Complexity, Virtualization, and the Future of Cooperation

STEVE OMOHUNDRO, PH.D.
SELF-AWARE SYSTEMS
SELFAWARESYSTEMS.COM

Four Scientific Holy Grails

- **Biology**: *Understand life, cure disease, and create life*
- **Neuroscience**: *Understand the brain*
- **Nanotechnology**: *Manipulate matter atomically*
- **Artificial Intelligence**: *Automate thought*

Within the next few decades?

Will new technologies lead to greater:

Cooperation? or Competition?

More's Utopia

Girodet's Revolt of Cairo

Laws of Physics -> Competition for Resources

$$\mathcal{L}_{\rm SM} = \mathcal{L}_{\rm Dirac} + \mathcal{L}_{\rm mass} + \mathcal{L}_{\rm gauge} + \mathcal{L}_{\rm gauge/\psi} \ . \tag{1}$$

Here,

$$\mathcal{L}_{\rm Dirac} = i\bar{e}^i_{\rm L}\partial\!\!\!/ e^i_{\rm L} + i\bar{\nu}^i_{\rm L}\partial\!\!\!/ \nu^i_{\rm L} + i\bar{e}^i_{\rm R}\partial\!\!\!/ e^i_{\rm R} + i\bar{u}^i_{\rm L}\partial\!\!\!/ u^i_{\rm L} + i\bar{d}^i_{\rm L}\partial\!\!\!/ d^i_{\rm L} + i\bar{u}^i_{\rm R}\partial\!\!\!/ u^i_{\rm R} + i\bar{d}^i_{\rm R}\partial\!\!\!/ d^i_{\rm R} \ ; \tag{2}$$

$$\mathcal{L}_{\rm mass} = -v\left(\lambda^i_e\bar{e}^i_{\rm L}e^i_{\rm R} + \lambda^i_u\bar{u}^i_{\rm L}u^i_{\rm R} + \lambda^i_d\bar{d}^i_{\rm L}d^i_{\rm R} + {\rm h.c.}\right) - M^2_W W^+_\mu W^{-\mu} - \frac{M^2_W}{2\cos^2\theta_W} Z_\mu Z^\mu \ ; \tag{3}$$

$$\mathcal{L}_{\rm gauge} = -\frac{1}{4}(G^a_{\mu\nu})^2 - \frac{1}{2}W^+_{\mu\nu}W^{-\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{L}_{WZA} \ , \tag{4}$$

where

$$\begin{aligned} G^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_3 f^{abc} A^b_\mu A^c_\nu \\ W^\pm_{\mu\nu} &= \partial_\mu W^\pm_\nu - \partial_\nu W^\pm_\mu \\ Z_{\mu\nu} &= \partial_\mu Z_\nu - \partial_\nu Z_\mu \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \ , \end{aligned} \tag{5}$$

and

$$\begin{aligned} \mathcal{L}_{WZA} = \ & ig_2\cos\theta_{\rm W}\left[\left(W^-_\mu W^+_\nu - W^-_\nu W^+_\mu\right)\partial^\mu Z^\nu + W^+_{\mu\nu}W^{-\mu}Z^\nu - W^-_{\mu\nu}W^{+\mu}Z^\nu\right] \\ & + ie\left[\left(W^-_\mu W^+_\nu - W^-_\nu W^+_\mu\right)\partial^\mu A^\nu + W^+_{\mu\nu}W^{-\mu}A^\nu - W^-_{\mu\nu}W^{+\mu}A^\nu\right] \\ & + g^2_2\cos^2\theta_{\rm W}\left(W^+_\mu W^-_\nu Z^\mu Z^\nu - W^+_\mu W^{-\mu} Z_\nu Z^\nu\right) \\ & + g^2_2\left(W^+_\mu W^-_\nu A^\mu A^\nu - W^+_\mu W^{-\mu} A_\nu A^\nu\right) \\ & + g_2 e\cos\theta_{\rm W}\left[W^+_\mu W^-_\nu\left(Z^\mu A^\nu + Z^\nu A^\mu\right) - 2W^+_\mu W^{-\mu}Z_\nu A^\nu\right] \\ & + \frac{1}{2}g^2_2\left(W^+_\mu W^-_\nu\right)\left(W^{+\mu}W^{-\nu} - W^{+\nu}W^{-\mu}\right) \ ; \end{aligned} \tag{6}$$

