Chapter 3

Coevolution of Tapes and Machines

3.1 Introduction

3.1.1 Evolution of Genetic Code

Biological code in molecular levels are called *the genetic code*. It carries genetic information which can be passed from a generation to a generation or from cell to cell for self-reproduction. The genetic code is composed of 64 codon triplets assigned to specific amino acids or stop commands (Table 3.1).

Table 3.1: Universal genetic code: Each of 61 codon triplet is assigned to a specific amino acid and the other 3 are stop codons that terminate the synthesis of a protein molecule. The great portion of biological systems have same code system with this universal code. But some exceptions are found. The abbreviations referring amino acids are the followings: Ala = Alanine, Arg = Arginine, Asn = Asparagine, Asp = Aspartic acid, Cys = Cysteine, Gln = Glutamine, Glu = Glutamic acid, Gly = Glycine, His = Histidine, Ile = Isoleucine, Leu = Leucine, Lys = Lysine, Met = Methionine, Phe = Phenylalanine, Pro = Proline, Ser = Serine, Thr = Threonine, Trp = Tryptophan, Tyr = Tyrosine, Val = Valine.

Codon	Amino acid	Codon	Amino acid	Codon	Amino acid	Codon	Amino acid
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	stop	UGA	stop
UUG	Leu	UCG	Ser	UAG	stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

The evolution of the genetic code is one of the major subjects in the origins of life. Numerous attempts have been made to make clear how the code had evolved (for example [21, 107, 65, 73, 105, 6]). Here we are interested in how sequences of codons are interpreted to proteins and how the proteins act to serve self-reproduction.

The code in Table 3.1 was thought to be universal. But exceptions were found in human mitochondria [5] and in nuclear code [62]. These facts suggest that the genetic code is not frozen from the early stage of life. Evolution of biological systems in their forms and functions is accompanied by evolution of the genetic codes. Biases of the GC content in genome are related with changes of the genetic code. In eubacteria, the variation of the mean GC content is relevant tob variations in phylogeny [47, 48]. It suggests that there exist directional mutation pressures. The directional mutation may be caused by internal factors such as errors in the reproduction of DNA [89]. (Osawa et al reviewed the evidences of evolution of genetic code and directional mutation pressure in [78].)

A code is to present ways to interpret symbols. The genetic code directs the way to synthesize proteins from DNAs. Because the genetic code is not established by somebody but autonomously organized, it is essential to discuss evolutionary paths of the code for self-reproduction as co-evolutionary processes of information sequences and their decoders.

3.1.2 Problems in Self-reproduction

Origin of life is often attributed to the emergence of self-reproductive properties. Therefore the genetic code must be organized to serve self-reproduction. John von Neumann first has proposed an automaton model for the self-reproduction problem [102]. In his abstract modeling, the fundamental problem in reproduction is that to copy something we first have to observe the object, however, the observation in some cases disturbs the object.

In addition to the problem of observation and copying, the self-referential problem known as Richard's problem occurs for self-copying. A self-reproducing automaton should interfere with itself for replication of itself. This generates a self-referential paradox [42]. To avoid the difficulty, von Neumann separated a machine from its description tape as well as proteins and RNAs/DNAs. In Neumann's model, he defined a tape as a pattern of stationary states. Without external disturbances, the replication scenario is perfect. But external noise causes error actions, replication can become unstable.

Another problem is caused in the case that many machines come to interact. A description tape can be read by distinct machines in different manners. Several different machines can be produced from a tape. A machine disturbing reproduction of other machines may be produce.

We propose the additional problem of self-replication, as is also stressed by Neumann himself.

It is evolvability, that is, how to evolve more complex or functional machines from simple self-replicating machines. Random update of automaton states brings mutation into Neumann's automaton model. If mutation occurs in a machine itself, a machine may change its function. But such mutation should not be copied onto the next generation. If mutation occurs in a tape, it will be copied onto the next generation.

How to encode machines on tapes and how to acquire stability and evolvability of them in ensemble of tapes and machines under noisy environment will be discussed here.

3.1.3 Difficulties for Evolution of Stable Autocatalytic Networks

Eigen and Schuster have tried to elucidate the origin of genetic code based on their hypercycle model [24]. But they didn't discuss evolution from simple autocatalytic systems to a higher order hypercycle. Here we study evolution of genetic code as evolution of stable autocatalytic networks of machines and their description tapes.

Several drawbacks of hypercycle model are known. In a hypercyclic loop all connectivity among pairs of replicator and enzyme must be exactly equal. All couplings in a hypercycle are restricted to homogeneous in this sense. Niesert, Harnash and Bresch has pointed out the following three processes breaking down an hypercycle by problems of stability [76]. (1) If a selfish replicator emerges by accident, the minimal replication loop takes over the whole loop. A selfish replicator makes an enzyme not to help the replication of the sequence at next position in the loop but to help the reproduction of the replicator itself. (2) There is a danger of short circuit catastrophe. A short cut loop can appear when a mutant catalyze a sequence farther than normal one. Loops with large and short loops sharing partial components cannot be sustained. The whole hypercycle always contracts to the shorter circuit. (3) A large hypercycle dilutes each component in the system. Therefore the large a hypercycle becomes, the more possible that concentrations of components go down to zero by fluctuation. A hypercycle collapses if a critical component is lost.

In order to overcome these drawbacks, Kauffman and Rasmussen have discussed several improvements. Kauffman has proposed possible countermeasures for above problems [61]. He stated that these problems could be evaded by introducing both activating and inhibiting coupling. Redundancy of connections will increase stability of a catalytic system. He also presented another plan. If an enzyme can catalyze many replicators at a time, the whole system can be more stable.

Rasmussen has conceived the emergence of autocatalytic network as self-organization in a random directed graph by introducing mutually catalyzing strand-pairs [83]. Although he envisioned the emergence of autocatalytic network, the following problem has been pointed out by Kauffman [61]: the larger the number of kinds of strand-pairs, the more difficult to find catalyzing pair in a graph, since each strand in Rasmussen's model only has a single pairing target for catalyzing.

In spite of their improvements, the essential weak point in Eigen and Schuster's hypercycle is inevitable. It is the homogeneous coupling. One of the purposes here is to study how large autocatalytic network with inhomogeneous couplings is stabilized.

In our study, machines and description tapes are introduced and they are influenced by random external noise. While a machine reads a tape, both probabilistic and deterministic change occurs on tape symbols. The probabilistic change is caused by external noise and is called a probabilistic mutation or a passive mutation. On the other hand, the deterministic change is caused by machine action and is named a deterministic mutation or an active mutation. We refer these two processes as mutations since a tape is rewritten into new tapes by machines in place of replication. As mentioned before, mutation induced by internal factors has been suggested in the study of real genetic code [89].

We see evolution of a perfect replicating network composed of a pair of a machine and a tape in low external noise regime. For high noise region, a complex autocatalytic network sustaining deterministic mutation evolves. Self-replication not as an individual but as a whole network now becomes important. In the other autocatalytic models [28, 24], reproduction only by machines or by tapes has been discussed. No machine can reproduce itself, however, without the coding tapes and vice versa. By considering both machines and tapes, we can discuss the evolution and selection of code in terms of evolution of an autocatalytic network.

3.1.4 Organization of This Chapter

Sections in this chapter are organized as follows. Our basic model is introduced in §3.2. Population dynamics of machines and description tapes are introduced with the notion of active and passive mutation. In §3.3, evolution of a network composed of machines and tapes is discussed. If external noise is low, a simple pair of self-copying machine and tape is formed, where the active mutation is suppressed. But in the realm of high external noise, a complex network can emerge by sustaining high active mutation. A mechanism of forming this complex but stable network is attributed to the formation of core structure. A network acquiring noise stability is defined as a “core network”. The core network is studied in §3.4. In §3.5, classification of core networks will be presented. A core network consists of many autocatalytic loops. In §3.6, we study embedded autocatalytic loops in core networks. A network which both machines and tapes forming catalytic loops is found to stabilize large and complex networks. §3.8 is devoted to discussion of our results. Some possible connections with real biological systems will be discussed.

3.2 Modeling

Our system consists of two different objects, tapes and machines. They are simple analogues of the polynucleotide and polypeptide or ribosome, respectively. A tape has a bit string of a circular form. A machine which is an automaton like a Turing machine consists of 3 different parts, a head, a tail and a transition table. Each head and tail is expressed by a 4 bit string, whose pattern will be compared with binary patterns of tapes. A transition table consists of 4 mappings; $(\sigma^M, \sigma^T) \to (\sigma'^M, \sigma'^T)$, where σ^M and σ^T represents a current binary state of machine and tape. A tape and machine state will change to (σ'^M, σ'^T), respectively depending on a current state of machine and tape.

Introducing an ensemble of tapes and machines, we carry out the machine-tape reaction process as follows.

3.2.1 Interaction of Machines and Tapes

A machine $\mathbf{M}_i$ can react with a tape $\mathbf{T}_j$ iff the tape $\mathbf{T}_j$ has a head $\mathbf{h}_i$ and tail $\mathbf{t}_i$ of the machine $\mathbf{M}_i$ in a different site.

