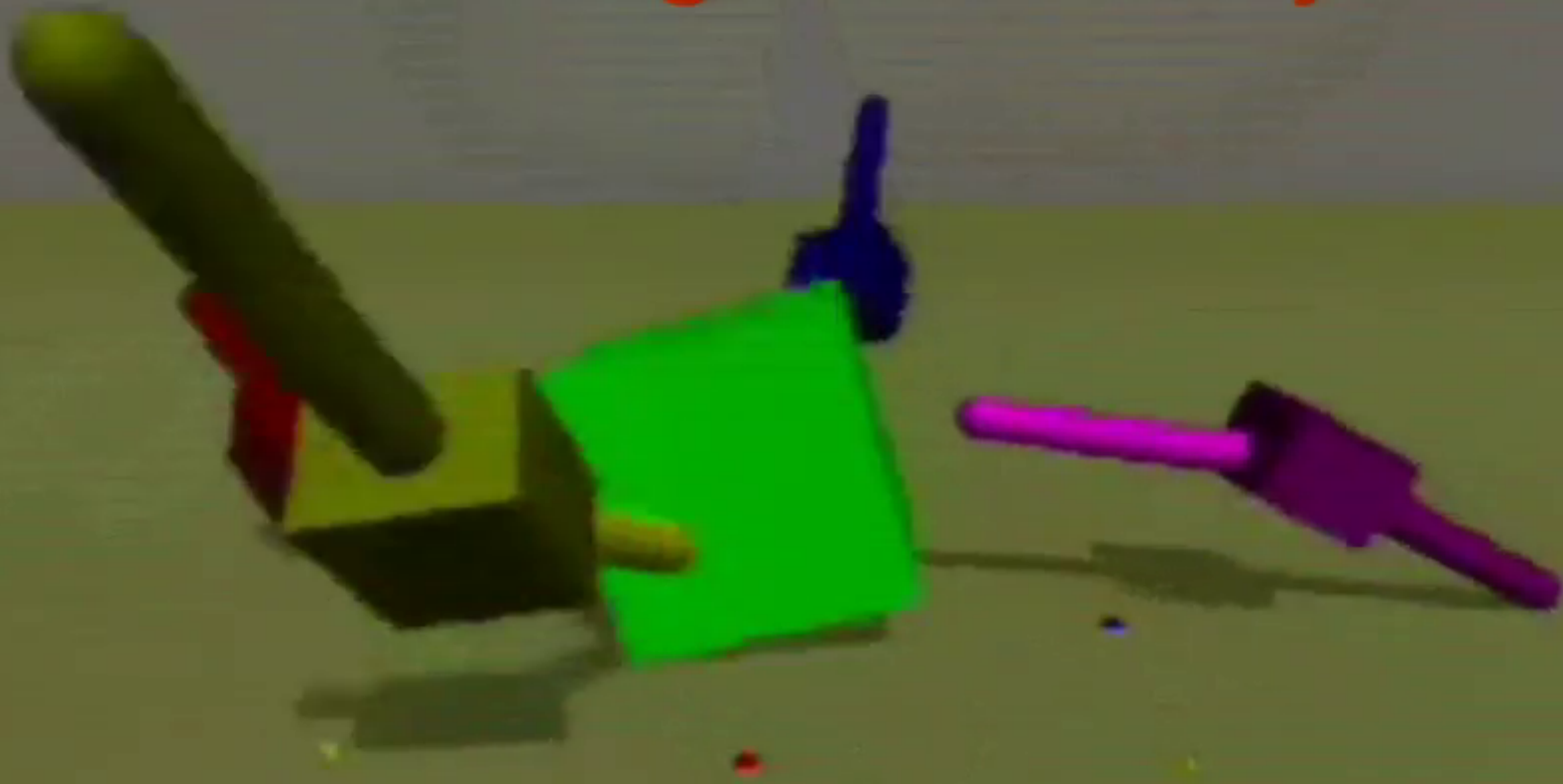


# Damage Recovery



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

# Damage Recovery

Other gaits generated for  
damaged morphology

# Damage Recovery

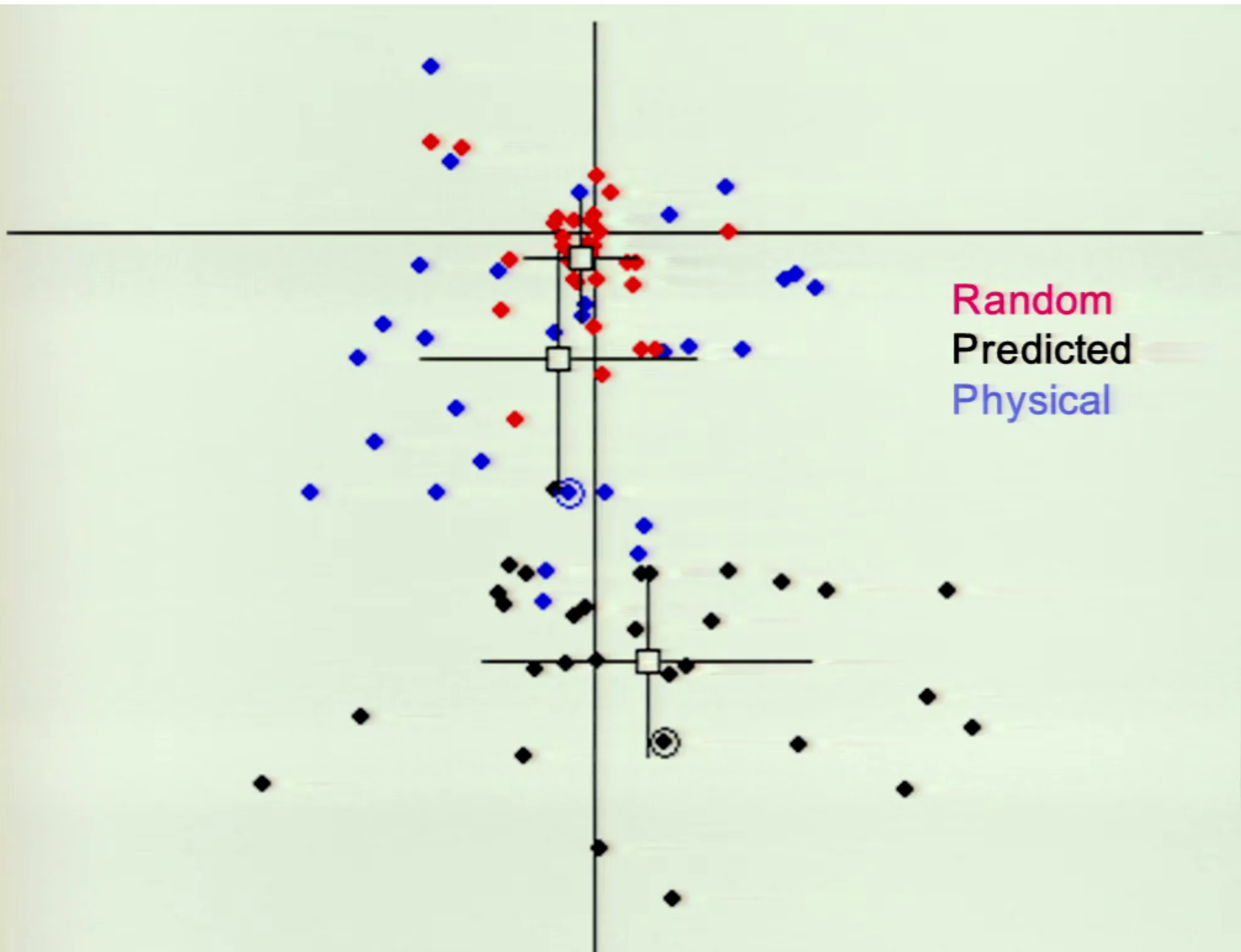


With Josh Bongard and Victor Zykov, Science 2006

# Damage Recovery

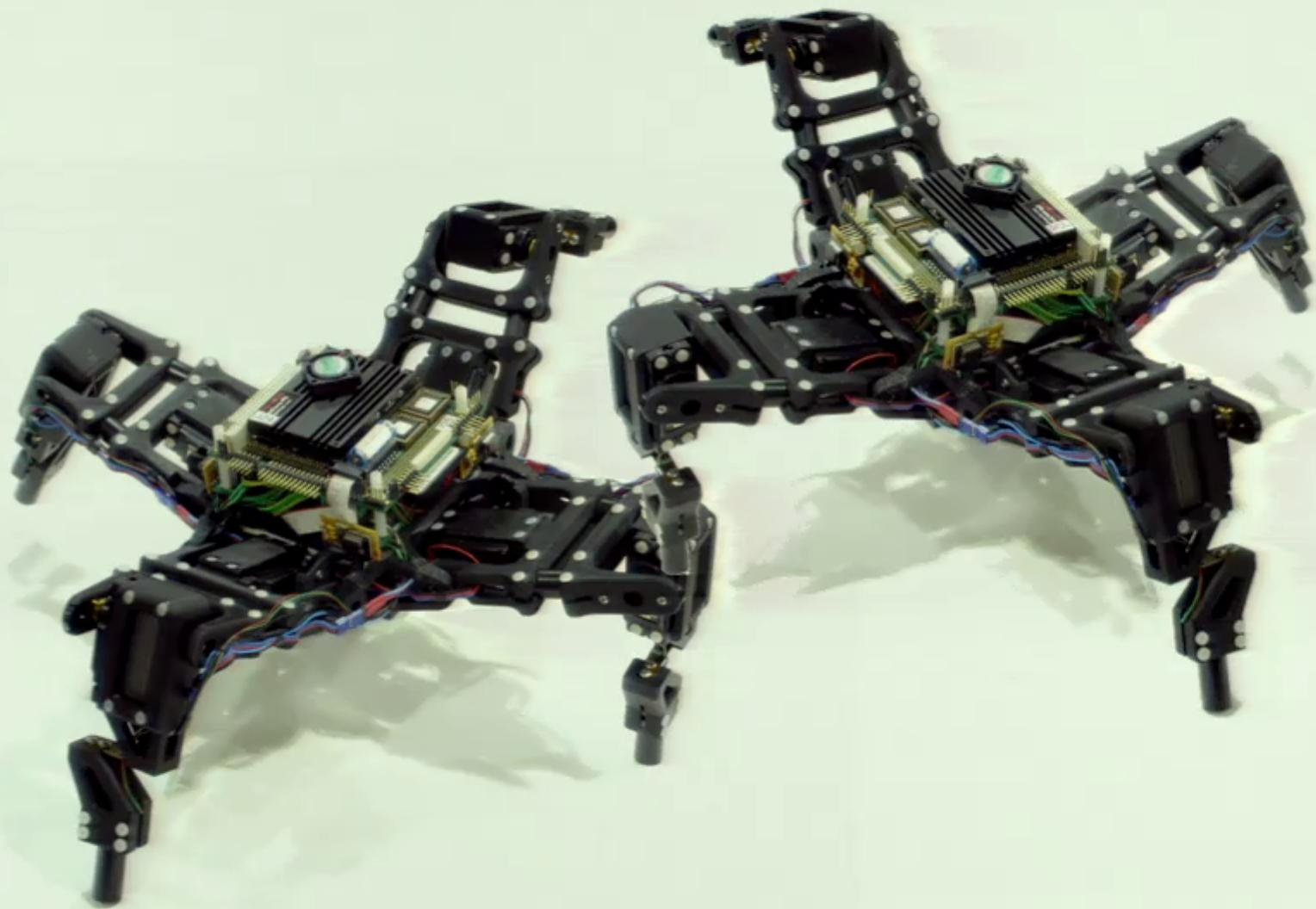


With Josh Bongard and Victor Zykov, Science 2006



Random  
Predicted  
Physical





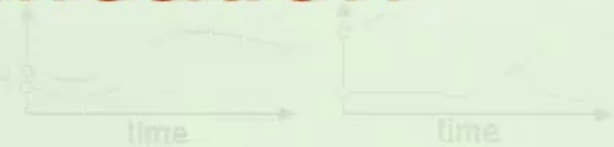
# System Identification

Candidate models

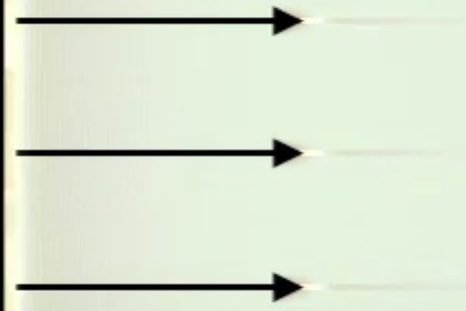
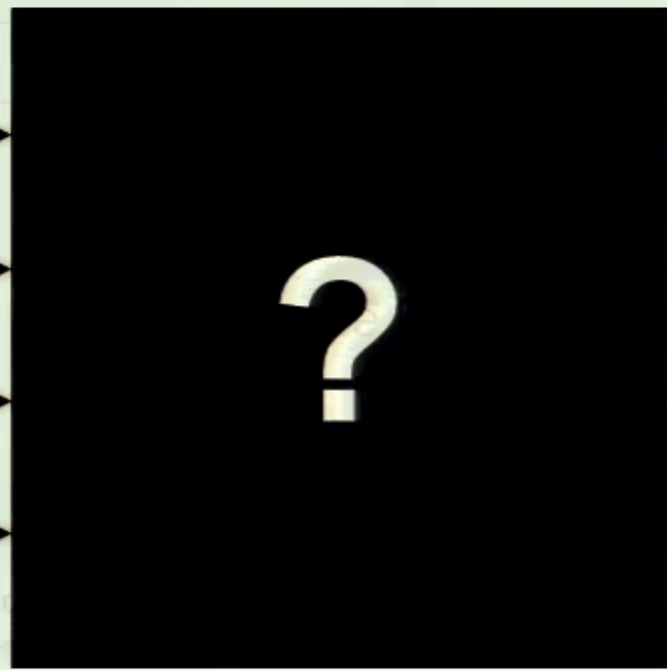
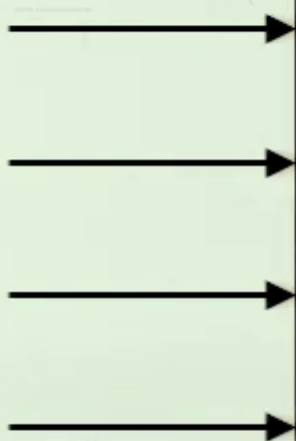
Candidate tests

$$\frac{dx}{dt} = -2x + \sin t$$
$$\frac{dx}{dt} = -x + \frac{1}{t}$$
$$\frac{dx}{dt} = -x + 1$$

$$\frac{dy}{dt} = -ay^b$$
$$\frac{dx}{dt} = -x + 1$$
$$\frac{dy}{dt} = -y + 1$$



Candidate Initial



Or  
(se

ial  
nditions  
tuators)



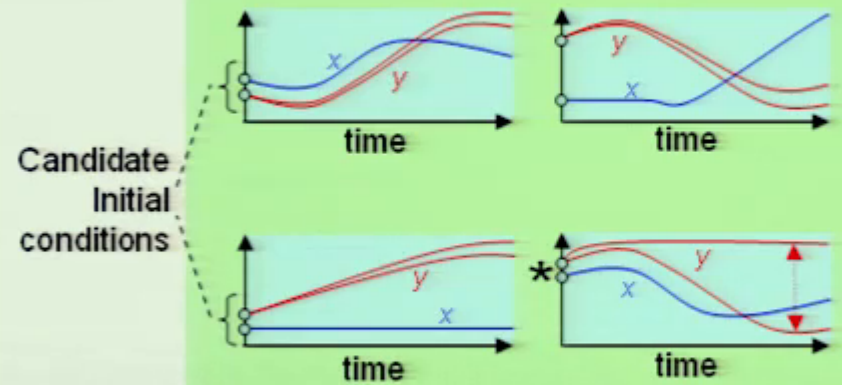
## Candidate models

$$\begin{cases} \frac{dx}{dt} = -2y^2 + \log x \\ \frac{dy}{dt} = -x + \frac{y}{6} \end{cases} \quad \begin{cases} \frac{dx}{dt} = -\sqrt{y} + \frac{x}{5} \\ \frac{dy}{dt} = -\sin y \end{cases}$$

?

$$\begin{cases} \frac{dx}{dt} = -3\frac{y+1}{y-1} \\ \frac{dy}{dt} = -\frac{x^2}{x^2+1} \end{cases} \quad \begin{cases} \frac{dx}{dt} = -y^{1.8} + \log x \\ \frac{dy}{dt} = -x + \frac{y}{4x} \end{cases}$$

## Candidate tests



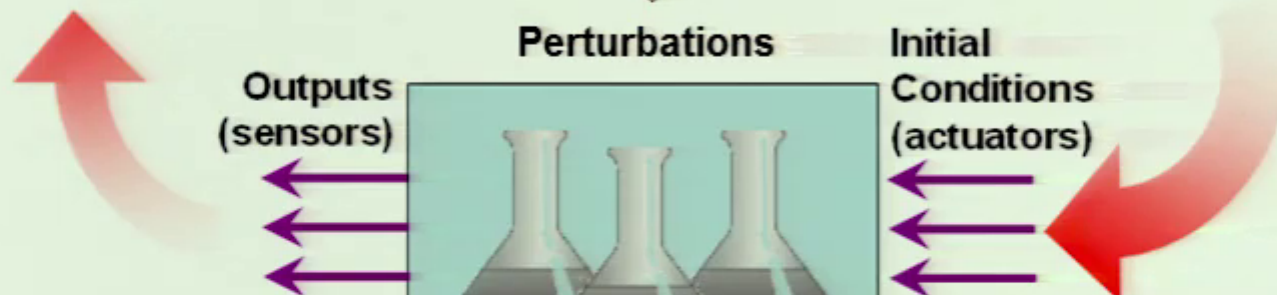
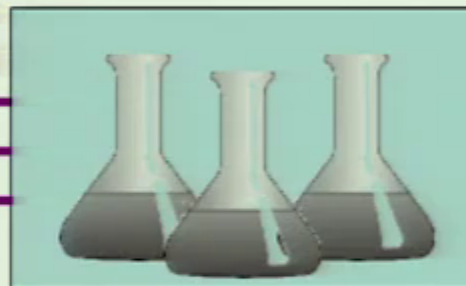
Inference Process



