Complexity and coarse graining: Can you divide and conquer anything?

Chris Stephens,
Instituto de Ciencias Nucleares, UNAM
LEVERHULME LECTURES, University of Essex
July-August 2006

Content

- Part 1: Complexity
  - What is it? Some definitions
  - Phenomenology and taxonomy
    - Physical complexity
    - Symbolic complexity
  - Modeling complexity
- Part 2: Coarse Graining
    - Why?
    - What?
    - How?
  - Divide and conquer – from the micro to the macro
    - Renormalization group
      - Examples in physics
      - Examples in evolutionary dynamics
    - Coarse graining in systems where the EDOF change
    - Coarse graining in complex systems

Complexity:
What is it?

“Definitions”

- in computer science…?

- **Computational complexity** – refers to the computational resources necessary to solve a given problem
- **Descriptive complexity -** of a string is the length of the string's shortest description in some description language

In both manifestations the most "complex" problems are random!

Some other "definitions"…

- A complex system is a highly structured system, which shows structure with variations (Goldenfeld and Kadanoff)
- A complex system is one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve (Whitesides and Ismagilov)
- A complex system is one that by design or function or both is difficult to understand and verify (Weng, Bhalla and Iyengar)
- A complex system is one in which there are multiple interactions between many different components (D. Rind)
- Complex systems are systems in process that constantly evolve and unfold over time (W. Brian Arthur)
- "Complex things exhibit complex behavior" (Parisi)

You always see: "many degrees of freedom" and "non-linear" – that just about covers everything! Even quantum field theory, where we have an infinite number of degrees of freedom.

Effective Complexity (Gell-Mann)

- Descriptive Complexity, but measured not on an observed phenomena but through a subjective interpretation of interest to the observer, i.e., a *model* - an algorithm for specifying a probability distribution over the observed data.

But this sounds very subjective, as it depends on our model, how good it is and how well we can test it. Also, what data?

Complexity

Phenomenology and taxonomy

“Physical” Complexity

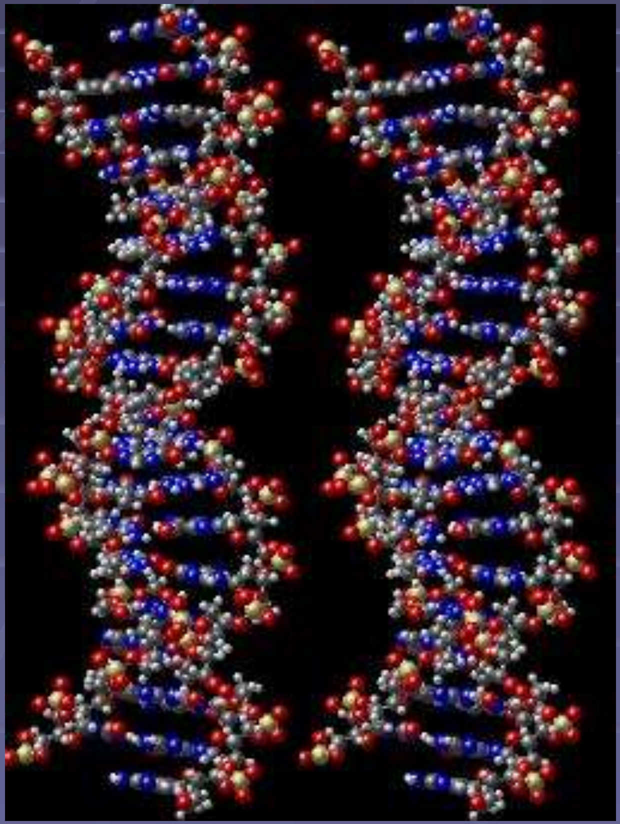
So what do we know for sure **is** complex and intelligent?

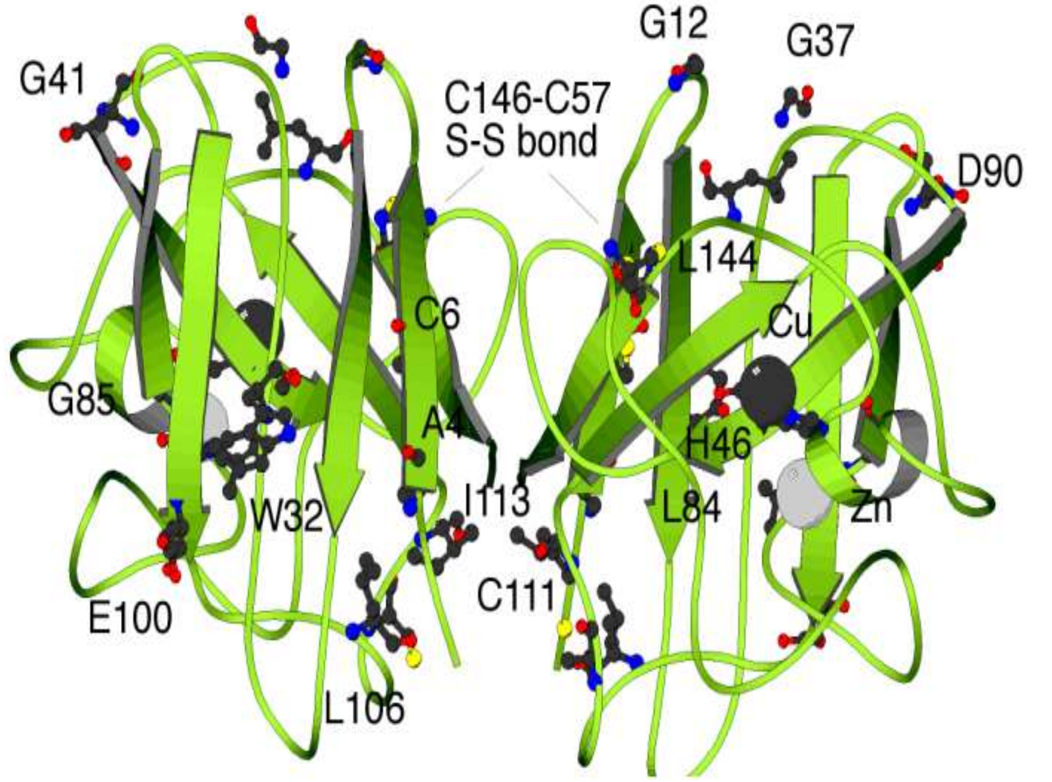
Well, what about this?

No?…this then…?

And this …?

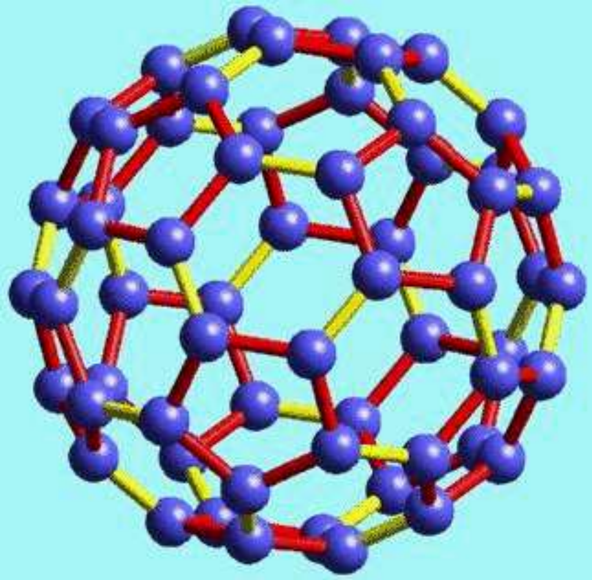
And these?



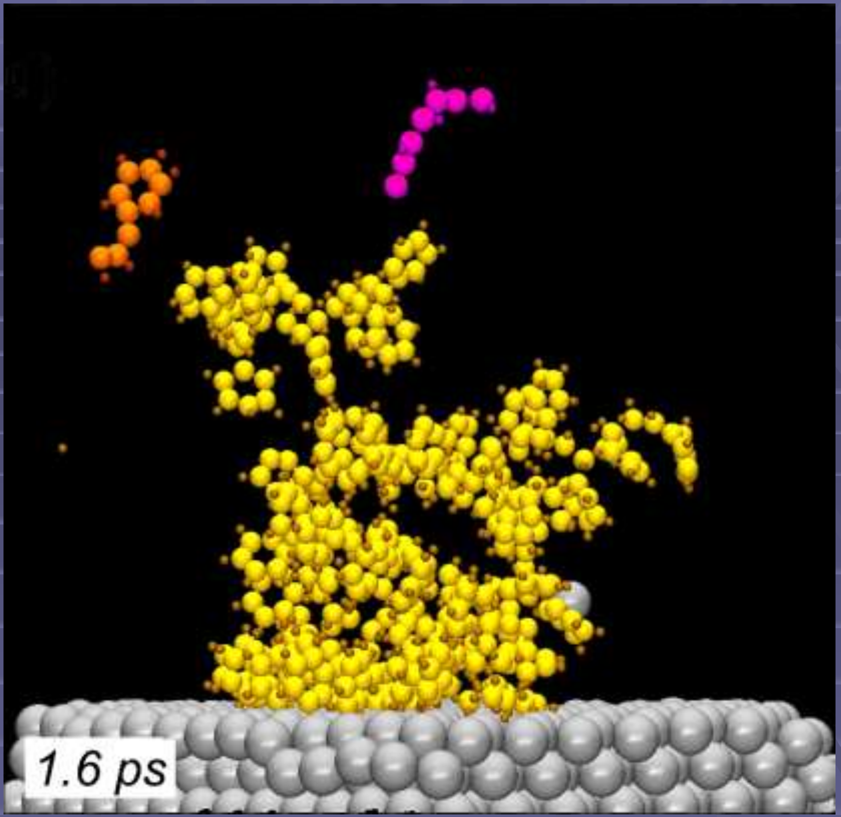


Model illustrating the formation of a misfolded species (M) from a folding intermediate (I). The region of the protein that misfolds is shown in red. The misfolded protein itself, or a self-assembled form, may be toxic to cells, leading to disease. The black arrows represent the relative rates of the various conformational events under native physiological conditions in the absence of mutation. The blue dash arrows represent the possible effects of mutation.

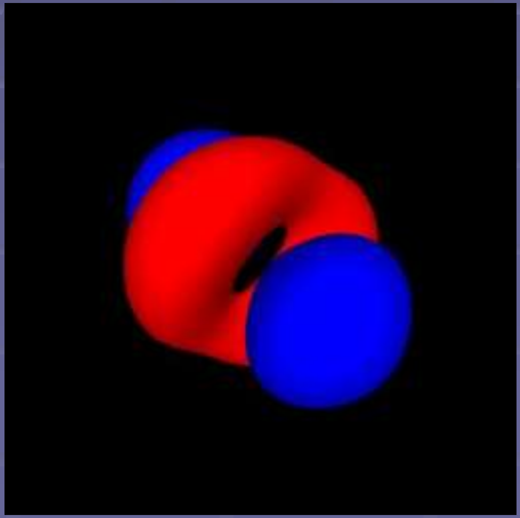
And what about these?



Buckyball C_{60}



Polystyrene on a silver surface



n=3, l=2 energy level of H

So, maybe we can agree on what is definitely complex, and what is definitely not complex. But where do we change from one to the other?

The "Edge of Chaos"?

On the "Edge of Chaos" in "micro"-physics?

Barkhausen effect – "avalanches" of magnetic domains

"Dirty"

Near critical

Typical critical phenomenon showing <u>collective behavior</u> and scaling $Y \sim X^a$

But…

- Only one important length scale – the correlation length – that governs the scale of “collectivity”; Scale invariant near critical point (phase transition) – maximal “collectivity”
- Only one type of effective degree of freedom – a magnetic domain “avalanche”, but …
- Complex? Once the spectrum of “avalanche” sizes is given then there’s nothing much more to be said. Not very interesting living on the “Edge” in physics!
- The same is true for other canonical critical or self-organised critical phenomena

So what does distinguish the phenomena that we "agree" are complex from those that we "agree" aren't complex?