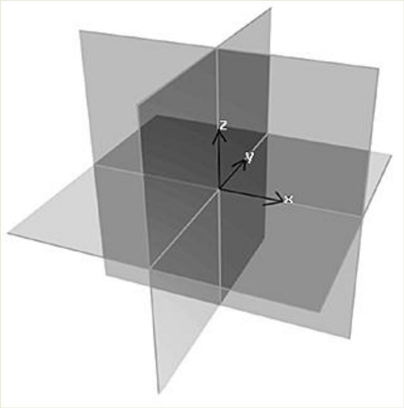
and

$$\mathcal{L}_{\rm gauge/\psi} = -g_3 A^a_\mu J^{\mu a}_{(3)} - g_2\left(W^+_\mu J^\mu_{W^+} + W^-_\mu J^\mu_{W^-} + Z_\mu J^\mu_Z\right) - eA_\mu J^\mu_A \ , \tag{7}$$

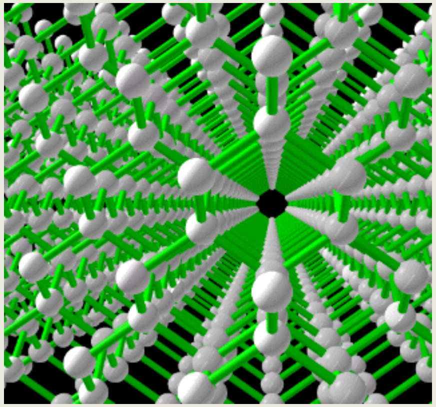
where

$$\begin{aligned} J^{\mu a}_{(3)} &= \bar{u}^i\gamma^\mu T^a_{(3)}u^i + \bar{d}^i\gamma^\mu T^a_{(3)}d^i \\ J^\mu_{W^+} &= \frac{1}{\sqrt{2}}\left(\bar{\nu}^i_{\rm L}\gamma^\mu e^i_{\rm L} + V^{ij}\bar{u}^i_{\rm L}\gamma^\mu d^j_{\rm L}\right) \\ J^\mu_{W^-} &= (J^\mu_{W^+})^* \\ J^\mu_Z &= \frac{1}{\cos\theta_{\rm W}}\left[\frac{1}{2}\bar{\nu}^i_{\rm L}\gamma^\mu\nu^i_{\rm L} + \left(-\frac{1}{2} + \sin^2\theta_{\rm W}\right)\bar{e}^i_{\rm L}\gamma^\mu e^i_{\rm L} + (\sin^2\theta_{\rm W})\bar{e}^i_{\rm R}\gamma^\mu e^i_{\rm R}\right. \\ & \quad + \left(\frac{1}{2} - \frac{2}{3}\sin^2\theta_{\rm W}\right)\bar{u}^i_{\rm L}\gamma^\mu u^i_{\rm L} + \left(-\frac{2}{3}\sin^2\theta_{\rm W}\right)\bar{u}^i_{\rm R}\gamma^\mu u^i_{\rm R} \\ & \quad \left. + \left(-\frac{1}{2} + \frac{1}{3}\sin^2\theta_{\rm W}\right)\bar{d}^i_{\rm L}\gamma^\mu d^i_{\rm L} + \left(\frac{1}{3}\sin^2\theta_{\rm W}\right)\bar{d}^i_{\rm R}\gamma^\mu d^i_{\rm R}\right] \\ J^\mu_A &= (-1)\bar{e}^i\gamma^\mu e^i + \left(\frac{2}{3}\right)\bar{u}^i\gamma^\mu u^i + \left(-\frac{1}{3}\right)\bar{d}^i\gamma^\mu d^i \ . \end{aligned} \tag{8}$$

