Study Notes for Introduction to Complexity

Lihao Yan

Winter 2022

This is my study notes for the free online course *Introduction to Complexity* from Santa Fe Institute.

1 Unit 1: What is Complexity?

There are many complex systems in our world, here are some examples of them:

- Ants
- Termite mound
- Neurons of the brain
- Immune system organs
- Human genome
- Food web
- A social network
- Cities

Property 1.1. Basic properties for the complex system:

i) Simple components or agents

ii) Nonlinear Interactions among components

iii) No central control

iv) Emergent behaviors

The last property is very important, and interesting, in condensed matter systems.

The field of complex system can be broken down into several core disciplines.

1. Dynamics
2. Information
3. Computation
4. Evolution

The Evolution is different from Dynamics in that the evolution is about how the system change according to the external environment. Two main goals: 1) Cross-disciplinary insights into complex systems 2) General theory Many people think the idea of an unified general theory is out of reach, but everyone would agree that this is the ultimate dream of all the scientists. Complex scientists use experimental, theoretical, and computational techniques to investigate different complex systems.

1.1 Defining Complexity

We can divide most of the problems in the world into three different categories 1) Problems of simplicity 2) Problems of organized complexity 3) Problems of disorganized complexity. Problems of simplicity are those concerning a few interacting objects. The problems of organized complexity are those concerning many objects but can be tackled with averaging tools. The problems of disorganized complexity cannot be tackled with averaging tools like statistical mechanics and therefore requires new sciences. There are many different definitions in characterizing complexity. We will only use two of them in the course.

Definition 1.1 (Complexity)**.**

1. Shannon's: TODO: see Definition 4.2
2. Fractal:

One reason may because those systems encode long histories. Some define not the systems are complex, but the questions are complex. May use the amount of information from Shannon to define complexity of a system. The system having unpredictable behaviors, emergent phenomenon, adaptations ... Many interacting parts and usually in a non-linear fashion. Cannot be characterized by a few simple equations.

1.2 NetLogo

This is a computational tool useful for modeling and studying complex systems. Skipped for now.

2 Dynamics and Chaos

Dynamics is the study of how systems change over time. Some examples are fluid, electricity, financial market .... One aspect of the dynamic system is its exponential growth. Consider the following example.

Example 1 (Population of rabbit in a forest)**.** The new baby rabbits in the population of rabbit is proportional to the total population of the total rabbit. $n_{t+1} = 1$. Therefore, there is an exponential growth in the population w.r.t. 2^t.

2.1 Logistic Map and Chaos

Definition 2.1 (Logistic Map)**.** A simple, completely deterministic equation that, when iterated can display chaos (depending on the value of R) The logistic map is defined as $x_{t+1} = R(x_t - x_t^2)$

Definition 2.2 (Chaos)**.** Seemingly random behavior with *sensitive dependence on initial conditions.*

Example 2. Let $R = 2, x_0 = 0.2$, so we have

$$\begin{aligned} x_1 &= 0.32 \\ x_2 &= 0.4352 \\ &\vdots \\ x_5 &= 0.499999961 \\ x_6 &= 0.5. \end{aligned}$$

We see that the values are "attracted" to the value 0.5. The points x will be drawn to the fixed point as $t \mapsto \infty$ and will stay there forever. The point 0.5 is called a *a fixed point attractor*, which will become very useful in the future. Check out the online lecture for animations.

