The Causal Structure of Complex Systems

Thomas Machon

Presented here is a short summary of computational mechanics and its applications and implications regarding the construction of macrostates and good observables for a given system. This work is based on papers Ref.[1] and Ref.[3]. Since these two papers inform the majority of this report, references in the text will be largely omitted.

Introduction

When one is studying a system with some eminently practicle purpose in mind, one is almost always concerned with aggregate quantities, or at least quantities that are defined on a coarser scale than the (presumed) fundamental constituents of the system (we denote these quantities, now and henceforth, as macrostates, distinguished from microstates which are the states of the system at the lowest level of description). Given this concern, it is natural to ask whether or not one can just consider the dynamics of the macrostates themselves, remain ignorant of the microstates, and still know (as much as possible about) the future of this system. Indeed from such a scheme one could hopefully make gains both in the computational effort required for simulation and in understanding of the system on scales of interest. In this short report we review the theory of computational mechanics, due to Crutchfield, Shalizi *et al.* and, in particular, relate it to the notion of macrostates of systems. We discuss the philosophical implications and insights from doing so and then, in Section III, apply the methods to crude models of epidemics on networks.

I: Computational Mechanics & Observables

I.I: Computational Mechanics

This section contains a selection of claims regarding the properties of causal states and related objects. For proofs of these, readers are invited to consult Ref.[3].

Consider a discrete time stochastic process s_t, where s takes values $s \in \mathcal{S}$ and time is infinite in both directions (informally we might say the process has always been running). We choose a time, t, and split the process into a past (which includes the current state), $\overleftarrow{s}$, and a future $\overrightarrow{s}$. Considering the set of all histories $\overleftarrow{S}$ we make the following definition.

Definition 1 (**Casual Equivalence**). *Two histories,* $\overleftarrow{s}_1, \overleftarrow{s}_2 \in \overleftarrow{S}$ *are causally equivalent* ($\overleftarrow{s}_1 \sim_\epsilon \overleftarrow{s}_2$) *iff*

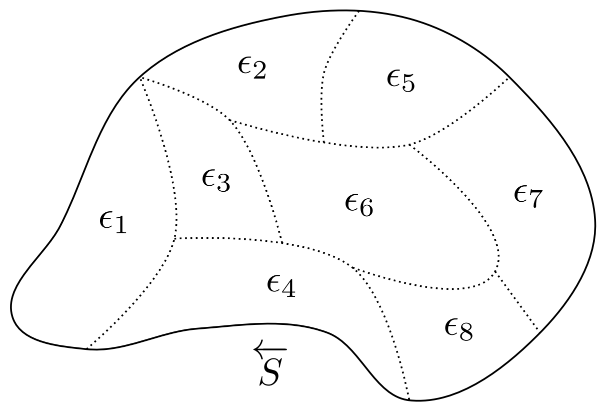


Figure 1: An example partition of histories into causal states (here there are eight such).

$$\mathbb{P}(\overrightarrow{s}|\overleftarrow{s}_1) = \mathbb{P}(\overrightarrow{s}|\overleftarrow{s}_2).$$

Now $\sim_\epsilon$ is symmetric, reflexive and transitive and so partitions $\overleftarrow{S}$ into equivalence classes, and we define the following.

Definition 2 (**Causal States**). *The causal state of a history, $\epsilon(\overleftarrow{s})$, is given by the equivalence class of $\overleftarrow{s}$ with respect to $\sim_\epsilon$:*

$$\epsilon(\overleftarrow{s}) = \{\overleftarrow{s}'|\overleftarrow{s}' \sim_\epsilon \overleftarrow{s}\}.$$

The theory of computational mechanics is built around the notion of causal states. The idea is captured by the following equation

$$\mathbb{P}(\overrightarrow{s}|\overleftarrow{s}) = \mathbb{P}(\overrightarrow{s}|\epsilon(\overleftarrow{s})). \tag{1}$$

That is, the dependence of the system on its past is completely captured by its causal state. To establish the causal states as objects worthy of consideration we will examine some of their properties. Consider some general partition $\mathcal{P}$ of $\overleftarrow{S}$. We have that

$$I(\overrightarrow{s};\mathcal{P}) \leq I(\overrightarrow{s};\overleftarrow{s}), \tag{2}$$

where $I[X;Y]$ is the mutual information between two random variables X and Y. Intuitively we say that the equivalence class to which a history belongs can never contain more information about the system's future than the history itself. Using (2) we motivate the following definition.

