# Hierarchical Thermodynamics

James P Crutchfield Complexity Sciences Center Physics Department University of California, Davis http://csc.ucdavis.edu/~chaos/

Workshop on Information Engines at the Frontier of Nanoscale Thermodynamics Telluride Science Research Center 3-11 August 2017

Joint work with Cina Aghamohammadi, Alec Boyd, Dibyendu Mandal, Sarah Marzen, Paul Riechers, and Greg Wimsatt



## Structural Hierarchy in Biology

In what ways does nature organize? (Phenomenology)

How does it organize? (Mechanism)

Are these levels real or merely convenient? (Objectivity)

Why does nature organize? (Optimization versus chance versus ....)

Does thermodynamics play a role?

Meaning, Purpose, & Functionality

#### Meaning, Purpose, & Functionality

Calculi of Emergence (1992) but with energetics

#### Meaning, Purpose, & Functionality

Calculi of Emergence (1992) but with energetics

### Punch Line

- Meaning, purpose, and functionality arise from
  - Organization
  - Thermodynamics

### **PROBLEM STATEMENT**



- = Ecological population dynamics of structurally complex adapting agents
  - + Reproduction (evolutionary population dynamics)

JP Crutchfield, "The Calculi of Emergence: Computation, Dynamics, and Induction", Physica D 75 (1994) 11-54. In Proceedings of the Oji International Seminar:

*Complex Systems—from Complex Dynamics to Artificial Reality* 5 - 9 April 1993, Numazu, Japan.

### **PROBLEM STATEMENT**



- = Ecological population dynamics of structurally complex adapting agents
  - + Reproduction (evolutionary population dynamics)

JP Crutchfield, "The Calculi of Emergence: Computation, Dynamics, and Induction", Physica D 75 (1994) 11-54. In Proceedings of the Oji International Seminar:

*Complex Systems—from Complex Dynamics to Artificial Reality* 5 - 9 April 1993, Numazu, Japan.

# Agenda

- Level Thermodynamics
- Level Organization
- Level Thermo-Semantics
- Hierarchical Thermodynamics
- Hierarchical Organization

#### Level Thermodynamics

- Level Organization
- Level Thermo-Semantics
- Hierarchical Thermodynamics
- Hierarchical Organization

Thermodynamics of Organization: Information Processing Second Law of Thermodynamics (IPSL)

### **INFORMATION RATCHETS** Beyond Maxwell+Szilard: Net Work Extraction!



A. Boyd, D. Mandal, and JPC, "Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets". New Journal of Physics **18** (2016) 023149.

D. Mandal and C. Jarzynski. "Work and information processing in a solvable model of Maxwell's demon". Proc. Natl. Acad. Sci. USA, **109**(29):11641–11645, 2012.

### **INFORMATION RATCHETS** Beyond Maxwell+Szilard: Net Work Extraction!



A. Boyd, D. Mandal, and JPC, "Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets". New Journal of Physics **18** (2016) 023149.

D. Mandal and C. Jarzynski. "Work and information processing in a solvable model of Maxwell's demon". Proc. Natl. Acad. Sci. USA, **109**(29):11641–11645, 2012.

### INFORMATION PROCESSING SECOND LAW OF THERMODYNAMICS

• Asymptotic IPSL:

$$\langle W \rangle \leq k_{\rm B} T \ln 2 \left( h_{\mu}' - h_{\mu} \right)$$

• Information is fuel:

(Ordered inputs are a thermodynamic resource.)

• Generalizes Landauer Principle; cf.:

$$Q_{\text{erase}} \ge k_B T \ln 2 \qquad (h_{\mu}' = 0; h_{\mu} = 1)$$
  
output input

A. B. Boyd, D. Mandal, and JPC, "Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets", New Journal of Physics **18** (2016) 023149.

### INFORMATION PROCESSING SECOND LAW OF THERMODYNAMICS

- IPSL constrains information processing done by a thermodynamic system.
- Upper bound on the maximum average work  $\langle W \rangle$  extracted per cycle.
- Lower bounds the amount −⟨W⟩ of input work required for a physical system to support a given rate of intrinsic computation.