Machine $\mathbf{M}_i$ reads a tape $\mathbf{T}_j$ from the first site of $\mathbf{h}_i$ to the first site of $\mathbf{t}_i$. And a new tape written according to the transition tables of $\mathbf{M}_i$ is produced. And then a new machine is produced from the new tape. The site from $\mathbf{h}_i$ to $\mathbf{t}_i$ will be called the reading frame. A half population of machine starts to read a tape with the internal state 1 and the other half does with the state 0. As the result, it generates a new set of machine $\mathbf{M}_k$ and tape $\mathbf{T}_l$ per each interaction. This process s written as a reaction equation,

$$\mathbf{M}_i + \mathbf{T}_j \to \mathbf{M}_i + \mathbf{T}_j + \mathbf{M}_k + \mathbf{T}_l. \tag{3.1}$$

During the read/write process, both probabilistic and deterministic changes on bits of tape pattern are assumed to occur. The probabilistic change is caused by external noise and is called a **passive mutation**. On the other hand, the deterministic change is caused by machine action, and is named an **active mutation**. We call it mutation since it does not replicate a tape but actively rewrites it. A rewritten tape can be taken as a mis-copy of the original tape. The rate of passive mutation denoted by μ_P is measured by the bit flip rate per bit. The active mutation per a reaction is measured by the rewriting rate per a length of reading frame when a machine

i reads a tape j. Namely, it is given by,

$$\mu_{Aij} = \frac{w}{L_{ij}}, \tag{3.2}$$

where a symbol w denotes the number of rewritten bits and L_{ij} denotes a length where machine i read a tape j. Average of active mutation rate is given by,

$$\langle \mu_A \rangle = \frac{\sum_{ij} \mu_{Aij} m_i t_j}{\sum_{ij} m_i t_j}, \tag{3.3}$$

where, m_i, t_i, is the population of machine and tape, respectively.

3.2.2 Translation of Tapes

Not only bits of a reading frame, but every bit of tape is repeatedly picked up to construct a new machine from a first site of the reading frame. If a length of a tape is not enough, the same bit is used for coding several different part of a new machine. In the present model, we use a fixed length of 7-bit tapes with 16-bit machines. A first 8 bits are mapped onto head and tail parts in order. The next 8 bits are mapped onto a transition table. In order to cover 16 bits by 7 bits, several bits are multiply used. This complicated mapping from bits of tape onto machine function is assumed to reflect the nonlinear mapping from one-dimensional DNA to 3-dimensional protein folding structure.

3.2.3 Population Dynamics

We presume two abstract spaces of the size N for tapes and machines. A total m machines and t tapes distribute over the respective space. By iterating the following procedures, we simulate the machine/tape reacting system.

1. Compute concentration of machines and tapes by dividing the population number of each object, denoted by m_i, t_i, respectively, by the capacity size N as following equations,

$$f_i^M = \frac{m_i}{N}, \tag{3.4}$$

$$f_j^T = \frac{t_j}{N}, \tag{3.5}$$

   where, f_i^M and f_j^T is concentration of i-th machine and j-th tape, respectively.

2. Make $c^{M/T}N$ numbers of new machines and tapes from the reaction of machine i and tape j. The rate of reaction f_{ij} is given by,

$$f_{ij} = \frac{c^{M/T} f_i^M f_j^T}{\sum_{k,l} f_k^M f_l^T}. \tag{3.6}$$

   if the i-th machine can read the j-th tape. Otherwise it takes zero value. The reaction coefficients $c^{M/T}$ take a positive constant $c^M (= c^T)$.

3. Remove d^M% of old machines and d^T% of old tapes.

4. Put the new machines and tapes back in each space. Hence the population of machine i and tape j of the next generation becomes,

$$m_i' = (1 - d^M)m_i + \sum_{k+j \to i} f_{kj} N, \tag{3.7}$$

$$t_j' = (1 - d^T)t_j + \sum_{k+i \to j} f_{ki} N. \tag{3.8}$$

   It should be noted here that each machine has its unique corresponding description tape but the inverse is not true. Generally each tape encodes several machines depending on which site the translation starts in.

5. Taking an integer part of the above population, we obtain the actual population of the next generation. The machine or tape whose concentration is lower than N^{-1} is removed from the population. Not to go beyond the total population size N, the reaction coefficient should satisfy the relation $c^{M/T} \leq d^{M/T}$.

3.2.4 Effect of External Noise

Out of $c^{M/T}N$ new tapes, a number of tapes as well as machines, are erroneously generated by external noise. It is assumed that the rate of error depends on the reading frame. Namely, the rate of error replication by external noise is given by,

$$\varepsilon = 1 - (1 - \mu_P)^L, \tag{3.9}$$

where the symbol L is the length of reading frame. We use Monte Carlo methods to get the mutant objects. At most $\varepsilon c^{M/T} N$ mutant populations are obtained by randomly flipping the bit within a reading frame.

3.2.5 Source to Start Searching on Tapes

Each tape has a source where an attached machine starts to search for the head and tail pattern. Starting from the site, patterns are searched for in the clockwise direction of a circular tape. When a head pattern is found, a tail pattern starts to be searched in the clockwise direction. The site of source of a tape can be updated randomly when the tape is newly generated after an extinction. Note that every translational invariant tape has the same source.

3.3 Destabilization of a Minimal Self-replicating Loop

About 10 randomly selected machines with 2 or 3 tapes are prepared as the initial configuration. A machine without description tape is unstable and smoothly removed from the system. Hence an initial configuration which does not include any description tapes of the initial machines will die out if there is no mutation process. External noise may produce description tapes by mistake.

Even without external noise, a machine can acquire its description tape by the other machines as a normal product. To sustain the tape, we have to have the description tape of the machine which generates the tape of the first machine. In order to make it stable, a successive reproducing process should form a closed loop; each machine on the loop reproduces a machine for the next position. (Refer Fig. 3.8-a).) Detailed discussion will be given in §3.6.

We will see how the replicating networks evolve by changing the amount of external noise. Examples of temporal evolution of population of machines and tapes are shown in Fig. 3.1.

By introducing a lower amount of external noise in a system, we see a minimal autocatalytic loop evolve. In this example, a machine $\mathbf{M_{1002}}$ reads a tape $\mathbf{T_1}$ to replicate the same tape and its own machine. The number attached to tapes and machines are hexadecimal number converted from its binary representation. Many initial configurations reach this minimal autocatalytic system for a lower noise regime.

A system with the minimal autocatalytic loop is said to be metastable since it remains stable after turning off external noise. However the minimal loop is destabilized by increasing external noise. Translation under external noise generates many machines, most of which rewrites existing tapes. Increasing of parasite machines destabilizes the original self-replicating loop. In Fig. 3.1-a), a parasitic machine $\mathbf{M_{1222}}$ invades the network with its description tape $\mathbf{T_{41}}$.

A greater variety of machines and tapes induces unstable oscillation in Fig. 3.1-b). An original self-replicating pair becomes unstable if too many parasitic machines attach to it. Population of each machine and tape show unstable oscillation in time. An oscillation of large amplitude is caused by the original self-replicating loop. Other oscillations are caused by parasitic machines and tapes.

This temporal oscillation spontaneously crashes by exhausting the self-replicating loop. The system restarts by having the minimal self-replicating loop, otherwise it become extinction.

The minimal loop shows zero active mutation as being depicted in Fig. 3.1-a). Periodically rising of active mutation rates are caused by the parasite machine $\mathbf{M_{1222}}$ and the tape $\mathbf{T_{41}}$.