There are several types of attractors:

1. Fixed point attractor: attracted to a fixed value
2. Periodic attractor: the system changes periodically around the attractor

Remark 1. Periodic attractors can be very sensitive to the initial conditions of the system. NetLogo provides a package called "Sensitive Dependence on Initial Conditions"

The Bifurcation Diagram shows how chaos is developed out of the logistic map:

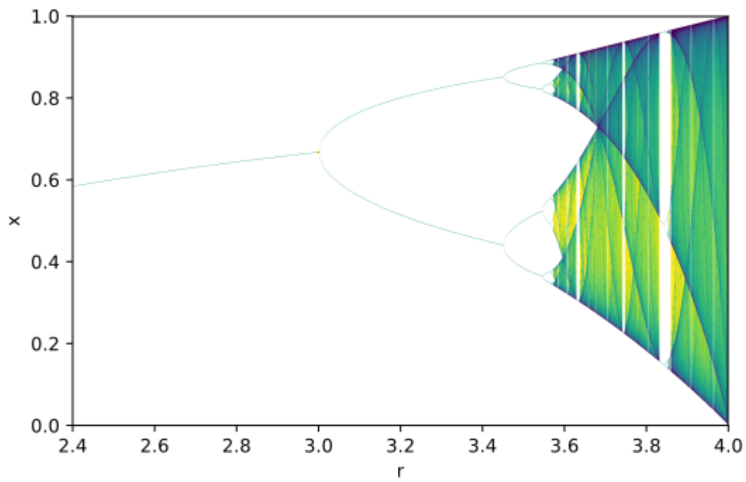

Figure 1: Bifurcation Diagram derived from the logistic map

Remark 2 (Determinitic chaos)*.* It is impossible in principle to determine all the future behaviors of a chaotic system because we do not know its initial conditions.

However, while we cannot solve for the entire system, many systems have classifiable and universal properties. For example, the Feigenbaum's constant in unimodal systems.

3 Introduction of Fratals

Definition 3.1 (Fractals)**.** Objects with "self-similarity" at different scales.

Some examples are trees, snowflakes ... This phrase was coined by Benoit Mandelbrot. He was trying to give a precise mathematical interpretation of roughness.

Example 3 (Measuring the length of the coastline of Great Britain)**.** What size ruler should you use for the task? The finer the ruler the more details of the coastline are involved. We get a different length! The question is then how do we properly define "distance" in the real world for fractal like objects like coastlines.

3.1 The Koch Curve

We consider a simple example of fractal: Koch Curve.

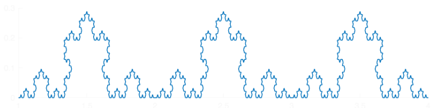

Figure 2: An graph for a Koch curve. The graph is very illustrative, each additional folding indicates a new iteration.

The curve length grows exponentially with respect to the iterations. The length "fills" the entire space, this is why Koch curve is called a *space filling curve*.

3.2 Fractal Dimension

We can divide the length of its sides by a number M, so number of copies $N = M^D$ is related to the dimension D. We can define dimension in this way.

Definition 3.2 (Fractal (Hausdorff) Dimension)**.** The dimension is given by,

$$D = \frac{\log N}{\log M},$$

where N is the number of copies of figure from previous level, and M is the size reduction factor of a side of the previous level.

Example 4 (Koch Curve)**.** Dimension of Koch curve is calculated as follows:

$$\begin{aligned} M &= 3 \\ N &= 4 \\ D &= \frac{\log 4}{\log 3} \approx 1.26 \end{aligned}$$

Fractal dimensions can be interpreted as "Quantifies the cascade of detail of an object.–Santiago Guisasola"

Example 5 (Calculated Fractal dimensions)**.**

- Bifurcation diagram Figure 1 has $D \approx 0.538$
- Cauliflower: $D \approx 2.8$
- Stock Market: a research group compared the fraction dimension of stock market to the fraction dimension of random walk. They found that they are not similar, which indicates that stock market is not completely random.

3.3 Box counting

Box counting is a method for calculating the approximated fractal dimension. It is most useful when calculating the real world objects because we do not have an analytical expression for the objects' fractal relations. Consider the coastline example, we grid the map of England's using boxes of size a. We count the number of boxes containing the coastline, then we scale the size a and the number of boxes is related to the dimension of the system following equation

$$\log N = D \log \frac{1}{a},$$

where N is the number of boxes, a is the size of the box and D is the dimension in the box counting method.

