Research Activity

(For the reference numbers, refer to List of Papers)

Teturo Kamae

1. Randomness and Algorithmic Complexity

Arround 1968, I was interested in the problem asking what is randomness. I approached this problem from both probability and mathematical logic sides. Before Kolmogorov's formulation of the notion of probability, von Mises defined collective as the notion of probability. It is defined for individual sequences of symbols and more concrete than Kolmogorov's formulation, so that it fits better for discussing random numbers. The normal numbers introduced by E. Borel in the begining of the 20th century are the infinite sequences of symbols such that each block of symbols appears with the frequencies excepted for the independent and identically distributed sequence of random variables. I discussed in 1973 in [6] how close the normal numbers to the collective. By the modified definition by Church, an infinite sequence of symbols is called a collective if any subsequence chosen from it in a way that whether to choose an occurrence of symbol is determined by the value of a computable function at the sequence of symbols appeared before it as the input has the same frequencies of symbols as the original one. In fact, I consider all the infinite subsets of indices to be chosen for defining the subsequence, independent of what symbols appeared before, and characterize which subsets of indices preserve the frequncies of symbols for any normal number. In fact, this holds if and only if the set of indices has 0 entropy together with a positive density [6]. [7][8] are related papers. This result is discussed even now to extend to various classes of random numbers.

Algorithmic complexity defined by Kolmogorov and Chaitin is an effective mean to catch the randomness from the point of view of algorithm. I discussed it as a quantity of information in 1973 in [5], and proved that all large natural numbers have certain amount of information with respect to anything. This fact was called "oracle" and discussed later.

This Kolmogorov-Chaitin complexity is a theoretically perfect, single valued criterion of randomness for finite sequences of symbols, but it has two shortcomings. First is that it is not a computable function, second is that it has an ambiguity up to adding an arbitrary constant, so that it is not at all practical. In [64], I proposed a new criterion which is enough satisfactory both theoretically and practically. Further developement of this research is expected.

2. Ergodic Theory and Fractal

At the same time when I was intereseted in randomness, I was also interested the opposite problem what is regularity and periodicity. To find the solution, I started studying Ergodic Theory. The basic question asked in the 19th century phisics whether the space average in the dynamical system is realized as the time average or not hatched Ergodic Theory as a field of mathematics in the 20th century. The fundamental theorem in this field is the individual ergodic theorem. I gave a completely new proof of it using Nonstandard Analysis in [18] in 1982. It was shown there that any measure preserving transformation can be realized as one step of the rotation in a hyper-finite circle and the individual ergodic theorem is nothing but the associative law of the addition, $(a + b) + c = a + (b + c)$. This idea was used for to prove different kinds of ergodic theorems without using Nonstandard Analysis. [32] is one of them.

I started to be interested in Fractal functions from the late 1980's. I reformulated them as homogeneous cocycles on a compact metric space Ω with a continuous additive action of $\mathbb{R}$ together with a continuous multiplicative action of $\mathbb{R}_+$ satisfying the associative law $\lambda(\omega + t) = \lambda\omega + \lambda t$ $(\omega \in \Omega,\ t \in \mathbb{R},\ \lambda \in \mathbb{R}_+)$. I further assumed that the additive action to Ω is minimal and admits a unique invariant probability measure P. An additive cocycle on Ω is a continuous function $F : \Omega \times \mathbb{R} \to \mathbb{R}$ satisfying that $F(\omega, t+s) = F(\omega+t, s)$. This F is called α-homogeneous, where $0 < \alpha < 1$, if $F(\lambda\omega, \lambda t) = \lambda^{\alpha} F(\omega, t)$ holds. It follows that a nontrivial α-homogeneous additive cocycle $F(\omega, t)$ considered as a function of t, fixing $\omega \in \Omega$, is α-Hölder continuous, but not $(\alpha + \varepsilon)$-Hölder continuous. We also consider $F(\omega, t)$ as this as a stochastic process on the probability space (Ω, P). It turns out to be a process with stationary and α-self similar increments. I constructed such compact metric spaces Ω and homogeneous additive cocycles systematically on the weighted substitutions, and studied the fractal properties, ergodic theoretical properties and probability theoretical properties of them in the papers [24][26][28][29][36][42][49][50]. In [37], I studied the stochastic processes as this from the point view of prediction and proposed to apply it for predicting stock markets. Furthermore, I discussed the relation between discreteness of the spectrum of the additive action and the weighted substitution to be of Pisot type in [54].