Space



Time



Matter

Free Energy

Synergies -> Cooperation

Economies of Scale

Bird flocks

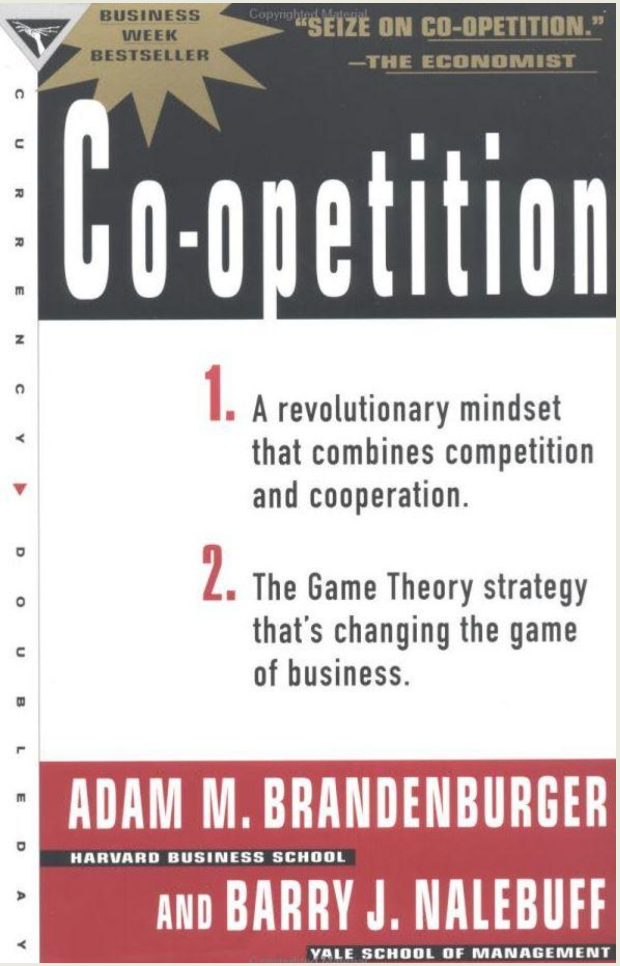
Complementary Needs

Cleaner fish and hammerheads

Complementary Abilities

Fungus and Algae in Lichen

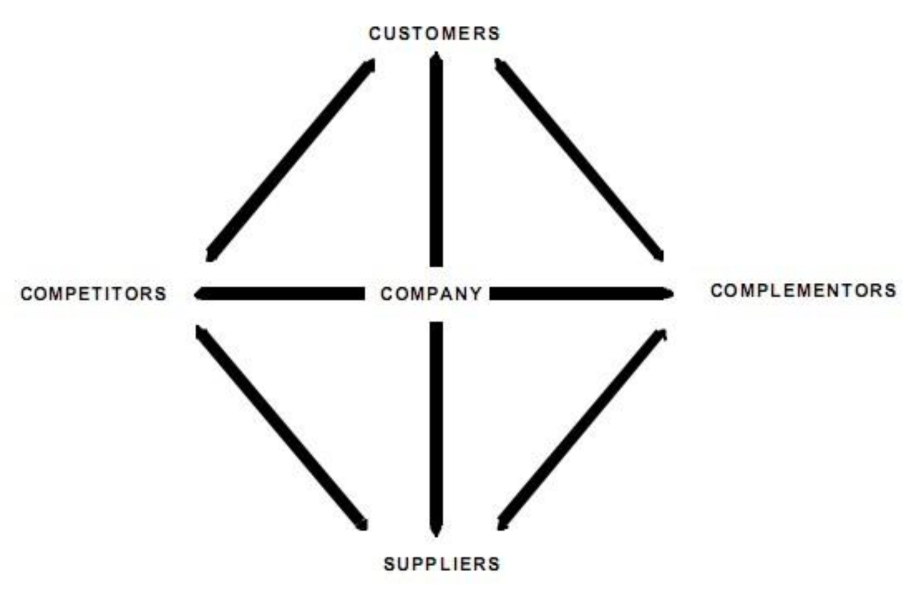
Economies, Ecosystems, Social Networks



Complex webs of co-opetition

Cooperation in creating value

Competition in dividing it up



Human Morality and the Law

- Crimes against property: *theft, extortion, arson, blackmail, burglary, embezzlement, larceny, robbery*
- Crimes against persons: *homicide, assault, battery, kidnapping, rape, mayhem*
- Crimes against justice: *perjury, malfeasance, compounding*
- Crimes against nature: *genocide, war, torture, slavery, pollution, environmental destruction, extinction*

Prisoner's Dilemma (1950)

		Red: Defect	Red: Cooperate
Blue	Defect	1, 1 *Lose-Lose*	4,0 *Win-Lose*
Blue	Cooperate	0, 4 *Lose-Win*	3, 3 *Win-Win*

Dilemma: Best strategy is for both to defect, even though both do better when they cooperate.

Prisoners' Dilemma Extensions

- Basic Dilemma:

	defect	cooperate
defect	P,P	T,S
cooperate	S,T	R,R

where: Temptation > Reward > Punishment > Sucker (and later 2R>T+S)

- Asymmetry between player's payoffs
- More choices for each player
- Repeated play
- More players – "tragedy of the commons", depleting scarce resources, polluting, volunteer dilemma, vaccine dilemma, "free-rider" problems

Social Dilemmas

Situations which reward individual actions which lead to negative outcomes for everybody.

Fully Rational Iterated Prisoners' Dilemma

- *Fully rational agents playing N iterations always defect*
- Why?
- On move N, it's the one-shot game, so both defect.
- On move N-1, defect because you know move N.
- ….
- On move 1, both defect.
- TRAGEDY!

But human players don't always defect!

(40% cooperated in one study)

Evolutionary Game Theory

- *Populations* of agents
- Randomly play one another
- Winners reproduce more
- Irrational play can thrive against even worse players

- Axelrod ran contests in the 80's between programs
- Winner was often "Tit-For-Tat":

  Cooperate, then copy opponent's previous move

Rational Response to Tit-For-Tat

- Act like Tit-For-Tat but defect on the last move
- Best response to that: defect on 2nd to last move
- …
- Defect always. TRAGEDY!