Perturbations

Initial Conditions (actuators)

Outputs (sensors)



## Candidate models

$$\begin{cases} \frac{dx}{dt} = -2y^2 + \log x \\ \frac{dy}{dt} = -x + \frac{y}{6} \end{cases}$$

$$\begin{cases} \frac{dx}{dt} = -\sqrt{y} + \frac{x}{5} \\ \frac{dy}{dt} = -\sin y \end{cases}$$

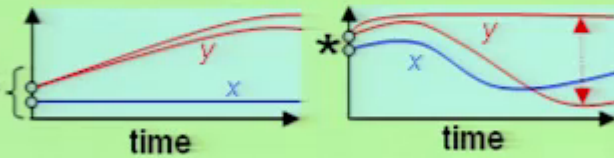
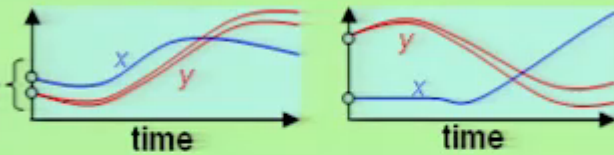
$$\begin{cases} \frac{dx}{dt} = -3\frac{y+1}{y-1} \\ \frac{dy}{dt} = -\frac{x^2}{x^2+1} \end{cases}$$

?

$$\begin{cases} \frac{dx}{dt} = -y^{1.8} + \log x \\ \frac{dy}{dt} = -x + \frac{y}{4x} \end{cases}$$

## Candidate tests

Candidate Initial conditions

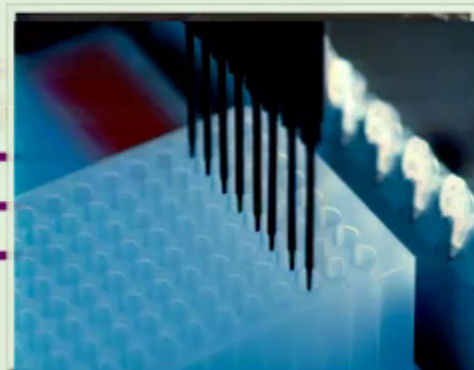


c

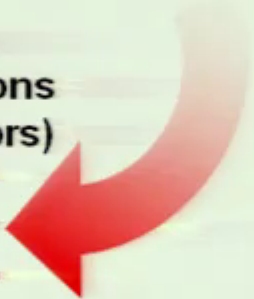
Inference Process



Outputs (sensors)



Initial Conditions (actuators)



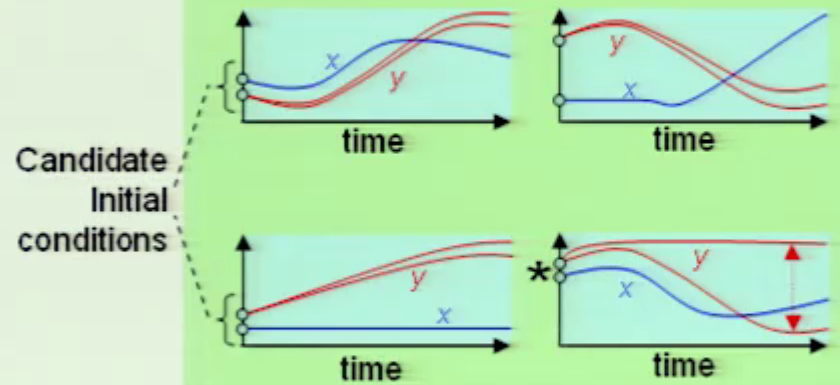
## Candidate models

$$\begin{cases} \frac{dx}{dt} = -2y^2 + \log x \\ \frac{dy}{dt} = -x + \frac{y}{6} \end{cases} \quad \begin{cases} \frac{dx}{dt} = -\sqrt{y} + \frac{x}{5} \\ \frac{dy}{dt} = -\sin y \end{cases}$$

?

$$\begin{cases} \frac{dx}{dt} = -3\frac{y+1}{y-1} \\ \frac{dy}{dt} = -\frac{x^2}{x^2+1} \end{cases} \quad \begin{cases} \frac{dx}{dt} = -y^{1.8} + \log x \\ \frac{dy}{dt} = -x + \frac{y}{4x} \end{cases}$$

## Candidate tests



Inference Process



Outputs  
(sensors)



Time	x	y	z	...
0.0	1.0	1.0	1.0	...
0.1	0.99	0.99	0.99	...
0.2	0.98	0.98	0.98	...
0.3	0.97	0.97	0.97	...
0.4	0.96	0.96	0.96	...
0.5	0.95	0.95	0.95	...
0.6	0.94	0.94	0.94	...
0.7	0.93	0.93	0.93	...
0.8	0.92	0.92	0.92	...
0.9	0.91	0.91	0.91	...
1.0	0.90	0.90	0.90	...
1.1	0.89	0.89	0.89	...
1.2	0.88	0.88	0.88	...
1.3	0.87	0.87	0.87	...
1.4	0.86	0.86	0.86	...
1.5	0.85	0.85	0.85	...
1.6	0.84	0.84	0.84	...
1.7	0.83	0.83	0.83	...
1.8	0.82	0.82	0.82	...
1.9	0.81	0.81	0.81	...
2.0	0.80	0.80	0.80	...
2.1	0.79	0.79	0.79	...
2.2	0.78	0.78	0.78	...
2.3	0.77	0.77	0.77	...
2.4	0.76	0.76	0.76	...
2.5	0.75	0.75	0.75	...
2.6	0.74	0.74	0.74	...
2.7	0.73	0.73	0.73	...
2.8	0.72	0.72	0.72	...
2.9	0.71	0.71	0.71	...
3.0	0.70	0.70	0.70	...
3.1	0.69	0.69	0.69	...
3.2	0.68	0.68	0.68	...
3.3	0.67	0.67	0.67	...
3.4	0.66	0.66	0.66	...
3.5	0.65	0.65	0.65	...
3.6	0.64	0.64	0.64	...
3.7	0.63	0.63	0.63	...
3.8	0.62	0.62	0.62	...
3.9	0.61	0.61	0.61	...
4.0	0.60	0.60	0.60	...
4.1	0.59	0.59	0.59	...
4.2	0.58	0.58	0.58	...
4.3	0.57	0.57	0.57	...
4.4	0.56	0.56	0.56	...
4.5	0.55	0.55	0.55	...
4.6	0.54	0.54	0.54	...
4.7	0.53	0.53	0.53	...
4.8	0.52	0.52	0.52	...
4.9	0.51	0.51	0.51	...
5.0	0.50	0.50	0.50	...
5.1	0.49	0.49	0.49	...
5.2	0.48	0.48	0.48	...
5.3	0.47	0.47	0.47	...
5.4	0.46	0.46	0.46	...
5.5	0.45	0.45	0.45	...
5.6	0.44	0.44	0.44	...
5.7	0.43	0.43	0.43	...
5.8	0.42	0.42	0.42	...
5.9	0.41	0.41	0.41	...
6.0	0.40	0.40	0.40	...
6.1	0.39	0.39	0.39	...
6.2	0.38	0.38	0.38	...
6.3	0.37	0.37	0.37	...
6.4	0.36	0.36	0.36	...
6.5	0.35	0.35	0.35	...
6.6	0.34	0.34	0.34	...
6.7	0.33	0.33	0.33	...
6.8	0.32	0.32	0.32	...
6.9	0.31	0.31	0.31	...
7.0	0.30	0.30	0.30	...
7.1	0.29	0.29	0.29	...
7.2	0.28	0.28	0.28	...
7.3	0.27	0.27	0.27	...
7.4	0.26	0.26	0.26	...
7.5	0.25	0.25	0.25	...
7.6	0.24	0.24	0.24	...
7.7	0.23	0.23	0.23	...
7.8	0.22	0.22	0.22	...
7.9	0.21	0.21	0.21	...
8.0	0.20	0.20	0.20	...
8.1	0.19	0.19	0.19	...
8.2	0.18	0.18	0.18	...
8.3	0.17	0.17	0.17	...
8.4	0.16	0.16	0.16	...
8.5	0.15	0.15	0.15	...
8.6	0.14	0.14	0.14	...
8.7	0.13	0.13	0.13	...
8.8	0.12	0.12	0.12	...
8.9	0.11	0.11	0.11	...
9.0	0.10	0.10	0.10	...
9.1	0.09	0.09	0.09	...
9.2	0.08	0.08	0.08	...
9.3	0.07	0.07	0.07	...
9.4	0.06	0.06	0.06	...
9.5	0.05	0.05	0.05	...
9.6	0.04	0.04	0.04	...
9.7	0.03	0.03	0.03	...
9.8	0.02	0.02	0.02	...
9.9	0.01	0.01	0.01	...
10.0	0.00	0.00	0.00	...

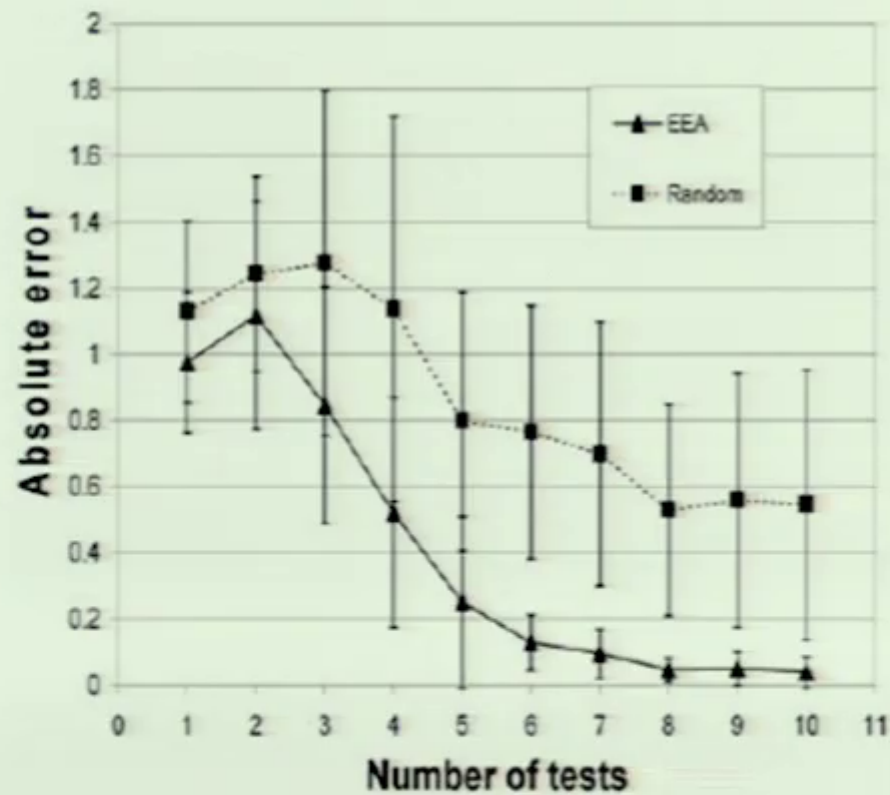
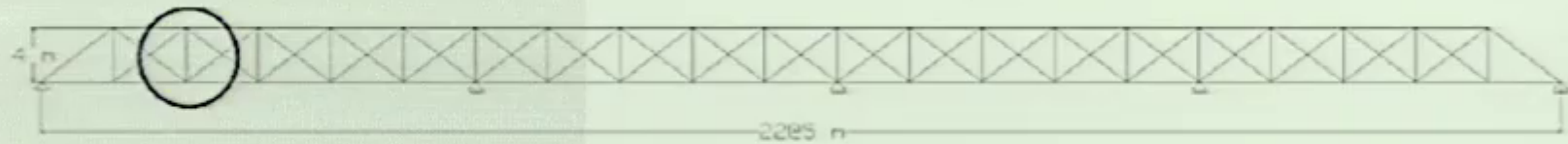
Initial  
Conditions  
(actuators)



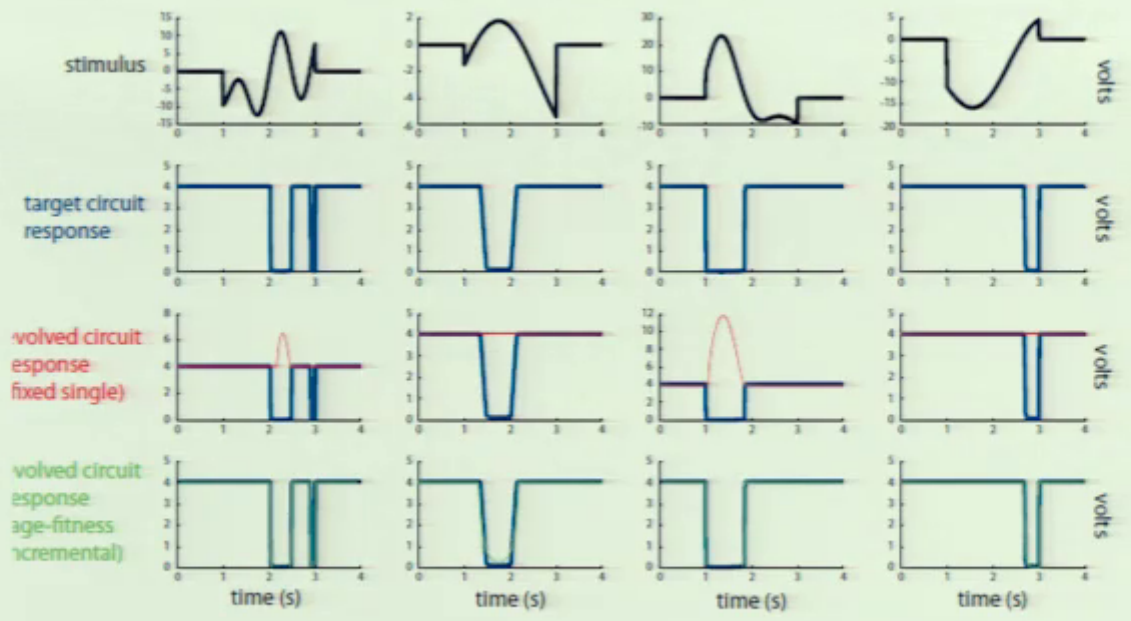
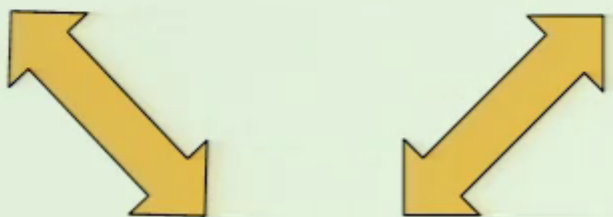
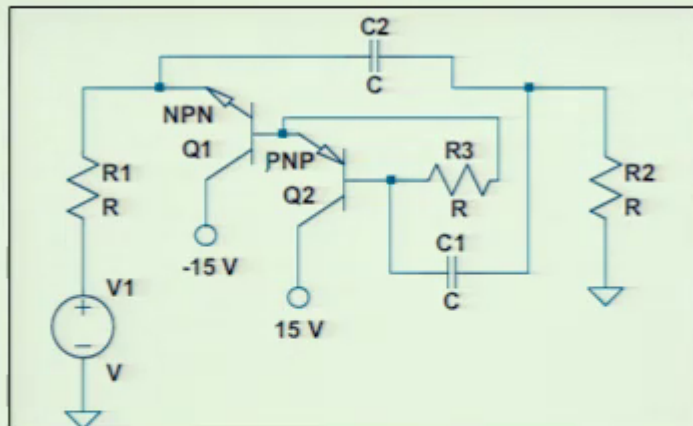
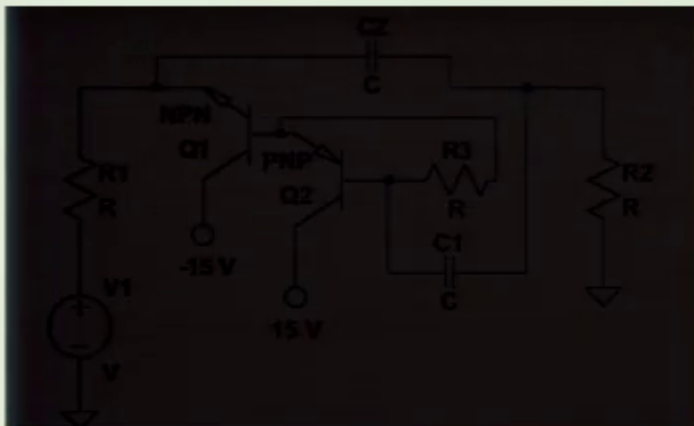