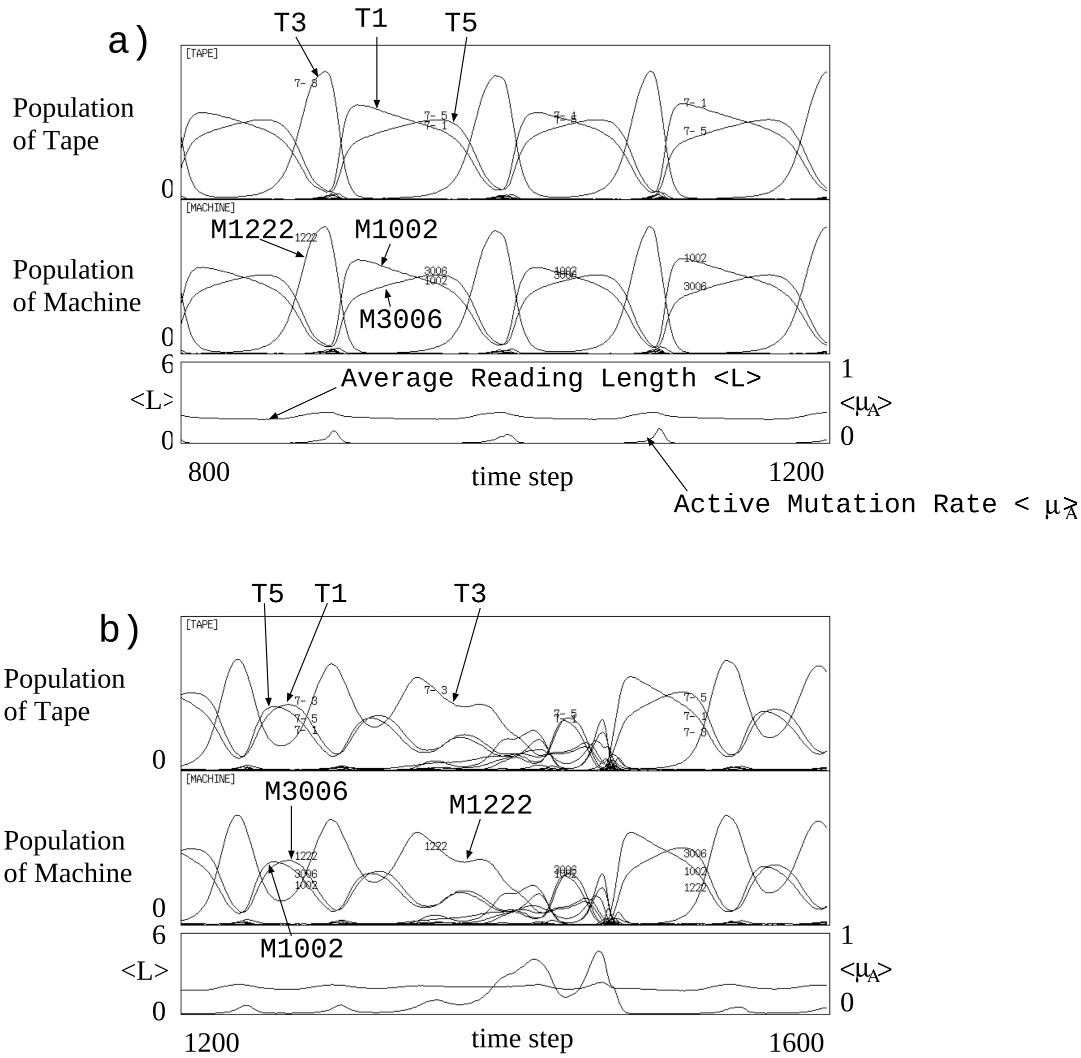


Figure 3.1: Temporal evolution of population of machines (the top row in both graphs) and tapes (the middle row). The bottom row displays temporal evolution of averaged active mutation rates and the averaged length of the reading frame. The parameters of population dynamics are $c^{M/T} = d^{M/T} = 0.6$ in the influence of, a) lower external noise ($\mu_P = 0.04$) and b) higher external noise ($\mu_P = 0.055$). Both start from the same initial states.

The active mutation rate intermittently bursts in time as shown in Fig. 3.1-b). It should be noted that a low rate of active mutations means that the network has more individual self-replication; each tape is self-replicated without mutation. But a high active mutation suggests that many machines read tapes and producing different machines with different tapes from the original ones. Namely, failure of self-replication is reflected in the amplitude of active errors.

3.4 Emergence of Core Network

In the region of high external noise ($\mu_p \geq 0.05$), a stable structure seems to be evolved. Unstable oscillation in population amplitude as we see in Fig. 3.1-b) is spontaneously stabilized around time step 600 in Fig. 3.2. At the same time, the active mutation rate is sustained at the high level.

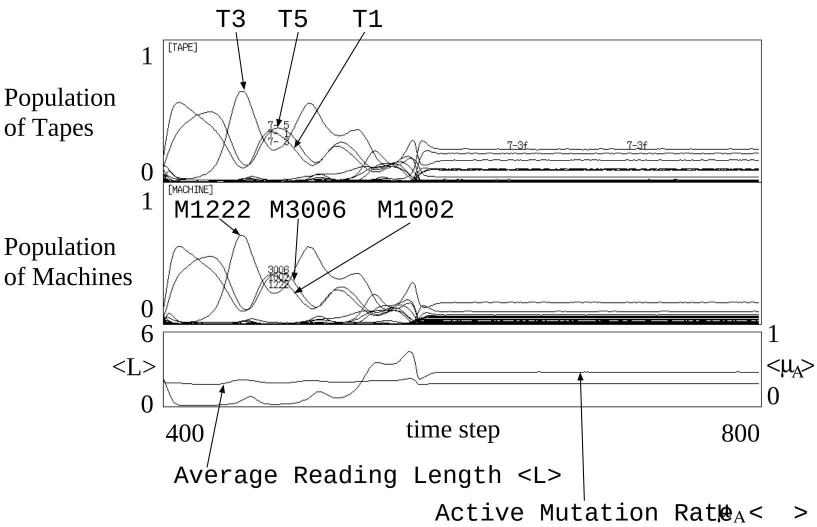


Figure 3.2: A spontaneous transition into a core network of a fixed point state. It displays temporal evolution of population of machines (the top row) and tapes (the middle row), and averaged active error rates and the averaged length of the reading frame (the bottom row). The rate of external noise is set at $\mu_P = 0.07$. The parameters of population dynamics are $c^{M/T} = d^{M/T} = 0.6$.

If we turn off external noise after time step 600, the numbers of machines and tapes remain stable. But if the noise is removed before time step 600, it will back to a minimal self-replicating loop. We call a network which acquired an implicit stable structure a core network.

Fig. 3.3 shows an example of temporal evolution of machines and tapes before and after a transition to a core network and after turning off external noise. After the transition the amplitude of total number of machines and tapes become less rugged than that of before the transition. Generally, machines disappear from the system when turning off the noise. Then a true core network is left in the system.

A minimal self-replicating loop (Fig. 3.1-a)) is also an example of core net. It cannot sustain

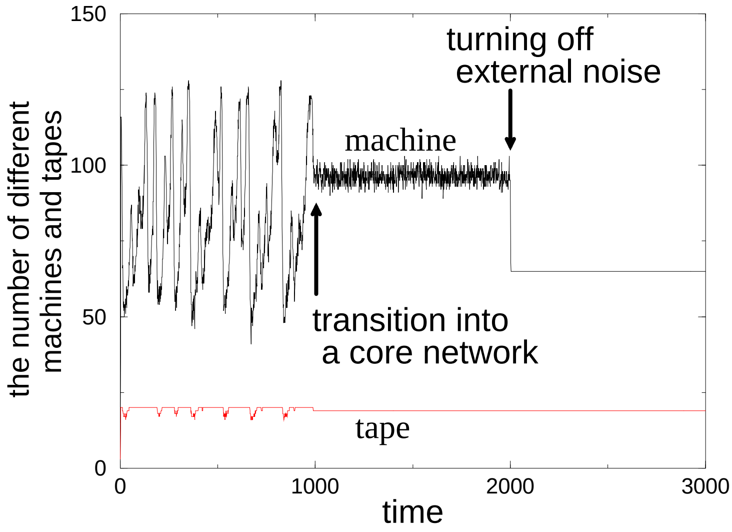


Figure 3.3: Temporal evolution of distinct numbers of machines and tapes. External noise(μ_P = 0.08) is turned off at time step 2000 after the emergence of a core network at time step 1100. The parameters of population dynamics are $c^{M/T} = d^{M/T} = 0.6$.

an active mutation rate by definition. All the examples of complex core networks which appear at middle external noise range ($0.05 \leq \mu_P < 0.1$) have high active mutation rates. By increasing the amount of external noise more, an attainable core net again loses complexity, becoming a minimal self-replicating loop.

In Fig. 3.4, we depict the number of distinct machines and tapes and an active mutation rate of core nets as a function of external noise, which are attained by turning off external noise at time step 2000. This diagram depends on the initial configuration of machines and tapes. Some initial configurations never attain any core networks. However the general tendency is that complex core networks emerge in he mid range of ($0.05 \leq \mu_P < 0.1$) of external noise. Namely, there exist upper and lower bound on external noise to evolve rich core networks. Core networks with a fixed point state are more attainable than ones with oscillating states. A core network at external noise of 0.05 contains a rather smaller number of machines than the other core networks. It has been found that this network uses more 0-rich tapes, which contain a bit of state 0 more than that of 1, in autocatalytic loops. Whereas the other core networks in a fixed point state use more 1-rich tapes, which contain a bit of state 1 more.