- A "hierarchy" of many <u>different</u> length scales
- Effective degrees of freedom ("collectivity") at different length scales are <u>qualitatively</u> different with different effective interactions

- Systems are <u>adaptive</u>
- Dynamical evolution depends on <u>many different rules/strategies</u>
- Systems "<u>learn</u>" (feedback from environment to system which is then used to update rules)

- More complex "<u>behavior</u>" (the "phenotype")

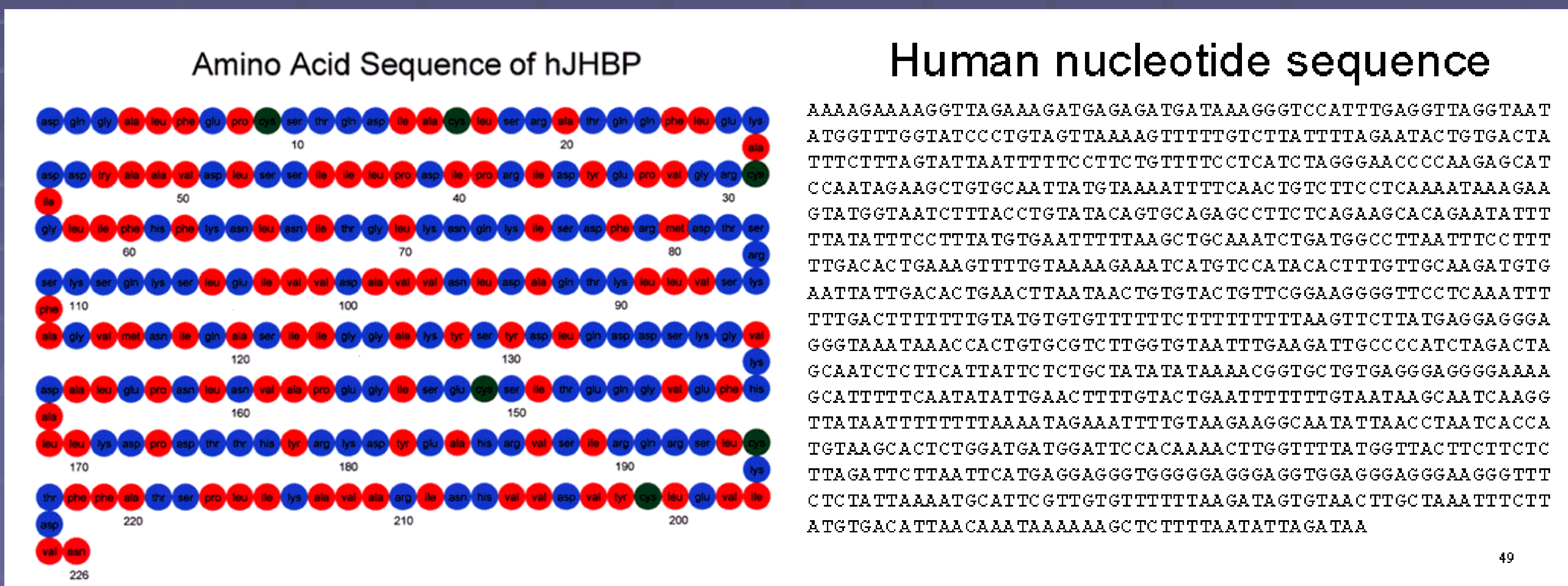
The trouble is that the definitions of complexity given before do not discriminate – too many false positives!

Perhaps contain some necessary conditions but certainly not sufficient

Symbolic Complexity

To be, or not to be--that is the question:
Whether 'tis nobler in the mind to suffer
The slings and arrows of outrageous fortune
Or to take arms against a sea of troubles
And by opposing end them. To die, to sleep--
No more--and by a sleep to say we end
The heartache, and the thousand natural shocks
That flesh is heir to. 'Tis a consummation
Devoutly to be wished. To die, to sleep--
To sleep--perchance to dream: ay, there's the rub,
For in that sleep of death what dreams may come
When we have shuffled off this mortal coil,
Must give us pause.

What about complexity in this case?



…and here?

aaaa aaaa aaaa aaaa aaaa aaaa aaaa… “crystalline”

asmjgre fj sdjf s rege geoiie rgeasdffi… “amorphous”

…_ _ _ … … _ _ _ … …_ _ _ … … _ _ _ … “layered”

1001 110 11001 1111 10101 1 10010 101 1101 1 10010 10010 … “?”

If you are married or are a man and woman living together as if you are married you must claim jointly … “complex”

How might we even recognize something as being “complex”?

What about a “symbolic” Edge of Chaos?

the the the the the the the the the the the the the the…. ordered

mercy proudly rush interrogative registered clansman therapeutic… disordered

Parameter to distinguish between the ordered and disordered states…
s – where:

Zipf's law may be stated mathematically as:

$$f(k;s,N)=\frac{1/k^s}{\sum_{n=1}^{N}1/n^s}$$

where N is the number of elements, k is their rank, and s is the exponent characterizing the distribution. In the example of the frequency of words in the English language, N is the number of words in the English language and, if we use the classic version of Zipf's law, the exponent s will be equal to unity. $f(k; s,N)$ will then be the fraction of the time the kth most common word occurs.

In Hamlet (and more generally in natural language) s is about 1

So, natural language is on the “Edge of Chaos”!

Is that now an adequate description of Hamlet? That the frequency distribution of words is scale invariant with exponent s?

NO!

So what’s in Hamlet that isn’t in a “sandpile”?

Language from the point of view of a martian statistical physicist

$$< \alpha_i > \quad i \in a, b, ..., z$$

t is much higher frequency than x

$$< \alpha_i \alpha_j > \quad i, j \in a, b, ..., z$$

t and h are much more correlated than x and q; high peak in adjacent positions

$$< \alpha_i \alpha_j \alpha_k > \quad i, j, k \in a, b, ..., z$$

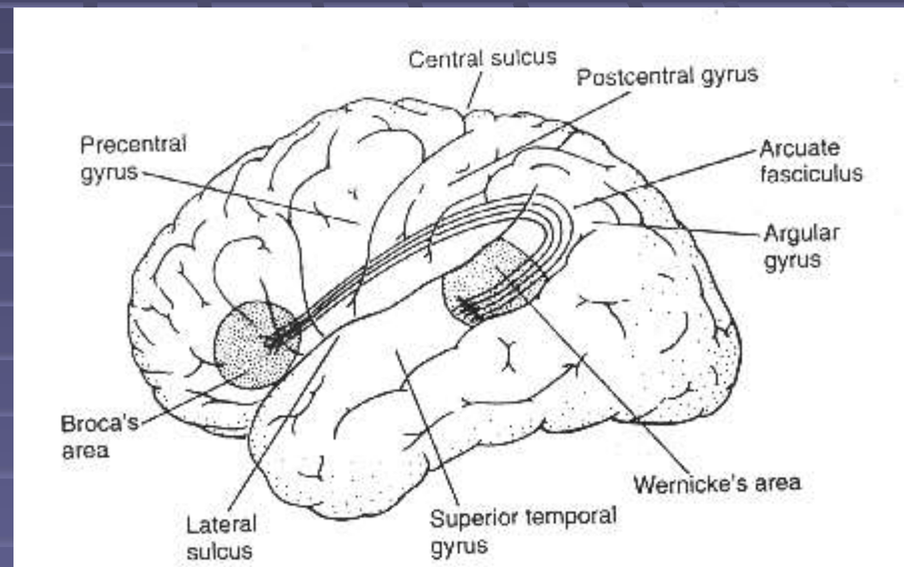
t, h and e are highly correlated in adjacent positions; detection of the "word" effective degree of freedom. Can then look at correlation functions between these new EDOF.

-
-
-

So Hamlet will show highly non-trivial correlation functions that show neither order nor disorder, but much more structure than Edge of Chaos. The correlation functions are our "structure detector/measurement device". But are statistical correlation detectors sufficient?

What's a better measurement device?

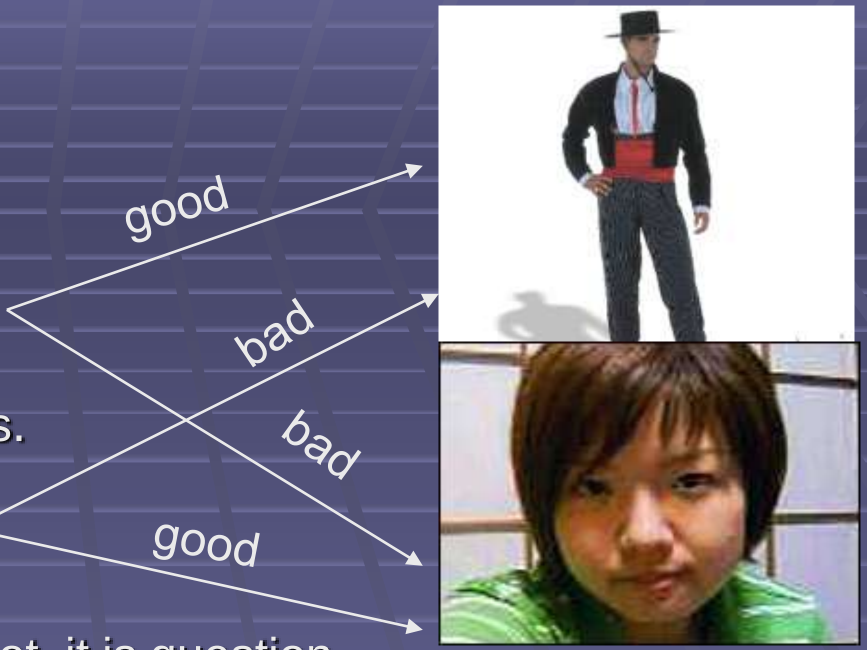
To be or not to be that is the question.



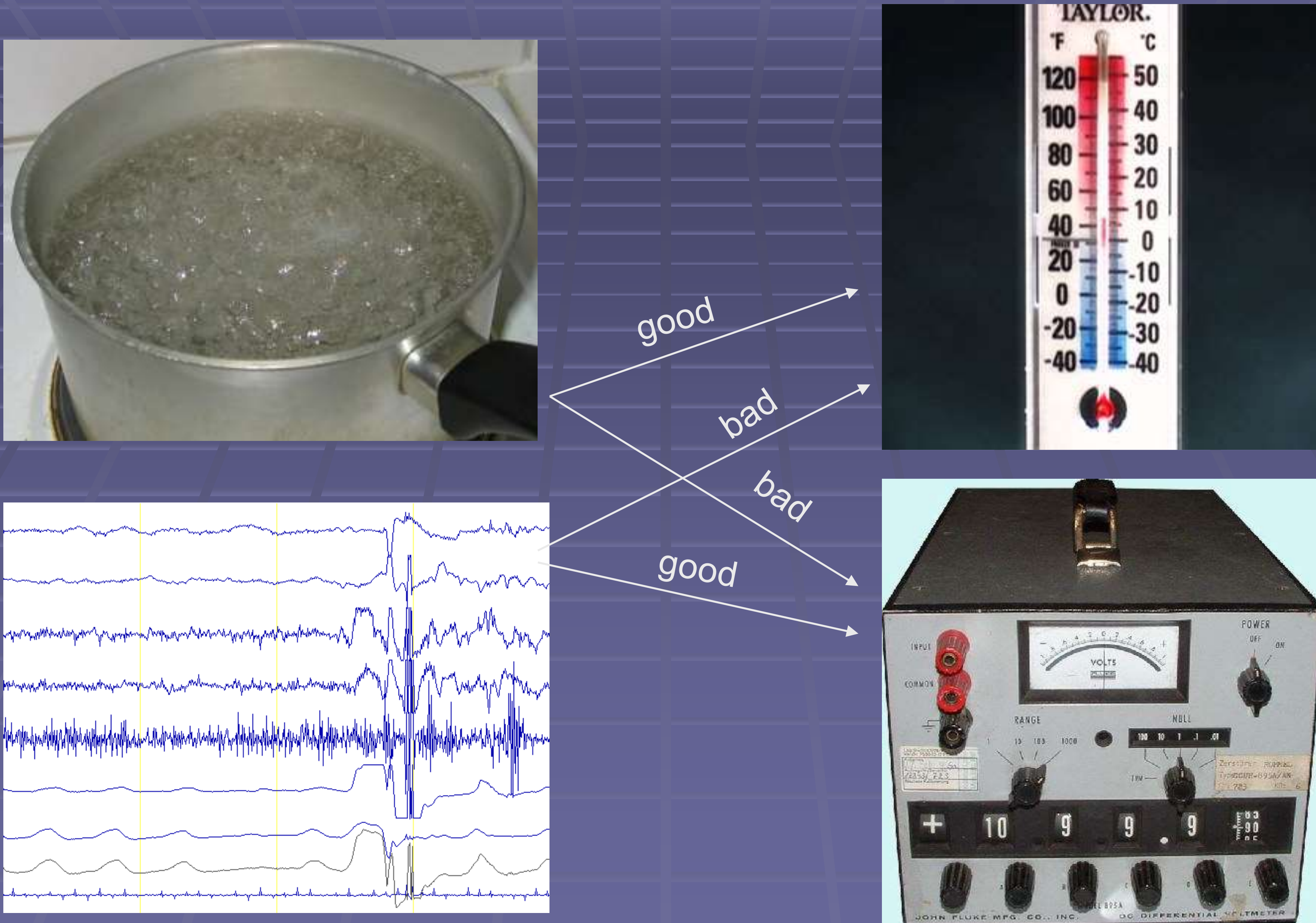
This measuring device certainly seems capable of measuring complexity.
Or does it…?