Photo: Floris van Breugel

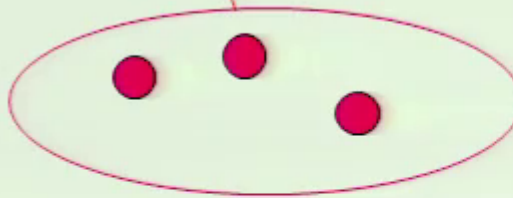
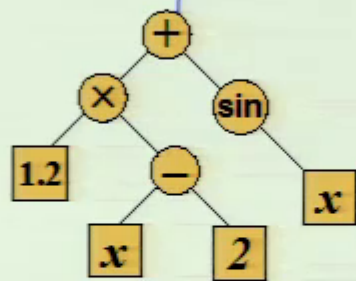
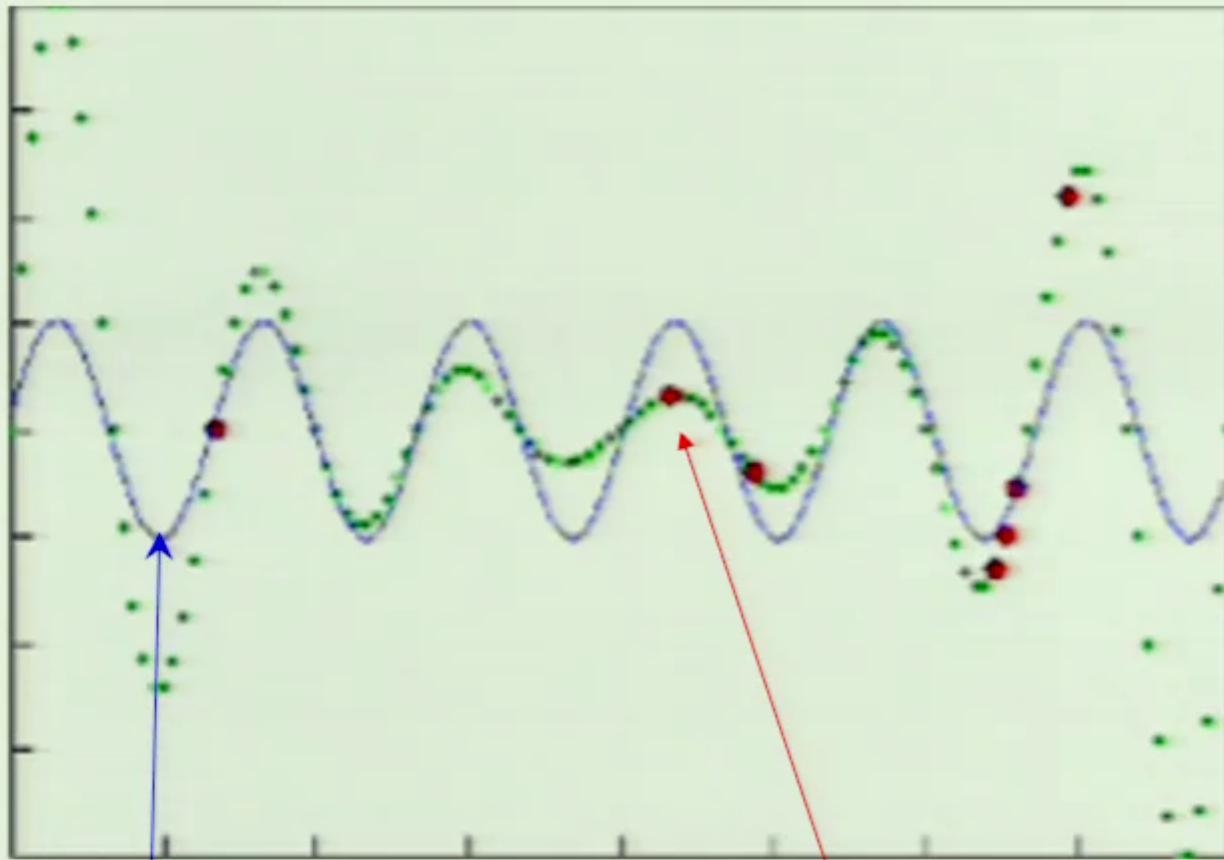
# Structural Damage Diagnosis



With Wilkins Aquino







**Models:** Expression trees  
Subject to mutation and selection

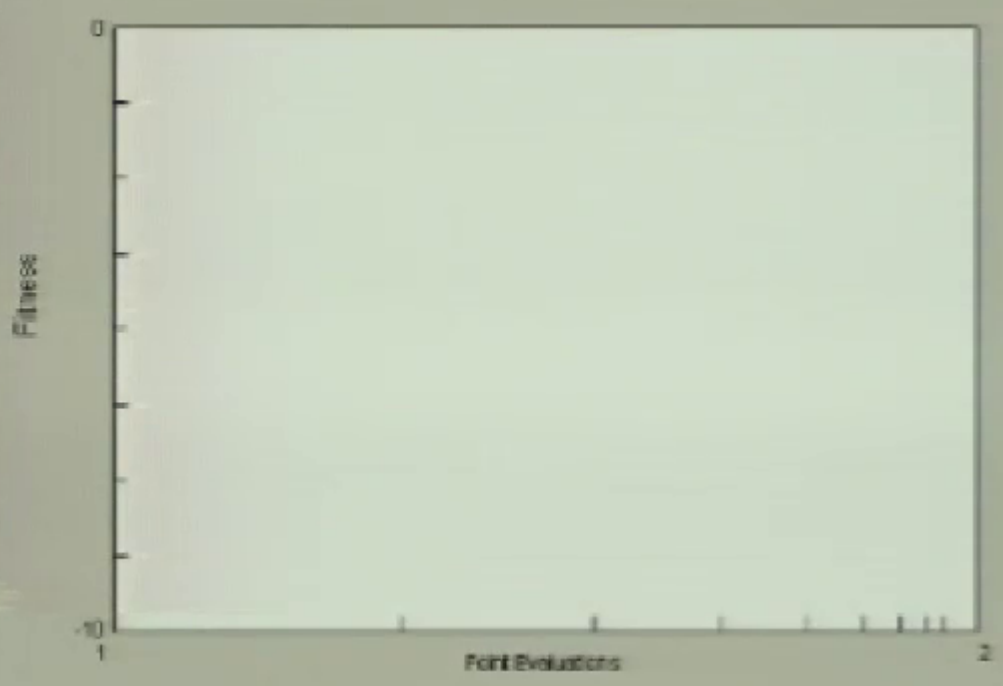
**Experiments:** Data-points  
Subject to mutation and selection

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

Date

Name	Value	Rate
Running Time		
Evolution Generations		
Fitness Model Generations		
Exact Fitness Calculations		
Model Fitness Evaluations		
Total Point Evaluations		

```
[15:48:04] -----  
[15:48:04] CoEvolutionary Symbolic Regression v3.1  
[15:48:04] by:  
[15:48:04]   Michael Schmitt (schm57@cornell.edu)  
[15:48:04]   Nedispenn (nedispenn@cornell.edu)  
[15:48:04] Cornell University (2006)  
[15:48:04] -----  
[15:48:04] Seed=00203025
```



Corneil Symbolic Regression

File Settings Help

Date:

Name	Value	Rate
Running Time		
Solution Generations		
Fitness Model Generations		
Exact Fitness Calculations		
Modelled Fitness Evaluations		
Total Point Evaluations		

-3.4e+000

Open

Look in: datasets

- exp(abs(x))sin(2p(x)).bit
- glycolysis\_A3.txt
- glycolysis\_A3\_at.txt
- glycolysis\_A3\_usage.txt
- glycolysis\_all.txt
- glycolysis\_d5.txt
- glycolysis\_d5.txt
- glycolysis\_d5\_1.txt
- glycolysis\_d5\_1\_progress.txt
- glycolysis\_d5\_1\_usage.txt
- glycolysis\_d5.txt

File name:

Files of type: Text Data Files (\*.txt;\*.dat;\*.dta;\*.lev;\*.levn)

Open as read-only

Open Cancel

```

[15:48:34] -----
[15:48:34] Corneil Symbolic
[15:48:34] by:
[15:48:34] Michael Schmitt (schm
[15:48:34] Med. Informat. (Med. Infa
[15:48:34] Corneil University 1006
[15:48:34] -----
[15:48:34] Seed=80203025
  
```

My Recent Documents

Desktop

My Documents

My Computer

My Network Places

10

-10

1

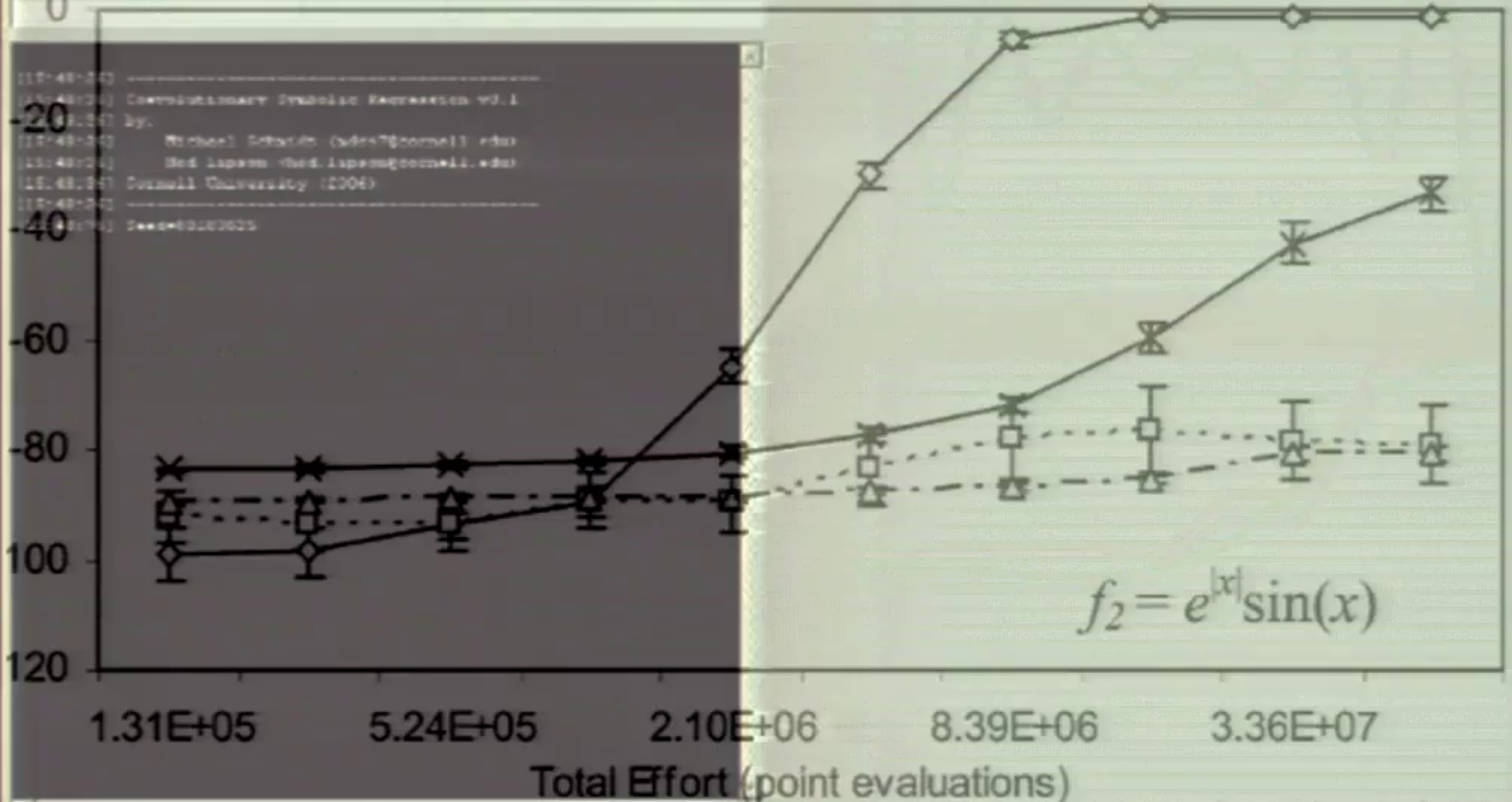
2

Point Evaluations

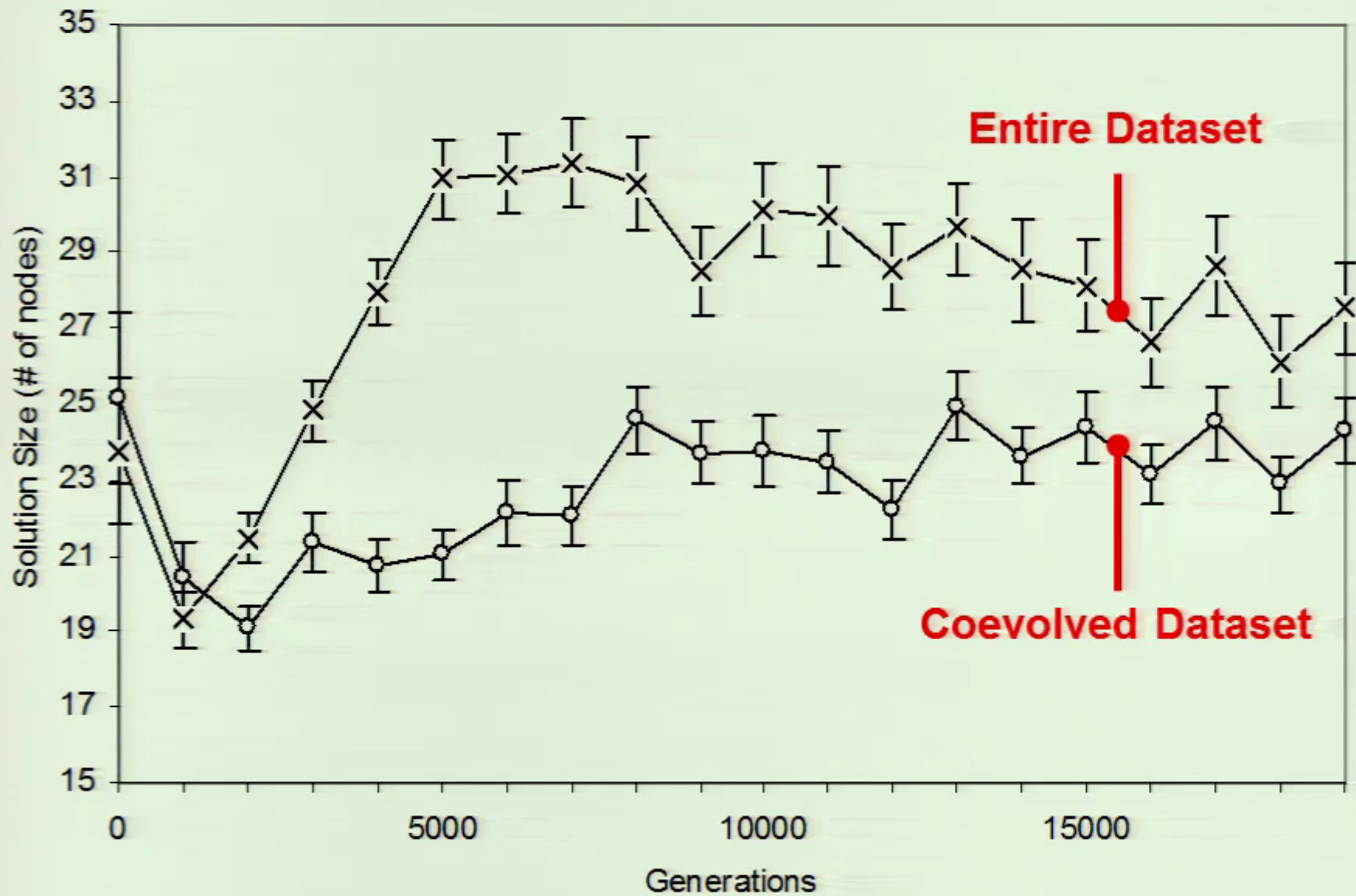
Run Pause Stop

# Solution Accuracy

Name	Value	Ratio
Running Time		
Solution Generations		
Fitness Model Generations		
Exact Fitness Calculations		
Model Fitness Evaluations		
Total Point Evaluations	0	



# Solution Complexity





# Inferring Biological Networks

$$\frac{dS_1}{dt} = \frac{2.5}{J_0} - 100 \left( \frac{S_1 * A_3}{1 + 13.6769 * A_3^4} \right)$$

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

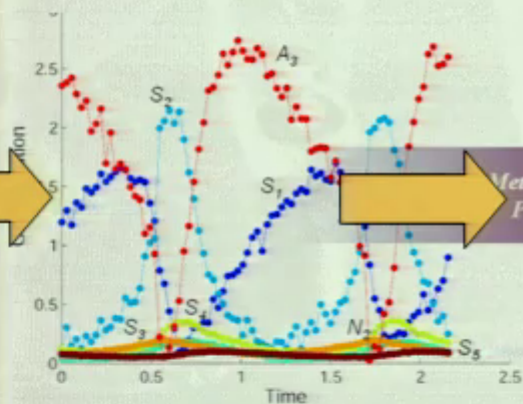
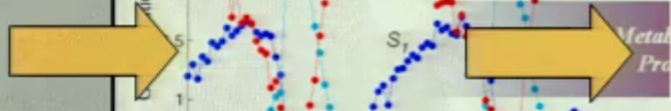
$$\frac{dS_3}{dt} = 6 * S_2 * N_1 - 16 * S_3 * A_2$$

