[붙임]

KAIST Grand Challenge 30 Project 제안서

<table>
<tr><td rowspan="2">①연구제목</td><td colspan="2">국문</td><td colspan="3"></td></tr>
<tr><td colspan="2">영문</td><td colspan="3"></td></tr>
<tr><td rowspan="5">②제안자</td><td rowspan="2">성명</td><td>국문</td><td></td><td rowspan="2">소속
학부/학과</td><td rowspan="2"></td></tr>
<tr><td>영문</td><td></td></tr>
<tr><td colspan="2">주요
연구분야</td><td colspan="3"></td></tr>
<tr><td colspan="5">연구 impact 내용 서술</td></tr>
<tr><td colspan="5"></td></tr>
</table>

③제안내용
(1) 연구주제에 대한 개요, 일반적인 해결방안 등을 소개하는 세계적 연구현황, 본인의 독창적 해결법을 소개하는 연구방법을 제시 (2) 연구의 성격이 공고문의 신청자격에서 명시된 내용에 해당하는 이유를 명시 (3) 연구의 필요성, 기대효과 등은 필요 없음. (4) 고등학생도 이해할 수 있는 평이한 언어로 2-3쪽 정도 작성

제안자 : _____ (인)

Proposal for KAIST Grand Challenge 30 Project

<table>
<tr><td rowspan="2">1.Title</td><td colspan="2">Korean</td><td colspan="3">지식의 물리적 토대.</td></tr>
<tr><td colspan="2">English</td><td colspan="3">The Physical Basis of Knowledge</td></tr>
<tr><td rowspan="5">2. Principal Investigator(PI)</td><td rowspan="2">Name</td><td>Korean</td><td></td><td rowspan="2">Department</td><td rowspan="2">Bio and Brain Engin.</td></tr>
<tr><td>English</td><td>Fiorillo, Christopher D.</td></tr>
<tr><td colspan="2">Major Research Field</td><td colspan="3">Neuroscience</td></tr>
<tr><td colspan="5">Impact of research project</td></tr>
<tr><td colspan="5">Science and brains deal with information (knowledge), and yet the physical basis of information is still not understood. We seek to characterize a state of knowledge through the mathematics of probability, and then use that mathematics to describe the knowledge in physical systems. Our ultimate goal is to lay the foundation for a theory of brain function, as well as machine intelligence. The more immediate benefit will be improved statistical methods for describing the knowledge of scientists (data analysis).</td></tr>
<tr><td>3. Project Summary</td><td colspan="5">Our goal is to describe the knowledge in a physical system with probabilities, in analogy to the way that calculus can describe the motion of a physical system. We plan to first describe the knowledge in a geometrical configuration of points, before using the same methods to describe the knowledge in a physical system (such as the position and velocity of a particle, or the voltage and current across a neuron's membrane).</td></tr>
<tr><td colspan="6">a. Provide the global research trends concerning the proposed research theme including an overview and general solutions.
b. Specify how the nature of the proposed research satisfy the requirements of eligibility(b).
c. No need to write research necessity and expected outcome.
d. Write in plain language within 2~3 pages.</td></tr>
</table>

Applicant : _____ (signature)

Science is based on knowledge and reason (information and logic). However, science lacks an understanding of these concepts in relation to physical systems. What we call "the laws of physics" is a model that science derived from knowledge and reason. Until these laws are understood in relation to knowledge and reason, science does not have a proper and strong foundation. The inadequate foundation may explain why some aspects of quantum mechanics appear paradoxical and absurd. As a neuroscientist, I believe that we must understand the physical basis of knowledge in order to understand the brain, and to advance machine intelligence. A general theory of brain function currently appears out of reach to most neuroscientists because they have virtually no understanding of how information relates to physical systems.

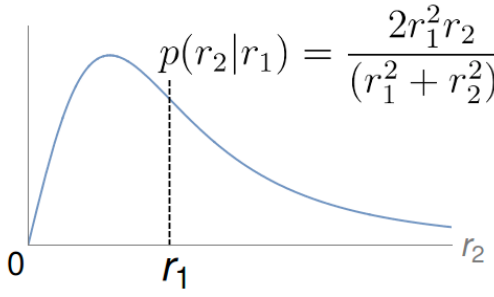
E.T. Jaynes, a physicist, developed a unified mathematical framework for knowledge and reason[1]. Unfortunately it is not well known to most scientists. I summarize it as "logic is objective and universal, knowledge is subjective and local." He used logic (objective Bayesian probabilities) to describe subjective states of knowledge. His mathematical methods have already been very useful in statistics and machine intelligence. But their full potential has not been realized. They provide an objective means of characterizing subjectivity, a topic that has caused fear and confusion for scientists. His framework has become the basis for an increasingly popular view that the general function of the brain is prediction (or inference)[2,3]. My lab's research has contributed to this area, both with respect to theory[4-8] , experimental tests of theory[9-11], and patents on machine learning algorithms[12-15].

However, none of this past work has adequately explained the physical basis of knowledge. Scientists have focused on characterizing *their knowledge about physical systems* (and sometimes they have confused their knowledge with that of the physical systems they study)[6]. In contrast, I want to characterize *the knowledge in physical systems*. What does one physical system know about another? What does it mean for a physical system (like a neuron or brain) to "be a subjective observer?" Science has already made tremendous progress in developing excellent physical models. Given a physical model of a neuron, I want to "take the neuron's point of view." This is similar to the way that we try to understand other people by considering their point of view. This may not appear scientific, but we know it works. I believe this can be objective and scientific given the mathematics of Jaynes together with physical models.