How good is your measuring apparatus?

- To be or not to be that is the question.
- Para ser o no ser que es la pregunta.
- Om te zijn of te zijn niet dat de vraag is.
- あるためまたはないため質問である。
- Because of a certain or because it is not, it is question.
- Because or it is not for the sake of, that having asked and being convinced.
- Being not to be for the sake of, or that that, you ask, are convinced.
- It is that without having for the sake of, or, you ask, are convinced.



But is this any different than the physical world?



- So, is complexity more a property of a system **and** a measuring apparatus together rather than something intrinsic to a system itself?
- Different measuring devices measure it in different ways and some are more appropriate than others

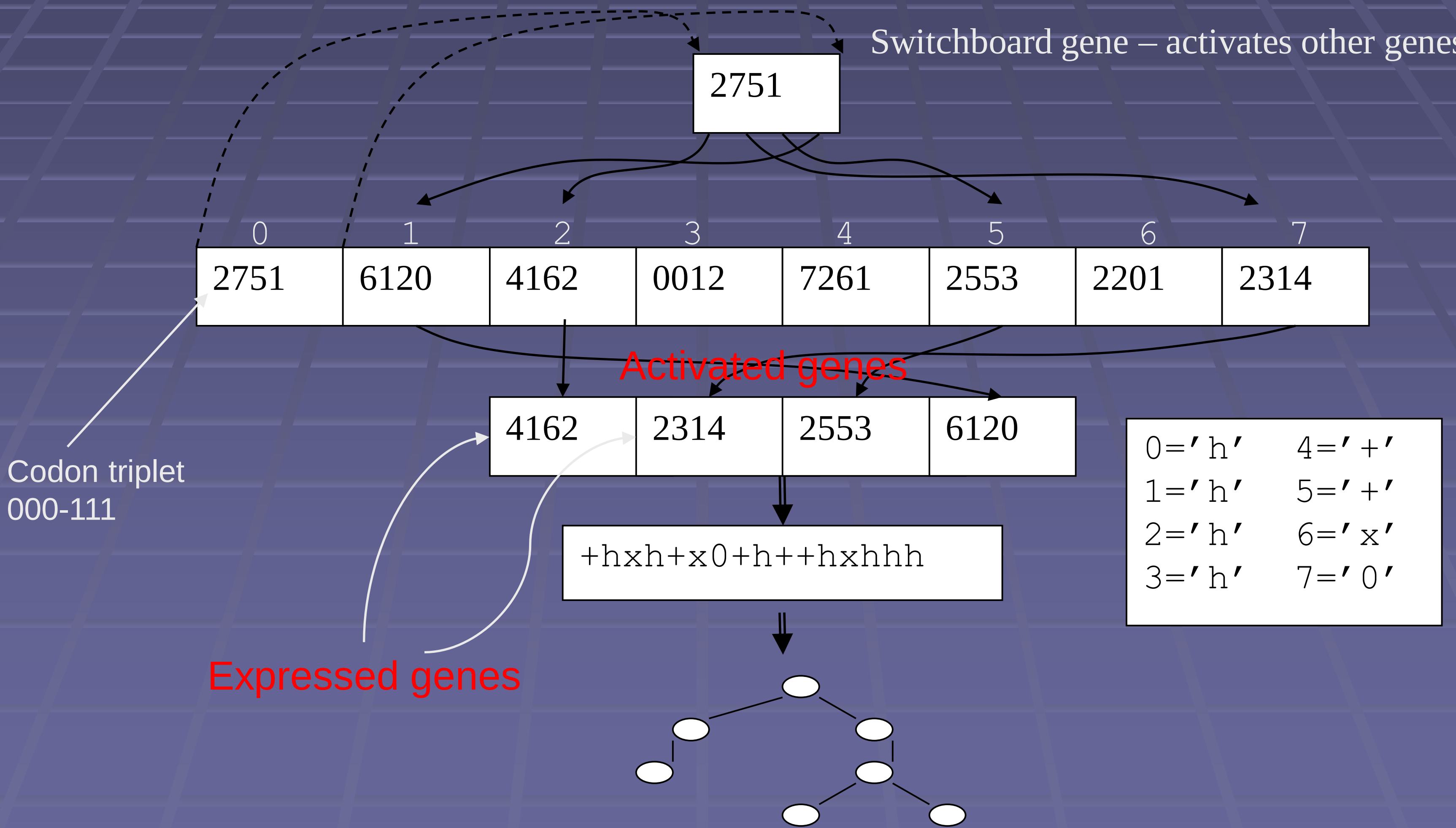
How do we know if our measuring device is good?
How do we distinguish between something that is complex but viewed with an inappropriate device versus something that is not complex?
In other words, do we know if we'd recognise it if we saw it?

So let’s see how we might look for it…

…by considering a model genetic system that has features of both genetic programming and genetic regulatory networks, showing how various forms of degeneracy in the genotype-phenotype map can induce complex and subtle behavior in the dynamics that lead to enhanced evolutionary robustness and can be fruitfully described in terms of an elementary “algorithmic language”.

The “gene expression inspired” Genotype-Phenotype map

(Angeles, Stephens, Waelbroeck – Biosystems 47, 1998)



Codon mapping

Eight codons (0-7) that code for the production rules of a closed grammar

```
<e> :: = tanh(<e>)       (0)
         tanh(<e>)       (1)
         tanh(<e>)       (2)
         tanh(<e>)       (3)
         add(<e>,<e>)    (4)
         add(<e>,<e>)    (5)
         X               (6)
         0               (7)
```

Example Mapping

```
4567 1623 0021 4401
```

Activated genes

Each codon will *always* make the same choice. Thus, codon 0 will always perform the mapping `e -> tanh(e)`, while codon 7 will always perform the mapping `e -> X`. The mapping steps are as follows:

```
4 -> add(e,e)
5 -> add(add(e,e),e)
6 -> add(add(X,e),e)
7 -> add(add(X,0),e)
1 -> add(add(X,0),tanh(e))
6 -> add(add(X,0),tanh(X))
```

"sentence"

```
++X0 hXhh hhhh ++hh
```

"word"

"letter"

"Switchboard" gene governs "syntax"

Multiple hierarchical levels of degeneracy

- Codon level – when there are more codon values than production rules
- Mapping level – when a production rules is represented multiple times (e.g. h and +)
- Gene activation – only a limited subset of genes are activated by the switchboard
- Gene expression – not all activated genes are expressed
- Gene expression – different ways of expressing the same thing (e.g. 4X=+++X+++XXX000 or +X+X+XX)

Are all these degenerate genotypes equally (effectively) fit?

Strategy – choose simple problem where structure may be readily interpreted

Symbolic regression: f(x) = 4x

Fitness function: $error = \sum_i \left(f(x_i) - v(x_i) \right)^2$

Program output

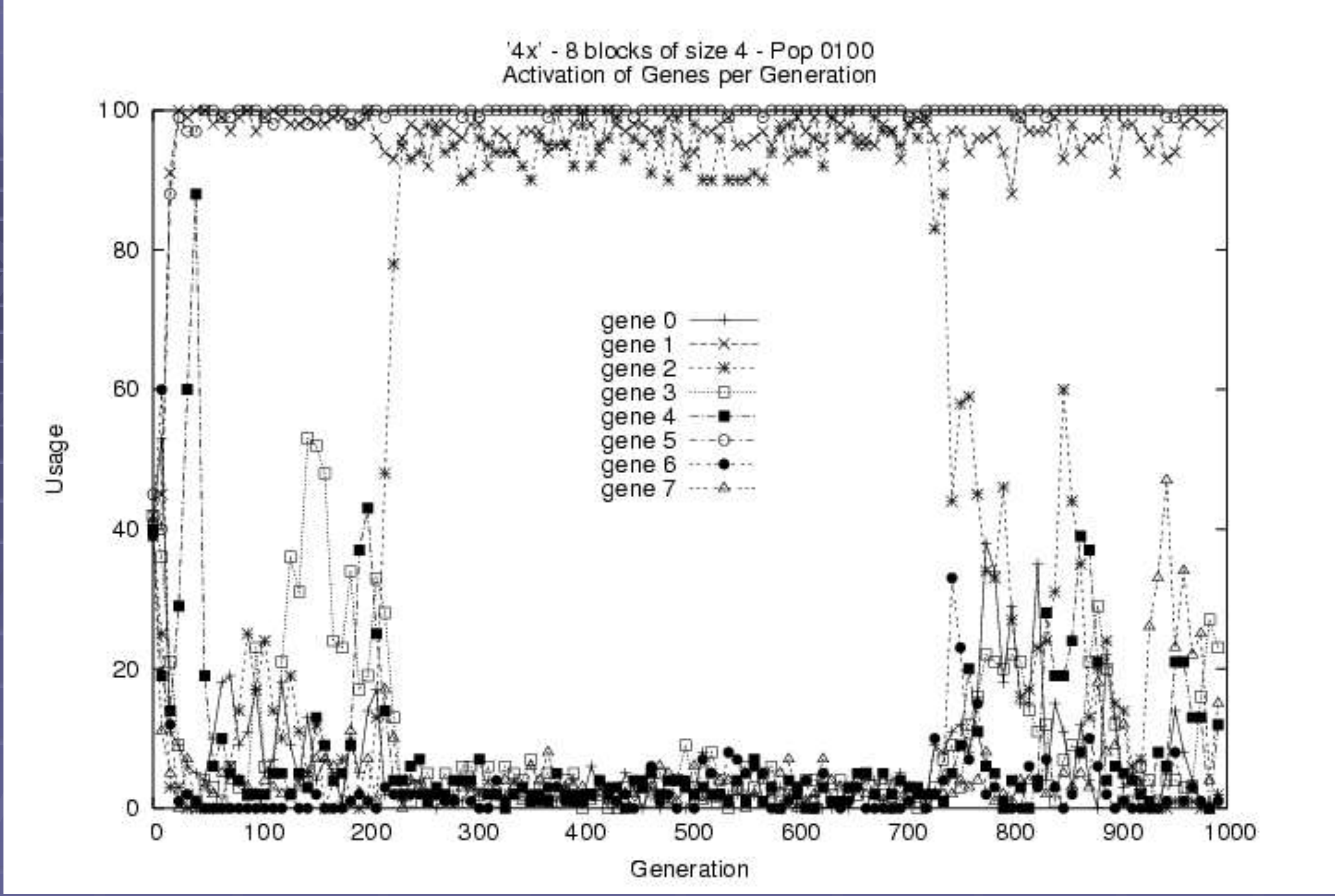
Population size - 100

Rank selection

Mutation (bit level) 0.01, Recombination (1pt, restricted to boundary to genes) 0.9

Runs – 30

- Interest is in studying emergent structures that are associated with the hierarchy of degeneracy of the Genotype-Phenotype map
- These can be contingent in their specific form



- If random expect to see a gene activated in 50% of the population
- By gen 25 > 95% of switchboard genes are of form 5***, > 60% of form 55**; 92% have **1* and 83% **11
- Check contents of genes 5 and 1!