3. Number Theoretical Functions and Symbolic Dynamics

During my visit to France from 1976 to 1977, I was interested in the function $f(n)$ of $n \in \mathbb{N}$ defined through the digits of r-adic representation of n and the symbolic dynamics associated to it. That is, the topological dynamics of the lift Ω of $f \in \mathbb{A}^{\mathbb{N}}$ with respect to the shift, where Ω is the closure of $\{T^n f \in \mathbb{A}^{\mathbb{N}};\ n \in \mathbb{N}\}$. I specially interested in the spectral properties of it. For an example, Thue-Morse sequence $f(n)$ is defined through the 2-adic representation of $n \in \mathbb{N}$ in the way that if the 2-adic representation contains an odd number of 1, then $f(n) = 1$, otherwise $f(n) = 0$. This sequence has been studied well from various point of views. From the spectral point of view, it is known that the symbolic dynamics associated to it has a partially continuous and singular spectrum. Also, in the case of Rudin-Shapiro sequence, where $f(n)$ is defined to be 1 if the number of the occurrence of the block 11 in the 2-adic representation of n is odd, otherwise 0. Rudin-Shapiro sequence is well known in Fourier analysis, and the spectrum of the symbolic dynamics associated to it is known to be partially Lebesgue. In the papers [10][11], we developed the spectral analysis of the symbolic dynamics associated to the functions $f(n)$ defined through the digits of n with respect to movable bases. Specially we succeeded in representing the continuous part of the spectrum as an infinite product converging in the weak sence of measure. I discussed the mutual singularity of the spectrals using these infinite products [25]. Also, I proved in [12] that the mutual singularity of the spectrals implies the disjointness of dynamical systems. Furthermore, in [17], we proved that the formal power series of $\sum_n f(n)x^n \in (\mathbb{Z}/r\mathbb{Z})[[x]]$ is algebraic and the real value $\sum_n f(n)(1/r)^n$ is transcendental for these function $f(n)$ of digits with some mild condition.

4. Maximal Pattern Complexity and Pattern Recognition

Let $\mathbb{A}$ be a finite set with $\#\mathbb{A} \geq 2$ and Σ be an infinite set. Denote by $\mathcal{F}_k(\Sigma)$ the family of subsets S of Σ with $\#S = k$, and denote $\mathcal{F}(\Sigma) = \cup_k \mathcal{F}_k(\Sigma)$. Given $S \in \mathcal{F}(\Sigma)$ and a nonempty subset Ω of $\mathbb{A}^\Sigma$. We denote $p_\Omega(S) = \#\Omega|_S$, which will be called the complexity of Ω at S. That is, $p_\Omega(S)$ is the number of distinct mappings $\omega|_S$, the restriction of $\omega \in \Omega$ to S. The maximal pattern complexity of Ω is the function $p^*_\Omega(k)$ of $k = 1, 2, \cdots$ defined as $p^*_\Omega(k) = \sup_{S \in \mathcal{F}_k(\Sigma)} p_\Omega(S)$.

The maximal pattern complexity is introduced in the papers [39][40] in 2002. At first, it was defined for an infinite sequence $\omega \in \mathbb{A}^\mathbb{N}$ as ,

$$p^*_\omega(k) = \sup_{\{s_1 < \cdots < s_k\} \subset \mathbb{N}} \#\{\omega(n+s_1)\cdots\omega(n+s_k) \in \mathbb{A}^k;\ n \in \mathbb{N}\}.$$

If Ω is the closure of $\{T^n\omega;\ n \in \mathbb{N}\}$, then this $p^*_\omega(k)$ coincides with the above $p^*_\Omega(k)$. It was proved in [39] that $p^*_\omega(k)$ is bounded in k if and only if ω is eventually periodic, and furthermore, these properties are equivalent to that $p^*_\omega(k) < 2k$ holds for some $k = 1, 2, \cdots$. Hence, $\omega \in \mathbb{A}^\mathbb{N}$ such that $p^*_\omega(k) = 2k$ for any $k = 1, 2, \cdots$ is an aperiodic sequence having the smallest complexity among aperiodic sequences. We studied such infinite sequences in detail. We also proved that if the dynamical system (Ω, T) havs a partially continuous spectrum under some invariant probability measure, then $p^*_\Omega(k)$ increases exponentially in k.

After 2006, I studied the maximal pattern complexity of a nonempty set $\Omega \subset \mathbb{A}^\Sigma$, sometimes with a general indice set Σ other than $\mathbb{N}$, collaborating

with Chinese mathematicians [41][43][45][47][48]. I developed this problem to the problem of pattern recognition in [55][59][60]. That is, considering $\mathbb{A}$ as the set of digital information possessed by each point in the space Σ, $\omega \in \mathbb{A}^\Sigma$ can be considered as a picture drawn on the space Σ. Hence, Ω is considered as a set of pictures. To distinguish the pictures in Ω, take a set $S \in \mathcal{F}_k(\Sigma)$ of sampling points and use the information $\omega|_S$ for $\omega \in \Omega$. The maximum number of pictures in Ω distinguished by the sets of of sampling points of size k is $p^*_\Omega(k)$. To know this number together with $S \in \mathcal{F}_k(\Sigma)$ which attains this maximum is an important problem from the point view of pattern recognition. We obtained the answer to this problem for some mathematical settings of Ω. On the other hand, about the limit as k tends to the infinity of the quantity of infomation per one sampling point, it is known [59] that
"there exists $h(\Omega) := \lim_{k\to\infty} \log p^*_\Omega(k)/k$, called entropy of Ω, and it takes only restricted values either $\log 1, \ \log 2, \ \cdots$, or $\log \#\mathbb{A}$."