Specifically, we would like to characterize knowledge by finding the probability distribution over states of physical system 'A,' conditional on the state (knowledge) of system 'B.' For example, what is the probability distribution over external states conditional on a neuron's membrane voltage and current (at a

moment in time)? As a simpler example, what is the distribution over location of particle A, conditional only on the position and velocity of particle B?

In trying to answer these questions, we found that there are currently no known solutions (probability distributions) for simple but realistic states of knowledge. Therefore we chose to begin with a problem of pure geometry before considering physical systems. Given the known locations of N points, where is an unknown point? It may appear that every location is equally probable, and thus there is no interesting solution. However, since the unknown distance must be finite, it is more likely to be closer rather than further from the known locations. We have introduced the problem to the literature, emphasizing that we are seeking an exact solution with no assumptions[8]. Since then we have derived the exact solution to the simpler but related question "given N known distances, what is an unknown distance?" The figure shows the distribution over distance r_2 given knowledge of distance r_1. The known distance is the median of the distribution, and thus the unknown distance is equally likely to be more or less than the known distance. This could be published alone, but we are hoping to solve the entire N-point problem soon.



Although our ultimate interest is knowledge in physical systems, the N-point problem is at the foundation of statistics, and its solution should be of widespread utility. Because it could be useful in estimating the position of stars, I gave an invited seminar at Korea Astronomy and Space Science Institute in November 2015.

Specific Aims

1. Solve the 'N-point problem.'
2. Given the position, velocity, and acceleration of one particle, where is another particle? 'Position and velocity' is similar to (or the same as) the special case of N=2 in the 'N-point problem.'
3. Given a neuron's membrane voltage and current, what is the external state of the world? We suspect that this may be similar to the case of the position and velocity of a particle (above), but on a macroscopic level in which many particles move in a highly coherent manner.

Eligibility

This proposal meets all of the six criteria.

1. Understanding the physical basis of knowledge is directly relevant to the "global challenge" of a successful theory of the brain, and machine intelligence.
2. It is the most fundamental issue in science in my opinion. The brain remains among the greatest mysteries in science. Our knowledge led us to "the laws of physics."
3. It is far from my main focus on the theory and physiology of the brain.
4. It would obviously be a challenge to get funding for this project, since it is highly novel and far from my established expertise. A PhD student, Sunil Kim, has assisted me on this project without any dedicated grant support.
5. This is clearly not a "hot topic." I cannot name one scientist working on this problem. The closest may be Ariel Caticha, a physicist. His work is quite different from mine, and not "hot."
6. As applied mathematics, it lacks commercial potential within 10 years.

Concluding Remarks

This project is far from my established area of research (my most recent and notable success was a 2013 single-author paper in Science on the representation of reward in the brain of monkeys). Pursuit of this project is a risk to my career. I am taking the risk because I believe that this project is extremely important, and it will succeed. It is the biggest and best of my ideas. In addition, the cost of this research is minimal, just salary and personal computers. The potential long-term benefits are very large.

References (the two most relevant are underlined)

1. Jaynes, E. T. *Probability Theory: The Logic of Science*. (Cambridge University Press, 2003).
2. Friston, K. J. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11,** 127–138 (2010).
3. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* **36,** 181–204 (2013).
4. Fiorillo, C. D. Towards a general theory of neural computation based on prediction by single neurons. *PLoS ONE* **3,** e3298 (2008).
5. Fiorillo, C. D. A neurocentric approach to Bayesian inference. *Nat. Rev. Neurosci.* **11,** 605–author reply 605 (2010).
6. Fiorillo, C. D. Beyond Bayes: On the need for a unified and Jaynesian definition of probability and information within neuroscience. *Information* **3,** 175–203 (2012).
7. Fiorillo, C. D., Kim, J. K. & Hong, S. Z. The meaning of spikes from the neuron's point of view: predictive homeostasis generates the appearance of randomness. *Front. Comput. Neurosci.* **8,** 49 (2014).
8. Kim, S. L. & Fiorillo, C. D. Describing realistic states of knowledge with exact probabilities. *Amer. Inst. Physics Conf. Proc.* **1757,** 060008: 1–8 (2016).
9. Hong, S. Z., Kim, H. R. & Fiorillo, C. D. T-type calcium channels promote predictive homeostasis of input-output relations in thalamocortical neurons of lateral geniculate nucleus. *Front. Comput. Neurosci.* **8,** 98 (2014).

10. Kim, H. R., Hong, S. Z. & Fiorillo, C. D. T-type calcium channels cause bursts of spikes in motor but not sensory thalamic neurons during mimicry of natural patterns of synaptic input. *Front. Cell. Neurosci.* **9,** 879 (2015).
11. Kim, J.K. & Fiorillo, C.D. Theory of optimal homeostatic balance predicts and explains the amplitude and decay time of synaptic inhibition (under revision for Nature Communications).
12. Fiorillo CD (2016) "Reinforcement learning method and apparatus with elementary value signals" Republic of Korea Patent 10-1603940 (application 10-2014-0025967, issued 2016.03.10) (originally 10-2014-0023249; filed 2013.03.05).
13. Fiorillo CD (2013) "Prediction by single neurons" U.S. Patent 8,504,502 (issued 2013.08.06, application 12/762,120, filed 2010.04.16).
14. Fiorillo CD (2012) "Prediction by single neurons and networks" U.S. Patent 8,112,372 (issued 2012.02.07, application 12/271,282, filed 2008.11.14)
15. Fiorillo CD (2012) "Prediction by single neurons and networks" Republic of Korea Patent 10-1122158 (issued 2012.02.23, application 2010-7011168, filed 2010.5.20).