**Reading Notes: Alexander B. Boyd, Dibyendu Mandal, and James P. Crutchfield (2018). Thermodynamics of Modularity: Structural Costs Beyond the Landauer Bound, PHYSICAL REVIEW X 8, 031036**

**Uwe Ehret**

**Summary**

This paper addresses thermodynamical aspects of information processing. Specifically, the authors investigate the differences between globally integrated and modular (locally connected) implementations of computational operations, and establish that modular implementations, while more easily/efficiently to establish and process than global implementations, always come with an added thermodynamic cost (modularity dissipation). This loss is different from the well-known Landauer bound, which states that erasing a single bit of stored information at temperature T must be accompanied by the expense of at least kB ∙T∙ ln2 amount of heat released, thus establishing a fundamental link between thermodynamics and information processing. The authors show that the dissipative loss from modular implementations can be quantified by the differences in the change of global and local free energy, and that this loss can be reduced by preserving relevant global correlations in modular implementations (which are ubiquitous in nature, indicating that they are evolutionary favorable over globally integrated processing). As in the Landauer principle, the authors make their arguments based on a dual view (physical-thermodynamical and statistical-information theoretical) of the systems of interest.

They start by introducing the concept of composite information reservoirs, where only parts are available for external control (thus establishing modularity), but all are linked to an external heath bath. For that system, the authors show that the modularity dissipation is proportional to the reduction in mutual information between the controlled and uncontrolled subsystem, and illustrate it at the example of an information ratchet (an special form of information engine which can implement a universal Turning machine). Then, the authors demonstrate at the example of pattern extractors (which consume structure in the input to create work, producing unstructured output) and pattern generators (which consume work to create structure in the output, using unstructured input), how a purposeful design can minimize modularity dissipation, and that efficient pattern generators are time-reversed versions of pattern extractors. The authors conclude that the concept of modularity dissipation is distinct from Landauers principle, and that it can support the design of efficient modular information processors.

**Open questions**

* The focus in this paper is on thermodynamic aspects of the mechanics of computation. Can we use the insights also to guide development of high-level earth science models, where the focus is more on macroscopic (model-level) rather than microscopic (processor-level) computational efficiency? E.g. to decide whether we should build globally integrated ML models, or build hierarchical modular models (compare Lin et al. 2017), and for the latter to decide how and to which degree global correlations (spatial, temporal, causal) should be represented in modular subsystems?
* Are transformers an efficient way to preserve important global correlations in dynamical systems?
* Inspired by the statements about designing optimal pattern extractors and generators ("extractor transducers whose states are predictive of their input are optimal", and "generator transducers whose states are retrodictive of their output are optimal"): Can we learn something from an analysis of environmental system observables about the required architectural or computational design of a model thereof?
* "Pattern generation may be viewed as the time reversal of pattern extraction": This corresponds to the discussion about the relation of hierarchical generative processes and classification in Lin et al. (2017). Should we then train models for both classification and generation simultaneously? Would this yield more realistic models?
* In which sense are our typical earth science models information ratchets (see Fig. 3), and is this comparison helpful?