Jonathan Miller, Assistant Professor
Biochemistry and Molecular Biology
Baylor College of Medicine
One Baylor Plaza, BCM125
Houston, TX 77030
Tel. (713)-798-3542
email: millerj@bcm.tmc.edu

Dec. 11, 2003

Biocomplexity Faculty Search Committe
c/o Prof. Rob de Ruyter van Steveninck
Biocomplexity Institute
Indiana University
Swain Hall West 117
Bloomington IN, 47405-7105

Dear Rob,

It is my pleasure to write in support of Ilya Nemenman's application for a faculty position in Biocomplexity at Indiana University's Biocomplexity Institute.

I've known Ilya for five years, starting in 1998-2001 when he was first a graduate student and then a post-doc of Bill Bialek at NEC Research Institute (NECI) in Princeton, NJ.
At the time, NECI's basic research programs in physics and biophysics were still second-to-none; since then most of the senior research staff, including Bill, have moved to tenured faculty positions at Princeton University and elsewhere.

Ilya's thesis involved the appropriation of physics concepts and techniques to reformulate and address important problems in inference and information theory, and is thus part of a broader vision of Bill's exemplified by the work of Bialek, Callan and Strong on inferring entropy of continuous distributions from undersampled data.

Ilya has forcefully and effectively advanced this program.
Bill relied on him exclusively for numerics, and strongly for analytics.
Ilya's calculations were continually revealing ways in which the two of them needed to reformulate and reinterpret their beliefs and conclusions. Ilya had a strong impact on the research he did with Bill.

One of the most exciting outcomes of Ilya's work with Bill was their characterization of complexity as the subextensive contribution to the entropy of a time-series. Ilya and Bill have revealed that the language of physics is natural for problems of this kind.

Although the characterization of complexity, a classical problem, is in itself increasingly important in applications of information theory, Bill and Ilya's more recent achievement seems likely to have a direct and timely impact on the biological sciences.

In particular, they developed a new method, based on an old idea of S.K. Ma, of estimating the entropy of discrete undersampled distributions.

The challenge here is a ubiquitous one in the natural sciences: to infer the randomness or non-uniformity of a distribution from which are drawn binned samples, when the number of samples is so small that the naive estimate is unreliable. Biological data seems to be generically in this regime. Although it is of course a necessary step toward the more ambitious and manifestly important goal of deducing mutual information in the undersampled regime, entropy inference is frequently a critical task in its own right.

I was following this development closely; Ilya's role was pivotal and decisive. I've studied the method carefully and it is both subtle and effective. Bill and Ilya have organized a workshop on the subject in this year's NIPS meeting. It will surely be generalized eventually to mutual information, the reliable estimation of which, with current methods, is typically impossible in interesting biological contexts. For the moment, it has natural application to estimating entropies of neural spike-trains; these quantities play a fundamental role in interpreting experiments, and Ilya and Bill have recently used it in this context.

Lately Ilya has been developing ways to infer the structure of biological regulatory networks. Bayesian methods are rife in this field and Ilya can provide a badly needed perspective.
I have not yet had the chance to study his work with Chris Wiggins at Columbia University in this direction, but from my general knowledge of the field and of what my colleagues here at Baylor College of Medicine are themselves trying to accomplish, I believe he is addressing an important and timely problem with the correct tools.

Some comparisons may be helpful. I have known well several physics students and post-docs who have gone on to faculty positions as theorists in physics, biology and applied mathematics departments. From Caltech: Ranjan Mudnapymurky, now on the physics faculty at Clark University. From U. Chicago: A. Rojo, physics faculty at U. Michigan, Ann Arbor; Hao Li, now faculty at UCSF, one of the premier biological research institutions in the U.S.; Jane Wang, now on the Cornell applied mechanics faculty; Michael Brenner, now on the applied mathematics faculty at MIT. From NECI: Eldon Emberley, now on the physics faculty at Simon Fraser University in Vancouver.

Of these, Ilya is most comparable, in his originality, creativity, and intellectual ability, to Michael Brenner, and surpasses the others. Furthermore, like Michael, he is energetic and productive, and interacts readily and enthusiastically with colleagues.

Ilya's greatest weakness is his tendency, not unknown among Russian theorists, to forcefully articulate complex and technically difficult arguments as to why the alternatives in which he believes are the correct ones, when eventually, but only after persistent questioning, he would concede that the outcomes in which he doesn't believe are also quite plausible.

Also, his work has yet to show a decisive impact on important experiments.
I am certain this will change in coming years.

In conclusion, as a physicist and as a biologist, I can imagine no better liaison to the field of information theory than Ilya. As just one example, the genomes of many organisms have now been "completely" sequenced (a glib characterization of the state of the data, in some important cases), and the biological community is facing the challenge of extracting every drop of meaning from these sequences. Sooner or later, and probably sooner, sophisticated methods for data analysis, of the kind that Ilya has taken a leading role in developing, will be a principal route toward this end.

If you want to be a player in the emerging "age of biological information"
Ilya would be a fine choice for your faculty.

Sincerely,

Jonathan Miller
Assistant Professor
Biochemistry and Molecular Biology
Baylor College of Medicine