4 Information theory

4.1 Introduction

We have discussed self-organization of complex systems, and now we move onto another common property for them: information.

> Although [complex systems] differ widely in their physical attributes, they resemble one another in the way they handle information. That common feature is perhaps the best starting point for exploring how they operate. — Murray Gell-Mann, *The Quark and the Jaguar*, 1995

The idea of information is related to entropy in thermodynamics. Entropy can be interpreted as the orderedness of the system.

Axiom 4.1 (The second law of thermodynamics)**.** In an isolated system the entropy always increases until it reaches its maximum value.

A famous paradox regarding the second law of thermodynamics is the *Maxwell's Demon*, which is eventually resolved by connecting information with entropy by Leo Szilard. This intuition eventually leads to the field of physics of information.

4.2 Entropy in Statistical mechanics

Definition 4.1 (Entropy)**.** Measures the number of possible micro states that lead to a macro state. It is given by

$$S(\text{macro state}) = k \log W,$$

where W is the number of micro states corresponding to the macro state.

This is a more general formalism comparing to the thermodynamics theory. So we can rewrite the second law of thermodynamics as:

Axiom 4.2 (The second law of thermodynamics)**. Statistical Mechanics Version** In an isolated system the entropy always progress to a macro state that corresponds to the maximum number of micro states.

4.3 Shannon Information Content

Suppose there is a message source consisting of all kinds of possible messages, informally the level of surprise of the receiver is related to the entropy of the system.

Definition 4.2 (Shannon Information Content)**.** Let M be the number of possible messages, and p_i be the probability of message i. Then

$$H(\text{message source}) = -\sum_{i=1}^{M} p_i \log_2 p_i$$

We choose $\log_2$ because we want the final answer to be in units of computer bits.

Example 6 (Fair coin).
"Heads": probability 0.5
"Tails": probability 0.5

$$
\begin{aligned}
H(\text{fair coin}) &= -\sum_{i=1}^{N} p_i \log_2 p_i) \\
&= -\left[(0.5\log_2 0.5) + (0.5\log_2 0.5)\right] \\
&= -\left[0.5(-1) + 0.5(-1)\right] \\
&= 1\,\text{bit},
\end{aligned}
$$

which is measured on average, per message.

Check out the course for more interesting examples regarding information content of texts, etc.

5 Genetic Algorithms

5.1 Introduction

Adaptation and Evolution is universal, even in computation environments.

John H. Holland adopted Darwin's natural selection. Genetic algorithm evolves a fixed program to its desired state.

Example 7 (Genetic algorithms).

- Hyperparameter optimization
- Optimizing factory assembly line

1) Generate random initial strategies
2) For each strategy, calculate fitness (average reward minus penalties earned on random environments)
3) The strategies pair up and create offspring via "sexual recombination" with random mutations — the fitter the parents, the more offspring they create
4) keep going back to step until a good-enough strategy is found

5.2 Genetic Programming

John Koza proposed genetic programming in 1990s, which uses the genetic algorithm to develop programs.

Example 8 (Genetic Programming)**.**

- Karl Sims applied genetic programming to computer graphics and draw beautiful images. One extraordinary aspect of his project is that he used humans to determine the fitness of the computer graphics. Therefore, the graphics he produced are children of both humans and AIs. This maybe the future of 21st century: a collaboration between humans and computers.
- Another project by Karl Sims is his Evolving Virtual Creatures

6 Cellular Automata

6.1 Introduction: Game of Life

Cellular Automata (CA) will bring together many concepts we have encountered so far in the course. **The Game of Life** is the world's most famous cellular automation, but it is not really a game.

6.2 Elementary Cellular Automata

Developed by John von Neumann with his colleague, Stanislaw Ulam. This is a simple system, but it can lead to lot's of complex behaviors. For example, it shows strip like behavior or even oscillation behavior or fixed point behavior.