Definition 3 (**Presience**). *A partition, $\mathcal{P}$, of $\overleftarrow{S}$ is termed prescient iff*

$$I(\overrightarrow{s};\mathcal{P}) = I(\overrightarrow{s};\overleftarrow{s}).$$

Given this definition we claim that the causal states form a prescient parition, that they are the coarsest such partition and that any other prescient partition is a refinement of the causal states. Again, for proofs of these claims, see Ref.[3]. Furthermore, the causal states are the least complex prescient partition. We first define the statistical complexity.

Definition 4 (**Statistical Complexity**). *The statistical complexity of a partition, $\mathcal{P}$, is defined as*

$$H(\mathcal{P})$$

where H is the information entropy.

We say the statisical complexity of a partition is the complexity of the predictor the partition corresponds to and moreover it can be shown [3] that the causal states form the unique least complex prescient partition in this sense.

I.II: The Causal States of Observables

We now consider a different system. We say only that it is a dynamical system in discrete time with phase space U. Given the state of the system at time t, $q_t \in U$, we define a time evolution operator $\mathcal{L}$ such that

$$q_{t+1} = \mathcal{L} q_t. \tag{3}$$

We insist that the form of $\mathcal{L}$ is such that $\{q_t\}$ form a Markov chain, furthermore we consider an ensemble of systems, and by Q we denote the random variable for the current state. We consider measurements made on the system. More formally we define a surjection, f,

$$f : U \to \Gamma, \tag{4}$$

where Γ is our space of observables. Furthermore, f, partitions U and we denote the partition induced by f as the observational partition. Given f we can define the observed process $S_t = f(Q)$ (note that whilst Q is Markovian, S_t is not necessarily so).

Now consider the causal states, $\mathcal{X}$, of the observed process S_t. Since $\mathcal{L}$ is Markovian, any point in U will give a probability distribution over the future of Q, and hence of S_t. It follows then, that $\mathcal{X}$ induces a partition of U (which we call the causal partition). We then have two partitions, the causal and the observable. Furthermore we have four possibilities concerning their relationship.

1. The observational and causal partitions are the same.
2. The causal partition is a refinement of the observational partiton.
3. The observational partition is a refinement of the causal parition.
4. The two partitions are incomparable.

The issues resulting from these four situations are discussed in section II.

II: Philosophical and Structural Implications

II.I: Four Possibilities

1. The observational and causal partitions are the same. This is a good situation, the macrostates, by definition, allow observational quantities of interest to be computed, and their dynamics are Markovian and we are in a situation where we can describe the evolution of the Macrostate purely in its own terms. This is exactly what we set out to do in the introduction. An example of this situation is the ideal gas law.
2. The causal partition is a refinement of the observational partiton. Here the current macrostate does not give us enough information to predict the future dynamics. We must know some of the history also. However, it can be shown[1] that for any choice of observable partition $\mathcal{F}$ there is another partition $\mathcal{G}$ such that the product of the two $\mathcal{F} \cdot \mathcal{G}$ gives the causal states. Thus we can observe f and g, and we return to the situation where the causal partition and the observational partition are the same.
3. The observational partition is a refinement of the causal parition. In this situation several different values of the macrostate will have the same consequences for the evolution of the macrostates. Hence the distinction between them is meaningless. An example could a macrostate, specified by n variables, where specifying $n-1$ variables determines the last (this can happen unintentionally) and so there is a redundancy in the description.
4. The two partitions are incomparable. In this situation, neither partition is a refinement (or coarsening) of the other. There is not much to be done here. All that can be concluded is that the choice of observable(s) (and thus the observable partition) was poorly made.

II.II: What Constitutes a Good Macrostate?