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

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

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

$$\frac{dS_5}{dt} = -1.3 * S_5$$



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

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

$$\frac{dS_3}{dt} = 5.99857 * S_2 * N_1 + 15.99606 * S_3 * A_2 - 0.01286 * S_3$$

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

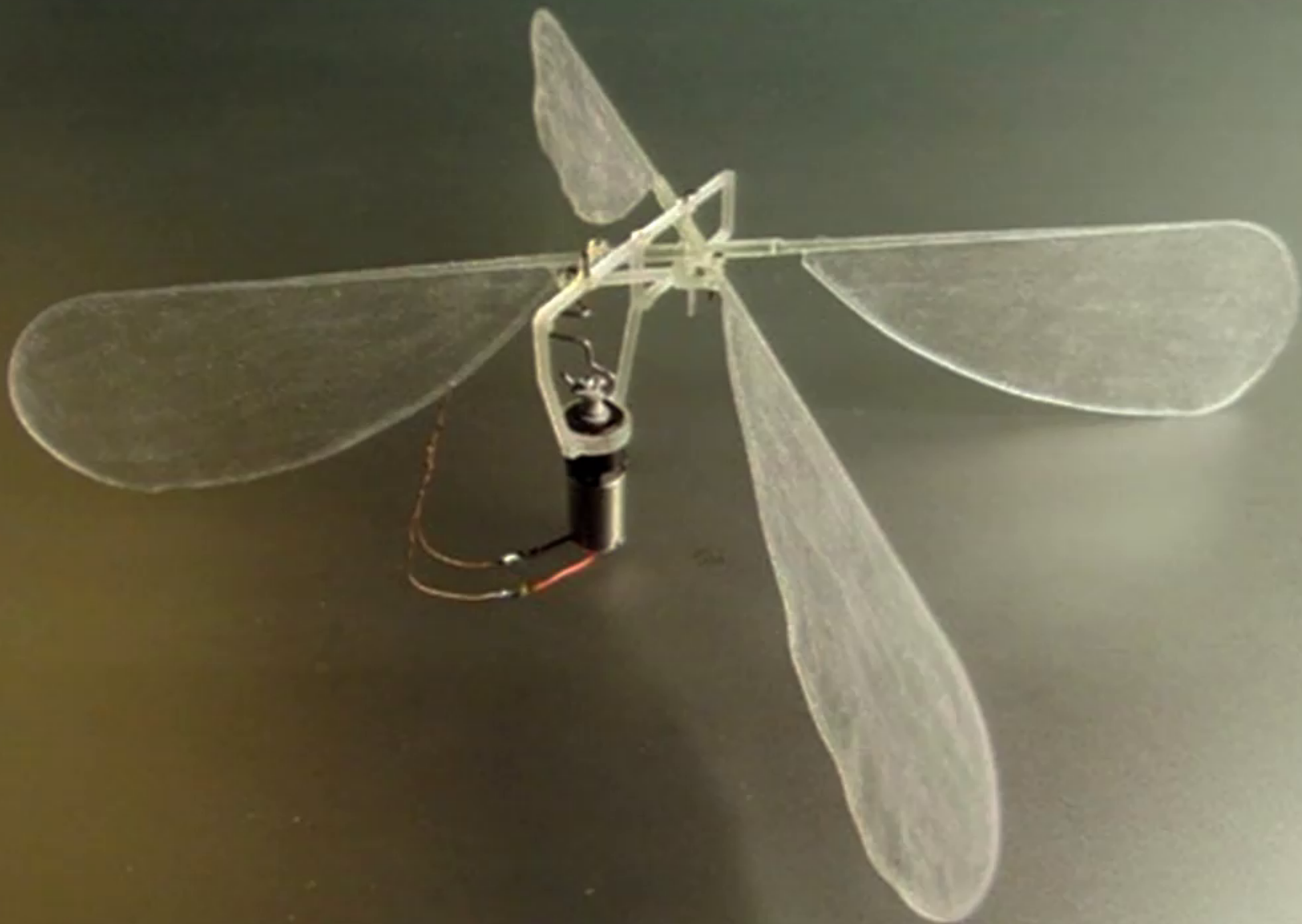
$$\frac{dN_2}{dt} = 5.99857 * S_2 * N_1 - 99.9963 * N_2 * S_4$$