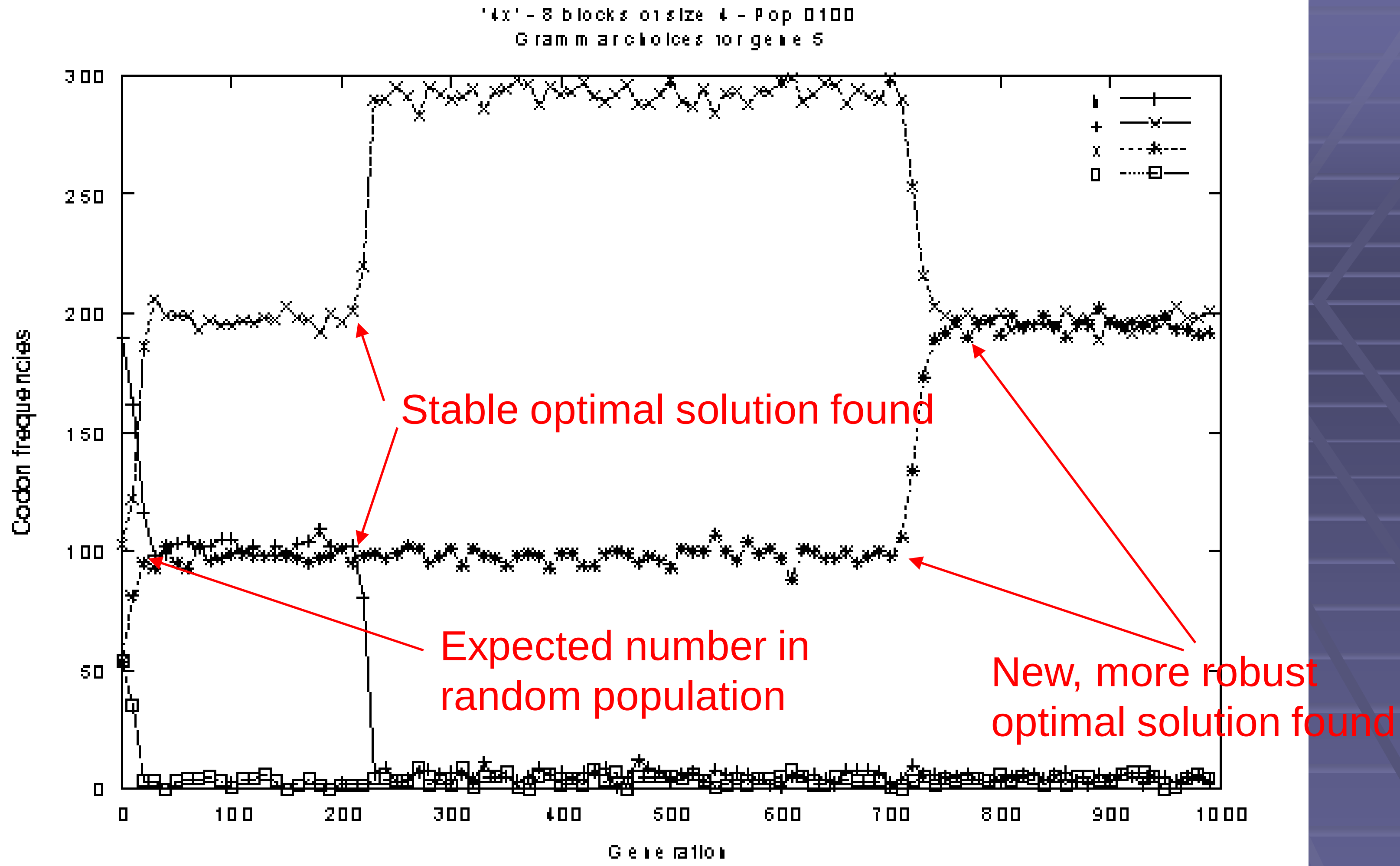


Fig. 4. Grammar choices for gene 5 per generation.

Gene 5 – important “core” gene – stability implies strong selection pressure
Gene 1 – contains 90% more X than random; more variation than gene 5

Description of the Dynamics

- Early dynamics (t < 213 gen) – based on 551* switchboard, i.e. two activated and three expressed genes; 5=++hX, 1=XX00 = 2X + 2tanhX
- Intermediate dynamics (213<t<704) – stable optimal solutions found; first (5533), has two activated and four expressed genes; no. of + in 5 gene increases by 50%; based on previous 55* template but achieved by single point mutation of 5 gene ++hX→+++X (3=X000); 5512 is dominant switchboard with three active and four expressed genes; 60% of optimal solutions have a + in gene 1 early on – this drops to zero (reduces number of activated codons needed)

Description of the Dynamics

- Late dynamics (t>704) – new mutant core gene 5=+X+X found – requires less codons but final terminating 0 in other gene; initially associated with gene 1
- Later though 41% (random = 15%) of optimal individuals have 0s (Zero is not a four letter word!) as the first codon in at least one of genes 4, 6 and 7 (“NON-SELECTED” GENES) → enhanced robustness to mutation of third codon of switchboard, e.g. 551* → 554* - GENETIC RESERVE
- System “prepares” itself via self-organization of the Genotype-Phenotype map to be more robust

The Algorithmic Language

Table 1. Description of the algorithmic language that has emerged by generation 1000

Words of the Language		
Gene	Codon content	Logic
Gene 5	$+X+X$	Gives possibility of at least 2X when expressed once and 4x when expressed twice
Gene 1	$0***$	Terminates the twice expressed $00**$ 5 gene with 0 codons that do not affect the 4X expression
Genes 4,6,7	$0***$	Genetic Reserve - predominantly not expressed but backup in case of mutation in third switch-board codon
Genes 2,3	$****$	"Non-coding" regions - either do not form part of the language or have unknown functionality
Syntax of the Language		
Switchboard gene	$551*$	Puts producing gene 5, expressed twice, before consuming gene 1 to given optimal ordering

Preference for Repeated sequences/Building Blocks

- Evolutionary competition between different optimal solutions: +X+X+X+X0 - 9 expressed codons but only 5 activated codons used rather than +X+X+XX with only 7 expressed but 7 activated
- Repeated switchboard codons ubiquitous → repeated gene use – Building Blocks – more mutationally robust → one mutant neighbours should be fitter…

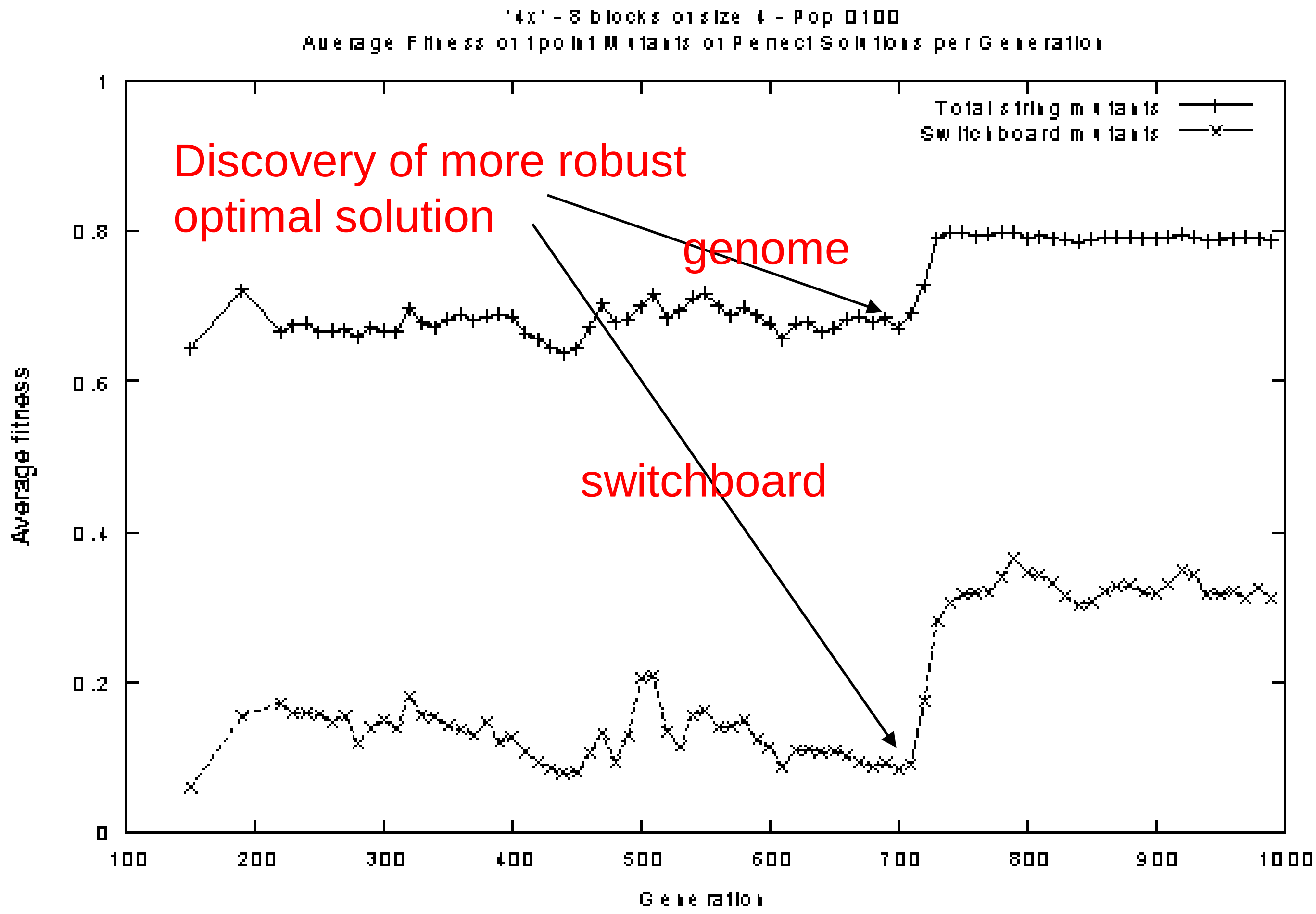


Fig. 5. Average fitness of 1-point mutants of perfect individuals, and of 1-point mutants of each perfect individual's switchboard.