Stephan Wolfram then tries to use the elementary cellular automata to explain complicated things in the real world. Wolfram would argue for explaining all the physics as emergent behavior from the Celluar Automata rules.

Definition 6.1 (Wolfram number)**.** Let black blocks represent one and white blocks represent zero. Then the mapping rule represent a binary number.

Example 9.

- Rule 232 is a majority voting rule that has a fixed point
- Rule 30 is a complex rule that gives complicated patterns, which is studied extensively by Stephan Wolfram

Wolfram classifies CA into four categories according to their behaviors.

1. Almost all initial configurations relax after a transient period to the same fixed configuration, e.g. rule 128

2. Almost all initial configurations relax after a transient period to some fixed point or some period cycle of configurations, but which one depends on the initial configuration, e.g. rule 10
3. Almost all initial configurations relax after a transient period to chaotic behavior, e.g. rule 22
4. some initial configurations result in complex localized structures, sometimes long-lived, note whether some patterns belong to this case can be subjective

The Game of Life is a Class 4 CA.

6.3 Cellular Automata as dynamic systems

CA is a type of dynamic systems. Chris Langton proposed *Lambda* as a control parameter for CAs (like R in logistic maps).

Definition 6.2 (Langton's Lambda)**.** Fraction of black output states in CA rule table

Check out this demonstration website called the *edge of chaos* to see how *Langton's Lambda* predicts the behavior of CA.

6.4 Cellular Automata as computers

Wolfram hypothesized that all Class 4 CAs are capable of "universal computation"/

Definition 6.3 (Universal computation)**.** Computer that can run any program on any input.

Remark 3. Only a small set of logical operations is needed to support universal computation!

Neumann develops the first CA that is capable of universal computation. In 2002, Matthew Cook showed that rule 110 CA is a universal computer. However, they are too difficult to program and too slow to run them.

6.5 Evolving Cellular Automata with Genetic Algorithms

The content of this section is derived from the paper of the same number by Melanie Mitchell. In this section, we seek to design a CA to decide whether or not the initial pattern has a majority of black cells. This is very interesting because this is the case for a decentralized system to determine a global property of the system with only local interactions. We use CA of an updating rule with 6 neighbors, so there are 128 different configurations and total $2^1$28 different rules.

We use genetic algorithm to find the rules that performs our task.

1) Create a random population of candidate CA rules
2) The "fitness" of each cellular automaton is how well it performs the task
3) The fittest CA get to reproduce themselves, with mutations and crossovers
4) This process continues for many generations

The details for calculating fitness need more care, you can find it in the course lecture notes.

7 Self-Organization Systems

Self-Organization is commonly seen in nature and it is known as the emergent phenomenon in the world. Here we give an informal definition.

Definition 7.1 (Self-Organization)**.** Production of organized patterns, resulting from localized interactions within the components of the system, **without any central control**.

Craig Reynold describes flock or school of birds from very simple rules. There is a NetLogo example on the course website. This is an example of synchronization in nature, here are more examples.

Example 10 (Synchronization in nature)**.**

- Fireflies flashing (NetLogo model available)
- Crickets chirping
- Cicadas development and emergence

Example 11 (Task allocation in ants collony)**. Question:** How does an individual ant decide which task to adopt in response to nest-wide environmental conditions, even though no ant directs the decision of any other ant, and each ant interacts only with a small number of other ants? For example, if food supply is large and high-quality, number of foragers will increase.
Answer: An ant can tell what job another ant has been doing by sensing chemical residues on the other ant. The more frequent a specific task, the more likely that tasks become high priority. Therefore, a larger ant colony will have better statistics and better "decision making."

7.1 Information Processing

We can compare information processing in computer science and in biology.