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

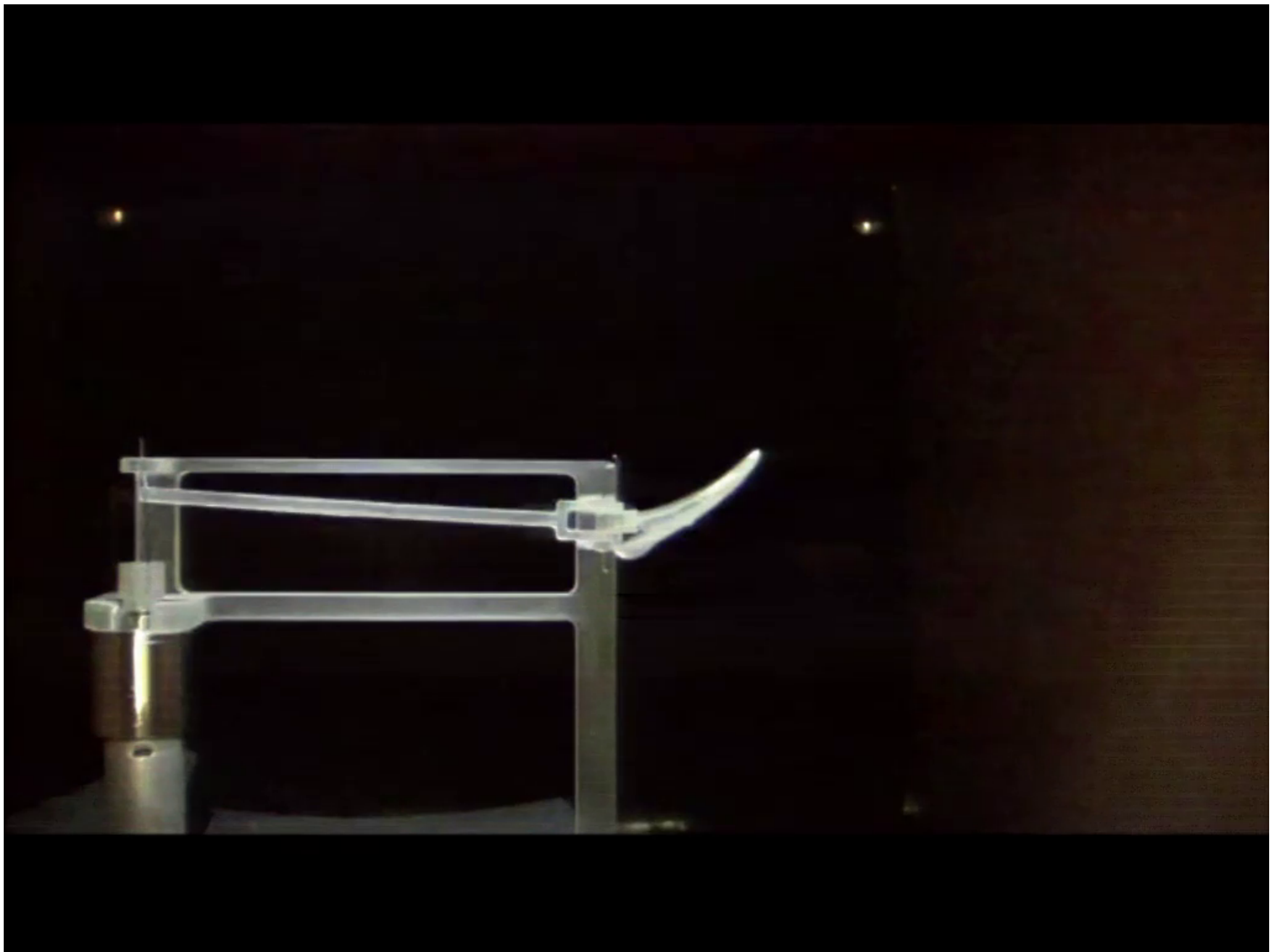
$$\frac{dS_5}{dt} = -1.29626 * S_5$$

Original Equations

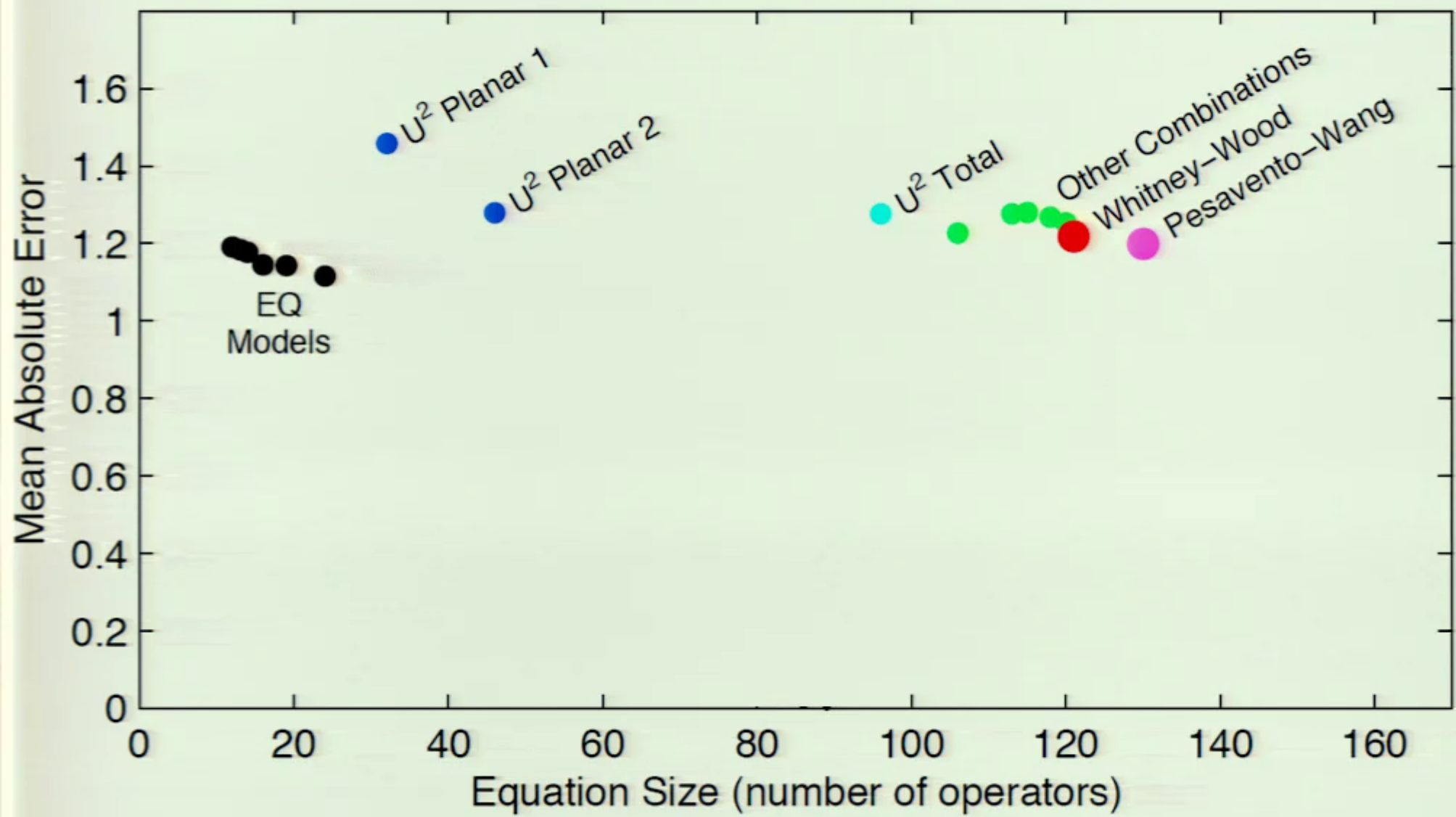
Inferred Equations

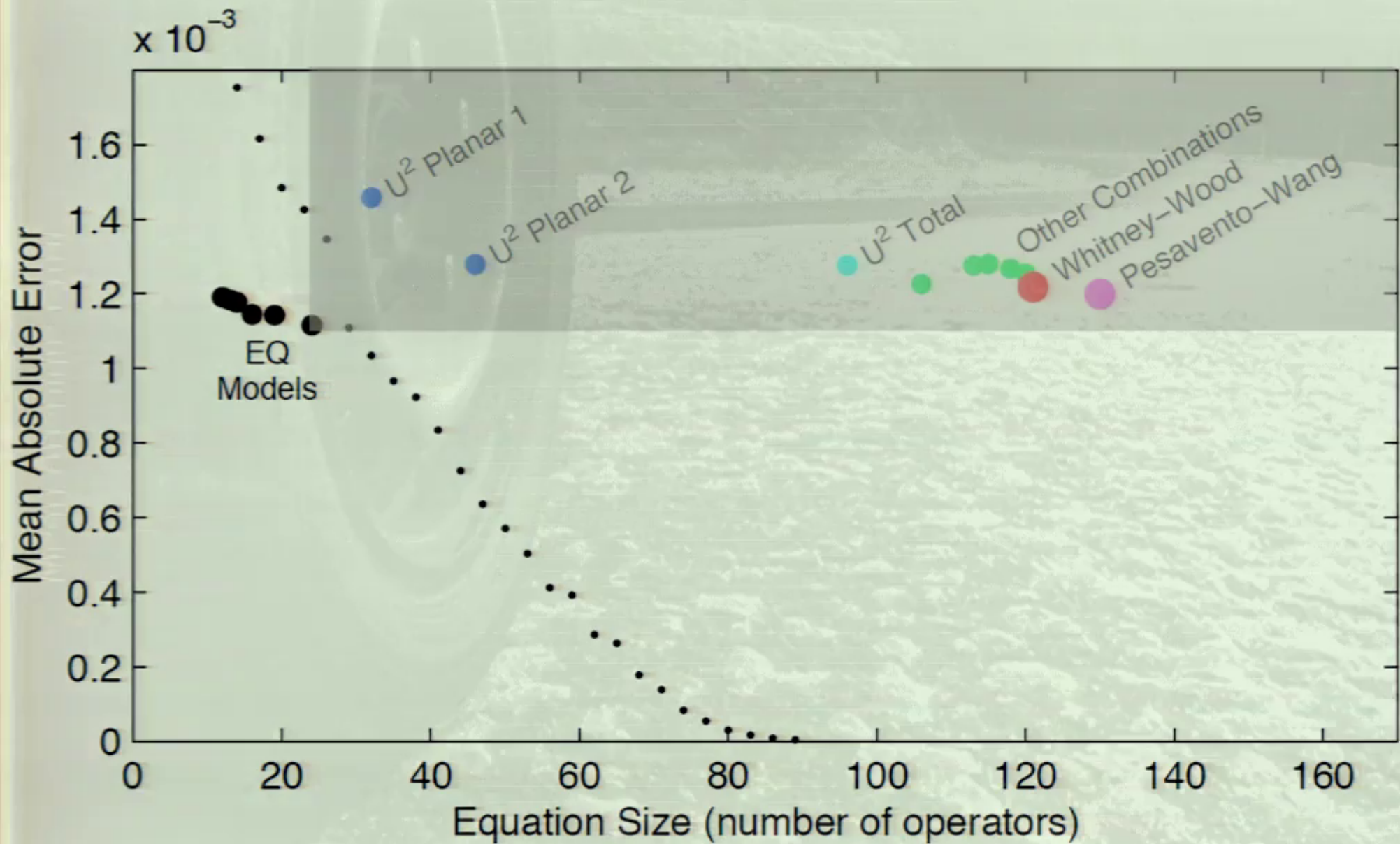






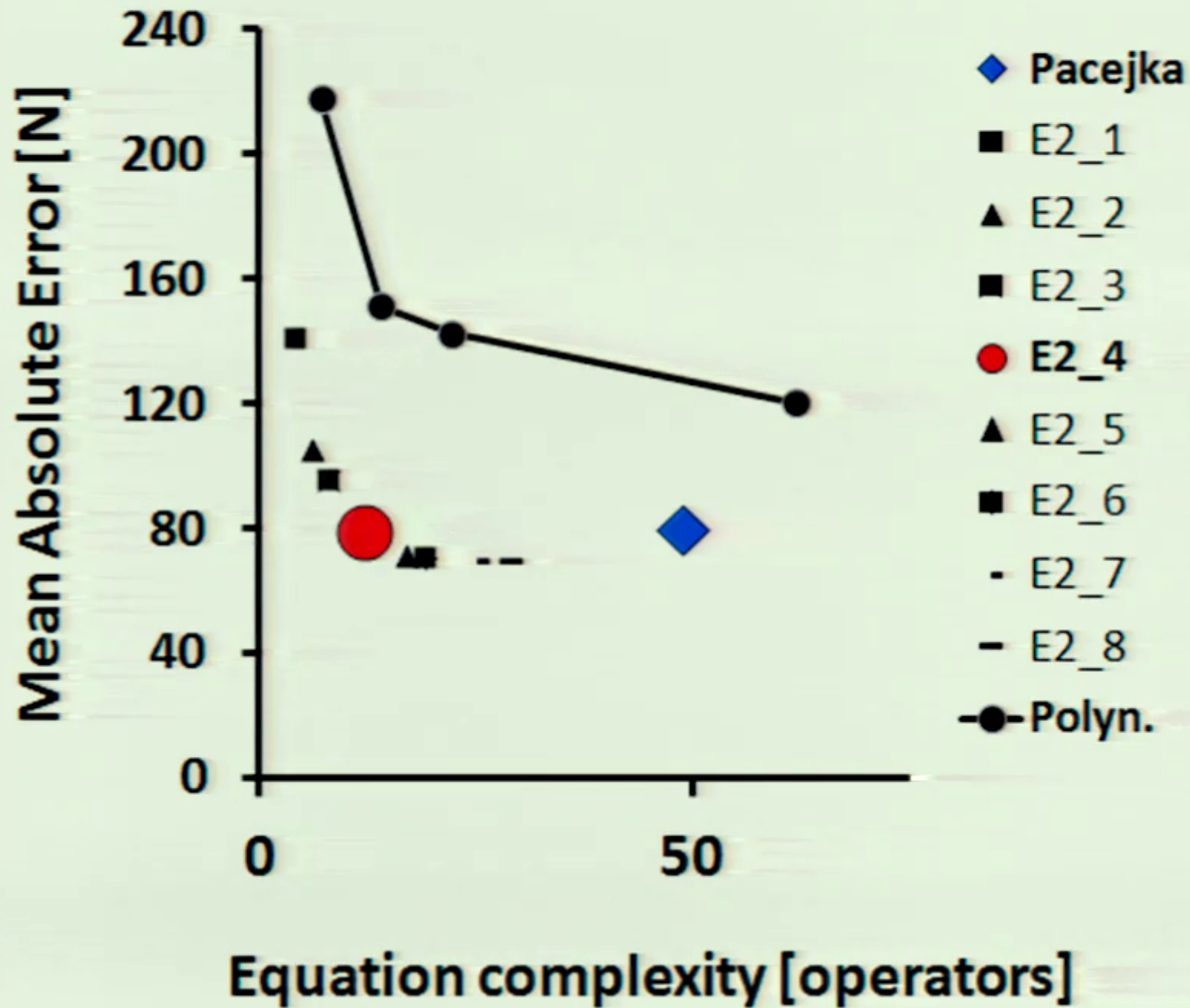
$\times 10^{-3}$



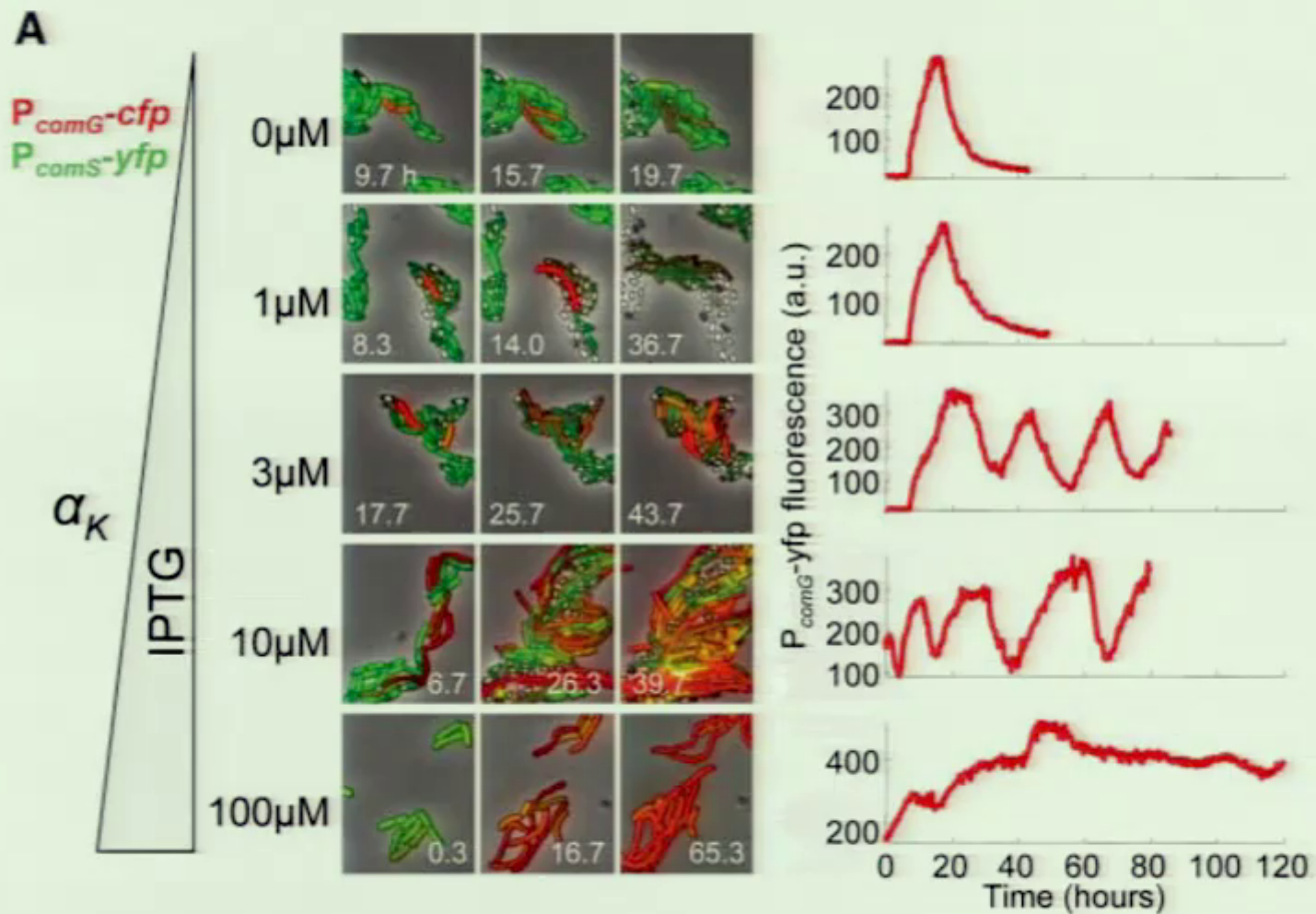




$FY_N$



# Wet Data, Unknown System



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

Symbolic Regression Inferred *Time-Delay* Model:

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

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

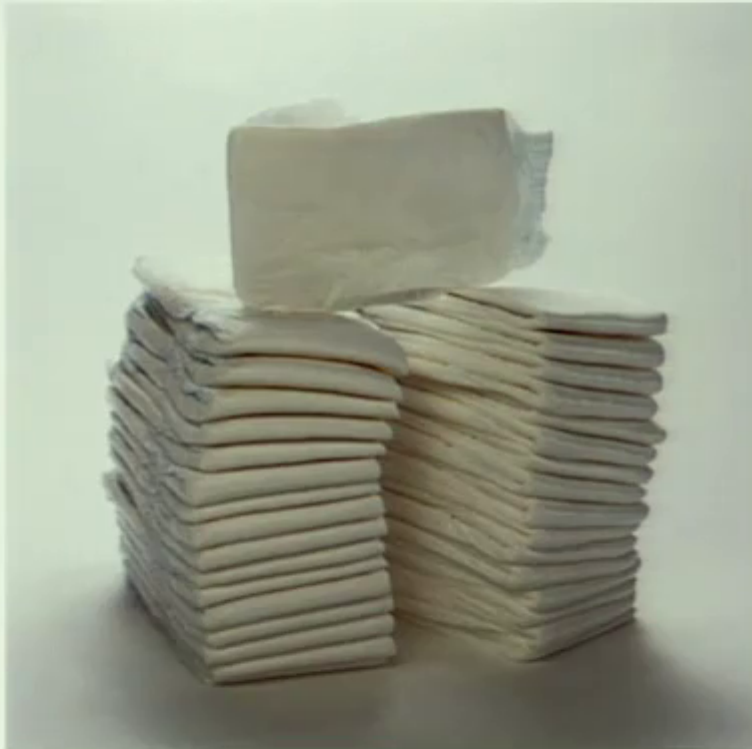
Searching for meaning



# Searching for meaning

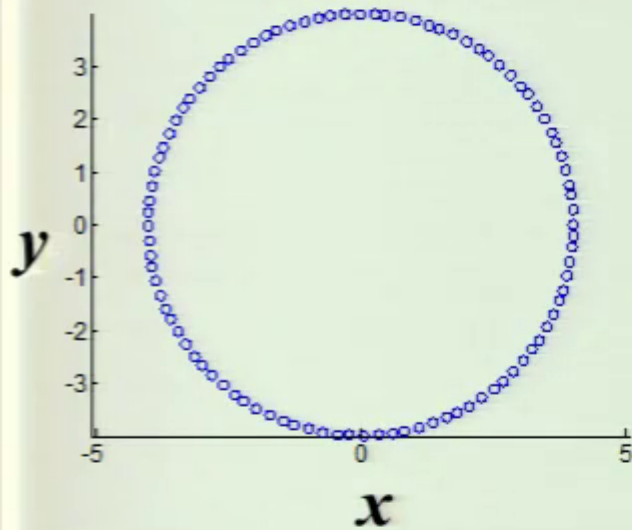


# Correlations



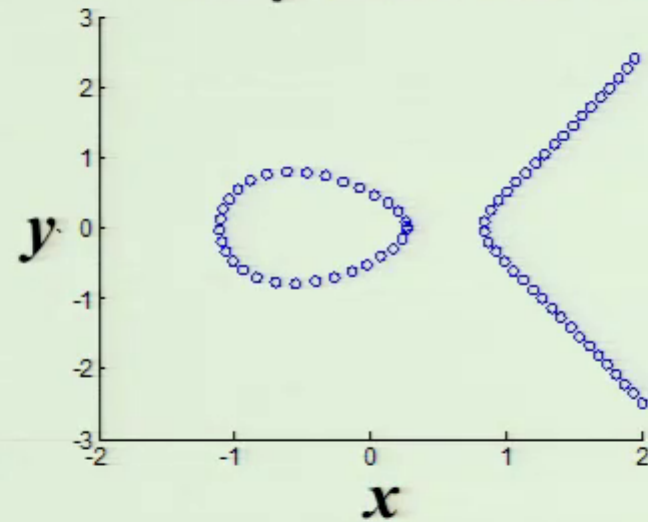
# Homework

Circle



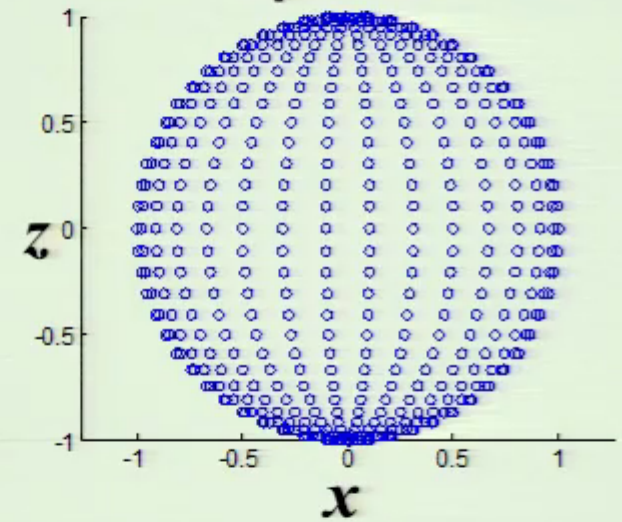
$$x^2 + y^2 - 16 = 0$$

Elliptic Curve

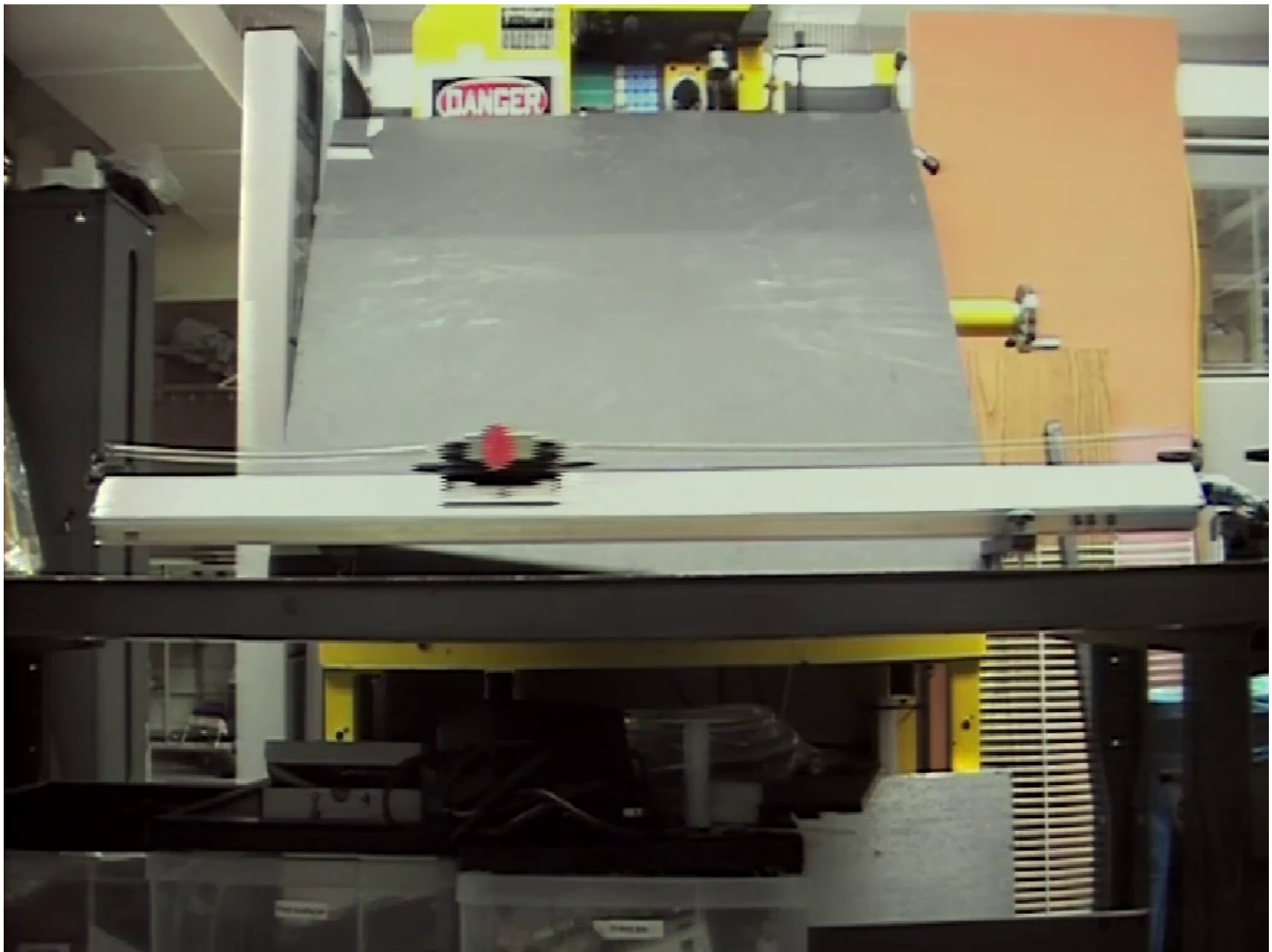


$$x^3 + x - y^2 - 1.5 = 0$$

Sphere



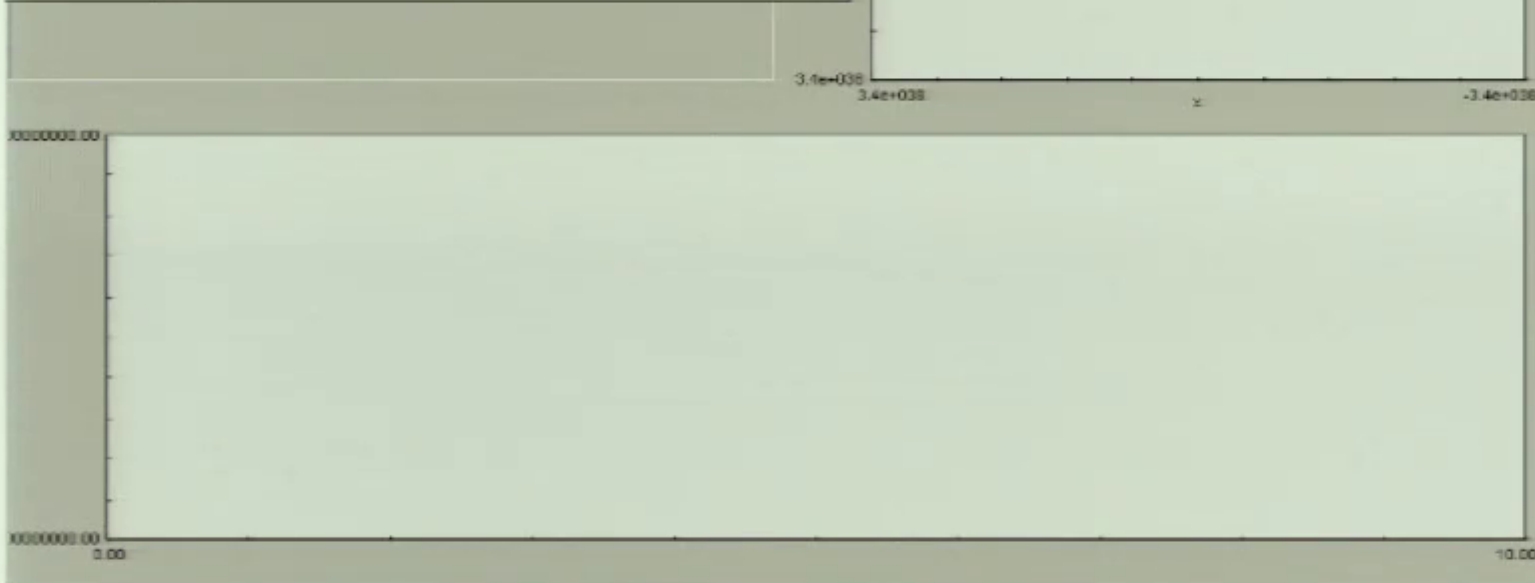
$$x^2 + y^2 + z^2 - 1 = 0$$

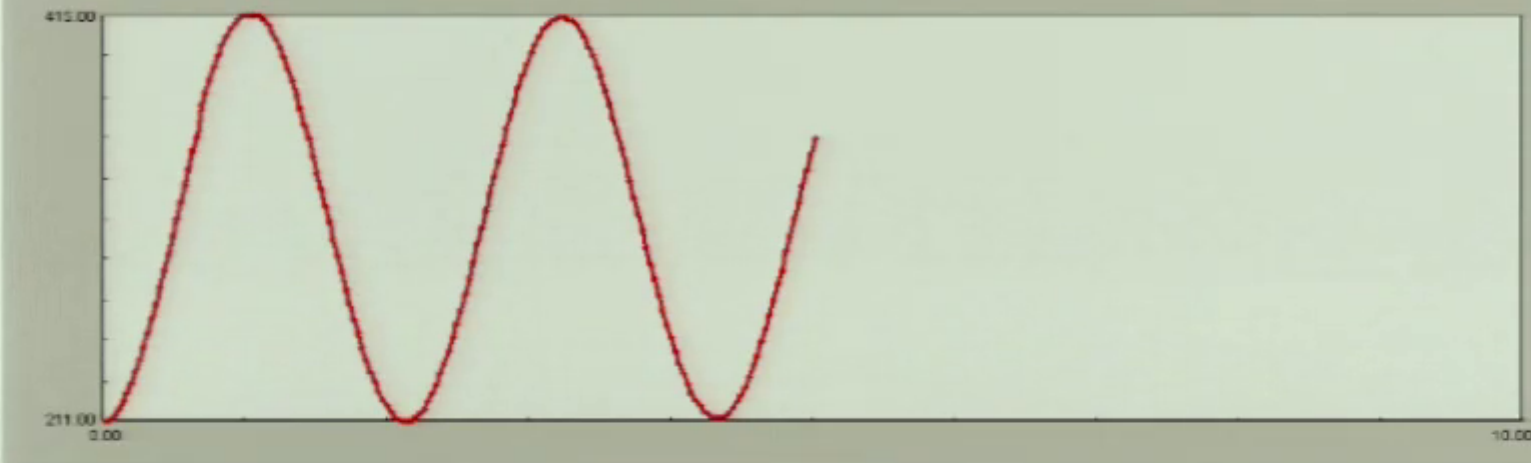
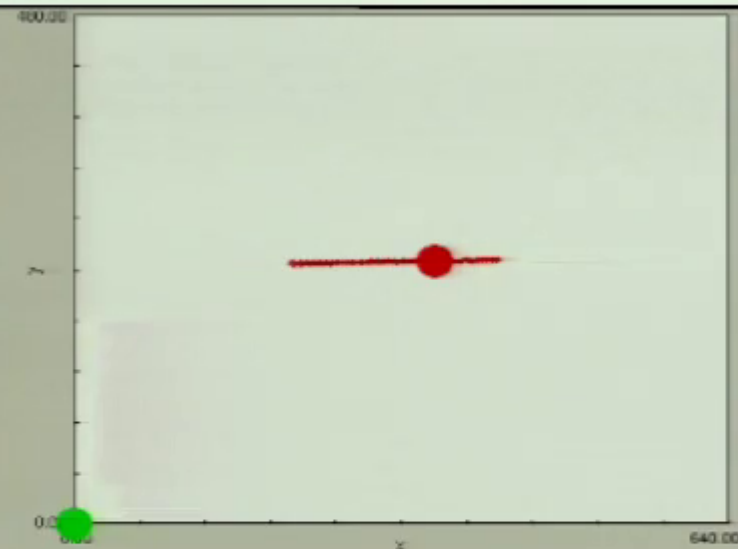
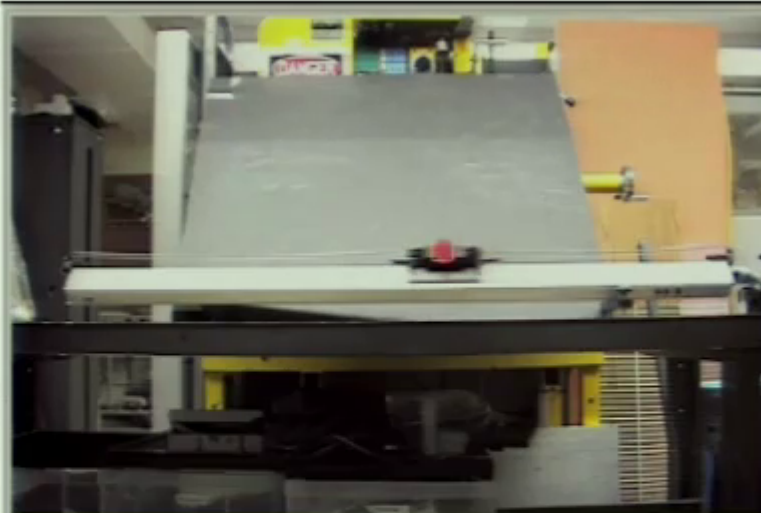


A File Explorer window is open, displaying a list of files in a folder named "Default\_3.avi". The files listed are:

- ar\_1.bit
- ar\_1\_h.bit
- ar\_1\_h\_2.bit
- ar\_1\_h\_3.bit
- ar\_2.bit
- ar\_2\_h.bit
- ar\_3.bit
- ar\_3\_h.bit
- capture\_data.bit
- data.bit
- Default\_0.avi
- Default\_1.avi

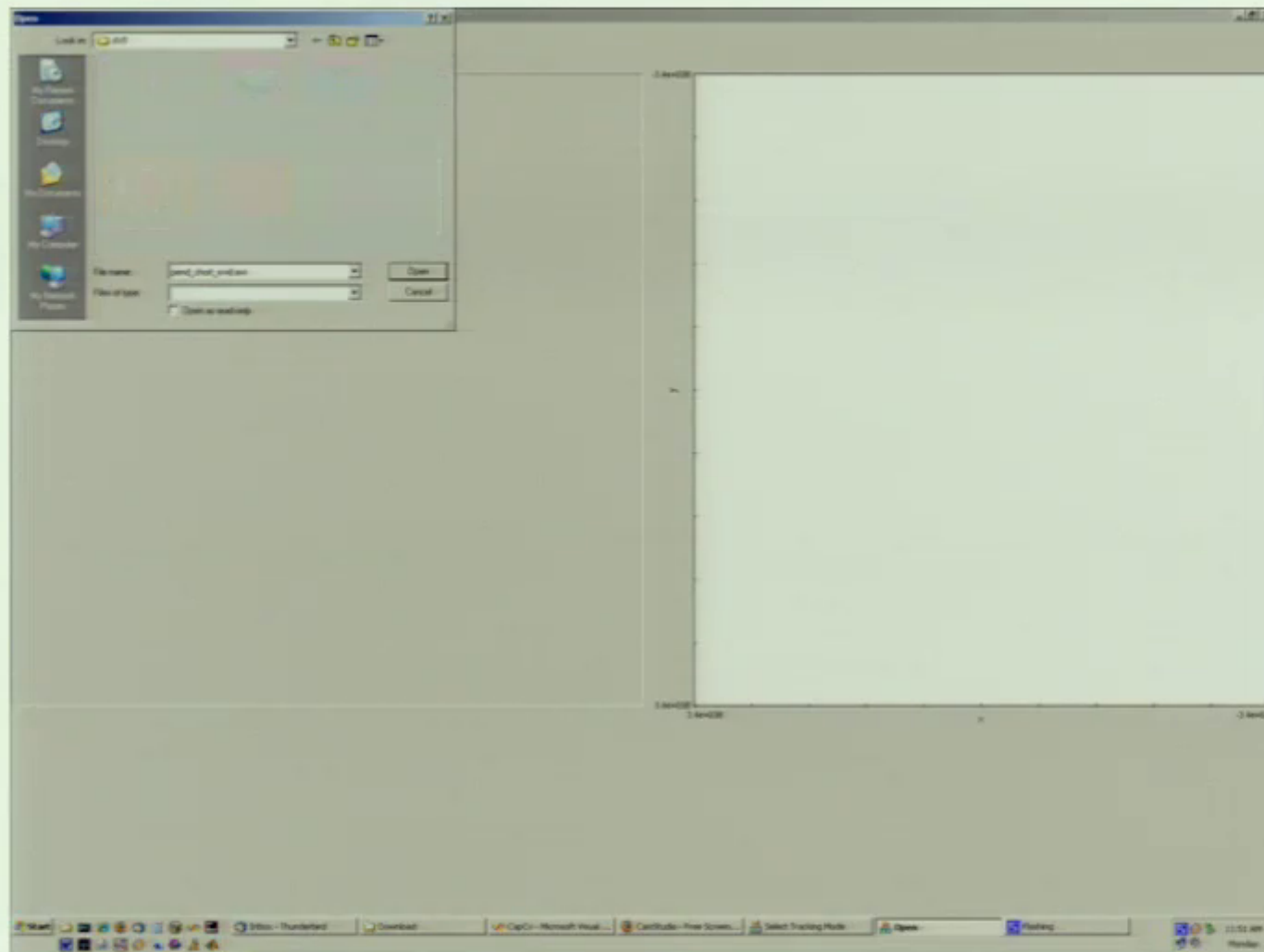
An "Open" dialog box is overlaid on the File Explorer. The "File name" field contains "Default\_2.avi". The "Files of type" dropdown is set to "All files". The "Open as read only" checkbox is unchecked. The "Open" and "Cancel" buttons are visible.





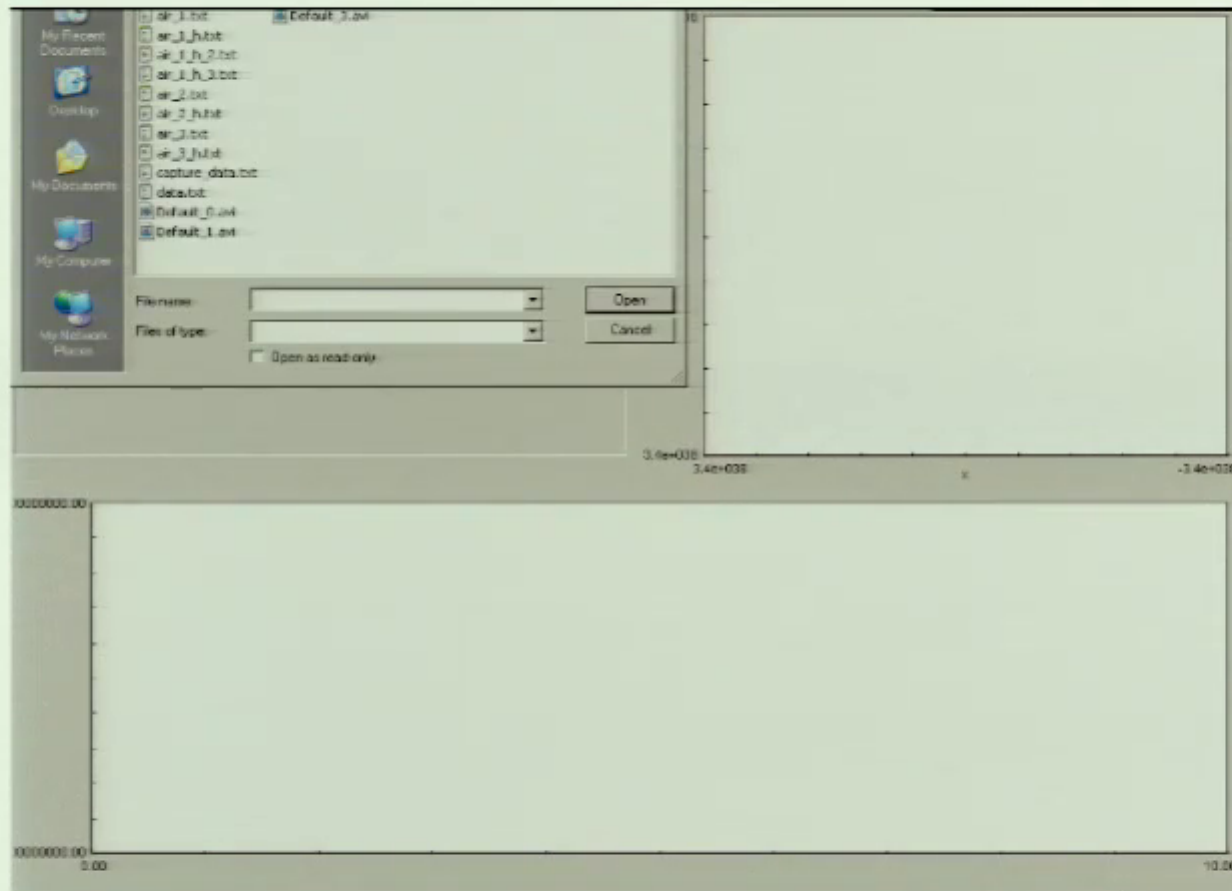
$$H = 114.28 * \left( \frac{dx}{dt} \right)^2 + 692.322 * x^2$$
$$L = 61.591 * \left( \frac{dx}{dt} \right)^2 - 369.495 * x^2$$

• 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

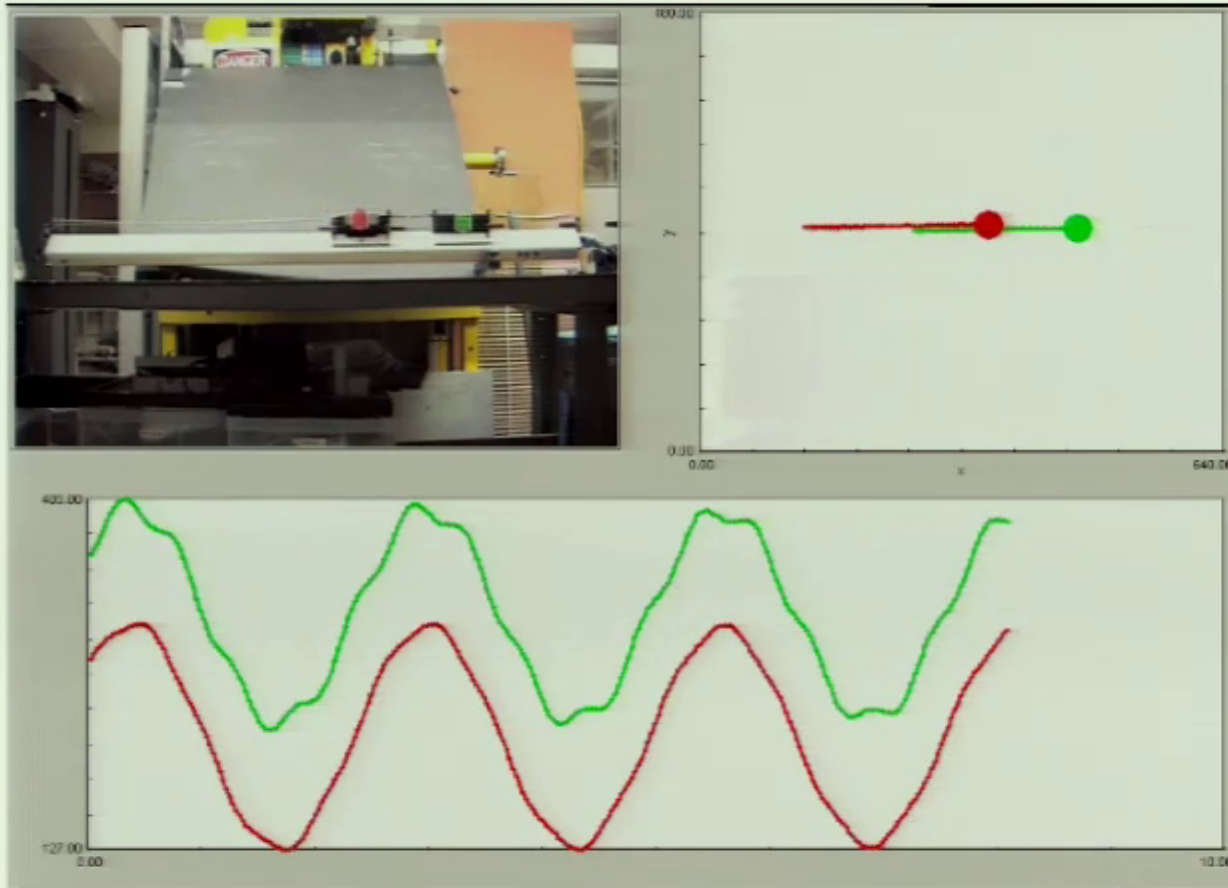


$$H = -14.691 * x_1^2 - 15.551 * x_2^2 - 21.676 * x_1 x_2 + 8.3808 * \left( \frac{dx_2}{dt} \right)^2 + 2.6046 * \left( \frac{dx_1}{dt} \right)^2$$