### INFORMATION PROCESSING SECOND LAW OF THERMODYNAMICS

#### IPSL determines Thermodynamic Functionality

Feature	Operation	Net Work	Net Computation
Engine	Extracts energy from the thermal reservoir, converts it	$\langle W \rangle > 0$	$h'_{\mu}-h_{\mu}>0$
	into work by randomizing input information		
Eraser	Uses external input of work to remove input	$\langle W \rangle < 0$	$h'_{\mu} - h_{\mu} < 0$
	information		
Dud	Uses (wastes) stored work energy to randomize output	$\langle W \rangle < 0$	$h'_{\mu}-h_{\mu}>0$

### **INFORMATION RATCHETS** Second Law for Intrinsic Computation

$$\langle W \rangle \leq k_{\rm B} T \ln 2 \left( h_{\mu}' - h_{\mu} \right)$$

Thermodynamic Functions



### **INFORMATION RATCHETS** Second Law for Intrinsic Computation



Entropy Rates  $\Delta H(1) = H'(1) - H(1)$   $\Delta h_{\mu} = h_{\mu}' - h_{\mu}$ 

#### Lesson:

Kolmogorov-Sinai entropy rates  $(h_{\mu}' \text{ and } h_{\mu})$  account for all correlations.

### **INFORMATION RATCHETS** Second Law for Intrinsic Computation



### **REQUISITE COMPLEXITY**

	Input Process	Ratchet Transducer	Output Process	Thermal Relations
memoryless input memoryless ratchet		$(A) \bigcirc_{\substack{1 0:q \\ 0 1:p \\ 1 1:1-p}}^{0 0:1-q}$	$ \underbrace{DA}_{1:bq+(1-b)(1-p)} 0:b(1-q)+(1-b)p \\ 1:bq+(1-b)(1-p) $	$0 = H_1 - h_\mu = H'_1 - h'_\mu$ $\langle W \rangle \le \Delta h_\mu = \Delta H_1$
memoryless input memoryful ratchet		$A \underbrace{\begin{smallmatrix} 0 0:1-p \\ 1 0:p \\ 1 1:1 \\ B \\ 0 1:q \\ 0 0:1 \\ B \\ $	$\underbrace{DA}_{0:b+q(1-b)}^{0:b(1-p)} \underbrace{DB}_{0:b+q(1-b)}^{0:b+q(1-b)} \\ 1:(1-b)(1-q)}$	$0 = H_1 - h_\mu \le H'_1 - h'_\mu$ $\langle W \rangle \le \Delta h_\mu \le \Delta H_1$
memoryful input memoryless ratchet	$ \begin{array}{c}                                     $	$ \underbrace{ \begin{array}{c} A \\ \end{array} }_{1 0:q} \overset{0 0:1-q}{\underset{0 1:p}{1 0:q}} \\ _{1 1:1-p} \end{array} $	$\begin{array}{c} 0:p/2 \\ 1:(1-p)/2 \\ FA \\ \hline \\ 0:1-q \\ 1:q \end{array} \begin{array}{c} 0:(1-q)/2 \\ 0:p \\ 1:1-p \\ EA \\ 0:1-q \\ 1:q \end{array}$	$0 \le H_1' - h_\mu' \le H_1 - h_\mu$ $\langle W  angle \le \Delta H_1 \le \Delta h_\mu$
memoryful input memoryful ratchet	$1: 0.5 \qquad D \\ 0: 0.5 \\ F \\ 0: 1.0 \\ E \\ 0: 1.0 \\ E$	$A \underbrace{\begin{matrix} 0 0:1-p\\1 0:p\\1 1:1\\0\\0 1:q\\0 0:1\end{matrix}}_{0 0:1}B$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$H_{1}' - h_{\mu}' \stackrel{?}{=} H_{1} - h_{\mu}$ $\langle W \rangle \leq \Delta h_{\mu}$ $\langle W \rangle \leq \Delta H_{1}$

- Lessons: Information processing thermodynamic systems should match the complexity of their inputs/environment:
  - Memoryless ratchets optimal for uncorrelated environments.
  - Memoryful ratchets optimal for correlated environments.

A. B. Boyd, D. Mandal, and JPC, "Leveraging Environmental Correlations: The Thermodynamics of Requisite Variety", Journal of Statistical Physics (2017) in press. <u>arxiv.org</u>:1609.05353.