As exemplified at a range of $\mu_P = 0.6$ in Fig. 3.4, a core network is not necessarily a fixed

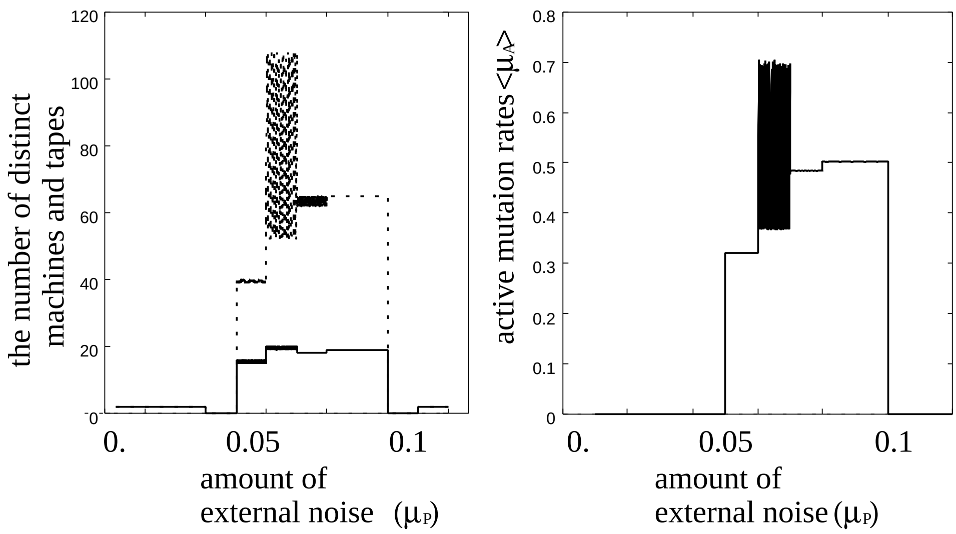


Figure 3.4: Temporal evolution of the distinct numbers of machines and tapes (left figure) and those of active mutation rates (right figure) during time step $[2500, 3000]$ are depicted as a function of external noise. The parameters of population dynamics are $c^{M/T} = d^{M/T} = 0.6$.

point state. It may start to oscillate after turning off the noise. We call it an oscillatory core network. Temporal evolution of the number of machines and tapes, and μ_A is described in Fig. 3.5. The topology of a network changes in time. Note that a network is called a fixed core net, if there is constant number of machines and tapes. Emergence of temporal oscillation of a network depends on the sustained network topology.

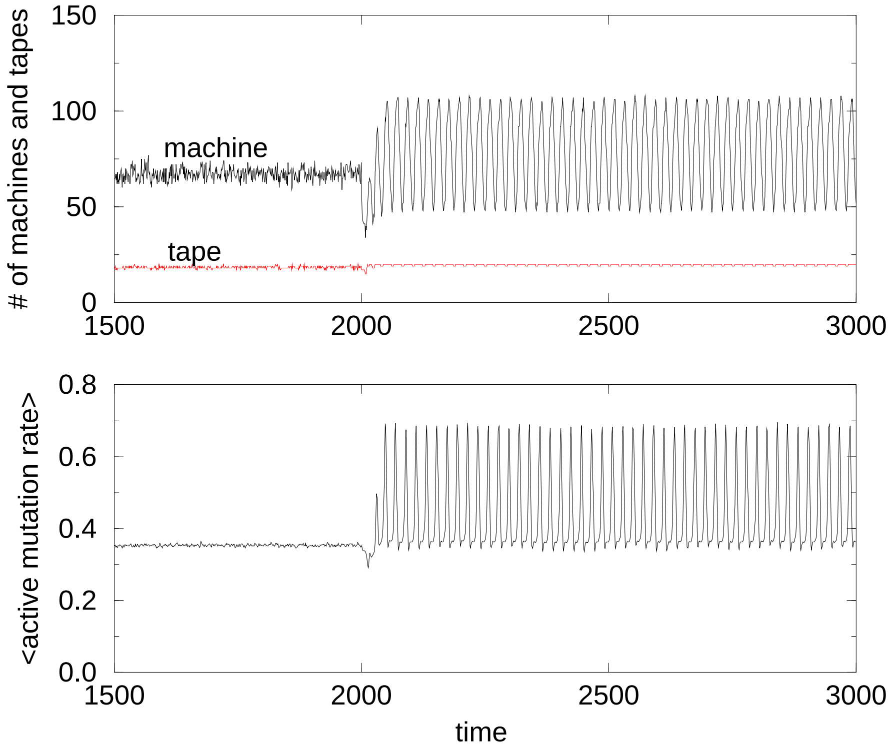


Figure 3.5: Temporal evolution of the distinct numbers of machines and tapes, and average of active mutation rate of oscillatory core. After turning off external noise ($\mu_P = 0.06$) at time step 2000, we observe oscillation of them. The parameters of population dynamics are $c^{M/T} = 0.6, d^{M/T} = 0.4$.

3.5 Taxonomy of Core Networks

As mentioned above, there are some attainable states of core networks depending on initial configurations of machines and tapes. They can be phenomenologically classified with respect to the dynamical states of active mutation rate.

Roughly, there are 3 distinguishable states – active mutation rate with zero, high constant and oscillating state. The zero active mutation rate state is a minimal self-replicating network which contains only a pair of a machine and a tape. There are two strong self-replicating machine-tape pairs, $\mathbf{M_{1002} - T_1}$ and $\mathbf{M_{2004} - T_1}$. If the latter type emerges, the system cannot develop a complex core network regardless of the amount of external noise.

There exist diverse kinds of network topologies in the core networks with high constant active mutation rate. They can be divided into two grades, one with lower number of machines ($\simeq 40$) and the other with higher number of machines ($\simeq 60$).

Many cores are in dynamically oscillatory state. These are fall into four subgroups: 1) Periodic oscillation of active mutation rate with varying its machine number ranging from 50 to 110. The period of oscillation is about $20 \sim 22$. 2) Quasi-periodic oscillation. The number of machines is same as that of group 1). 3) Amplitude of machine number is bounded between 80 and 110 with quasi-periodic oscillation. 4) Quasi periodic oscillation with a small amplitude. The average machine number is roughly 110.

These groups are also discriminated from return maps of the averaged active mutation rate. The typical figures are in Fig. 3.6. The width of amplitude are the same in Fig. 3.6-a) and -b). We can observe frequency locking in Fig. 3.6-a). The oscillatory core networks as depicted in Fig. 3.6-c) show higher and narrower oscillation in the average of active mutation rate than that of in Fig. 3.6-a) and -b). The oscillation of the number of machines is also high and narrow. the full number of tapes, on the other hand, are sustained in the networks of -c), 20 tapes exist all the time. Fig. 3.6-d) has also high and very narrow amplitudes of both the active mutation rate and the number of machines. The time series of the active mutation rate displays low frequency modulations and its power spectrum has many peaks.

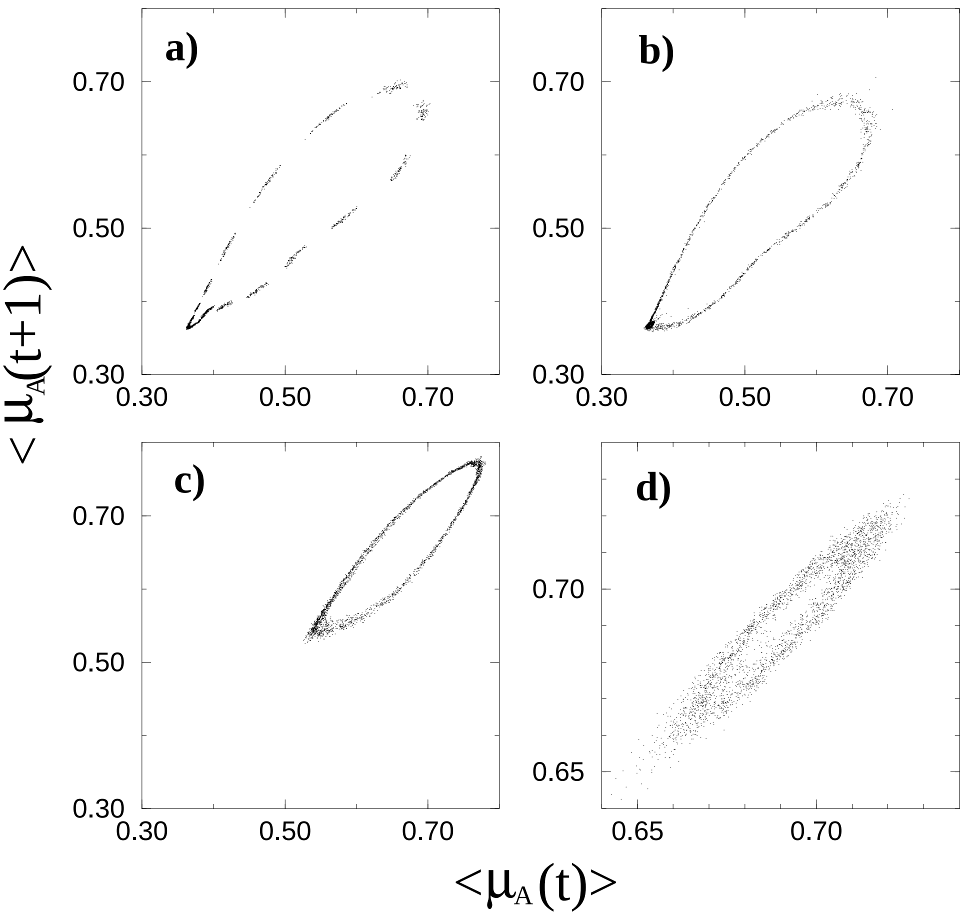


Figure 3.6: Return map of the averaged active mutation rate ($\langle \mu_A(t) \rangle$ v.s. $\langle \mu_A(t+1) \rangle$) for four distinguished oscillatory core networks. a) The widest amplitude, frequency locking. Period of main oscillation is around 21 or 22. b) the widest amplitude, c) middle amplitude, d) the narrowest amplitude, note that the range of this graph is different from the others. These are constructed from data taken after the cutting off the external noise. The rate of external noise is a) $\mu_P = 0.04$, b) $\mu_P = 0.04$, c) $\mu_P = 0.06$ and d) $\mu_P = 0.07$. The parameters of population dynamics are $c^{M/T} = 0.6$, $d^{M/T} = 0.4$ for all graphs.