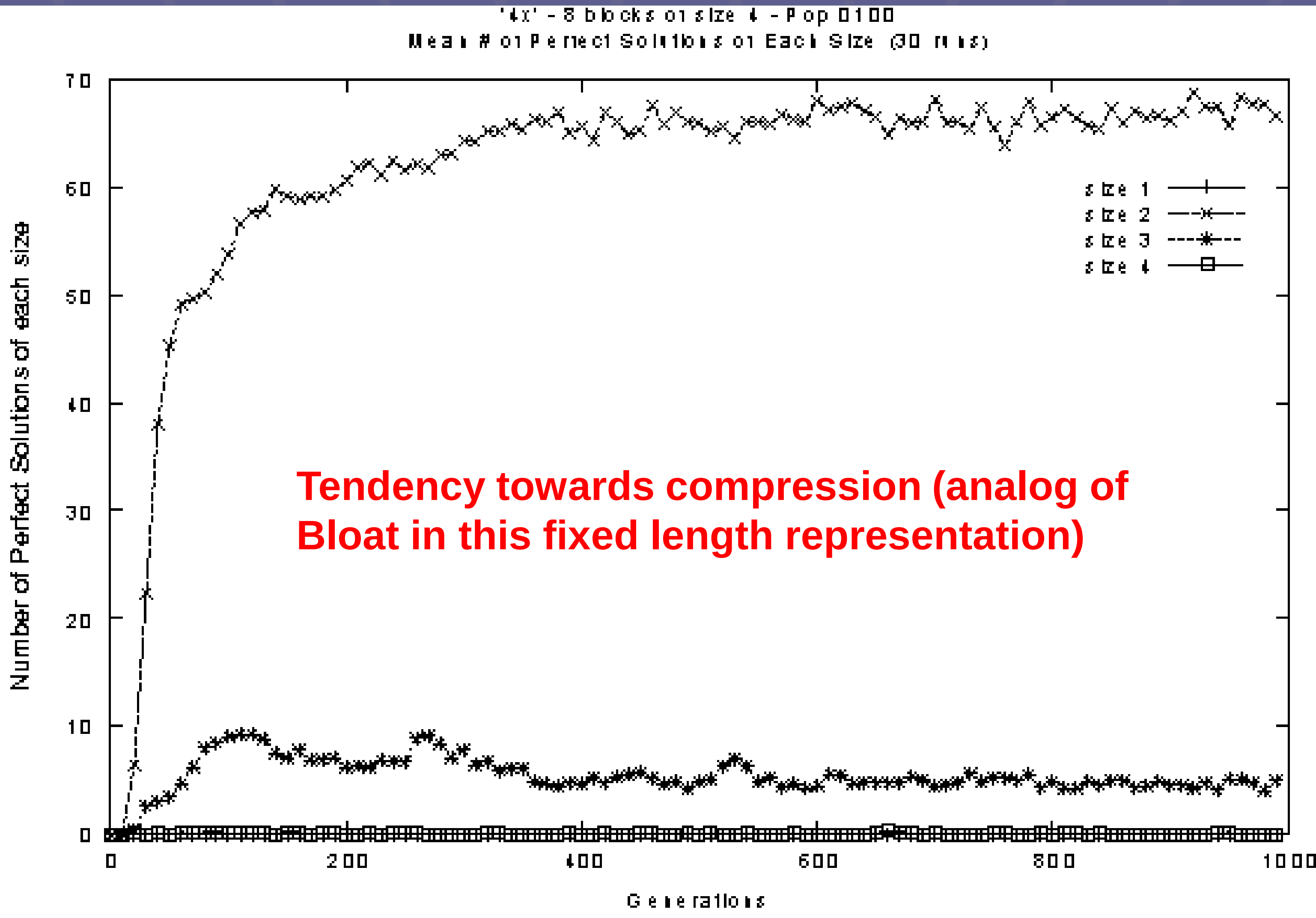


Fig. 6. Average size of perfect solutions per generation (30 runs).

So what does this example tell us?

- The system is trying to do <u>more</u> than just find an optimal solution.
- The distribution of optimal solutions is NOT random due to competition between them in the presence of mutation and recombination
- Collectivity/structure emerges on many “scales” – codons, genes and gene-products
- A simple “algorithmic language” emerges where the above structures can be interpreted as letters, words and sentences with an associated syntax determined by one master gene

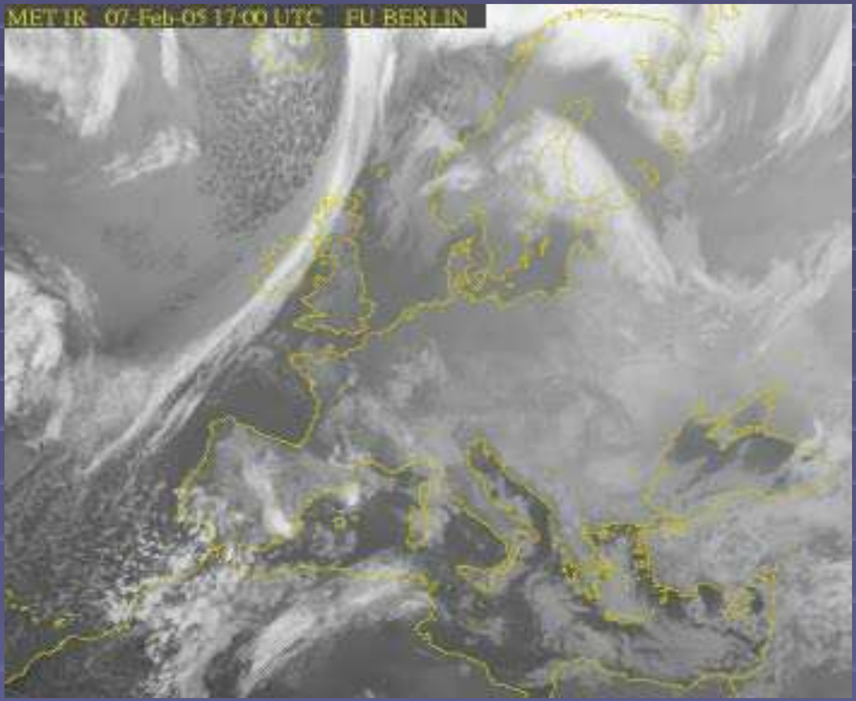
So what does this example tell us?

- All the interesting effects seem to have their origin in the multiple levels of degeneracy in the system
- If we had considered this as simply a search algorithm for doing symbolic regression we would have been happy with the output NOT how we got it. The interesting structure was found by examining in DETAIL individual runs.
- Complexity in other than the standard contexts of language and biology could be very difficult to detect! Here to understand the structure it’s also important to understand what the system is for.

So if symbolic complexity is a property of both a system and the measuring device, what about physical complexity in systems like…?

...the weather

47
49
50
50
51
...



Complex?

...or just complicated?

A **physical** phenomenon; underlying dynamics governed by Navier-Stokes equation (non-linear PDE)
Chaotic beyond 15 day horizon
"No" biological (human) component - Physics paradigms appropriate

Complexity – Subjective or Objective?

How Weather Affects Your Life	
Health	Health Forecast, Allergies, Skin Protection, Air Quality, Aches & Pains, Cold & Flu, Fitness
Travel	Travel Forecast, Business Traveler, Vacation Planner, Aviation
Driving	Driving Forecast, Interstate Forecast, Scenic Drives, Auto Advisor, Green Vehicles, Vehicle Safety
Events	Events Forecast, Sporting Events, Special Events
Recreation	Recreation Forecast, Golf, Boat & Beach, Outdoors, Ski
Home & Garden	Home & Garden Forecast, Home Planner, Lawn & Garden, Scotts Lawn & Garden Center, Schoolday
World	World Weather Forecasts & International Sites
News	News Center, Storm Watch, Tropical Update, Storm Stories, Road Crew
Weather Tools	My Page, Desktop, Email, Phone, PDA, Pager, My Site
Interact	Photo Gallery, Boards & Forums, Contact Us
Education	Weather Classroom, Dave's Dictionary, Weather Encyclopedia, Glossary, SafeSide, Rays Awareness
Multimedia	Video Forecasts
Shopping	The Weather Channel Store, Hot Offers and Cool Deals
TV - What's On	Storm Stories, Schedule, Road Crew, Personalities, Music, Forecast Earth
Mobile	Downloads, Messaging, PDAs

What's complex? The underlying phenomenon or our description of it?

The underlying phenomenon is not complex but its effects at the human level and our description of them are!

So what about something with a human component? Like…

…a stock market?

2086.66
2057.64
2075.06
2068.70
2062.41
2035.83
2047.15
…

Complex?

…or just complicated?

Geometric Brownian motion
Black-Scholes equation – Diffusion equation type PDE

Rational agents
Market efficiency
Equilibrium economics

Still very “physicsy” paradigms

Complexity – Subjective or Objective?

What's complex? The underlying phenomenon or just our description of it?

If markets are “efficient”, then they’re described by a “random” process and “predicting” the market seems to be then no different than counting elephants in the clouds or seeing people’s faces in a fire!

Financial markets: “Complex” Human Behavior → “Simple” Random Price Movements

Weather: “Simple” Navier-Stokes dynamics → “Complex” Human Behavior

So, what are we to make of all this?

- **Is it useful to distinguish between <u>intrinsic</u> (the system only) versus <u>extrinsic</u> (system and measuring apparatus)?**
- **I think so – can talk about correlations intrinsic to a system and correlations between a system and a measuring device**

Modeling complexity and complex systems

Consider the following "simple" dynamical model…

$$\mathbf{d}_i(t+\Delta t) = \sum_{j \neq i} \frac{\mathbf{c}_j(t) - \mathbf{c}_i(t)}{|(\mathbf{c}_j(t) - \mathbf{c}_i(t))|} + \sum_{j=1} \frac{\mathbf{v}_j(t)}{|\mathbf{v}_j(t)|}$$

Competition between effective repulsion and attraction between "particles"

$$\hat{\mathbf{d}}_i(t+\Delta t) = \mathbf{d}_i(t+\Delta t) / |\mathbf{d}_i(t+\Delta t)|$$

$c_i(t)$, $v_i(t)$ – position/direction vectors of a "particle"

$$\mathbf{d}_i{}'(t+\Delta t) = \frac{\hat{\mathbf{d}}_i(t+\Delta t) + \omega \mathbf{g}_i}{|\hat{\mathbf{d}}_i(t+\Delta t) + \omega \mathbf{g}_i|}$$

Equation for "charged" particles following an external force vector g_i

Couzin, I.D., Krause, J., Franks, N.R. & Levin, S.A. (2005) *Nature*, **433**, 513-516.

Does this represent a "complex" system?

- In this mathematical model there are only two scales:
  - The “micro-” associated with individual fish and their typical distances
  - The “macro-” associated with the school or shoal itself (remember the “sandpile” on the Edge of Chaos)
- Saying shoaling is an “emergent” phenomenon is like saying boiling is an emergent phenomena

- So, we are using a non-complex model to describe a complex system
- The complexity is associated with a range of behaviors and functions
- The model only describes statistically one restricted aspect of this rich complexity
- Need a more complex model to describe more complex behavior

Moral: It's important to distinguish between a description of complexity and a non-complex description of a phenomenon or behavior associated with a complex system.

So, what properties should mathematical models have if they are to model complexity?

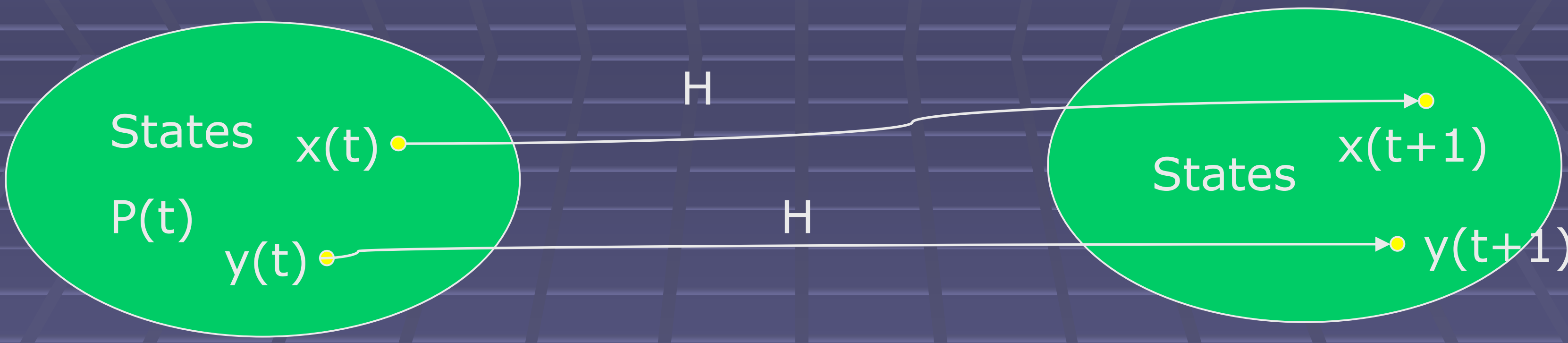
In biological, economic and social systems, organisms exhibit a rich array of (survival) **STRATEGIES (rules/models)**

The dynamical state of an individual at t+1 depends on not only on the state of the individual and others at t but also on which strategy (update rule) is chosen at t, which in turn depends on the update rules of others at t
→ need to work in the space of states and strategies/rules/models - sounds like game theory, but …

We don’t a priori know what that space is!