Given the previous we can give three criteria for a good macrostate:

A. Observational quantities of interest are computable from any given state.

B. The dynamics are Markovian.

C. The partition of phase space induced by the macrostates is the coarsest one satisfying items A and B.

That is, our choice of macrostate is driven by two considerations. Firstly, the choice of what observable(s) we would like to consider and secondly, the objective criterion that the dynamics be Markovian. In terms of the four possibilites of the previous section we can say the following for each possibility

1. A, B & C are satisfied, the macrostate is good.
2. A is satisfied, we introduce a new observable to the macrostate (as described in the previous section) such that B & C are satisfied.
3. A & B are satisfied, however, we can further coarsen the partition induced by our observables to satisfy C.
4. None of A, B or C are satisfied, and our choice of observable needs to be remade.

II.III: Multi-level Descriptions and Incompatibility

Computational mechanics has some interesting observations to make regarding the compatibility of different descriptions of the same system. We first consider the mathematical description of an ideal gas on three different levels, that of the particles themselves (the microstates), at the level of hydrodynamics and with thermodynamics both of which involve macrostates. For clarity, hydrodynamics deals with locally averaged quantities whereas thermodynamics deals with globally averaged (or summed) quantities. In an ideal gas the hydrodynamic description involves the Navier-Stokes (or, more properly, Euler) equations, whereas the thermodynamic description is simply the ideal gas law. Furthermore it is possible to derive, in full, the ideal gas law from both the microstates and the hydrodynamic description[5]. This situation is illustrated in Figure 2, with arrows representing a possible coarsening of description.

We now analyse this situation from the perspective of computational mechanics. Suppose in an alternate universe two scientists, Stokes and Boyle, say, are considering the properties of an ideal gas for the first time. Stokes comes up with the Navier-Stokes equations and Boyle the ideal gas law. Now in this universe the laws of physics are slightly different and it turns out that one cannot coarsen the Navier-Stokes equations and find the ideal gas law (and so ensues a bitter conlflict over who is correct). Is this situation possible? The answer is yes. If the causal states of the observables induced by the two perspectives are incomparable then it is simply not possible to refine one description into the other. However, both descriptions remain valid coarsenings of initial system (and may both produce insight).

The relevance of this to the discussion at hand is the simple observation that it is possible to have two high level descriptions of a system, that are both valid but that cannot be translated (*i.e.* one cannot be mapped to the other).

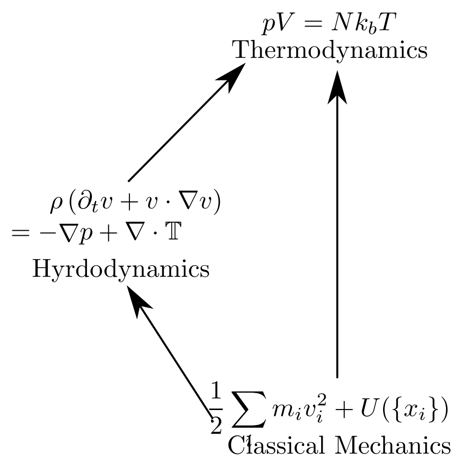


Figure 2: Different level descriptions of an ideal gas, arrows indicate possible coarsenings.

III: An Example: Contact Processes on Graphs

We now apply the tools of Section I to a simple contact process. Our strategy is simple, we construct the causal states of the observable N that counts the number of infected people (in the frequently used interpretation of the contact process as an epidemic) for two graphs of differing topologies. Having done this we evaluate the appropriateness of N as a macrostate in each case.

Definition 5 (Contact Process). *A contact process[4] $(\vec{\eta}_t : t \geq 0)$ is a continuous time stochastic process on a lattice Λ with state space $S = \{0,1\}^{|\Lambda|}$ and transtion rates, $R(\vec{\eta} \to \vec{\eta^x})$, defined as*

$$R(\vec{\eta} \to \vec{\eta^x}) = \eta(x) + \lambda(1 - \eta(x)) \sum_{y \sim x} \eta(y), \tag{5}$$

where $\eta(x)$ denotes the x^{th} component of $\vec{\eta}$, $\vec{\eta^x}$ denotes $\vec{\eta}$ with its x^{th} component flipped and $y \sim x$ means that the y^{th} component of Λ is connected to the x^{th} component.