W. Ross Asbhy, "An Introduction to Cybernetics." John Wiley and Sons, New York, second edition, 1960.

# Functional Fluctuations

J. P. Crutchfield and C. Aghamohammadi, "Not All Fluctuations are Created Equal: Spontaneous Variations in Thermodynamic Function. arxiv.org:1609.02519.

#### FUNCTIONAL FLUCTUATIONS

#### FUNCTIONAL FLUCTUATIONS

#### When is an Engine an Eraser?

#### FUNCTIONAL FLUCTUATIONS

#### When is an Engine an Eraser?

### Information Processing Second Law + Large Deviation Theory

### FLUCTUATIONS?

Biased Coin (60% Heads/40% Tails)



### FLUCTUATIONS?

#### Biased Coin: Fluctuations in sequence probabilities



### FLUCTUATIONS IN INTRINSIC COMPUTATION







### **INFORMATION RATCHETS**

(A. Boyd, D. Mandal, and JPC, "Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets", New Journal of Physics **18** (2016) 023149.)



#### FLUCTUATIONS IN

#### **THERMODYNAMIC FUNCTION**



#### FLUCTUATIONS IN

#### **THERMODYNAMIC FUNCTION**



Feature	Operation	Net Work	Net Computation
Engine	Extracts energy from the thermal reservoir, converts it	$\langle W \rangle > 0$	$h'_{\mu} - h_{\mu} > 0$
	into work by randomizing input information		
Eraser	Uses external input of work to remove input	$\langle W \rangle < 0$	$h'_{\mu} - h_{\mu} < 0$
	information		
Dud	Uses (wastes) stored work energy to randomize output	$\langle W \rangle < 0$	$h'_{\mu}-h_{\mu}>0$

#### Informational Second Law ⇒ Thermodynamic Function

#### FLUCTUATIONS IN

#### THERMODYNAMIC FUNCTION



### **FLUCTUATIONS IN THERMODYNAMIC FUNCTION**

#### Observable? Length 100 input sequences: Engine: 80% ★ Simulated result $I(U_{\rm in})$ $\Pr(u_{100})$ •---• $\star$ 0.4Dud: 17.8% 3 EngineEraser Dud Eraser: 2.2% I [bits/symbol] 20.3 $\Pr(u_{100})$

1

0

 $U_{\rm in}^{\rm min}$ 

Input Typical Set

1

0.1

0

 $U_{\rm in}^{\rm max}$ 

3

[bits/symbol]

 $\mathbf{2}$ 

 $U_{\rm in}, u_{100}$ 

### FLUCTUATIONS IN INTRINSIC COMPUTATION

• What's new: Determine S(U) and other measures for any structured process from



FIG. 22.  $\epsilon$ -Machine that generates the non-Markovian Nemo Process; its Markov order is infinite. The Nemo Process makes this perhaps clearer, however, since the recurrent states permute into each other upon observing a 0. The transient structure captures this explicitly: *ABC* maps back to itself on a 0.



## Lessons

• Function emerges from the

Information Processing Second Law of Thermodynamics

- Computing fluctuates
- Function fluctuates

- Level Thermodynamics
- Level Organization
- Level Thermo-Semantics
- Hierarchical Thermodynamics
- Hierarchical Organization

#### INFORMATION-THEORETIC ANALYSIS OF COMPLEX SYSTEMS ...

• Process  $Pr(\overleftarrow{X}, \overrightarrow{X})$  is a communication channel from the past  $\overleftarrow{X}$  to the future  $\overrightarrow{X}$ :



#### INFORMATION-THEORETIC ANALYSIS OF COMPLEX SYSTEMS ...

• Process  $Pr(\overleftarrow{X}, \overrightarrow{X})$  is a communication channel from the past  $\overleftarrow{X}$  to the future  $\overrightarrow{X}$ :



#### INFORMATION-THEORETIC ANALYSIS OF COMPLEX SYSTEMS ...

• Process  $Pr(\overleftarrow{X}, \overrightarrow{X})$  is a communication channel from the past  $\overleftarrow{X}$  to the future  $\overrightarrow{X}$ :



• Channel Utilization: Excess Entropy  $\mathbf{E} = I[\overleftarrow{X}; \overrightarrow{X}]$ 