3.6 Embedded Autocatalytic Loops

The embedded autocatalytic loops is an important criterion to distinguish network states. Since Eigen and Schuster's pioneering work [24], a notion of autocatalyticy has been known as a useful razor. We use this notion to dissect core networks.

Machines to be sustained in a system should possess their description tapes. Evolution from a simple self-replicating loop to complex ones are depicted in Fig. 3.7, where we express a system by actual machine-tape reaction graphs. In Fig. 3.7-a), a machine $\mathbf{M_{1002}}$ copies itself by reading its description tape $\mathbf{T_1}$. This minimal self-replicating loop exemplifies Eigen-Schuster's autocatalytic network type. Namely, a tape is self-replicated without error. In high noise regime, the minimal loop is gradually destabilized by side chains. Fig. 3.7-c) and -d) show a successive appearance of parasitic networks. A structure obtained by subtracting a network described in -c) from that in -d) gives a parasite network. This parasite network depends two input tapes ($\mathbf{T_3}$ and $\mathbf{T_5}$) on a network c). Similarly, a structure obtained by e) minus d) gives a hyper parasite, which depends a tape $\mathbf{T_b}$ on the first parasite network. Without a host network in -c), successive parasite networks will be extinguished.

A finally established core network is a combination of several loops which are autocatalytic. With respect to machine sets only, a closed loop is defined as a chain of k machines where a machine $\mathbf{M}_j$ generates a machine $\mathbf{M}_{j+1}$, the final machine $\mathbf{M}_k$ generating $\mathbf{M}_1$ as schematically displayed in Fig. 3.8-a).

A tape set attached to this machine loop can be divided into two sets. One is a tape set which need to sustain the machine loop, that is an input tape set $\{\mathbf{T}_j^{\mathrm{I}}\}$ and the other is a set produced by the machine loop, that is an output tape set $\{\mathbf{T}_j^{\mathrm{O}}\}$. Each machine reads at least one of the tapes in $\{\mathbf{T}_j^{\mathrm{I}}\}$ and generates one belonging to $\{\mathbf{T}_j^{\mathrm{O}}\}$.

A set $\{\mathbf{T}_j^{\mathrm{I}}\}$ is prerequisite for sustaining machine loops. If this set is automatically produced by the corresponding machine loop, it is called an autocatalytic loop. Namely, tape set in a network must be satisfy the following condition for the network to be independently autocatalytic,

$$\{\mathbf{T}_j^{\mathrm{I}}\} \subseteq \{\mathbf{T}_j^{\mathrm{O}}\} \ . \tag{3.10}$$

Otherwise a network has to use the products of other loops to compensate for necessary tapes. It is a necessary condition for any independent autocatalytic loops that a tape set needed to sustain the machine loop is self-reproduced by the machine loop.

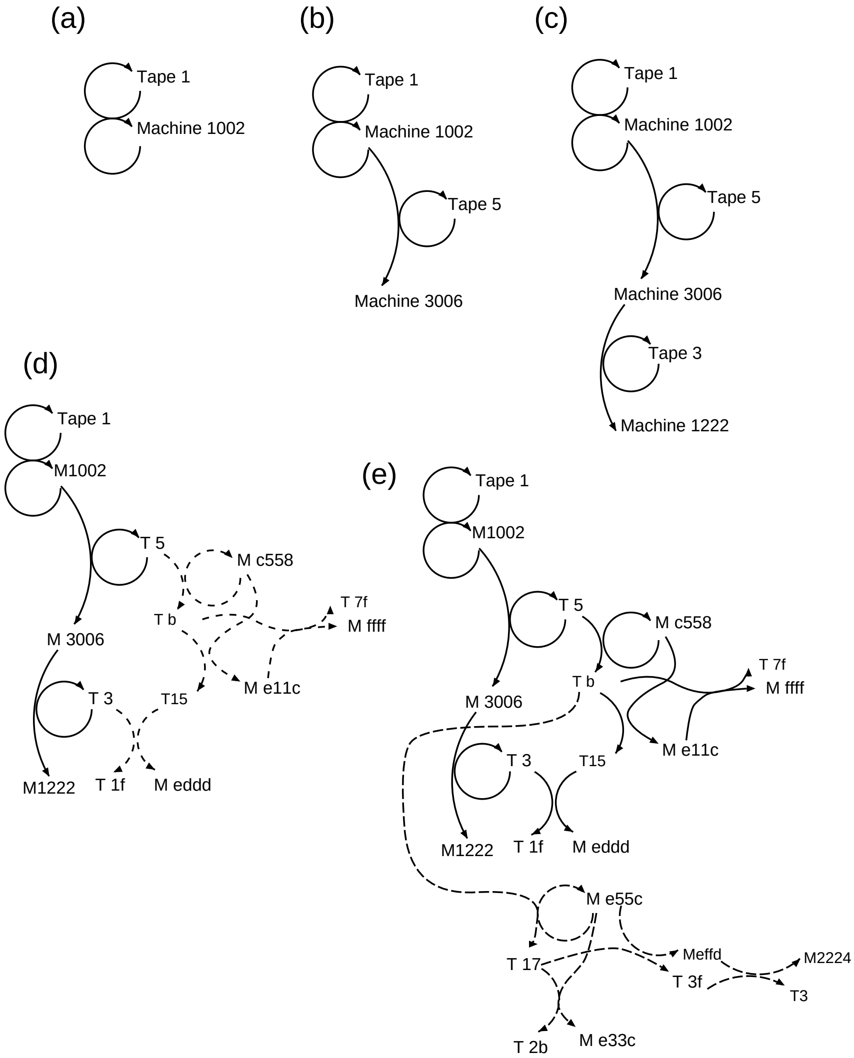


Figure 3.7: Evolution from a simple self-replicating loop to a complex network is described. a) An initially existing a minimal self-replicating loop, where a machine $\mathbf{M_{1002}}$ replicates itself by reading the tape $\mathbf{T_1}$. A net (a) will be successively exploited by $\mathbf{M_{3006}}$ with $\mathbf{T_5}$ (b)then by $\mathbf{M_{1222}}$ with $\mathbf{T_3}$ in (c). The net which is depicted on d) is a net c) with a parasitic network hanging on it. This is named a parasitic net, since as a network it depends on a net c) for the input tapes $\mathbf{T_5}$ and $\mathbf{T_3}$ to maintain its structure. Connections in the parasite network are drawn by broken arrow. In a network e), a hyper-parasite (drawn also by broken arrows) hanging on a parasite net in d) is depicted.

As special case, the autocatalytic condition is locally satisfied, hence each tape is self-replicated. A network of this type is called the Eigen-Schuster type (Fig. 3.8-b)). However, an autocatalytic condition is not locally satisfied in general. In the noisy environment, a network which globally satisfies the condition will be emerge (Fig. 3.8-c)).

In the general case of Fig. 3.8-c), not only machines but also tapes form a catalyzing loop structure. Hence we call the autocatalytic networks emerging in high noise regime a double loop network. In the low noise regime, a network can sustain its structure by Eigen-Schuster type, depicted in Fig. 3.8-b). For the high noise regime, in contrast, it is difficult to maintain prefect replication. It then switches to the double loop structure (Fig. 3.8-b)). Since Eigen-Schuster type only allows replication without error, it can be called a replicating system. On the other hand, a double network type locally changes tapes. It can be called an editing system, as tapes are used by being edited.

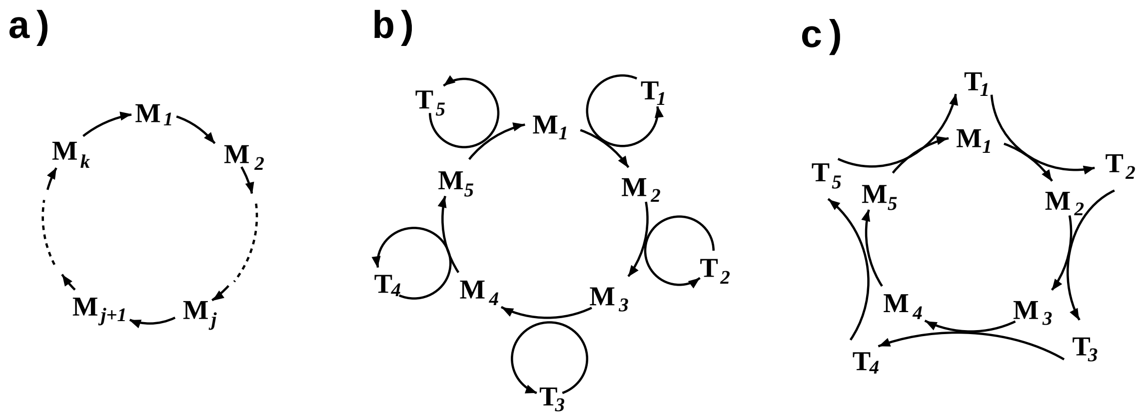


Figure 3.8: a) A loop of machines is schematically described. A machine produces a subsequent machine. b) and c) are illustrations of two different types of autocatalytic loops. They are Eigen-Schuster type b) and the double loop type c). In Eigen-Schuster type, tapes are replicated without mutations. In the double loop type, both each machine and tape are converted into subsequent one.