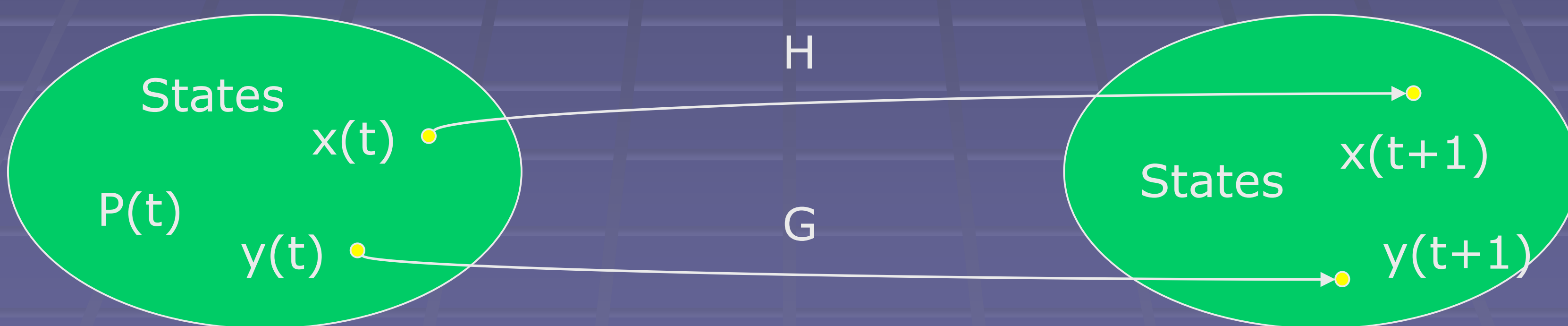
Also, the payoff/fitness for a strategy is RELATIVE not absolute – depends on the strategies used by others → a fixed fitness landscape is inappropriate; Fitness should be an emergent property. Imagine at the beginning of evolution specifying a priori the fitness of a lion!

Theoretical physics as it currently stands does not contain the mathematical and conceptual elements necessary to understand these issues…



Democratic evolution – one law for all (Physics)

$x(t+1)=H(x(t))$
$y(t+1)=H(y(t))$



Not all states are created equal (Not Physics)

$x(t+1)=H(x(t))$
$y(t+1)=G(y(t))$

G not equal to H ⇒ G(y(t)) not equal to H(y(t))

Theoretical Challenges for Modeling Complex Systems

- Develop frameworks within which one can work in the space of "<u>laws</u>" <u>and</u> states
- Understand what are "necessary" and "sufficient" conditions for complexity
- Statistical inference problems of observing complexity – can we speak the lingo?
- Work in a "game" where the rules change all the time and we don't know the payoffs
- Fitness as an emergent phenomenon
- Modularity – how to understand how different parts of a system can do different things then join together as "building blocks" to form more complex things
- Better understand the genotype-phenotype map
- Understand how to coarse grain (renormalization group) to see the emergence of effective degrees of freedom

Studying the following "experimentally" would help

- Develop systems that can do multi-tasking <u>adaptively</u>
- Develop systems where fitness is not specified
- Develop systems where modularity within a population emerges naturally – how do teams arise?
- Would these give us <u>open ended</u> evolution, i.e., continuous innovation?

Coarse Graining

Coarse Graining

Why?

What?

How?

Coarse Graining: Why?

1. **Emergence of "Effective Degrees of Freedom" (EDOF)/Collectivity/Universality**
2. **Curse of dimensionality/intractable dynamics**

Coarse-grained degrees of freedom are combinations of the underlying "microscopic" degrees of freedom. EDOF are those coarse-grained degrees of freedom that are important for the dynamics

If the world wasn't naturally "coarse grainable" there would have been no science!

Imagine having to do a quantum field theory calculation to describe a pendulum!

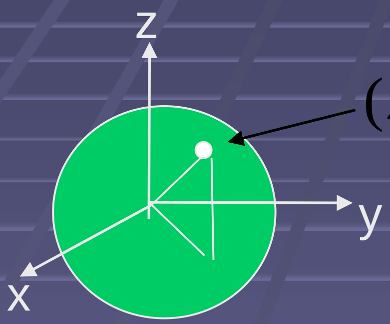
Coarse Graining: What?

Typical examples in the “physical” world:

- Rigid body motion
- Thermodynamics
- Waves

Rigid body motion

e.g. N point particles rigidly attached on a sphere of radius R



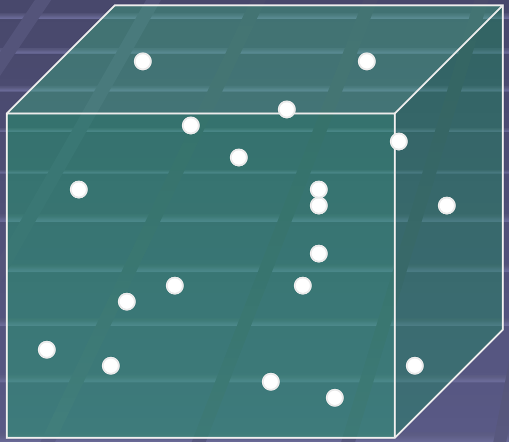
$$(x_i(t), y_i(t), z_i(t)) \longrightarrow (R, \theta_i(t), \phi_i(t))$$

Could exploit spherical symmetry to reduce # degrees of freedom, 3N → 2N but ….rigid body constraints imply only three effective degrees of freedom associated with the center of mass

$$(X(t), Y(t), Z(t))$$

So symmetries can affect what are the appropriate degrees of freedom and constraints, such as rigid body, even more so, reducing them to a much smaller number of effective degrees of freedom.

Thermodynamics



N non-interacting point particles in a container of volume V

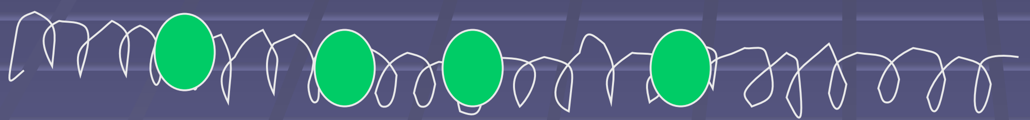
T - average kinetic energy of the particles
P - average momentum exchange at the container

Opposite situation from that of rigid body. No constraints, but similar in dimensional reduction, from 3N degrees of freedom to two effective degrees of freedom – P and T.

Thermodynamic description is "macroscopic". Statistical mechanics relates micro and macro

Waves

N point particles arranged on a 1-dimensional line and interacting with harmonic springs



$u(x_i) \longleftrightarrow u(k) = \frac{1}{\sqrt{2}} \sum_{x_i} \exp(ikx_i) u(x_i)$

In terms of interactions between the particles, intermediate between the rigid body and thermodynamic situations

Transformation to N new “coordinates” - Fourier amplitudes. Why? Harmonic motions are common – means that only a few Fourier modes may be excited, i.e. only a few of these “coordinates” are non-zero.

- Natural coarse grainings can emerge for many reasons – symmetries, constraints, the adequacy of a more “macroscopic” description, collectivity etc.
- Coarse grainings might not be understood as such without a knowledge of an underlying more microscopic description
- In essence, all these coarse grainings to a very reduced number of effective degrees of freedom are approximate
- Would like to relate (approximate) coarse grained description to underlying microscopic one
- One VERY powerful tool to do that is the RENORMALIZATION GROUP

Coarse graining – How? Using the Renormalization Group

Why call it the Renormalization Group?

Introduce a general coarse-graining operator $\mathcal{R}(\eta, \eta')$

Which coarse grains from the variables $\eta \in X_\eta$ to the variables $\eta' \in X_{\eta'} \subset X_\eta$

Given two such coarse grainings:

$$\mathcal{R}(\eta, \eta')P_\eta(t) = P_{\eta'}(t) \qquad \mathcal{R}(\eta, \eta'')P_\eta(t) = P_{\eta''}(t)$$

but $$\mathcal{R}(\eta', \eta'')P_{\eta'}(t) = P_{\eta''}(t)$$

hence $$\mathcal{R}(\eta, \eta'') = \mathcal{R}(\eta, \eta')\mathcal{R}(\eta', \eta'')$$

i.e. coarse grainings form a semi-group – **“Renormalization Group”**

Coarse Graining by Projection - “Divide and Conquer”

- Iterated map takes you to a problem with fewer degrees of freedom – NOT associated with “trivial” symmetries.
- Linearization around the fixed points of the equations give the asymptotic behaviour in space and/or time
- Can understand “universality” of behaviour
- Can coarse grain in both “space” and “time”
- Coarse graining can almost never be done exactly
- Have to decide what coarse graining is most appropriate for a given model

Coarse graining – RG for one dimensional Ising model

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	…

$$Z(T,H,N) = \sum_{\{s_i\}} \exp(-E/T)$$

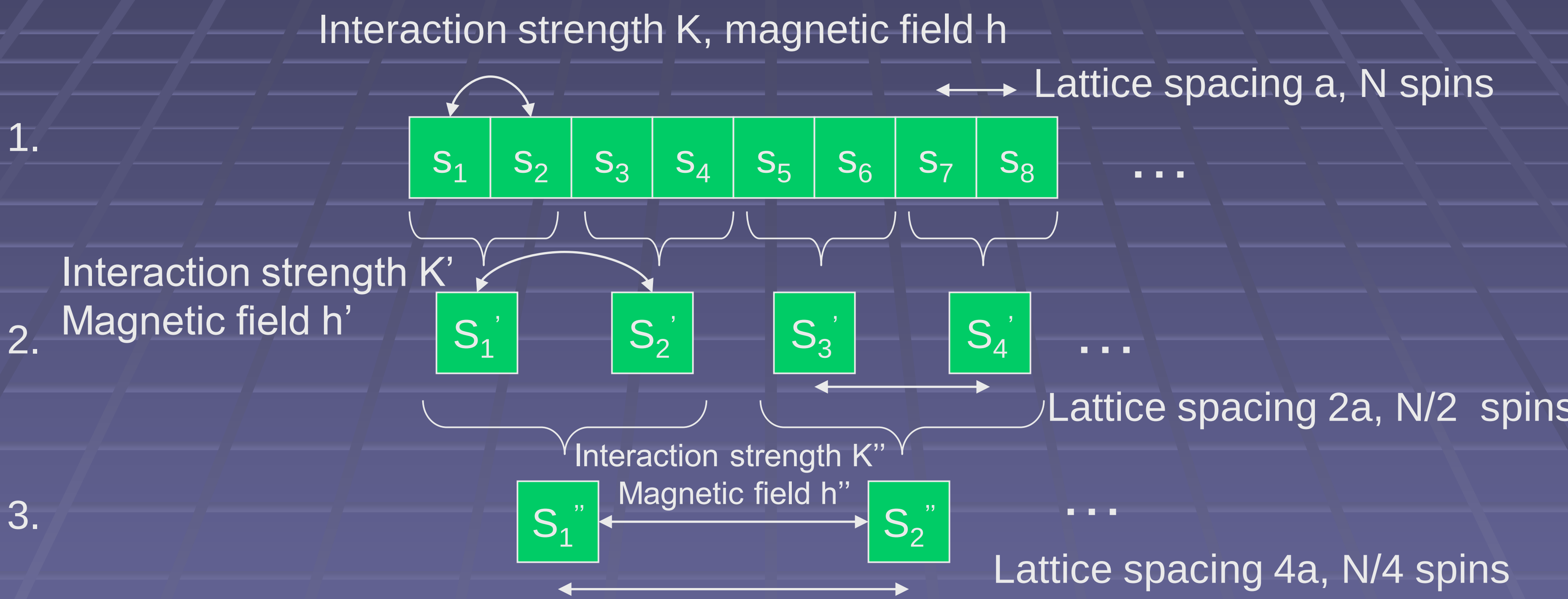
Partition function - Sum over over all possible microscopic configurations
$\{s_i\} = \{s_{\_1}=-1,1;\ldots s_N=-1,1\}$

$$E = \frac{1}{2}\left(\sum_i J s_i s_{i+1} + \sum_i H s_i\right)$$

Energy of the configuration

Governed by two paremeters K = J/T (temperature) and h = H/T (magnetic field)

Coarse graining – RG for one dimensional Ising model



Change from 1. to 2. by in Z doing the explicit sum over s_2, s_4, …
Change from 2. to 3. by in Z doing the explicit sums over s_2', s_4', …
Etc.

