KOÇ UNIVERSITY

Science – Math Seminar

Speaker: Dr. Muhittin Mungan
Department of Physics, Bogazici University
and Feza Gursey Institute

Date: Thursday, Dec. 15, 2005

Time: 16:45 (Tea and cookies will be served at 16:30)

Place: Science building, Room Z42

Title: String Matching and 1d Lattice Gases

Abstract:

I consider the probability distribution for the number
of occurrences of a given word inside a random string whose letters
have been generated by a stationary stochastic process. The
problem is non-trivial due to the possibility of overlapping
occurrences and has applications in communication theory and
analysis of biological sequences. I will show that this problem can

be cast into that of determining the partition function of a 1d lattice gas with interacting particles. Using this analogy, the probability distribution can be obtained from a virial expansion. It turns out that the properties of the stochastic process as well as the specifics of the word to be searched only serve to furnish the effective interactions between the particles of the lattice gas, which turn out to be rather generic: they are pairwise, of nearest neighbour type and characterized by a relatively strong core of size of the word length that is followed by a weak and exponentially decaying tail. Such interactions are known as Kac Gamma potentials and have been shown to give rise to phase transitions in 1d in the limit when their exponential tails decay infinitely slowly. I will discuss the implications of this on the string matching problem at hand as well as possible variants thereof.

_____