An example of core net in fixed point found in our simulation is depicted in Fig. 3.9. There are five double autocatalytic loops here. Each loop is extract in Fig. 3.10. Loops of Fig. 3.10-a) and b) belong to the Eigen-Schuster type and a loop of d) is the double loop type, which has no self-replicating tape at all. A loop of c) is also a double loop type but they have both replicating and editing reactions.

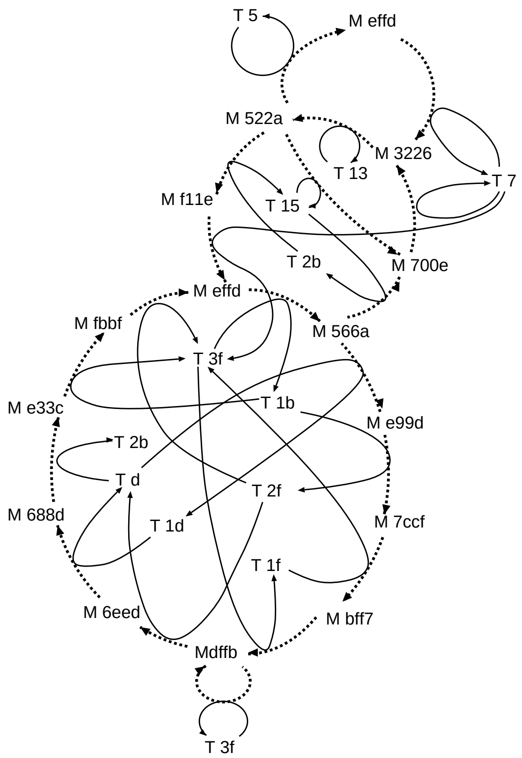


Figure 3.9: Embedded autocatalytic loops which are found in the core network after turning off external noise are depicted. We can see five independent autocatalytic loops here. Three are of the Eigen-Schuster type and the other two are of the double loop types.

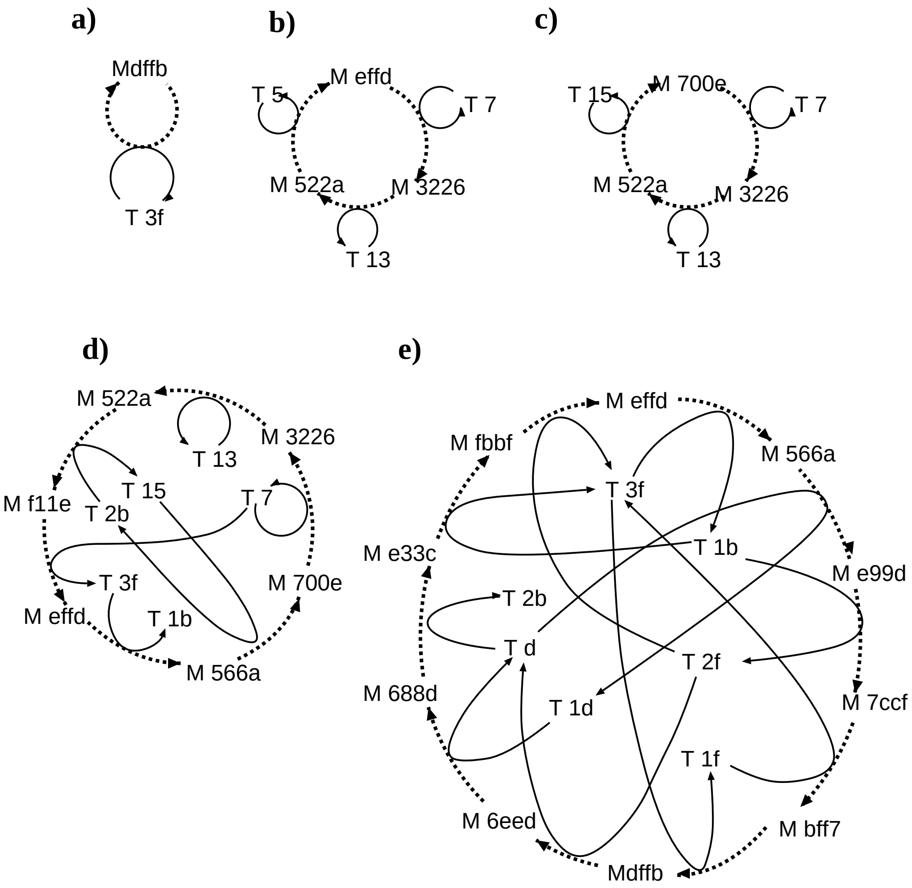


Figure 3.10: 5 loops in a fixed point core network in before figure are separately displayed. a), b) and c) are Eigen-Schuster type. There are both replicating and editing reaction in d) and e) consists of only editing connection.

Fig. 3.11-a) $\sim$ g) is examples of loops consisting of an oscillating core network. The oscillation is found to be the first group classified in §3.5, active mutation rate periodically oscillates, the number of machines ranges from 50 to 110. The network topology changes in time. After networks depicted in Fig. 3.11-g), very long and many autocatalytic networks appear. The largest length in the loops is about 60 and the maximum number of loops attains 200. Thereafter, it goes back to the same loops in Fig. 3.11-a). The period of oscillation is around 22 steps. Though not all networks have same forms, it can be said to have a recursive structure. The recursive structure is also found in the quasi-oscillatory cases of other types 2) $\sim$ 4) classified in §3.5.

a)

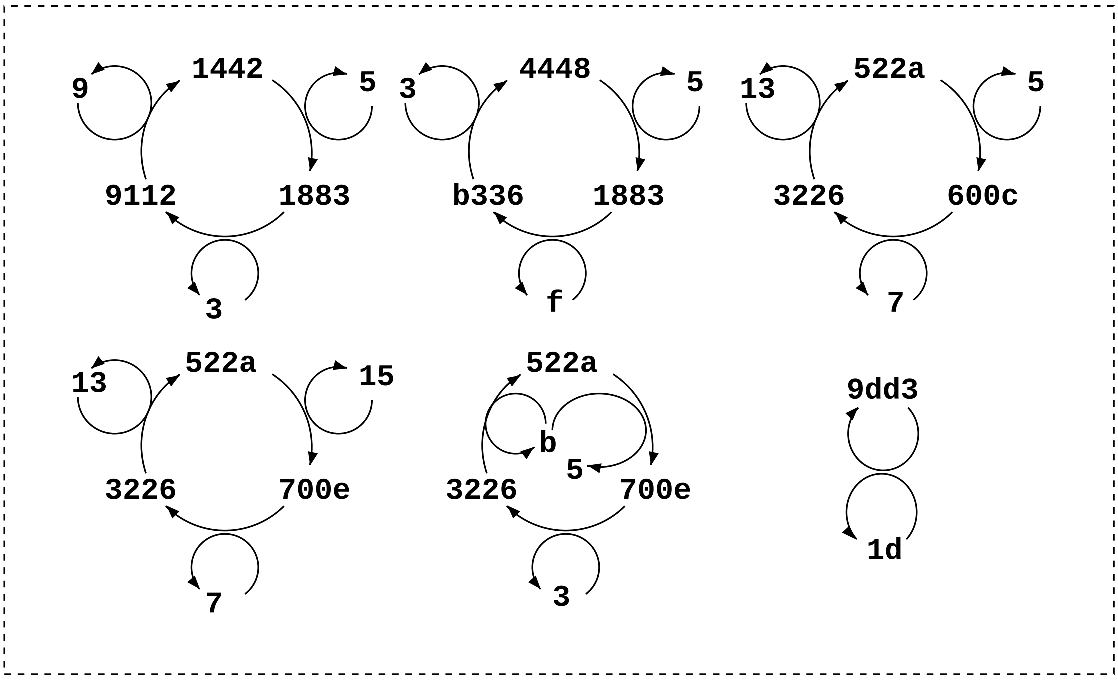


Figure 3.11: These figures are examples of embedded autocatalytic loops found in a oscillatory core network of first type. After g), very long and many loops appear. They come back to the same loops in a) after about 22 time steps.