\	Computer Science	Biology
Role of information	Digital, static, "passive"	Statistical, patterns, "active" (in space and time)
How is information processed?	Via deterministic, error-free, centralized rules	Via decentralized, local, fine-grained stochastic actions, and with **self-feedback, randomness**...
Meaning or purpose	Human interpretation	Natural selection for adaptive function

Table 1: Information Processing

Computer scientists are trying to develop biologically inspired "Self-Organized" computing.

8 Models of Cooperation in Social Systems

Models of cooperation are the models that make assumption of biology, social science, and economics that individuals act in order to maximize their own **utility**. The two examples we are going to cover in this section is *The Prisoner's Dilemma* and *The El Farol Problem*. One central question in this field is to understand how does cooperation come about in societies of selfish individuals.

8.1 The Prisoner's Dilemma

This idea was invented by mathematical game theorists Flood and Dresher in 1950 during the cold war, it has become one of the most famous and influential idea in the social science.

The Prisoner's Dilemma in the game formalism by Robert Axelrod can be described by the following table:

Alice, Bob	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

His main question is **Under what conditions will cooperation emerge in a world of egoists without central authority?**. The winning strategy of this game is the simplest TIT FOR TAT strategy defined below.

Definition 8.1 (TIT FOR TAT)**.** it has four guiding strategies

- **Be Nice**: never be first to defect
- **Be Forgiving**: be willing to cooperate if cooperation is offered
- **Be Retaliatory**: be willing to defect if others defect against you
- **Be Clear**: be transparent about what your strategy is — make it easy to infer

NetLogo has a model simulating multiple agents with various strategies one can choose from.

8.2 The El Farol Problem

Earliest idea was the rational people idea. *The Invisible Hand* by Adam Smith is one of the early examples of an emergent phenomenon in economics. However, the idea is too simple so people introduce the idea of complex economics that does not assume the equilibrium dynamics of the system under consideration. It treats the system as a chaos system that does not have enough time to reach the equilibrium. Consider the El Farol Model proposed by Brian Arthur.

Example 12 (El Farol Model)**.** El Farol is the name of a local bar. They have Irish music on Thursday nights. A hundred people want to go but the bar can only fit 60 people comfortably. There is no prior communication among people, but they know the number of people attended the Irish nights of the last M Thursdays. Assume everyone uses the same M, how does an agent decide whether he or she will go in or not.

They have different strategies in predicting how many people will show up this Thursday and their strategy will determine whether they will go in or not. Everyone is doing the same thing.

The strategies can be formed as below

1) Let $A(t)$ be the number of people attended at week t.
2) Strategy S: $S(t) = 100\left[\sum_i w_i A(t_i) + c\right]$, where$w_i \in [-1, 1]$ gives the weight for each day and c is an arbitrary constant.
3) Each person has N such strategies.
4) We define an error function that calculates the difference between S's prediction and the actual attendance for each week in the agent's memory.
5) The agent choose to act according to the strategy S^* that minimizes the error function. S^* is called the Best Current Strategy.
6) Each person makes a decision that if $S^*(t) >$ overcrowding-threshold they don't go, otherwise they go.
7) After they learn the attendance of the current week, each person update their Best Current Strategy S^* according to the new information.

From my past experience, the updating rule can be formulated using more complicated learning mechanisms. One example is to use reinforcement learning.

The guest Brian Arthur talked about the complexity economics, he talked about how we can observe similar phenomenon in economics systems as in the physical systems. He believes what's amazing in today's word is that now people are empowered by computers. We have a lot more computational power to tackle issues that are not well-defined. Complexity economics is doing the task as before, but doing it with modeling and with the aid of computer power.

9 Networks

Networks are natural in the real world. For example, we have neural networks, transportation networks, food web. We are interested in the common properties to all complex networks so we can find a general theory of the structure, evolution, and dynamics of networks. Network scientists have proposed some common properties are 1) Small world property, 2) Long-tailed degree distribution, 3) Clustering and community structure, 4) Robustness to random node failure, 5) Vulnerability to targeted hub attacks, 6) Vulnerability to casacding failures. We would not be able to cover all the cases here, but we will illustrate some as examples.