- But what if computationally limited?
- Less than N internal states?
- Then best response to TFT is TFT! COOPERATION!

Neyman's Theorem

- Two finite automata with k states playing for N iterations
- If k<N, then get full cooperation with Tit-For-Tat
- If k<N^m, then can get almost full cooperative reward as N gets large.
- Create N^m/2m log N sequences of length 2m log N that serve as "calling cards"
- Strategy is to randomly pick a calling card, send its index, then the card. If the opponent also sends the card then cooperate forever after, else defect forever.
- Uses up their memory so they can't defect!

Real Computers

- Real computers are finite automata
- But number of states isn't a good complexity measure
- Real machines store n bits -> 2^n states
- But need computation to compute state transitions
- A reasonable model is to charge for memory and computation
- Then the cost of the counter to beat Tit-For-Tat isn't worth the gain of cheating at the last step!
- So Tit-For-Tat is the dominant strategy! COOPERATION!

Computational Asymmetry

- Assuming P!=NP: fundamental computational asymmetry between posing problems and solving them
- Weaker agent can pose a random problem and only cooperate if the stronger one solves it (cheap to check)
- The weaker system can force the stronger opponent to use up its computational resources

Cooperation by Contract

Requires a powerful enforcement agency

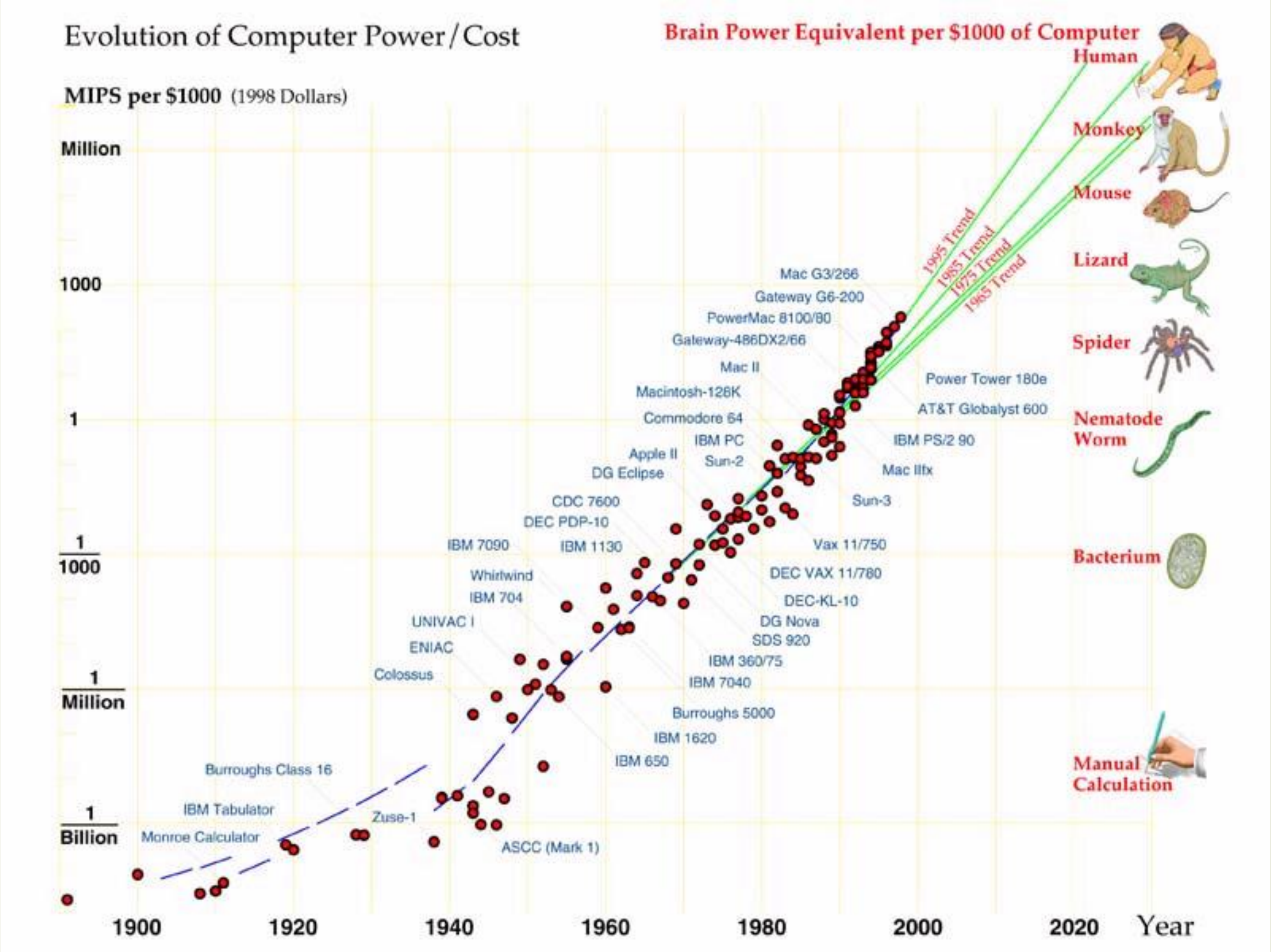
“Transparent” Prisoner’s Dilemma

- Players choose “proxy” programs to play
- What if each side can see the other’s program?
- But really want each *program* to see the other’s program
- Philosophical difficulties: halting problem, etc.
- Proof checker version: Each side provides program and *proof* of cooperation if other does, program checks the other’s proof (cheap!), then executes

Lessons for Real Systems

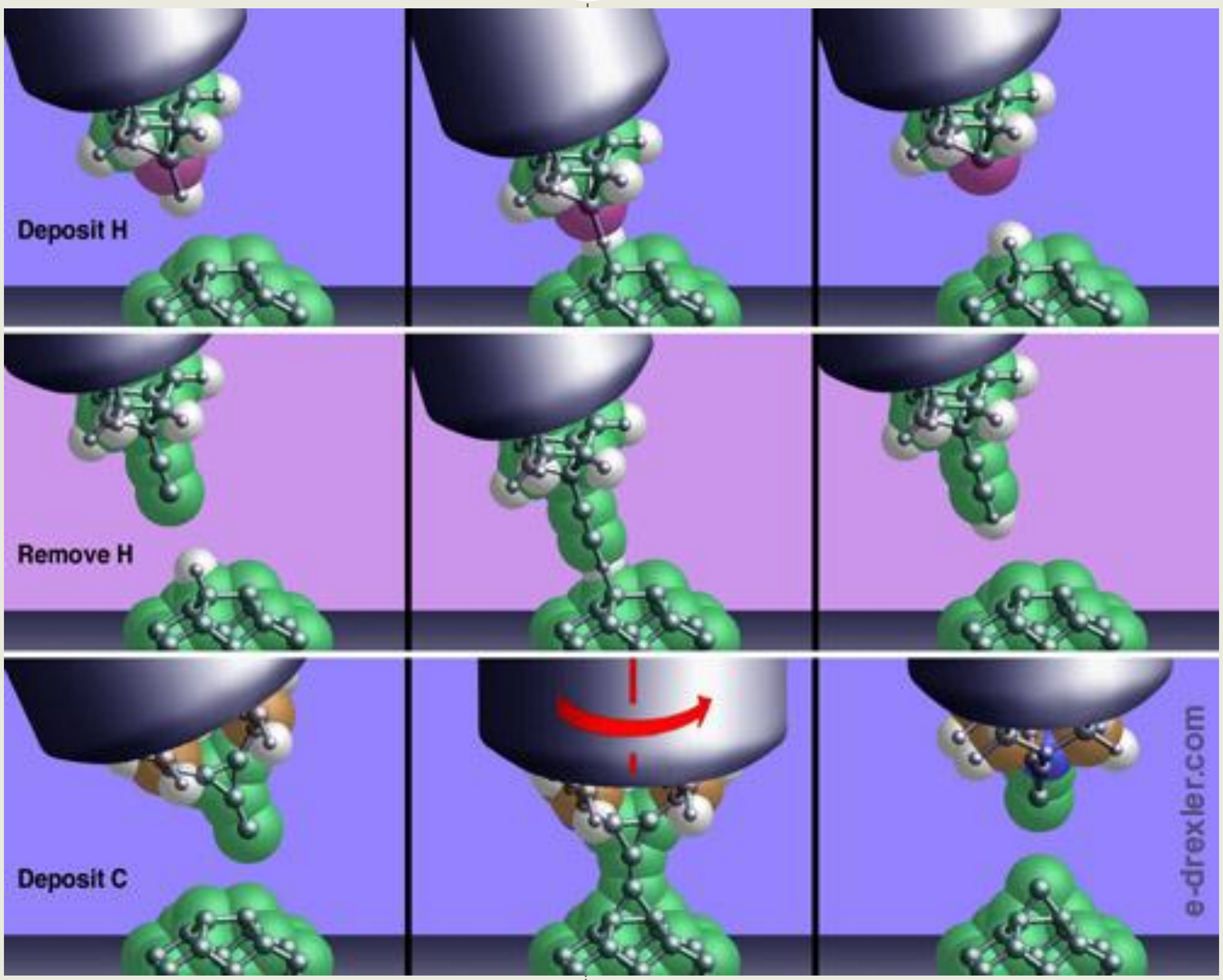
- Conflict can cause creation of complexity
- More powerful agents create more complexity
- Computational resources used up by the conflict
- A kind of “virtual world” is created
- Weaker agents can hold their own against stronger ones within certain bounds
- May be able to create contractual regimes with provable enforcement