In general, a mapping $\psi : \Omega \times \Sigma \to \mathbb{A}$ with infinite subsets Ω, Σ and a finite set $\mathbb{A}$ is called a duality mapping between Ω and Σ. For example, for an infinite set $\Omega \subset \mathbb{A}^\Sigma$ in the above, a duality mapping $\psi(\omega, \sigma) \in \mathbb{A}$ is defined as $\psi(\omega, \sigma) = \omega(\sigma)$. In this case, we can define the complexity of Σ under this duality mapping as well as the complexity of Ω in a way that $p_\Sigma(T) = \#\{\psi(\cdot, \sigma)|_T; \ \sigma \in \Sigma\}$, where $T \in \mathcal{F}(\Omega)$ and $\psi(\cdot, \sigma)$ implies the mapping $\omega \mapsto \psi(\omega, \sigma)$, and hence, $\psi(\cdot, \sigma) \in \mathbb{A}^\Omega$. In this case, $p^*_\Omega(k) = p^*_\Sigma(k)$does not hold in general, but about the above entropy, we have $h(\Omega) = h(\Sigma)$ [60].

5. Uniform Sets and Super-stationary Sets

A nonempty closed set $\Omega \subset \mathbb{A}^\Sigma$ is called a uniform set if the complexity $p_\Omega(S)$ of Ω at $S \in \mathcal{F}(\Sigma)$ depends only on $\#S$, and in this case, the complexity $p_\Omega(k) = p_\Omega(S)$ considered as the function of $\#S = k$ is called the uniform complexity of Ω. It is an ineresting problem to ask what function of k becomes a uniform complexity, which is discussed in [59] in detail.

To study this problem, an important fact is that any uniform complexity can be realized by a super-stationary set. Here, a nonempty closed set $\Theta \subset \mathbb{A}^\mathbb{N}$ is called a super-stationary set if for any infinite subset $\mathcal{N} = \{N_0 < N_1 < \cdots\}$ of $\mathbb{N}$, we have $\Theta[\mathcal{N}] = \Theta$, where for $\omega \in \mathbb{A}^\mathbb{N}$, we define $\omega[\mathcal{N}] \in \mathbb{A}^\mathbb{N}$ as $\omega[\mathcal{N}](n) = \omega(N_n)$ $(\forall n \in \mathbb{N})$ and $\Theta[\mathcal{N}] = \{\omega[\mathcal{N}]; \ \omega \in \Theta\}$. In fact, for any $\Omega \subset \mathbb{A}^\Sigma$, there exists an injection $\varphi : \mathbb{N} \to \Sigma$ such that $\Omega \circ \varphi \subset \mathbb{A}^\mathbb{N}$ is a super-stationary set [56]. On the other hand, we know some characterizations of super-stationary sets. One of them is that it is written as $\mathcal{P}(\Xi)$ with a finite set $\Xi \subset \mathbb{A}^+$ of prohibited words satisfying the condition (#) [56]. Here, $\mathbb{A}^+ = \cup_{k=1}^\infty \mathbb{A}^k$ and $\xi \in \mathbb{A}^k$ is said to be prohibited in $\omega \in \mathbb{A}^\mathbb{N}$ if $\omega(s_1) \cdots \omega(s_k) = \xi$ does not hold for any $\{s_1 < \cdots < s_k\} \subset \mathbb{N}$, and $\mathcal{P}(\Xi)$ is the set of $\omega \in \mathbb{A}^\mathbb{N}$ such that any word in $\xi \in \Xi$ is prohibited. Using this characterization, we can obtain an inductive formula for the uniform complexity $p_\Theta(k)$ of a super-stationary set Θ.

It is proved in [61] that any nonempty set $\Omega \subset \mathbb{A}^{\mathbb{N}}$, not necessarily a uniform set, contains super-stationary sets in some sense, which is called the super-stationary factors of Ω. If (Ω, T) is a symbolic dynamics and the shift T is considered as the unit time lapse, then the super-stationary factors represent properties of the system which is independent of time scaling. These properties are interesting since they are completely oposit to what have been interested in so far, that is, properties which are sensitive to time scaling like entropy.