Coarse graining – RG for one dimensional Ising model

Z remains the same! But…**try** to write it after one iteration as…

$$Z(T,H,N) = Z(T',H',N/2) = \sum_{\{s_i\}} \exp(-E'/T')$$

Where

$$E' = \frac{1}{2}\left(\sum_{i=1}^{N/2} J' s_i' s_{i+1}' + \sum_{i=1}^{N/2} H' s_i'\right)$$

Can we do this? Take as example, H=0, then…

Coarse graining – RG for one dimensional Ising model

$$\sum_{s_i=-1,1} \exp(Ks_{i-1}s_i + Ks_i s_{i+1})$$

$$= \exp(Ks_{i-1} + Ks_{i+1}) + \exp(-Ks_{i-1} - Ks_{i+1})$$

$$= f(K)\exp(K' s_{i-1} s_{i+1})$$

Note a “new” effective interacion between s_{i-1} and s_{i+1}

Where considering the two cases: $s_{i-1}= s_{i+1}=1$; $s_{i-1}= -s_{i+1}=1$ gives

$$K' = \frac{1}{2}\ln(\cosh(2K)) \qquad f^2(K) = 4\cosh(2K)$$

Thus the partition function (up to a trivial multiplicative constant), and therefore all the corresponding thermodynamic properties, of a 1D Ising model with N spins, at temperature 1/K is equivalent to that of a “renormalized” system with N/2 spins at a temperature 1/K’=2/ln(cosh(2K))

Coarse graining – RG for one dimensional Ising model

- There are many ways to coarse grain. E.g. Instead of combining adjacent pairs of spins we could have chosen triplets (OK) or chosen random pairs (Why? Doesn’t respect the natural structure of the model).
- Each coarse graining leads to different RG equations. All lead to the same physical results though! We just have an equivalence relation on a space of models.
- Is one of these equivalent models easier to solve?

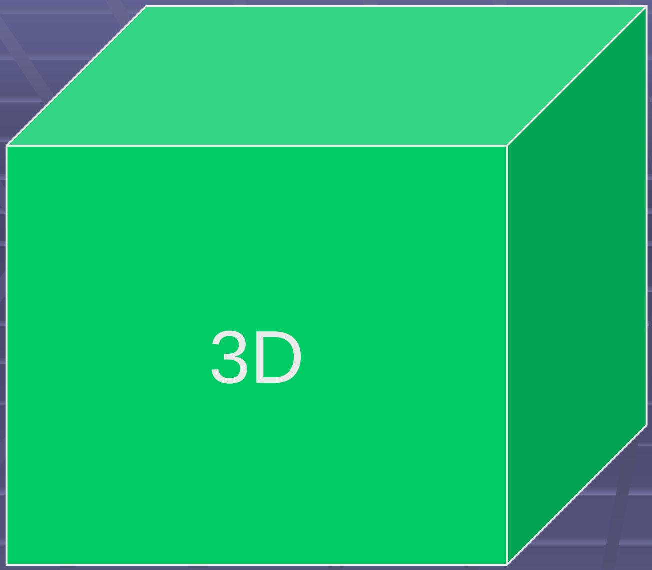
Coarse graining – RG for one dimensional Ising model

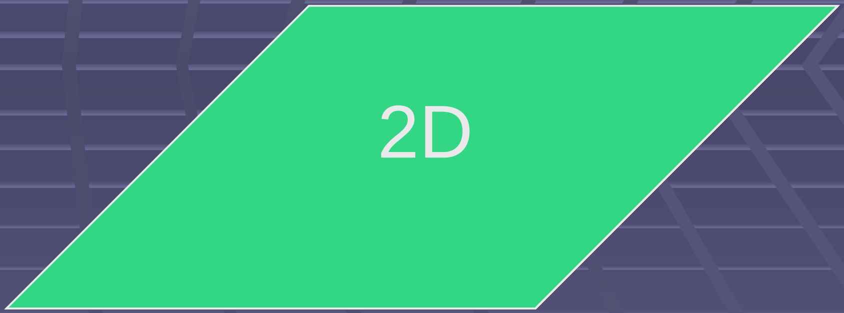
- RG puts the emphasis on the **relations** among a family (different iterations of the coarse graining) of equivalent systems
- The family is a RG flow trajectory in the space of parameters (e.g. in the 1D Ising model temperature and magnetic field)
- For this model there are two fixed points where K'=K: K=0 (stable, infinite temperature, zero correlation length) and K=infinity (unstable, zero temperature, infinite correlation length). At both these points the system is scale invariant.
- Different methods, e.g. linear perturbations around the fixed point, can be used once the structure of the space of RG flows is better understood.

Coarse graining – RG for one dimensional Ising model

- Associated with each point on the trajectory is an “effective degree of freedom”.
- These can change qualitatively, and should if that’s what is happening in the physical system we’re trying to model
- The 1D Ising model is simple in that the spins don’t change just the interactions between them. In other words the renormalized model is a model of interactions between elementary spin (+-) variables just with different interaction strength.
- What happens when that isn’t the case?

Environmentally Friendly Renormalization



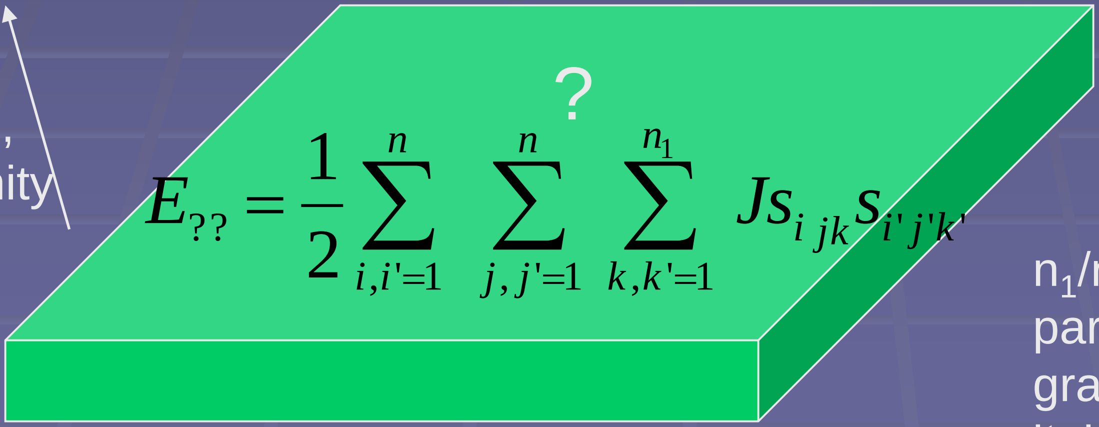


$$E_{3D} = \frac{1}{2}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{n}\sum_{k,k'=1}^{n} J s_{ijk} s_{i'j'k'}$$

$$E_{2D} = \frac{1}{2}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{n} J s_{ij} s_{i'j'}$$

$n_1/n \rightarrow 1$,
$n \rightarrow$ infinity

$n_1/n \rightarrow 0$,
$n \rightarrow$ infinity



$$E_{??} = \frac{1}{2}\sum_{i,i'=1}^{n}\sum_{j,j'=1}^{n}\sum_{k,k'=1}^{n_1} J s_{ijk} s_{i'j'k'}$$

n_1/n is an "environmental" parameter. A good coarse graining should depend on it, i.e., be "environmentally friendly"!

Environmentally Friendly Renormalization

$$m = A(D)(T - T_c(D))^{\nu(D)}$$

$$\Rightarrow \quad \nu(D) = \frac{d \ln m(D)}{d \ln t(D)}$$

Relationship between inverse Correlation length m and temperature, nu is the correlation length exponent and is “universal”

For the thin film of thickness L…using an environmentally friendly renormalization

$$\nu_{\text{eff}} = \frac{\left(\tanh\left(\frac{mL}{2}\right)\right)^{\frac{1}{3}} \int_0^{mL} \frac{x^{\frac{2}{3}}}{\left(\tanh\left(\frac{x}{2}\right)\right)^{\frac{1}{3}}} dx}{(mL)^{\frac{5}{3}}}$$

$$\nu_{eff} \xrightarrow{m \longrightarrow 0,\, Lm \longrightarrow \infty} 0.6$$

Characteristic of 3D

$$\nu_{eff} \xrightarrow{m \longrightarrow 0,\, Lm \longrightarrow 0} 0.75$$

Characteristic of 2D

So what about coarse graining and the RG in other than physics?

Typical coarse graining examples in genetic dynamics:

- **“Direct” dimensional reduction**
- **Phenotypes**
- **Schemata**
- **Hyperschemata**
- **Building Blocks**
- **Lowest cumulants of fitness distribution**
- **“Normal (e.g. Walsh) modes”**
- **Others**

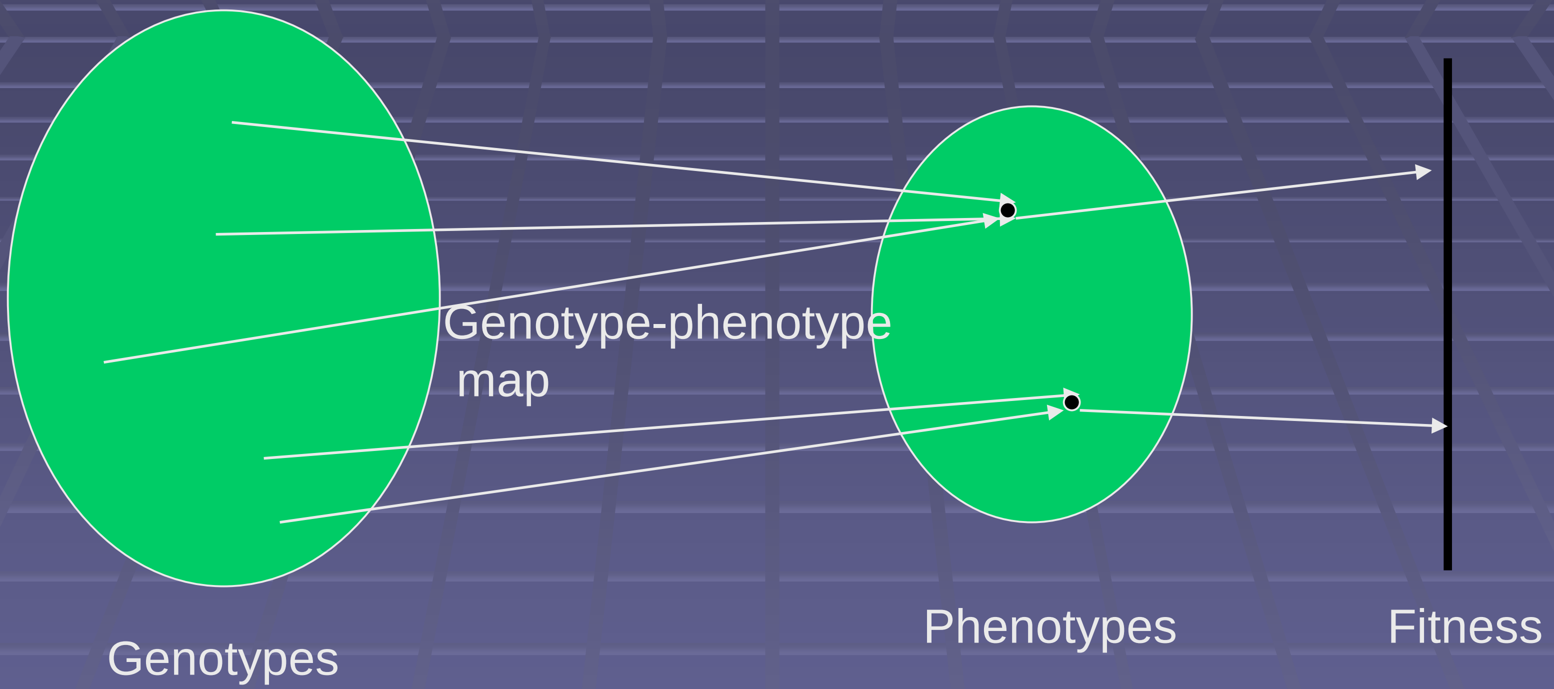
What is the most natural coarse graining depends on the operators and their corresponding parameters, the fitness landscape and the population.