#### FOUNDATIONS: COMPUTATIONAL MECHANICS



**STORED VERSUS GENERATED INFORMATION**  $C_{\mu} = -\sum_{\sigma \in \mathcal{S}} \Pr(\sigma) \log_2 \Pr(\sigma) \text{ VERSUS } h_{\mu} = -\sum_{\sigma \in \mathcal{S}} \Pr(\sigma) \sum_{\sigma' \in \mathcal{S}} \Pr(\sigma' | \sigma) \log_2 \Pr(\sigma' | \sigma)$ 

#### FOUNDATIONS: COMPUTATIONAL MECHANICS



**E-MACHINE: UNIQUE, MINIMAL, & OPTIMAL PREDICTOR** 

VERSUS **GENERATED** INFORMATION STORED



STRUCTURE VERSUS

RANDOMNESS



#### FOUNDATIONS: COMPUTATIONAL MECHANICS



**E-MACHINE: UNIQUE, MINIMAL, & OPTIMAL PREDICTOR** 

STORED VERSUS GENERATED INFORMATION

 $C_{\mu} = -\sum_{\sigma \in \boldsymbol{\mathcal{S}}} \Pr(\sigma) \log_2 \Pr(\sigma) \quad \mathbf{VERSUS} \quad h_{\mu} = -\sum_{\sigma \in \boldsymbol{\mathcal{S}}} \Pr(\sigma) \sum_{\sigma' \in \boldsymbol{\mathcal{S}}} \Pr(\sigma'|\sigma) \log_2 \Pr(\sigma'|\sigma)$ 

#### **INTRINSIC COMPUTATION:**

1. HOW MUCH HISTORICAL INFORMATION DOES A PROCESS STORE?

- 2. IN WHAT ARCHITECTURE IS IT STORED?
- 3. How is it used to produce future behavior?

J.P. CRUTCHFIELD & K. YOUNG, "INFERRING STATISTICAL COMPLEXITY", PHYSICAL REVIEW LETTERS 63 (1989) 105-108.

J.P. CRUTCHFIELD, "BETWEEN ORDER AND CHAOS", NATURE PHYSICS 8 (JANUARY 2012) 7-24.

### COMPUTATIONAL MECHANICS

• ε-Machine:

$$M = \left\{ \boldsymbol{\mathcal{S}}, \{ T^{(x)} : x \in \mathcal{A} \} \right\}$$

• Dynamic:

$$T_{\sigma,\sigma'}^{(x)} = \Pr(\sigma'|\sigma, x)$$
$$\sigma, \sigma' \in \mathcal{S}$$



### VARIETIES OF E-MACHINE



#### Denumerable Causal States



Fractal

J. P. Crutchfield, "Calculi of Emergence: Computation, Dynamics, and Induction", Physica D **75** (1994) 11-54.



Continuous

### Intrinsic Computation: Consequences

A system is unpredictable

if it has positive entropy rate:  $h_{\mu} > 0$ 

A system is complex

if it has positive structural complexity measures:  $C_{\mu} > 0$ 

A system is emergent

if its structural complexity increases over time:

 $C_{\mu}(t') > C_{\mu}(t), \text{ if } t' > t$ 

A system is hidden

if its crypticity is positive:  $\chi=C_{\mu}-\mathbf{E}>0$ 

# What is a Level?

- Pattern discovery:
  - Learn the world's hidden states  $\Pr(\mathcal{R} | X)$
- Causal shielding:

 $\Pr(\stackrel{\leftarrow}{X}\stackrel{\rightarrow}{X}) = \Pr(\stackrel{\leftarrow}{X}|\mathcal{R})\Pr(\stackrel{\rightarrow}{X}|\mathcal{R})$ 

- Search in the space of models:  $\mathcal{R} \in \mathcal{M}$
- Objective function

$$\min_{\Pr(\mathcal{R}|\overset{\leftarrow}{X})} \left( I[\overset{\leftarrow}{X};\mathcal{R}] + \beta I[\overset{\leftarrow}{X};\overset{\rightarrow}{X}|\mathcal{R}] \right)$$

Model: Map fromInfo states containReducehistories to statesabout historieshas ab

Reduce info history has about future

Long history:

N.H. Packard, J.P. Crutchfield, J.D. Farmer, R. S. Shaw, "Geometry from a Time Series", Phys. Rev. Lett. **45** (1980) 712-716. J. P. Crutchfield and B.S. McNamara, "Equations of Motion from a Data Series", Complex Systems **1** (1987) 417-452.