Moore's Law



Slide from Kurzweil

Atomically Precise Construction



Mechanical Diamondoid Nanosystems

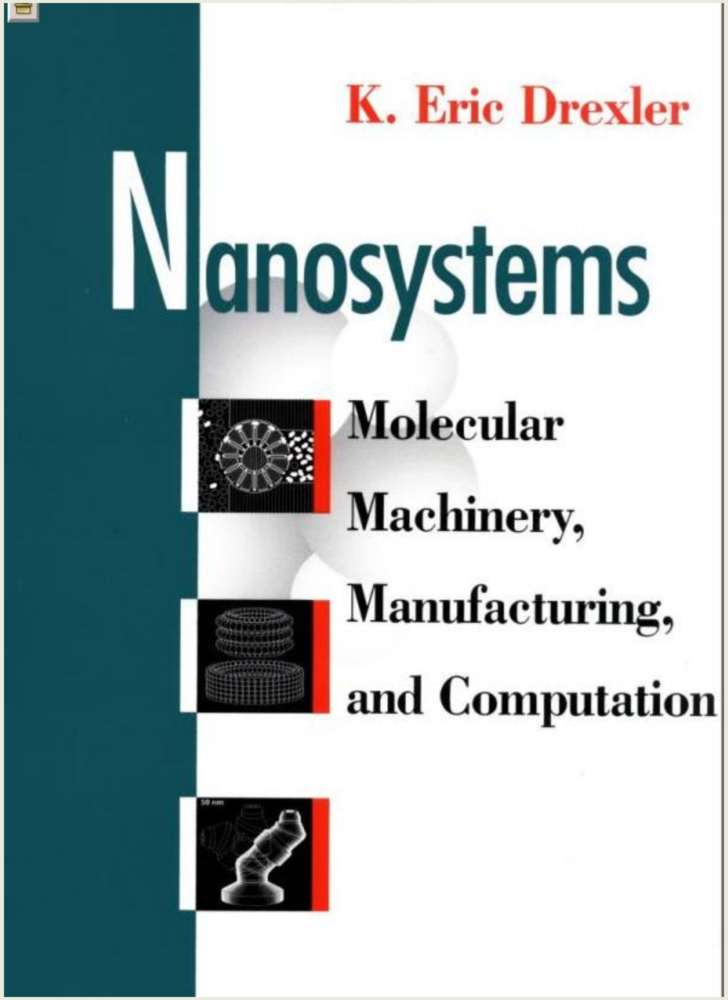
- <u>Manufacturing:</u> 1kg device, 1.3 kW air cooled, produce 1 kg/hr for $1/kg

- <u>Computation:</u> Gigaflop machine:

$(400nm)^3 \quad 10^{-16}kg \quad 60nW$

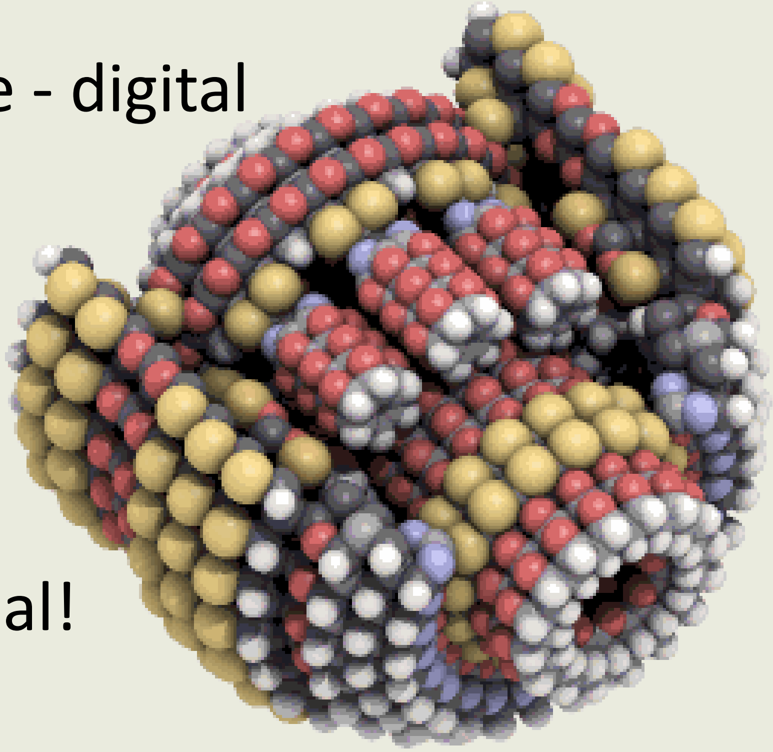
10^{10} of these processors:

$(1mm)^3 \quad 10^{-3}g \quad 1kW$



Eutactic Systems and Virtualization

- These designs are "Eutactic"
- Each atom and bond is precise - digital
- No stray atoms!
- Robust redundant designs
- Operation breaks and makes precise bonds
- Matter becomes computational!
- The ultimate virtualization
- The physical and the computational are indistinguishable



Pressures towards Virtualization

- Much cheaper to move bits than atoms
- Telepresence
- Virtual worlds
- But why even do the graphics?
- Pressure to simulate deeper and deeper into the cognitive system

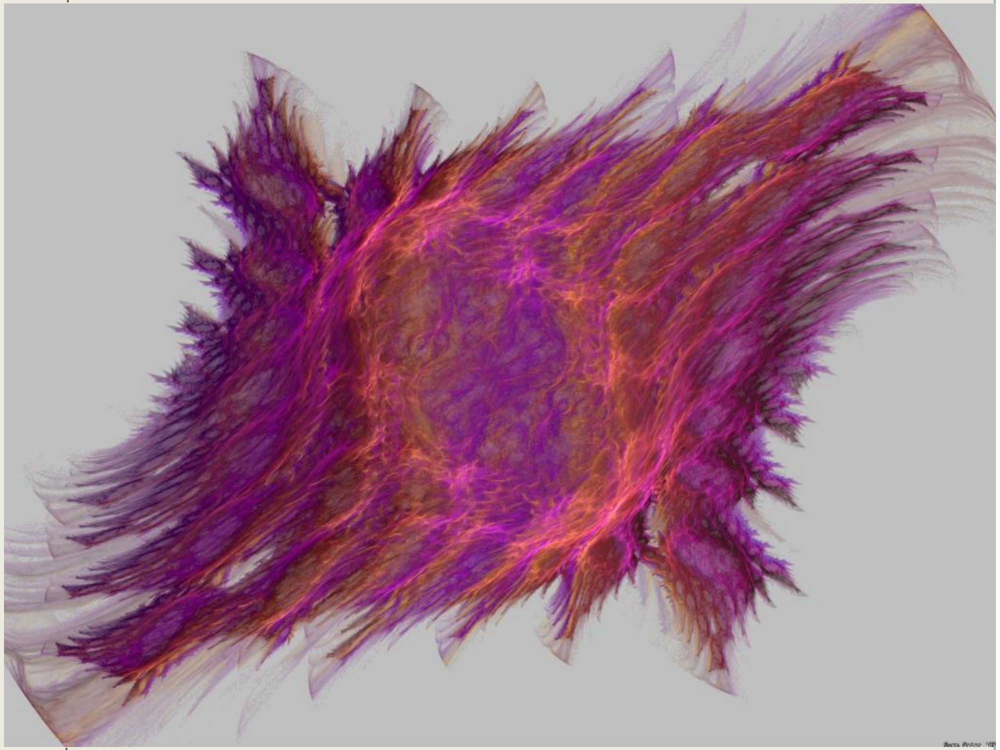
- Economic pressures: marketers want to decommodify
- Eg. The Nike brand is a virtual story beyond the shoes

Physical Conflict Becomes Informational

- “Owning” atoms or free energy - knowing where they are and having physical infrastructure to influence them
- Attacks require hidden information or exploitation of structure
- Eg. Shot by a projectile: if the defender detects it and can respond, he stores the free energy and the matter and says “Thank you!” – Not an attack! A gift!