We pick up in Fig. 3.11 loops of reaction paths which fulfill the autocatalytic condition defined by eq. (3.10). Note that networks depicted in Fig. 3.11 are not autocatalytic loops in a strict

b)

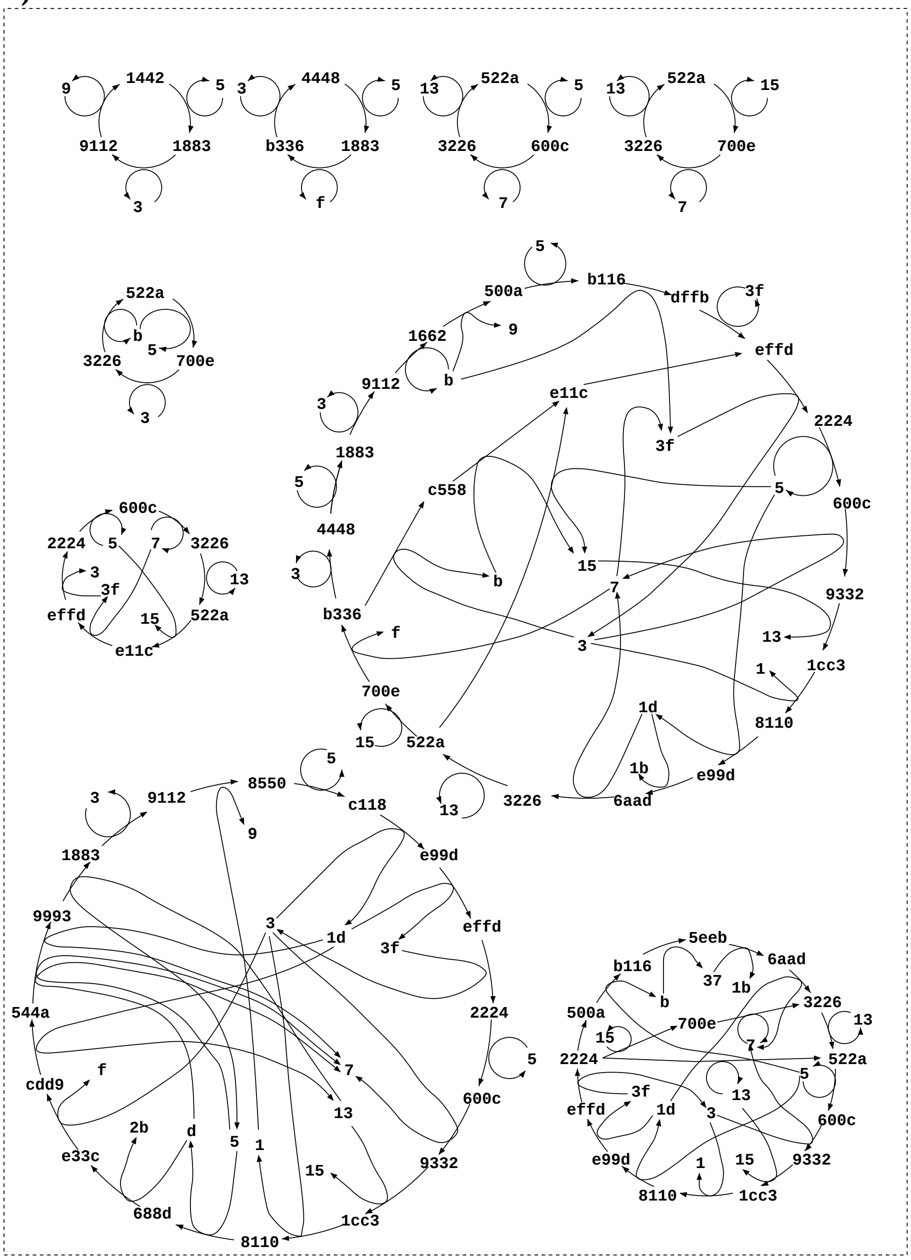


Figure 3.11: (continued)

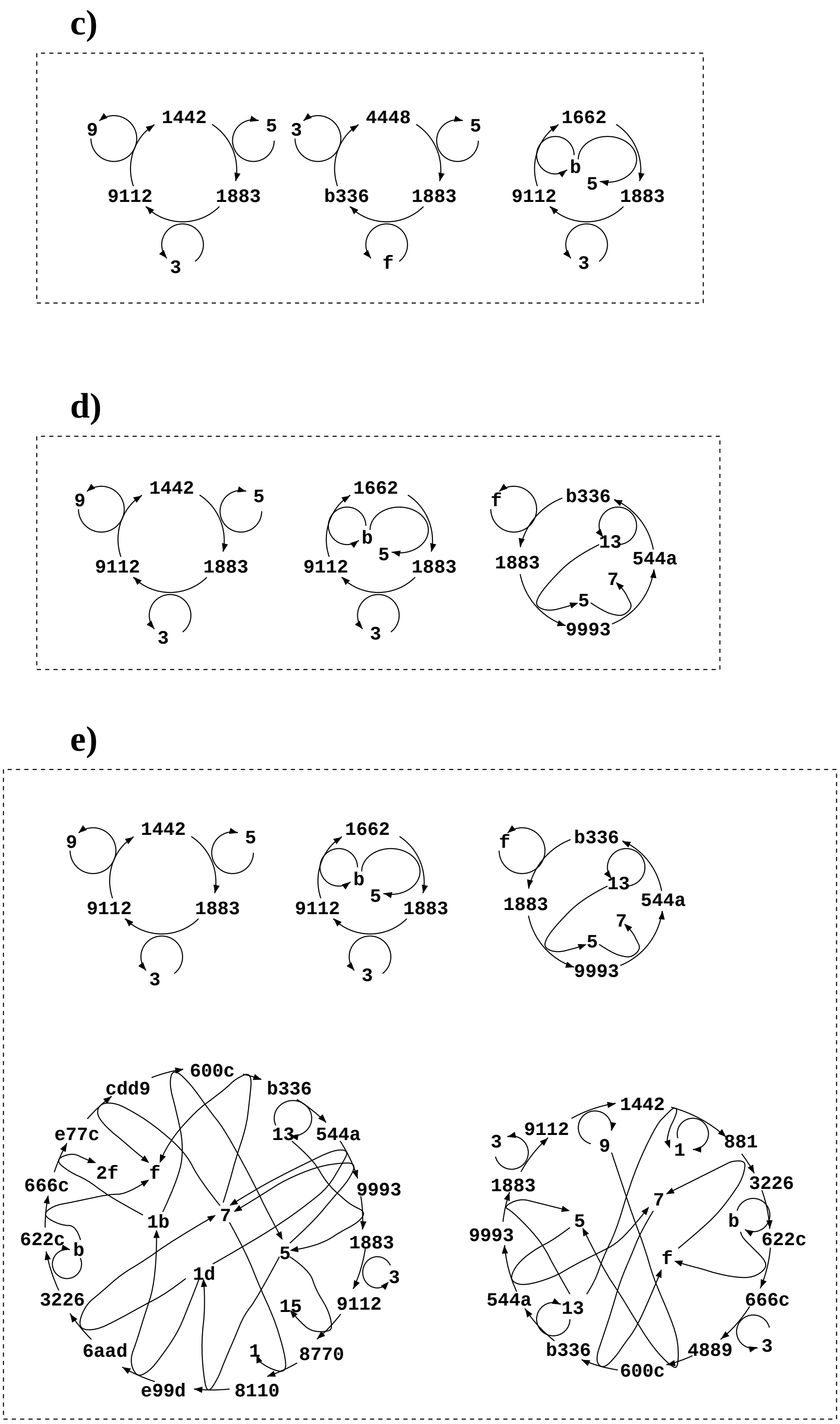
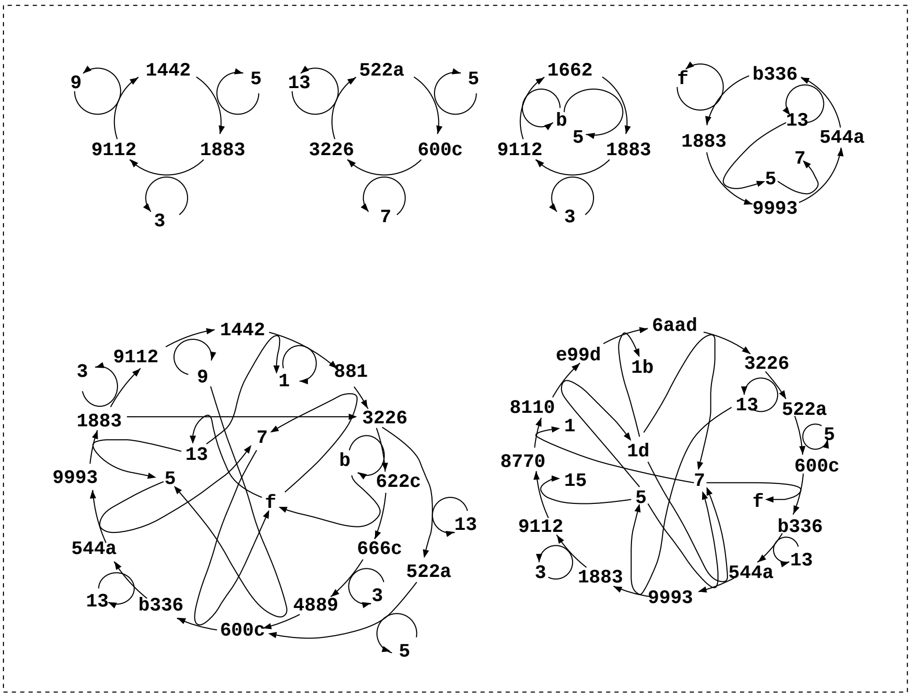


Figure 3.11: (continued)

f)



g)