9.1 Terminology

Here we define some basic terminologies necessary for network science literatures.

Definition 9.1 (Network)**.** Network consists of nodes connected by links. The links can be directed or undirected.

No we define some basic properties for a network.

Property 9.1 (Network properties)**.**

1. Degree of a node: number of links connected to the node. If the links are directed, the degree can be further classified into in-degree and out-degree. The diagram of the degree distribution of the network of interested can often provide crucial insights. E.g. a node with high degree could indicates its importance.

2. Hop: A hop is when one node can reach another node via one link.

3. Shortest path: we define path in terms of the number of hops between two nodes A and B, the shortest path is the path with the least number of hops.

4. Clustering: clustering defines to what extent the nodes in a network are connected. C_v is defined as the fraction of pairs of neighbors that are connected to one another. Clustering coefficient C is defined as $C = \frac{1}{n}\sum_n C_v$, where $\sum_n$ sums over all the nodes.

9.2 Small World Networks

Stanley Milgram, a professor at Harvard, famously did an experiment testing the connections between people in US society. He found that on average people are connected by "six degrees of separation".

Definition 9.2 (The Small-World Property)**.** THe network has relatively few "long-distance" links but there are short paths between most pairs of nodes.

In Duncan Watts & Steven Strogatz's paper *Collective dynamics of 'small-world' networks*, they studied how real-world networks of film actors, power grid, and C. elegans. They compare the real-world netowrk to a completely regular network and a completely random network. The results are summarized below. Refer to the lecture slides for more details.

Table 2: Small-World Network Summary

Regular network	**Small-World network**	**Random network**
high average distance	small average distance	small average distance
high clustering	high clustering	low clustering

9.3 Scale-Free Networks and Long-Tailed Distributions

Many real life networks exhibit the scale-free property defined below.

Definition 9.3 (Scale-Free Network)**.** Network with sacle-free (i.e., power-law) degree distribution.

For example, the number of nodes of degree k is given by the power law $N \propto \frac{1}{k^2}$.

However, the "scale free" hypothesis has been disputed to be an overestimation. The more generally acccepted statement is that **many real-world networks have long-tailed degree distributions,** where long tail means the distribution has many nodes with low degrees more than nodes with high degrees. One example for the long-tailed distribution is to explain how Google works. The search algorithm is to rank the websites according to their degrees. the WWW has this long-tailed distribution so that it is possible for Google to rank the websites.

The interesting question then is how does these networks form. They are not designed by any centralized agent! One hypothesis says that it is due to the *preferential attachment*, which argues that the new nodes are added to the network preferentially with a bias towards the existing nodes with high degrees. There is a NetLogo model for exploring this phenomenon.

Now we can discuss the robustness of long-tailed networks.

Property 9.2 (Robustness)**.**
1. **Vulnerable** to targeted "hub" failure. For example, if the front page of Yahoo is down, it may cause lot's of trouble. This is a result for the *clustering* effect in the networks.
2. **Robust** to random node failure. For example, an individual site on the website down won't cause a lot of trouble.

However, sometimes, nodes can cause other nodes to fail. These are called cascading failures.

Another interesting observation is that the events in tail are more likely than in normal distribution. Their probability do not fall out exponential.

10 Scaling in Biology and in Society

Scaling means how does a quantity (regardless whether it is a physical quantity or not) changes as the size of the size changes. This is very like the power law in physics.

Example 13 (Scaling of Area and Volume in 3D)**.** In 3D the area scales as $A = L^2$, while the volume scales as $V = L^3$.

Power laws are can be expressed in two forms:

$$y = cx^\alpha,$$

or equivalently

$$\log y = \alpha \log x + \log c.$$

Two very interesting examples excerpted from the Unit 10 slides are listed below. They are so interesting so I cannot resist to list them here.