S. Still, C.J. Ellison, J.P. Crutchfield: arxiv.org: 0708.0654 [physics.gen-ph] & 0708.1580[cs.IT]

• Optimal states  $\Pr(\mathcal{R} | \stackrel{\leftarrow}{X})$  are Gibbs distributions:

$$\Pr_{\text{opt}}(\mathcal{R} | \overleftarrow{X}) = \frac{\Pr(\mathcal{R})}{Z(\overleftarrow{X}, \beta)} e^{-\beta E(\mathcal{R}, \overleftarrow{X})}$$
  
where

$$E(\mathcal{R}, \overleftarrow{X}) = \mathcal{D}\left(\Pr(\overrightarrow{X} \mid \overleftarrow{X}) || \Pr(\overrightarrow{X} \mid \mathcal{R})\right)$$
  

$$\Pr(\overrightarrow{X} \mid \mathcal{R}) = \frac{1}{\Pr(\mathcal{R})} \sum_{\overleftarrow{X}} \Pr(\overrightarrow{X} \mid \overleftarrow{X}) \Pr(\mathcal{R} \mid \overleftarrow{X}) \Pr(\overleftarrow{X})$$
  

$$\Pr(\mathcal{R}) = \sum_{\overleftarrow{X}} \Pr(\mathcal{R} \mid \overleftarrow{X}) \Pr(\overleftarrow{X})$$

Now, solve these self-consistently

# Optimal balance structure & error At each level $\beta$ of approximation

#### In theory

Causal Rate Distortion Curve



#### Optimal balance structure & error At each level $\beta$ of approximation

#### In theory



#### Optimal balance structure & error At each level $\beta$ of approximation

#### In theory



- Level Thermodynamics
- Level Organization

#### Level Thermo-Semantics

- Hierarchical Thermodynamics
- Hierarchical Organization

### Beyond Structure to Meaning, Purpose, & Functionality

Semiotic Hierarchy of Information:

Syntax Semantics Pragmatics Function

For physical systems ...

### Beyond Structure to Meaning, Purpose, & Functionality

Semiotic Hierarchy of Information:

Syntax Semantics Pragmatics Function

For physical systems ...

#### **Semantics and Thermodynamics**

James P. Crutchfield

Physics Department<sup>\*</sup> University of California Berkeley, California 94720 USA

in Nonlinear Modeling and Forecasting, M. Casdagli and S. Eubank, editors, Santa Fe Institute Studies in the Sciences of Complexity XII Addison-Wesley, Reading, Massachusetts (1992) 317 – 359.

SFI 91-09-033

What does a particular measurement mean?



Measurement Channel

An  $\epsilon M$  Captures "Pattern":

Measurement semantics: Prediction level

What is the meaning of a particular measurement?

Shannon says the amount of "information" is

 $-\log_2 \Pr(\text{observing } s)$ 

Given  $\epsilon M$  (assuming you're sync'd):

 $-\log_2 \Pr(\text{observing } s) = -\log_2 \Pr(S \to_s S')$ 

An  $\epsilon M$  Captures "Pattern":

Measurement semantics: Prediction level ...

 $H(s_{11}|s_{10} = 1, s_9 = 1, ...) \approx h_{\mu} (\approx 0.585 \text{ bits})$ 

Degree of observer's surprise (predictability) Does not say what the event  $s_{11} = 1$  means to the observer!

Meaning: Tension between representations of same event at different levels; e.g.,

Level I is data stream and the event is a measurement Level 2 is the agent and the event updates it's model

Degree of meaning of observing  $s \in \mathcal{A}$ 

$$\Theta(s) = -\log_2 \Pr(\to_s S)$$

where  $\mathcal{S}$  is the causal state to which s brings observer.

Meaning content: State selected from anticipated palette.

Meaningless: Start state (all futures possible)

$$\Theta(s) = -\log_2 \Pr(\mathcal{S}_0) = -\log_2 1 = 0 \qquad s = \lambda$$

Action on disallowed transition:

Reset to state of total ignorance (start state) Disallowed transition is meaningless.