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

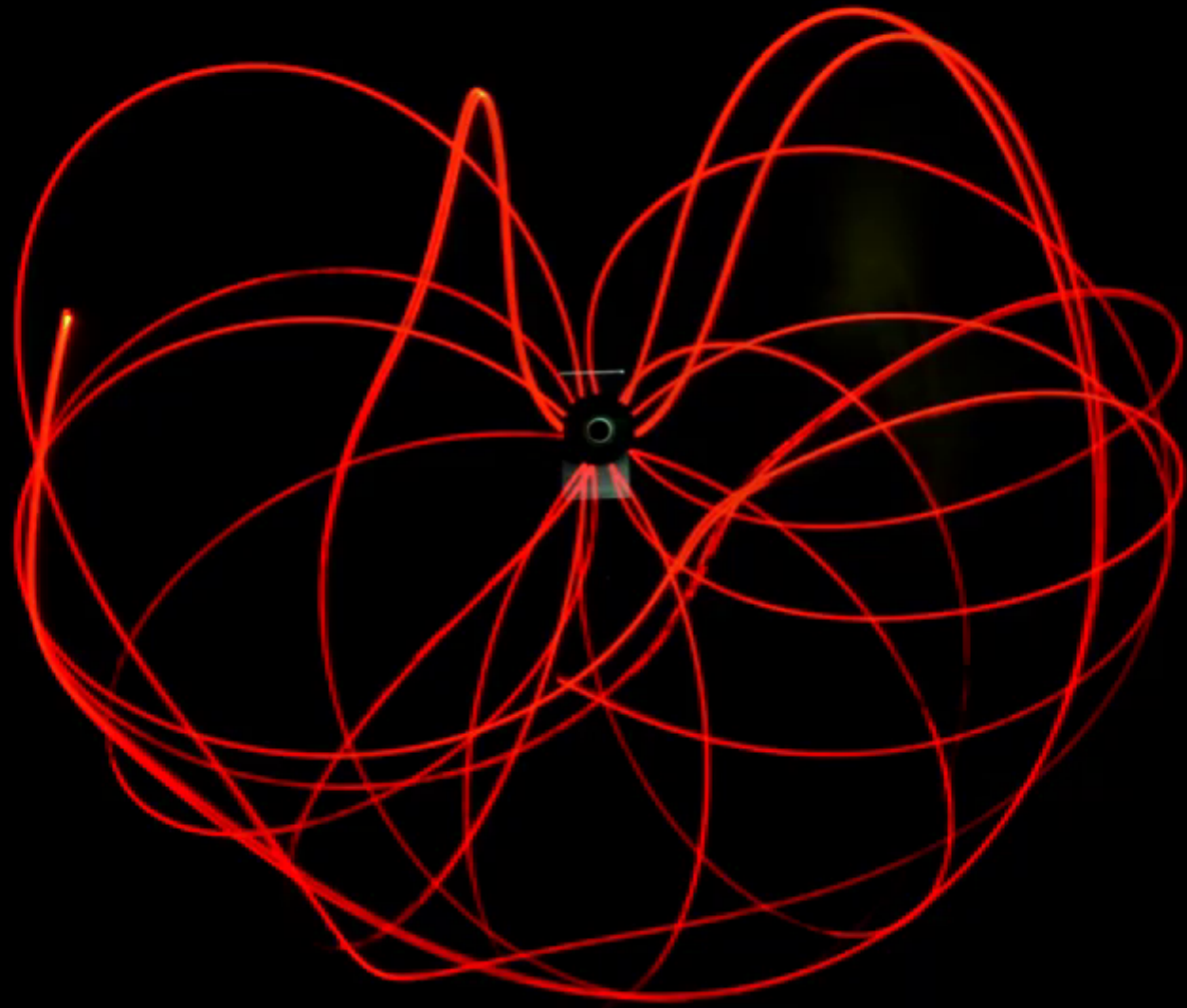


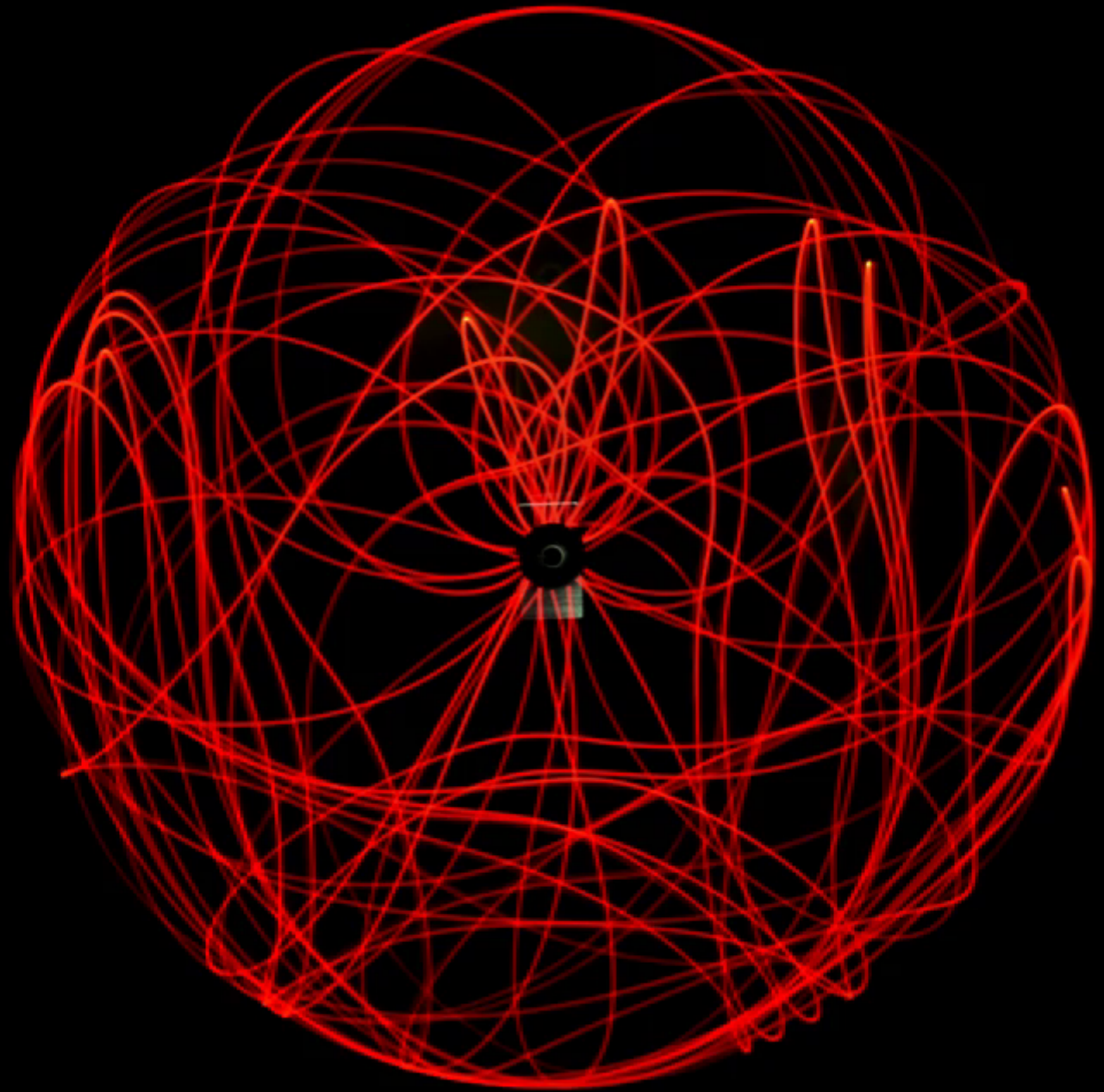




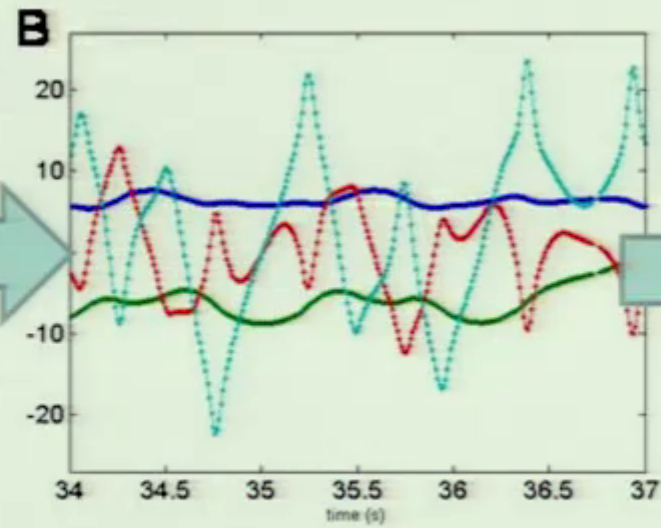












**C**

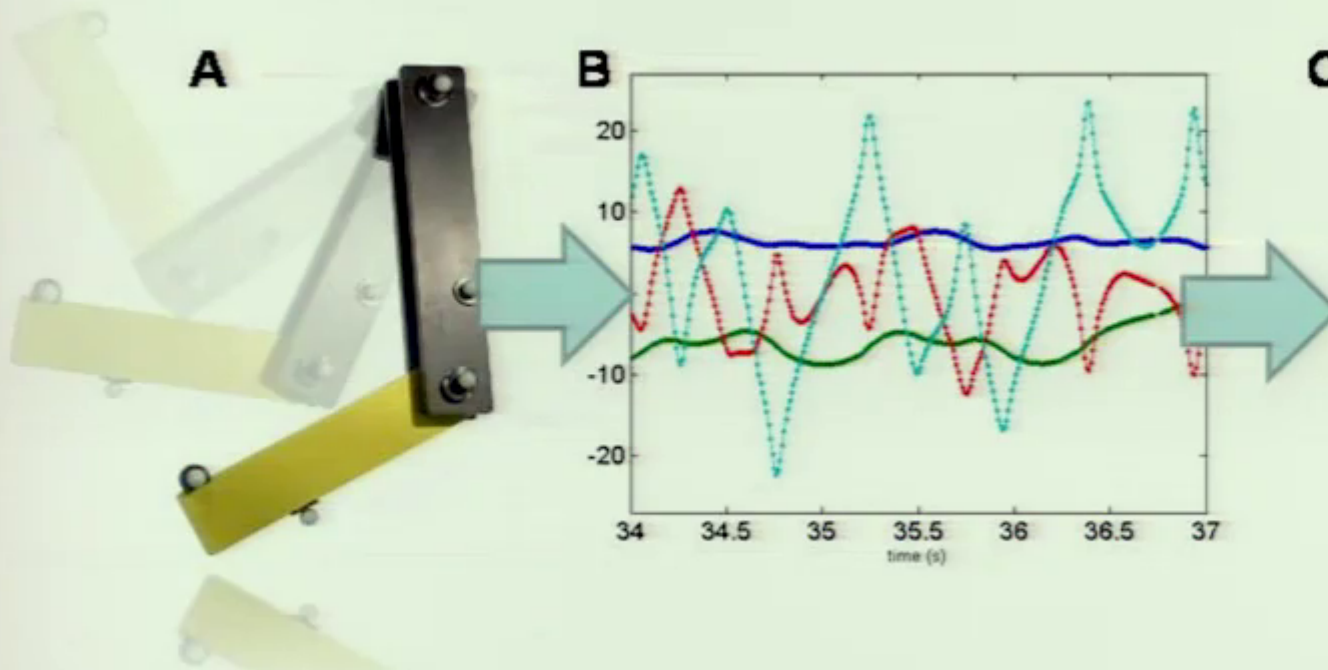
**Detected Invariance:**

$$L_1^2(m_1+m_2)\omega_1^2 + m_2L_2^2\omega_2^2 +$$

$$m_2L_1L_2\omega_1\omega_2\cos(\theta_1 - \theta_2) -$$

$$19.6L_1(m_1+m_2)\cos \theta_1 -$$

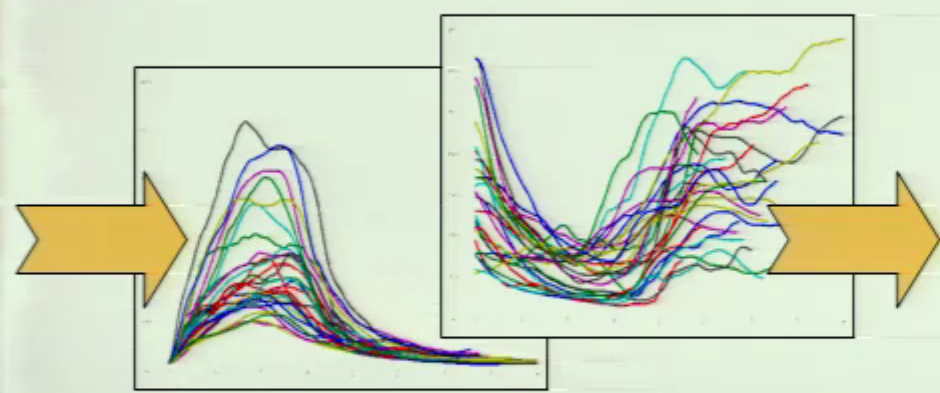
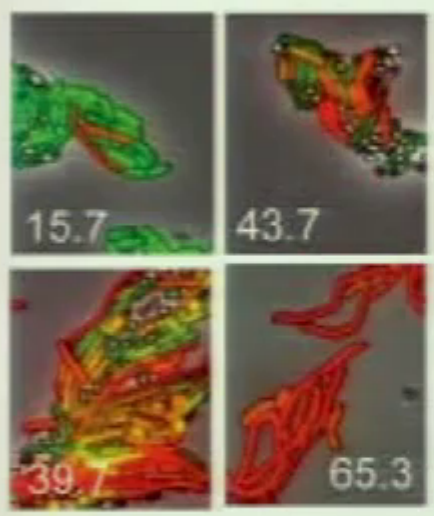
$$19.6m_2L_2\cos \theta_2$$



**C**

**Detected Invariance:**

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



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

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

Untitled - Eureka

File Edit Control Options Tools View Help

Enter Data  $f(x)$  Stack Modeling Stack  $\Delta$  Min/Max  $\Sigma$  Solution Statistics

	A	B	C	D	E	F	G	H	I	J
desc	some variable	some other variable	confidence in y							
var	x	y	w							
1	-3.00	0.82	1.00							
2	-2.94	0.48	0.85							
3	-2.89	1.25	0.91							
4	-2.82	1.98	0.93							
5	-2.70	2.51	0.94							
6	-2.70	2.98	0.92							
7	-2.64		0.89							
8	-2.58		0.84							
9	-2.52		0.82							
10	-2.46		0.75							
11	-2.40		0.83							
12	-2.34		0.86							
13	-2.28		0.71							
14	-2.22		0.81							
15	-2.16		0.70							
16	-2.10		0.86							
17	-2.04		0.91							
18	-1.98		0.84							
19	-1.92		0.85							
20	-1.86		0.80							
21	-1.80		0.89							
22	-1.74		0.82							

Eureka

69% 21 Models Found



Search Complete



Datasource searched:

Double Pendulum

Confidence that search is done



Performance details

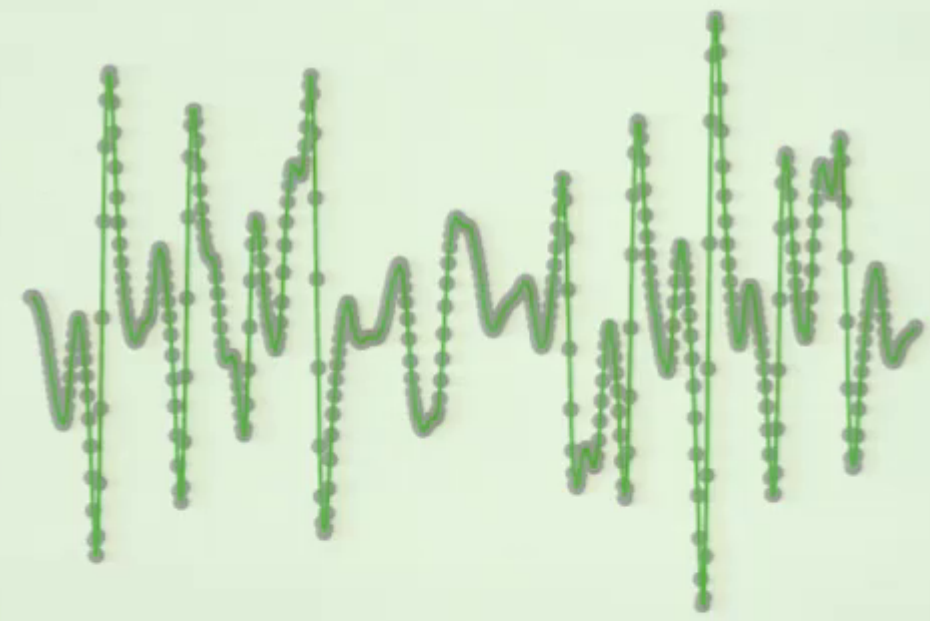
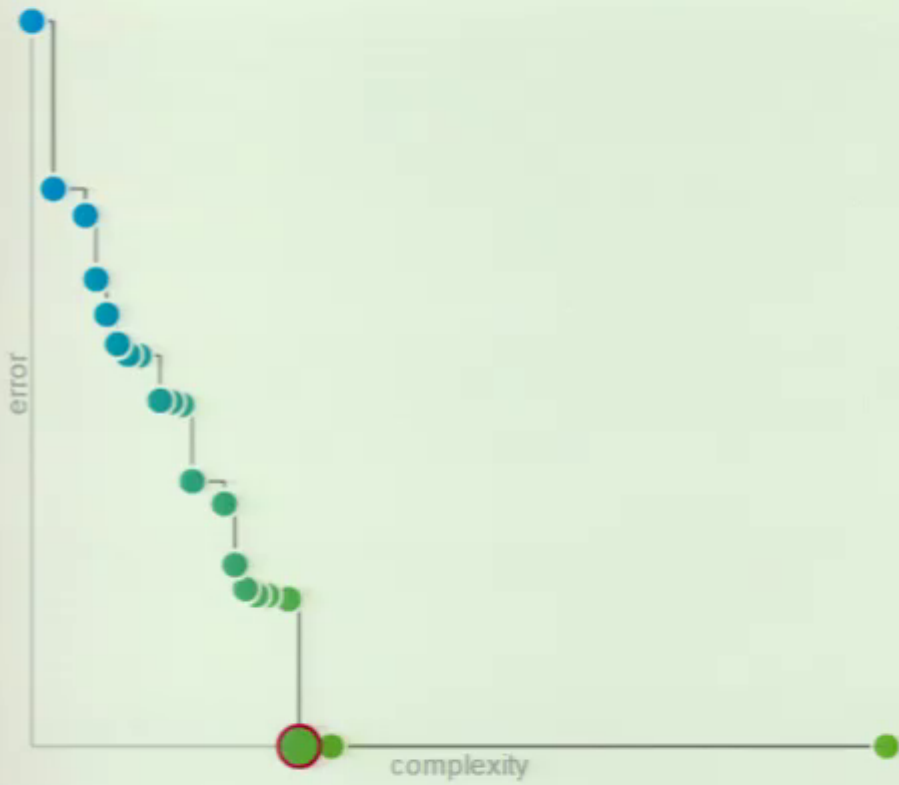
Evals/sec