10.1 Metabolic Scaling in Biology

It is often very intriguing why different animals have different sizes. The question we are trying to answer here is whether we can relate the size of the animal to their metabolic rates. Since the mass of each animal is (roughly) proportional to their sizes, we seek a relationship between their body mass and their metabolic rate. The famous result given by K. Schmidt-Nielsen is shown in figure 4.

In this course, we give the following definition for the metabolic rate:

Definition 10.1 (Metabolic rate)**.** Amount of energy expended by an organism per unit time. This can be measured as the amount heat emitted by the organism per unit time.

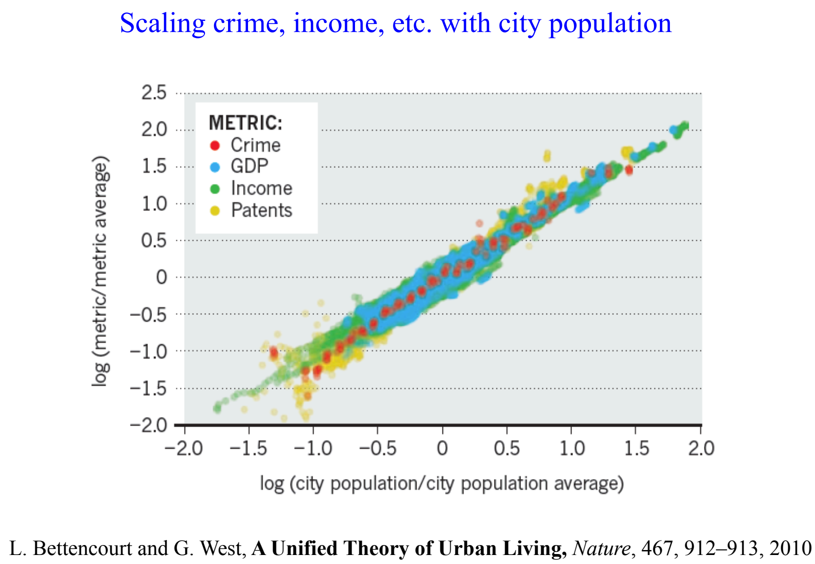

Figure 3: It is very surprising to see that even objects as complicated as a city may possess power laws. However, this finding is controversial.

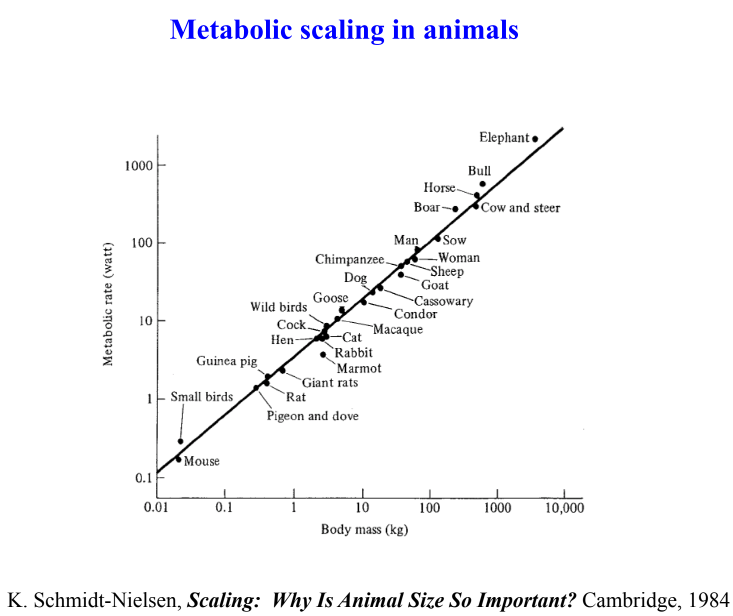

Figure 4: How the metabolic rate is related to the body mass of different animals.