Since we developed the machinery of computational mechanics to deal with discrete time processes, we will discretise the contact process by considering the jump chain. That is, the sequence of states that the system occupies. We can then compute well defined transition probabilities for these jumps. We will take the complete graph and the periodic one dimensional lattice, both with four nodes, as our example networks of which we will compute the causal states.

We consider the observable N, which counts the number of ones on the lattice (or the number of infected). If we only observe N, what causal structure will we see in the two cases? Given that we know the details of the underlying systems and that the systems themselves are small in size we can compute the states directly. We note that these two statements are not true for real systems of concern. Nevertheless this is an informative example and furthermore algorithms have been developed allowing for computation of causal structure from arbitrary time series data[2].

The actual computation was done by inspection. That is, all histories of observed Ns were considered, and their possible transitions (this is possible because of the small state space of the systems). Before proceeding we note that for even small increments in the system size, the effort required to compute the causal states increases by a large amount.

We show the results, illustrated as a Markov chain (recall the dynamics of the causal states are Markovian) in Figures 3 and 4 with the corresponding microstates shown in Figure 5.

We discuss first the complete graph. In this case, because of the network structure, the dynamics of N are Markovian, the macrostates are the same as the causal states and we have made a good choice in choosing N as our only observable (indeed this corresponds to situation 1 in our four possibilites of section II.I). The one dimensional lattice, however, is a little different, and we are in situation 2. The partition

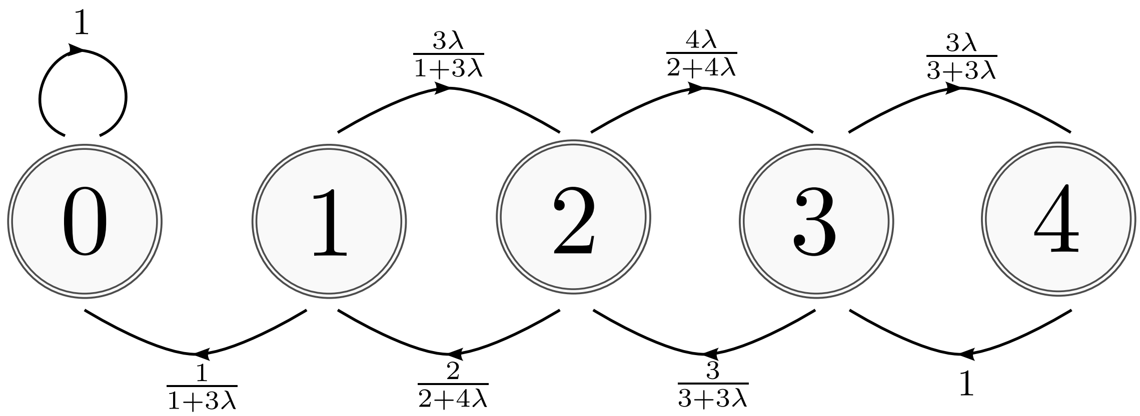


Figure 3: Causal states, and their transition probabilities, of the contact process for the complete graph with four nodes. As discussed in the text, the causal partition is the same as the observational one induced by counting the number of infected, and hence we label each casual state with the number of infected.

of the microstates induced by N is coarser than the causal states and hence the dynamics of N are not Markovian. Without introducing a new, additional observable (as was suggested in section II.I), we must, in order to satisy criterion A, look at the previous **two** values of N, to construct the causal states, and thus obtain Markovian dynamics. We can, instead, introduce the new observable G, which counts the number of connected groups of infected. Specifying the state with the pair (N, G) then satisfies our requirements.

What do we conclude from this? In this particular example we show that more than just the number of infected is required to optimally predict the future behaviour of the model.