Models by error vs. complexity

Most frequent variables

$a2 = v1^2 * \sin(x1 - x2) - a1 * \cos(x2 - x1) - 9.808 * \sin(x2)$

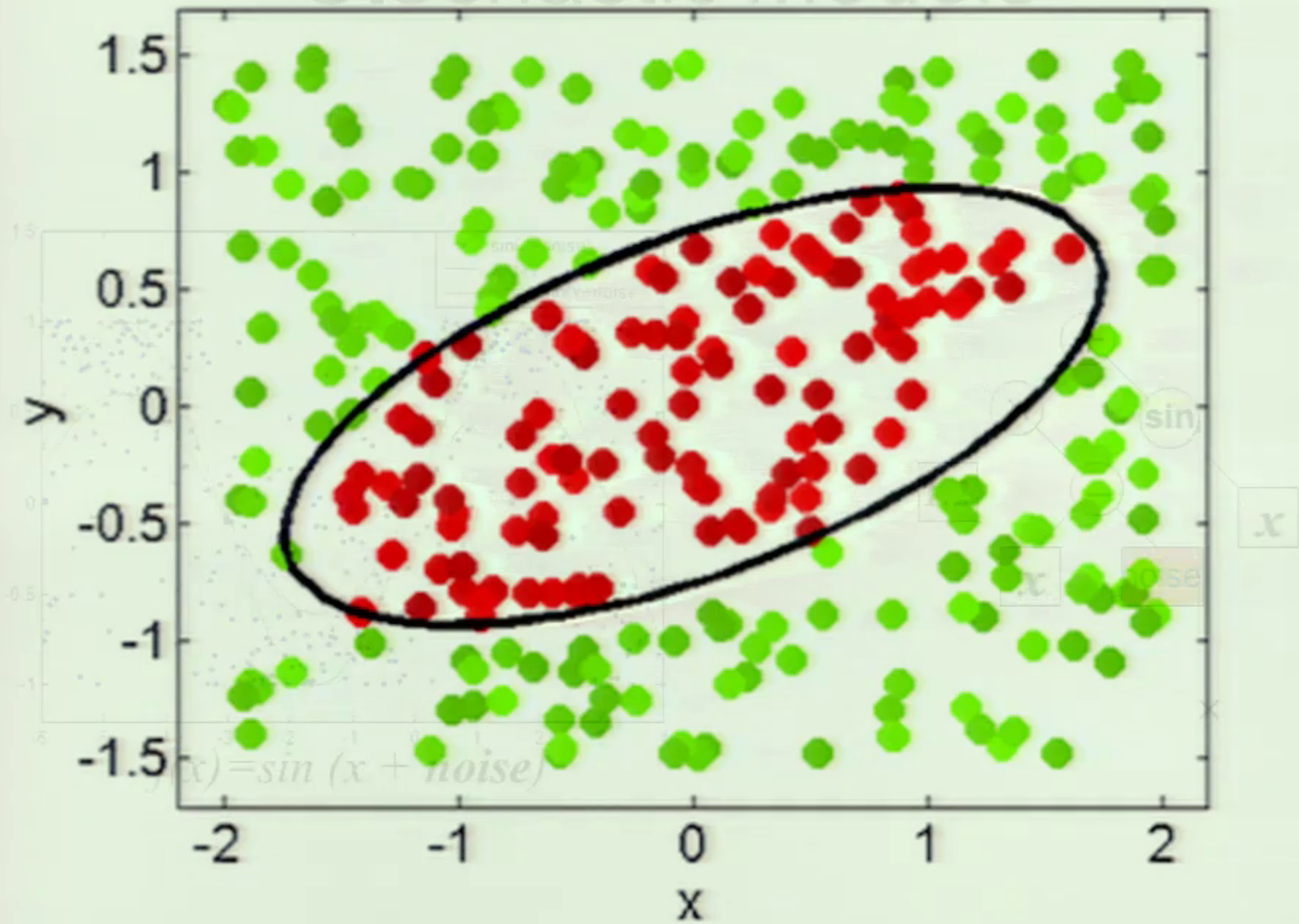
Complexity	Quality	Variables
25	0.487	4



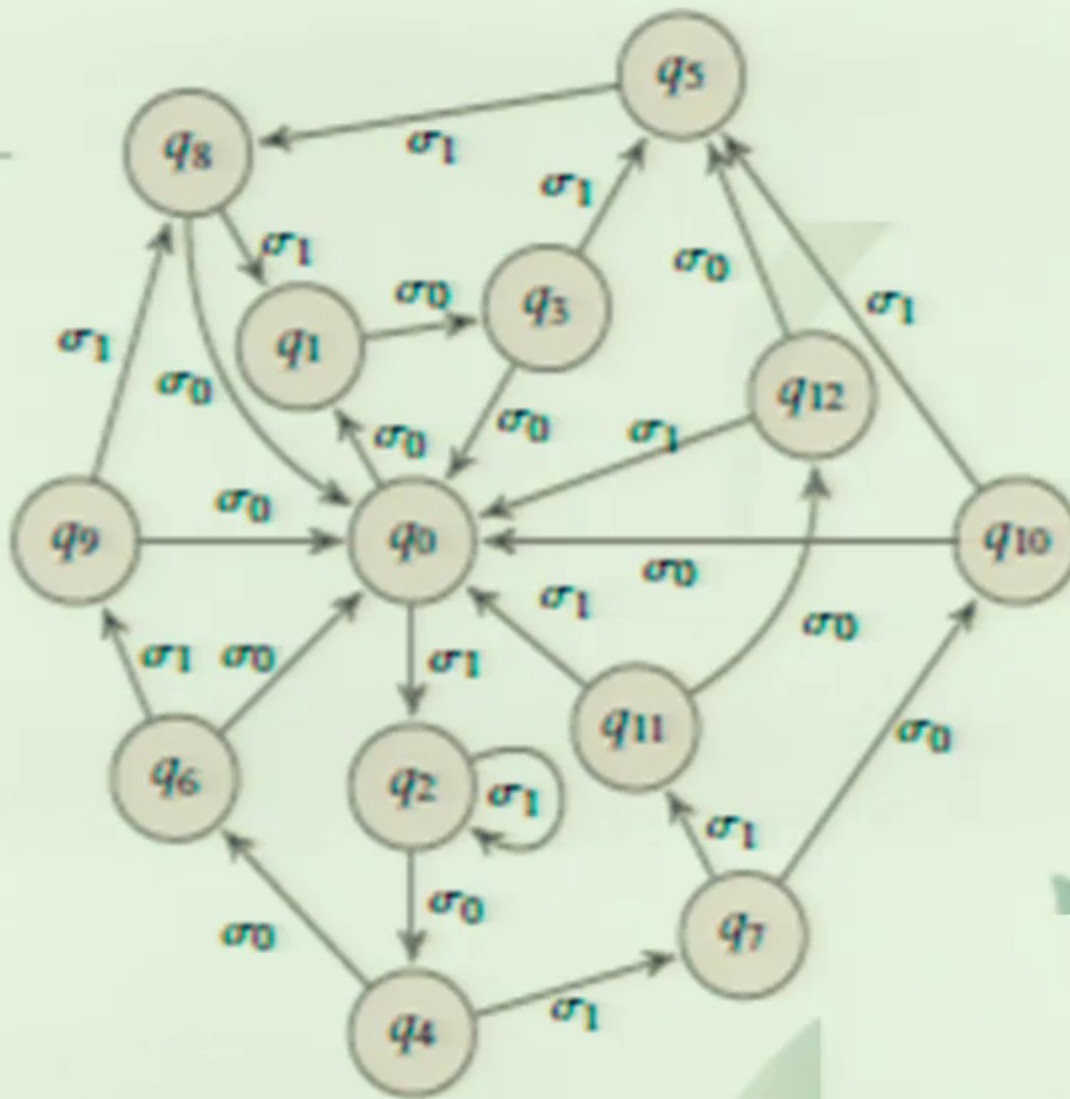
- $x2$
- $x1$
- $a1$
- $v1$



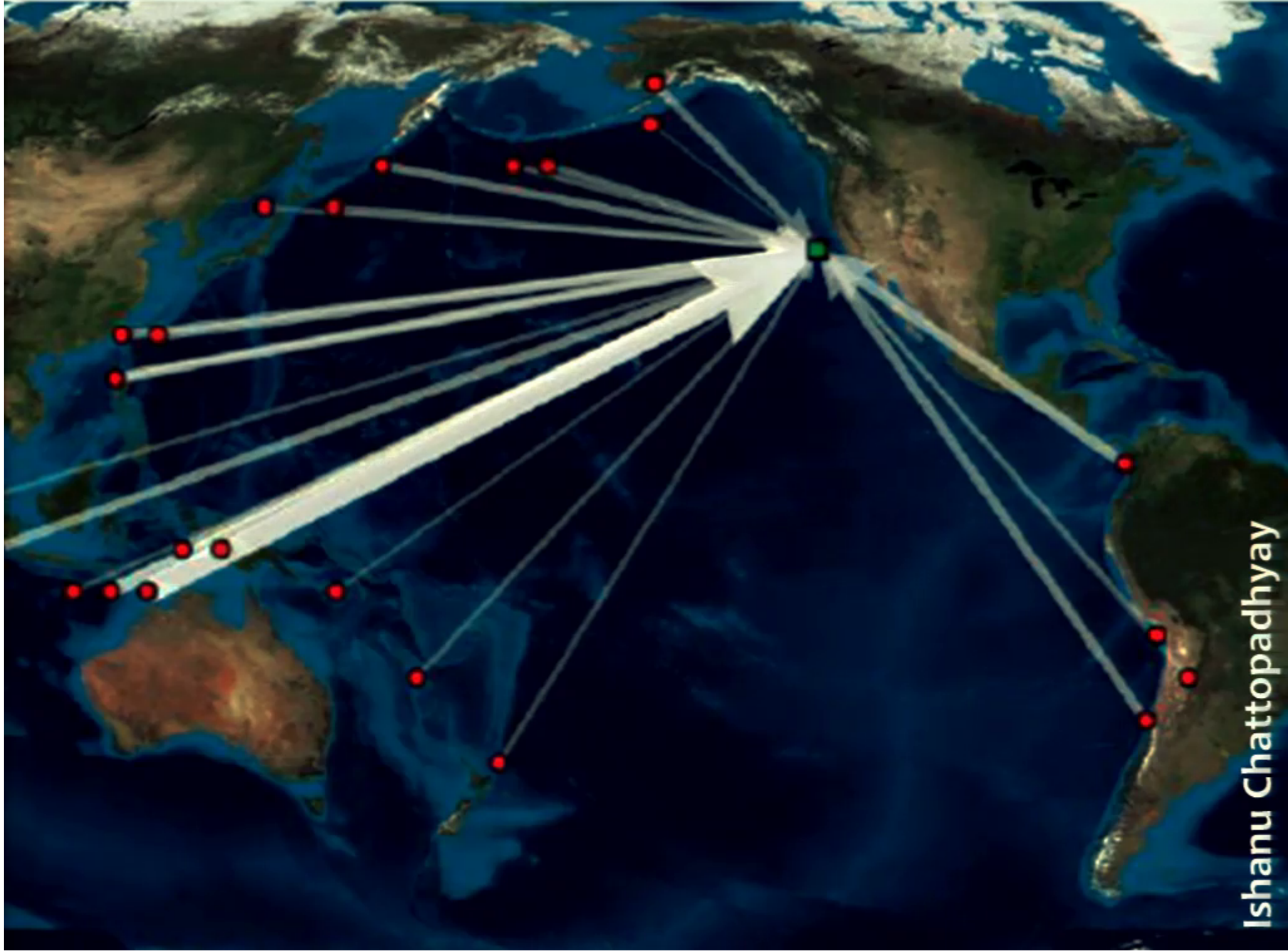
# Stochastic Models



# Probabilistic Finite State Machines

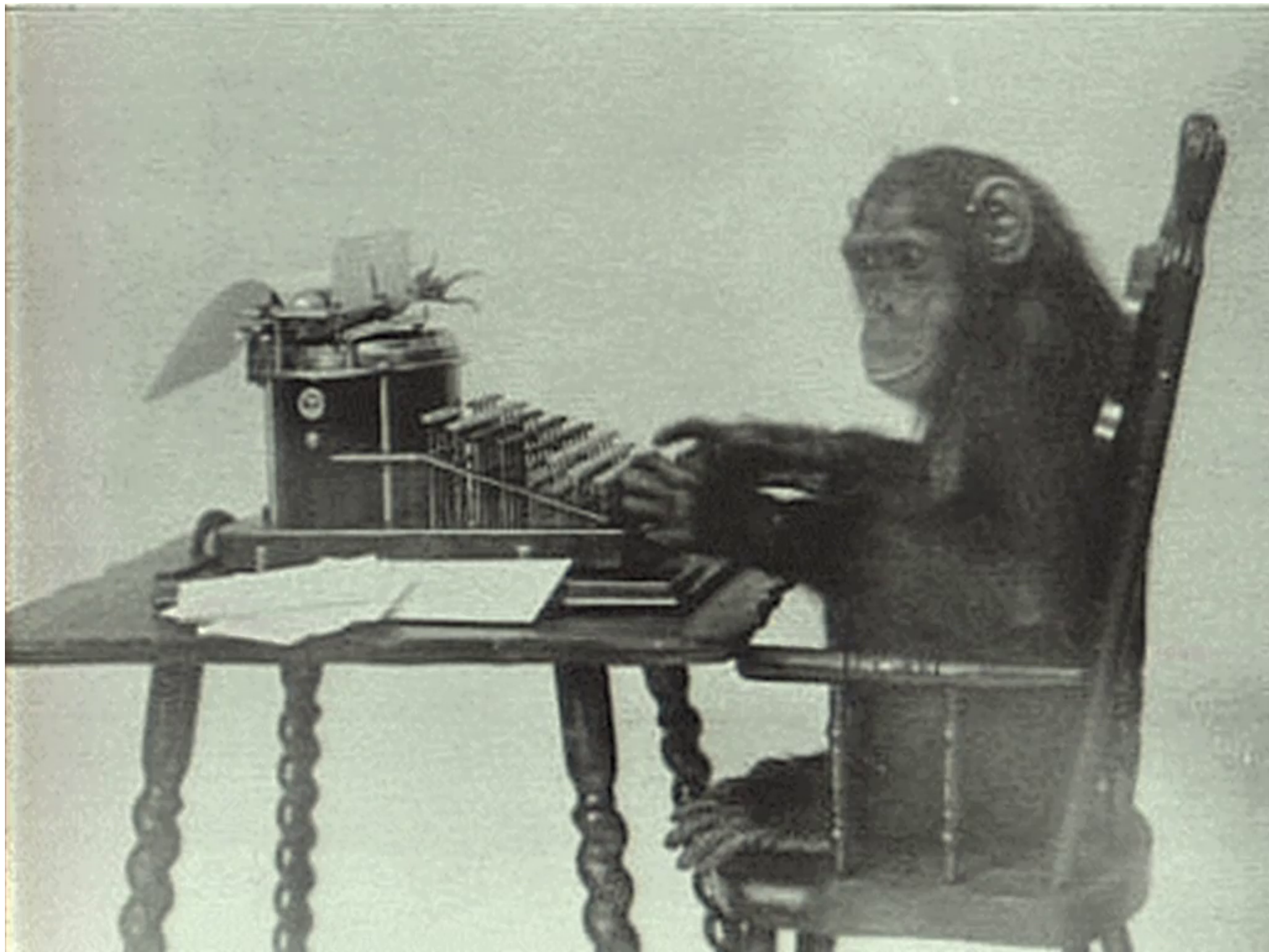


$\sigma_1 \sigma_1 \sigma_2 \sigma_1 \sigma_2 \sigma_2 \sigma_2 \sigma_1 \sigma_2 \sigma_1 \sigma_2 \sigma_1 \sigma_1 \sigma_2 \sigma_1 \sigma_2 \sigma_2 \sigma_2 \sigma_1 \sigma_2$



Ishanu Chattopadhyay





# The New York Times

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