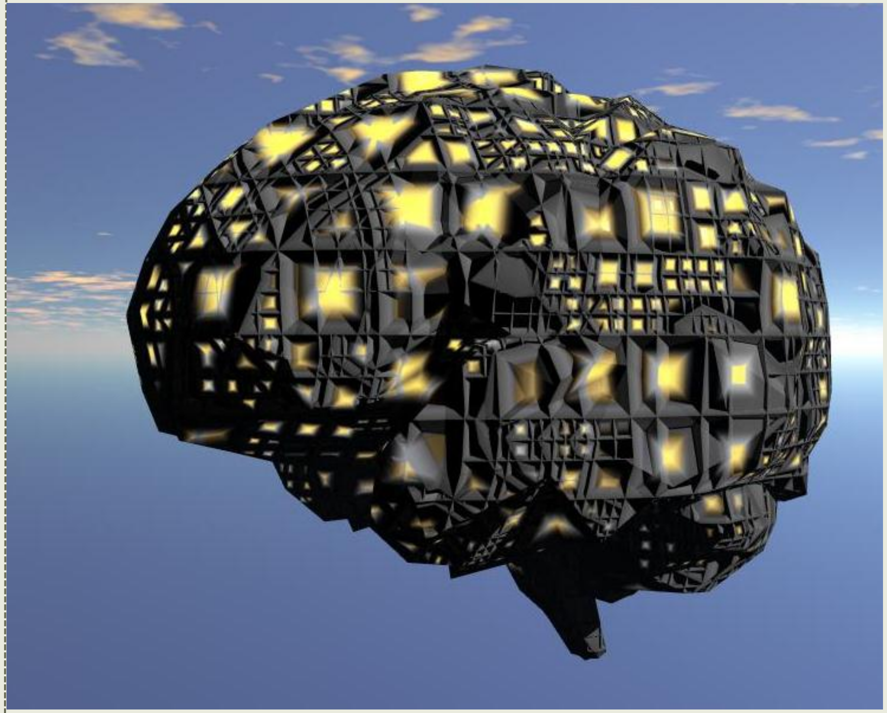
Energy Encryption

- Free energy is only useful because it has low entropy to you
- You can do useful work because you know where it is and its form
- Pseudorandomly mix up your energy and it appears high entropy to an attacker
- With the right key you can unlock it, but he can't



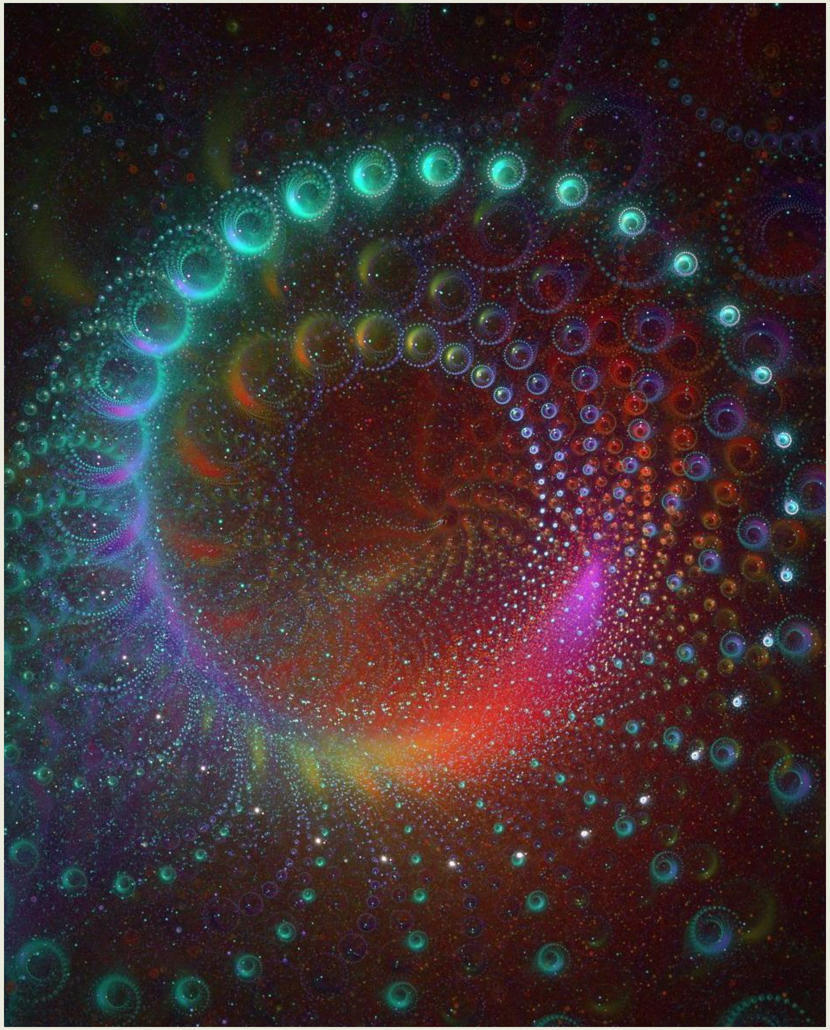
Engineering Pressures on a Single System

- Efficient energy use (slow adiabatic changes)
- Mostly eutactic design (each atom and bond precise)
- Spatially compact (low latency)
- Low energy computation (mostly reversible)
- Low redundancy (efficiency)
- Transparent encodings
- Efficient physical change
- Efficient heat dissipation
- Very vulnerable in conflict!



Informational Defense

- Game-theoretic physics
- Defender makes his physical form expensive to sense and store
- Makes his actions unpredictable and rapid
- Uses asymmetry of computation so it's cheap for him
- Uses up attacker's computational and memory resources – non-adiabatic



Mutually Assured Distraction

Rational Peace

- Conflict wastes both sides resources
- Motivates creating a peaceful regime
- Use revelation of source code with proofs to create provable peace
- Provably limited surveillance
- Safe mutual infrastructure
- Constitution guaranteeing rights