Direct Dimensional Reduction

Common in population genetics

Go from N degrees of freedom to 2 and postulate a (static) fitness landscape for the dimensionally reduced problem

Phenotypes



All genotypes that correspond to the same phenotype have the same fitness
Coarse grained in that once we pass to phenotypes we lose information about the individual genotypes that correspond to a phenotype

Schemata

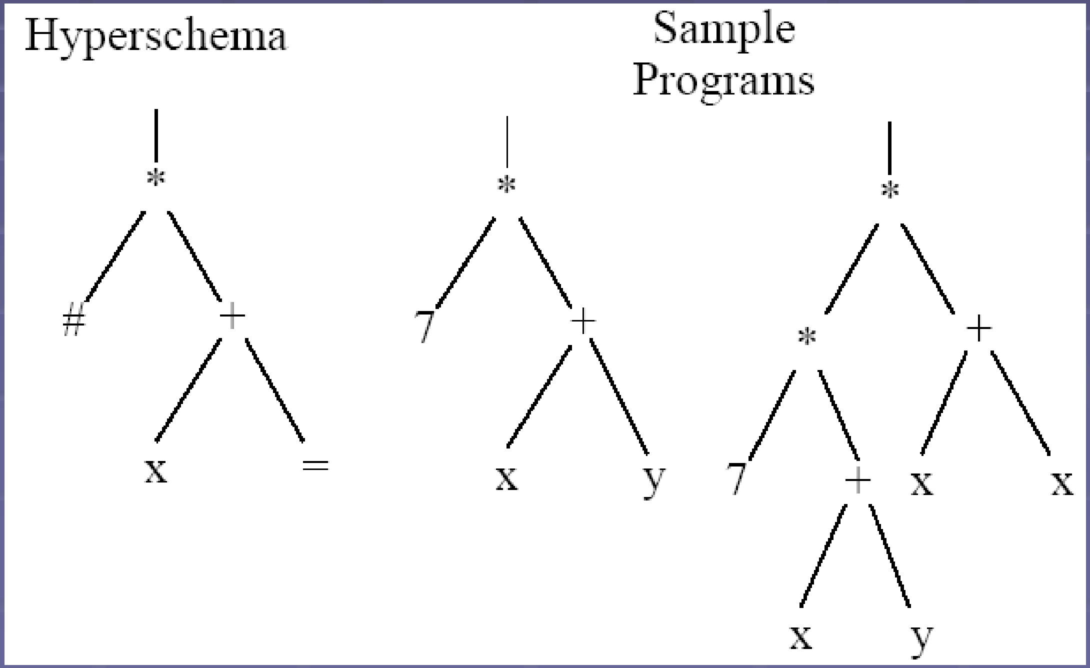
10110 + 11111 + 10111 + 11110 → 1*11*

Projection: $$P(1{*}11{*}) = \sum_{I_2=0,1} \sum_{I_5=0,1} P(1I_2 11I_5)$$

Hyperschemata

A *GP* *hyperschema* is a tree with internal nodes from $F \cup \{ = \}$ and leaves from $T \cup \{ =, \# \}$, where F is the function set and T the terminal set

= is a “don't care” symbol which stands for exactly one node, while # stands for any valid subtree.



Building Blocks

Particular schemata that contribute to the creation of another schema or string via recombination

1*1** + ***1* → 1*11*

Via one-point recombination between third and fourth loci

Coarse graining via Projections

Dynamics coarse grains via

$$\mathcal{R}(\eta,\eta')\mathcal{H}(\mathbf{p},\mathbf{f},\mathbf{P}_{\eta}(t))$$

If this can be written in the form

$$\mathcal{H}(\mathbf{p}',\mathbf{f}',\mathbf{P}_{\eta'}(t))$$

with suitable "renormalizations"

$\mathsf{f} \longrightarrow \mathsf{f}'$ and $\mathsf{p} \longrightarrow \mathsf{p}'$

then the dynamics is form covariant or invariant under the coarse graining. If $\mathsf{f} = \mathsf{f}'$ and $\mathsf{p} = \mathsf{p}'$ dynamics is "compatible"

Coarse graining via Projections

Examples: Compatible Coarse grainings

1. Selection and Phenotypes

- Unitation, e.g. 2^N genotypes → (N+1) phenotypes
- Eigen model (NIAH), e.g. 2^N genotypes → 2 phenotypes

2. Mutation and Crossover and Schemata

- 2^N genotypes → 2^{N_2} coarse-grained genotypes

Incompatible Coarse grainings

1. Selection, Mutation and Crossover and Schemata

- 2^N genotypes → 2^{N_2} coarse grained genotypes
- $f_\alpha = \mathcal{R}(x,\alpha) f_x = \sum_{x\in\alpha} f_x P_x(t) / \sum_{x\in\alpha} P_x(t).$ time-dependent

Coarse graining via Projections

In Building Block Basis for 1-point crossover…

$$P_{111}(t+1) = (1-p_c)P_{111}(t) + \frac{p_c}{2}(P_{1**}(t)P_{*11}(t) + P_{11*}(t)P_{**1}(t))$$

"Zap" (projection) $111 \rightarrow 11*$

$$P_{11*}(t+1) = (1-p_c)P_{11*}(t) + \frac{p_c}{2}(P_{1**}(t)P_{*1*}(t) + P_{11*}(t)P_{***}(t))$$

$$\Rightarrow P_{11*}(t+1) = (1-\frac{p_c}{2})P_{11*}(t) + \frac{p_c}{2}P_{1**}(t)P_{*1*}(t)$$

Note – coarse grained (projected) 3-bit equation same as "microscopic" 2-bit equation with "renormalization" $p_c \rightarrow \frac{p_c}{2}$

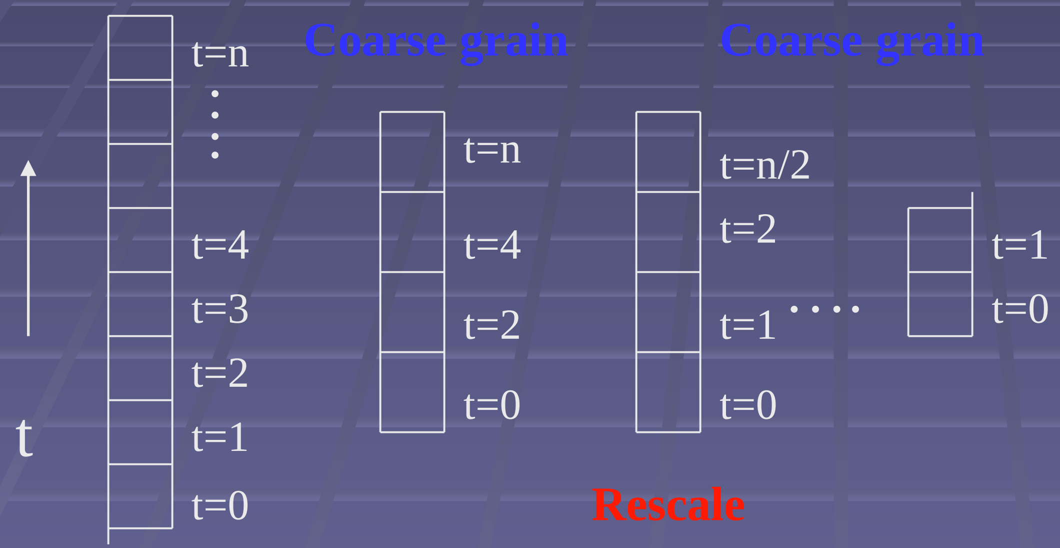
$$P_{11}(t+1) = (1-p_c)P_{11}(t) + p_c P_{1*}(t)P_{*1}(t)$$

FORM INVARIANCE

Coarse graining via Projections

- Generalizes to the case of variable-length GAs and GP; Building Block Schemata →Building Block Hyperschemata; “form invariance” of equations over different types of EA and form invariant upon coarse graining to schemata/hyperschemata;
- Gives exact form of the Schema Theorem and generalizes it to EAs other than GAs
- Neglecting the “construction” terms leads to standard Holland Schema Theorem as an approximation

Coarse Graining RG for one locus selection-mutation system



Example: 1-bit

Can we coarse grain an n generation problem to a one generation problem? Much easier to solve the dynamics over only one generation!

$X_1(t)$ – unnormalized incidence vector

p – mutation rate

$$\begin{pmatrix} X_1(t+2) \\ X_0(t+2) \end{pmatrix} = \underbrace{\begin{pmatrix} (1-p)f_1 & pf_0 \\ pf_1 & (1-p)f_0 \end{pmatrix}^2}_{} \begin{pmatrix} X_1(t) \\ X_0(t) \end{pmatrix}$$

Evolves bit two time steps in landscape f(1), f(0) with mutation p

Coarse Graining by Projection - "Divide and Conquer"

$$\begin{pmatrix} X_1(t'+1) \\ X_0(t'+1) \end{pmatrix} = \underbrace{\begin{pmatrix} (1-p'_1)f'_1 & p'_0 f'_0 \\ p'_1 f'_1 & (1-p'_0)f'_0 \end{pmatrix}} \begin{pmatrix} X_1(t') \\ X_0(t') \end{pmatrix}$$

Evolves bit one time step in "renormalized" landscape f'(1), f'(0) with asymmetric mutation rates p'(1) and p'(0)

Equivalent dynamics (all we did was "change names"!, i.e. "renormalize")

$$f'_1 = (1-p_1)f_1^2 + p_1 f_0 f_1$$

$$f'_0 = (1-p_0)f_0^2 + p_0 f_0 f_1$$

$$p'_1 = p_1 \left(\frac{(1-p_1)f_1 + (1-p_0)f_0}{(1-p_1)f_1 + p_1 f_0} \right)$$

$$p'_0 = p_0 \left(\frac{(1-p_0)f_0 + (1-p_1)f_1}{(1-p_0)f_0 + p_0 f_1} \right)$$

Coarse Graining by Projection - "Divide and Conquer"

Evolution of mutation/selection GA over n time steps with fitness landscape f(1), f(0) and mutation rates p(2) and p(1) is same as that of a GA with "renormalized" landscape and mutation rates, f'(1), f'(0), p'(2), p'(1) over n/2 time steps!

↓

UNIVERSALITY

Fixed points of Renormalization Group transformation:

$|\ln(f(1)/f(0))| = 0$, $p(1) = p(0) = 0$; no selection/mutation – "FERROMAGNETIC"

$|\ln(f(1)/f(0))| = \text{infinity}$, $p(1) = p(0) = 0$; strong selection – "FROZEN"

$|\ln(f(1)/f(0))| = \text{constant}$, $p(1) + p(0) = 1$; neutral evolution – "PARAMAGNETIC"

Coarse Graining by Projection - “Divide and Conquer”

- Iterated map takes you to a problem with fewer degrees of freedom – NOT associated with “trivial” symmetries.
- Linearization around the fixed points of the equations give the asymptotic behaviour in space and/or time
- Can understand “universality” of behaviour
- Can coarse grain in both “space” and “time”
- Coarse graining can almost never be done exactly
- Have to decide what coarse graining is most appropriate for a given model

- As a model that contains few degrees of freedom consider the lynx etc model

Fourier modes