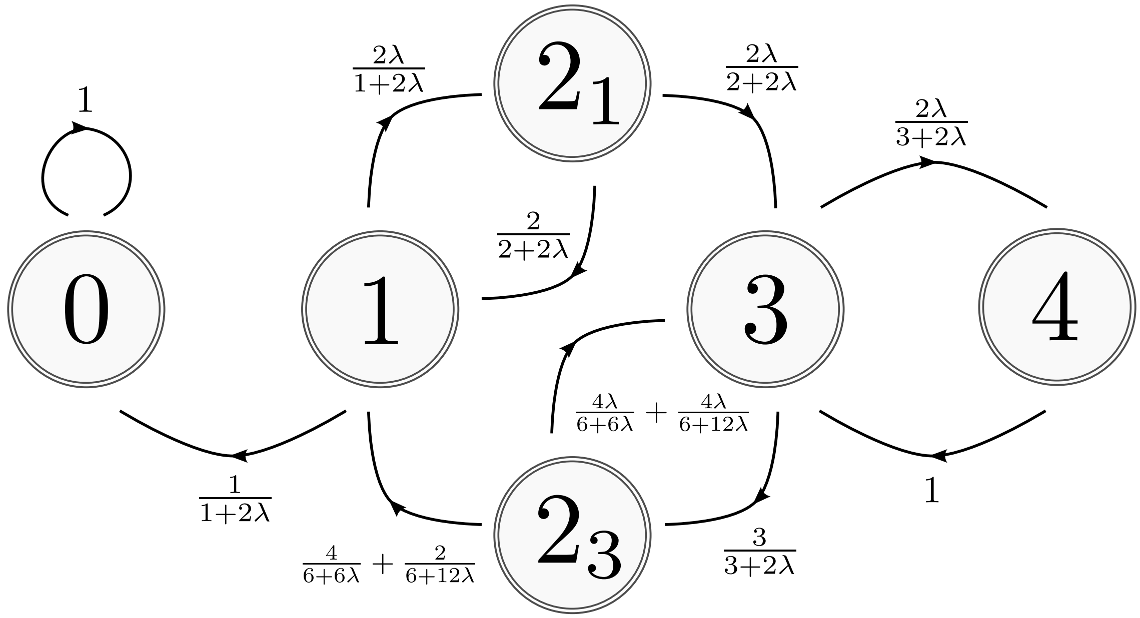
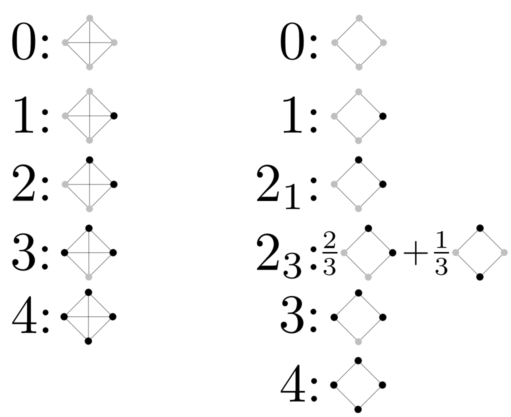


Figure 4: Causal states, and their transition probabilites, of the contact process for the complete graph with four nodes. As discussed in the text, the causal partition is finer than the observational partition induced by counting the number of infected. In particular we have two causal states 2_1 and 2_3, which have different transition probabilities, but the same value of $N = 2$. In particular 2_1 corresponds to observing $N = 1$ then $N = 2$, whereas 2_3 corresponds to observing $N = 3$ then $N = 2$.



(a) Complete Graph (b) 1d Lattice

Figure 5: Schematic of causal states for the contact process on (a) the complete graph and (b) the one dimensional lattice, both of degree four (black nodes correspond to infected).

IV: Conclusion

What have we acheived? Given a system and a subjective choice of observable(s), we give guidelines and criteria for choosing good macrostates of the system, and then finding their dynamics.

References

[1] C. R. Shalizi and C. Moore, What Is a Macrostate? Subjective Observations and Objective Dynamics, 2003, e-print cond-mat/0303625.

[2] C.R. Shalizi, K.L. Shalizi and J.P. Crutchfield, "An Algorithm for Pattern Discovery in Time Series", 2002, e-print cs.Lg/0210025

[3] C. R. Shalizi and J.P. Crutchfield, "Computational Mechanics: Pattern and Prediction, Structure and Simplicity", 1999, e-print cond-mat/9907176

[4] See University of Warwick module CO905, lectured by Stefan Grossinsky.

[5] Mehran Kardar, *Statistical Mechanics of Particles*, Cambridge University Press (2007).