Meaningless measurements are informative, though:

$$-\log_2 \Pr(\mathcal{S} \to_s \mathcal{S}_0) = -\log_2 0 = \infty$$

Theorem:

$$\langle \Theta(s) \rangle = C_{\mu}$$

Average amount of meaning is the Statistical Complexity.

Thermodynamic Cost of Extracting Meaning

## Thermodynamic Cost of Extracting Meaning

Two cases: o In NESS o Out of NESS

# Thermodynamic Cost of Extracting Meaning in NESS

- Learning about environment
- Agent predicts environment to leverage possible resources



(A. Boyd, D. Mandal, and JPC, "Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets", New Journal of Physics **18** (2016) 023149.)

## Thermodynamic Cost of Extracting Meaning in NESS

Thermo cost of implementation that predicts:

$$\langle Q^{\text{implement}} \rangle_{\text{min}} = k_{\text{B}}T\ln 2 \ \text{I}[\mathcal{S}; \overleftarrow{Y}']$$
  
=  $k_{\text{B}}T\ln 2 \ C_{\mu}$ 

Recall Theorem on Total Semantic Content

$$\langle \Theta(s) \rangle = C_{\mu}$$

Agent memory about environment.

Thermo-semantic cost:

$$\left\langle Q^{\text{implement}} \right\rangle_{\text{min}} = k_{\text{B}}T\ln 2 \left\langle \Theta(s) \right\rangle$$

Thermodynamics of Meaning and Function beyond NESS

## Thermodynamics of Meaning and Function beyond NESS



## Thermodynamics of Meaning and Function beyond NESS

- Semiotics of Information Engines:
  - Syntactic information = Measurements
  - Semantic information = Envt'l Phase, Sync/No Sync
  - Functional information = Error Correction

# Summary

- Level Thermodynamics
- Level Organization
- Level Semantics
- Hierarchical Thermodynamics
- Hierarchical Organization

#### Thanks!



## Thermodynamics of Organization

Bibliography

- Functional computation & optimal prediction: A.B. Boyd, D. Mandal, and J.P. Crutchfield, *Identifying Functional Thermodynamics in Autonomous Maxwellian Ratchets*, New Journal of Physics 18 (2016) 023149.
- Information creation: C. Aghamohammadi and J.P. Crutchfield, *Thermodynamics of Random Number Generation*, Physical Review E **95**:6 (2017) 062139.
- **Synchronization & error correction**: A.B. Boyd, D. Mandal, and J.P. Crutchfield, *Correlation-powered Information Engines and the Thermodynamics of Self-Correction*, Physical Review E **95**:1 (2017) 012152.
- Homeostasis: A.B. Boyd, D. Mandal, and J.P. Crutchfield, *Leveraging Environmental Correlations: Thermodynamics of Requisite Variety*, Journal Statistical Physics **167**:6 (2017) 1555-1585.
- Functional fluctuations: J. P. Crutchfield and C. Aghamohammadi, "Not All Fluctuations are Created Equal: Spontaneous Variations in Thermodynamic Function. arxiv.org:1609.02519.
- Controller structure: A.B. Boyd, D. Mandal, and J.P. Crutchfield, *Transient Dissipation and Structural Costs of Physical Information Transduction*, Physical Review Letters **118** (2017) 220602.
- Intelligent control: A.B. Boyd and J.P. Crutchfield, Maxwell Demon Dynamics: Deterministic Chaos, the Szilard Map, and the Intelligence of Thermodynamic Systems, Physical Review Letters 116 (2016) 190601.
- **NESS transitions**: P.M. Riechers and J.P. Crutchfield, *Fluctuations When Driving Between Nonequilibrium Steady States*, Journal of Statistical Physics (2017) in press. <u>arxiv.org</u>:1610.09444.
- Sensors: S. E. Marzen and J. P. Crutchfield, Prediction and Power in Molecular Sensors: Uncertainty and Dissipation When Conditionally Markovian Channels Are Driven by Semi-Markov Environments, arxiv.org: 1707.03962.
- **Modularity**: A.B. Boyd, D. Mandal, and J.P. Crutchfield, *Above and Beyond the Landauer Bound: Thermodynamics of Modularity*, <u>arXiv.org</u>, submitted.