## Complex Systems

I would like to answer this question by sectioning it from the four major perspectives that color my imagination, reflecting the approach of complexity science and also the institute.

As a student of **physics**, what interests me about complex systems is the question of **structure**. What concise mathematical laws are sufficient to explain the interaction dynamics in a system for a certain observer? Anything we perceive can be modeled as a complex system in some space, and these laws to me represent the essence of our understanding of the system at that level. The main drive for me as a physicist is to condense from different perspectives, a unified mathematical structure that may emerge in everything I see.

At the same time, I am also a student of **biology**. A walk in a garden, holding the **mechanisms** operating at each level reminds me of the infinities of **interactions** embedded therein. This sublates my drive for absolute unification; It gives me many more levels to explore structure at - from cellular metabolism and cancer to population ecology and the evolutionary fitness landscape, but also makes me quiet and contemplative of the sheer novelty of life.

The space for wonder engendered by biology is further deepened by my explorations in **philosophy**. Where the physicist would try to distill the essence of all that exists to a single drop in a desert, the ability to simultaneously hold the perspective of a biologist instead leads to the synthesis of a rainforest ontology, teeming with life at every level. This shift, and the ability to **hold together different epistemological frameworks** is the philosophical challenge in complexity science that interests me.

As an **engineer** studying computer science I am interested in the things we can **build** or the **practical problems** that we can solve. Recent developments make me optimistic about the abilities of computers and also make me want to explore just how far we can push them, either by using insights from other fields to make them more efficient or reverse-engineering new approaches to problems by interpreting how the computers look at them.

From my understanding Santa Fe is one of very few places that may call itself **truly transdisciplinary**. It appeals to me because The dynamism necessary for exploring my curiosity genuinely is something only such an environment could provide.

## Research Topic 1: Free Energy Principle- Interpretability in Neural Networks, Statistical Mechanics and Self Organization

The driving question here is: What is the analogy between the mathematical structures that come up while Neural Networks learn, in physical phase transitions and in Biological Self Organization with various different process theories e.g. Neuroscience or Evolution? How can we take insights from one of these fields and apply them in the context of the others?

As we become more reliant on compute for everyday tasks, and especially as the vision of AGI seems to be inching closer, it becomes important that the technology we engage with doesn't become incomprehensible under layers of abstraction. By understanding the workings of these models as they evolve through Developmental Interpretability, we can build robust and responsible AI systems.

I find it elegant that the mathematical language used in describing Developmental Interpretability of Neural Networks is very similar to that which describes Self-Organizing complex Systems, Neuroscience and Statistical Mechanics. I want to push the analogy between these as far as I can, incorporating the insights from each of these branches into others wherever applicable to see what we can learn about the process of our own thought and the functioning of these Deep Networks.

For this project I would like to explore similarities between work of Karl Friston on the Free Energy Principle which incorporates modeling with random dynamical systems/stochastic differential equations and a Variational Principle of sorts, together with the work of Sumio Watanabe on Singular Learning Theory which incorporates the work of Rene Thom in Catastrophe theory and Mathematical Morphogenesis. I would be using ideas from statistical mechanics,the specific biological process theories considered for the FEP, AI safety and interpretability, and contemporary deep learning.

## Research Topic 2: Entropy, Time and Thermodynamics from a Quantum Perspective

The question of entropy and the arrow of time are fundamental questions that have remained unanswered. They are important in themselves as answers to foundational problems in physics and philosophy and are also closely related to thermodynamics, so they have many practical applications such as finding the limits to the reversibility of computation. These will only become more important as quantum computers become closer and closer to reality.

Personally I am interested in these topics mainly for three reasons. First, the exotic physical and mathematical structures to be explored; Black Holes and the Geometry of Thermodynamics, Category theoretic formulations of Quantum Mechanics, Quantum chaos and its relation to random matrix theory are extremely powerful tools I have read about and want to explore more deeply. Second, the philosophical implications of these problems for our understanding of the universe, perception and time and third, the possibility of interpreting these ideas from mathematical and computational physics in the contexts of other complex systems.

Reading about basic concepts in this area such as differential geometry or quantum mechanics changed my approach to problems and thinking so much that I could not even have comprehended the ideas I would be interested in after reading them. Since there is so much I do not know, I would like to explore the landscape of available theories to see what resonates with me and makes me curious.

## Collaborative Work

My participation in hackathons and academic team projects has taught me the domain specific skills that are relevant for navigating through a project as a group, such as effectively negotiating contributions and communicating the whole process so that teammates understand your situation. However, in my opinion the most pertinent experience has been being a drummer for my college band. It has taught me how people who have different backgrounds, values and aspirations in life can be united by something they love. There are three important points that I have observed from this: At one point in time, I used to be intimidated by the sheer brilliance of the musicians I was working with, only slowly have I developed an appreciation for the fact that my perspective which feels limited to me might also appear insightful to others. The realization that all of us show up for practice in such an academically intensive environment engenders a degree of mutual trust about the dedication to the project, which is also essential for the functioning of any group in my opinion. Another thing I have observed that is necessary is a dynamic flowing of identity; For example, sometimes I may be the one motivating everyone to practice, but at other times when I feel down someone else may have to take up that role and they might even have to be a bit stern. Sometimes, when a creative block is hit it takes humility to realize that and take a break instead of continuing to practice till burnout. This requires being empathetic to the other members while also having clear sight of the goal the group was formed for. Even the music club for our college, which has been able to organize intercollege events with a footfall of over 700 people carries this feeling at its heart. I think groups function best when the hierarchy or organization is not enforced but rather arises spontaneously and intelligently if necessary, as a means to achieve a vision the members are committed to.