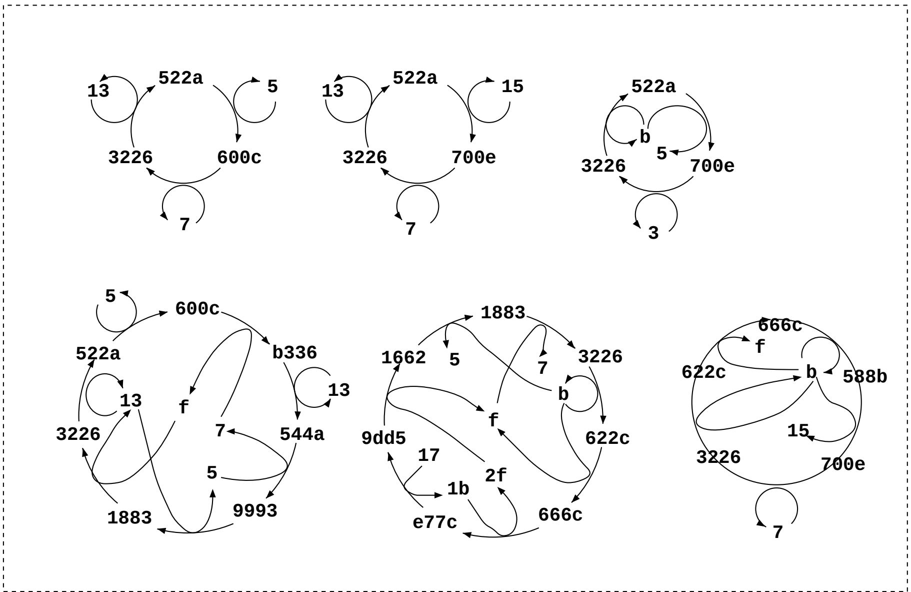


Figure 3.11: (continued)

sense, because they can not hold its structure stably.

Cores in oscillatory states change the number of autocatalytic loops. The number of Eigen-Schuster type autocatalytic loops shows less change than that of double loops type. In Fig. 3.12 we show temporal evolution of histogram of loop length.

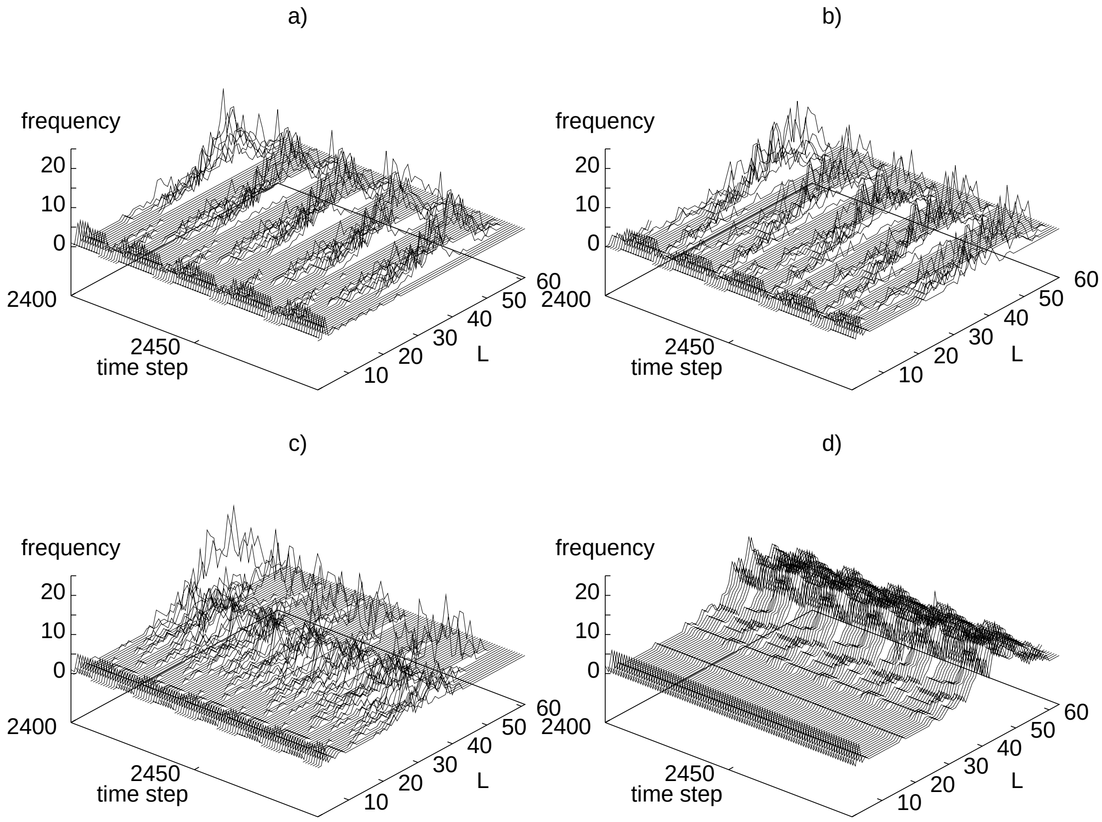


Figure 3.12: Histograms of lengths of autocatalytic loop change in the course of time as shown in these graphs. They are depicted in the interval between time step 2400 and 2500 after external noise is cut at time step 2000. They show relatively stable oscillations than immediately after cutting external noise. L is the length of a loop i.e. the number of machines in an autocatalytic loop. Four graphs are corresponding to four categories of oscillatory core networks described in Fig. 3.6.

The core nets shown in Fig. 3.12-a) and b) periodically switch their network structure from a network with few autocatalytic loops to one with many large loops. Fig. 3.12-c) shows many autocatalytic loops with length $L \simeq 30 \sim 40$ and periodic bursts to larger loops ($L = 45 \sim 61$).

Loop length around $L = 35$ grows when the burst occurs. These bursts are observed in a short period of time. Average of the loop length is around $L = 36$, which differs from the other networks, $L \simeq 42$. The core network depicted in Fig. 3.12-d) has two $L = 1$ loops, five $L = 3$ loops and a $L = 16$ loop for all time. The histogram of loop length shows small fluctuation. These facts suggest that the core network in Fig. 3.12-d) has a relatively stable structure similar to core network at fixed point state. Thus we observe that the return map of active mutation rate is torus with a small radius as described in Fig. 3.6-d).

Almost all core networks consists of autocatalytic loops with inhomogeneous couplings. Machines and tapes have the different input arrows from the output arrows apparently from Fig. 3.9. Besides, several autocatalytic loops shares part of their reaction paths. An example is Fig. 3.10-c) and d). They share the reaction paths $\mathbf{M_{700e}} \rightarrow \mathbf{M_{3226}} \rightarrow \mathbf{M_{522a}}$. Many cores are robust against perturbations that removes some tapes or machines from the network artificially. It is noted that we have introduced a machine to be able to create some different machines through interpretations of various tapes. A machine can catalyze several tapes. This redundancy contributes to complex and dynamical networks with sufficient stability.

Autocatalytic networks of double loop type are more stable than those of Eigen-Schuster type. For example, a loop with 16 machines of the double loop type is found in Fig. 3.12 for all time. In contrast, longer Eigen-Schuster type loops than three are not found. The stability of double loop network depends on the detailed topology of the network.

A tape can serve as templates of many different machines by being differently interpreted by different machines. But the core network evolves code, a set of interpretations. A core network selects some interpretations from diverse interpretations. We will describe the network from the viewpoint of “code” in the following section.

3.7 Diversity and Dynamics of Code

We assume an affinity between machines and tapes. The affinity is given by the condition that machines can read tapes if the head and tail patterns are found on tapes. Despite this affinity, there is a large arbitrariness by which machine a tape is read and as which machine the tape is interpreted. Selection of this interpretation ways constitutes a "code" of this system. A code of a network consists of interpretation rules as $(\mathbf{T}_i, \mathbf{M}_j) \rightarrow \mathbf{M}'_i$. For example, in the autocatalytic loop shown in Fig. 3.10-b), a machine $\mathbf{M_{effd}}$ reads a tape $\mathbf{T_7}$ and interprets it as a machine $\mathbf{M_{3226}}$. The interpretation rule for the process is $(\mathbf{T_7}, \mathbf{M_{effd}}) \rightarrow \mathbf{M_{3226}}$.

We depicted five autocatalytic loops in Fig. 3.10 consisting an example of fixed point core network. In Fig. 3.10-a), b) and c), there is a unique interpretation rule for each tape in the loops. And the loops b) and c) share their code, since interpretation ways for $\mathbf{T_7}$ and $\mathbf{T_{13}}$ is the same, $(\mathbf{T_7}, \mathbf{M_{effd}}) \rightarrow \mathbf{M_{3226}}$ and $(\mathbf{T_{13}}, \mathbf{M_{3226}}) \rightarrow \mathbf{M_{522a}}$. On the other hand, the other loops in Fig. 3.10 have more than two different ways to interpret a tape. A tape $\mathbf{T_7}$ in d), for example, is interpreted as a machine $\mathbf{M_{3226}}$ by $\mathbf{M_{700e}}$ and as $\mathbf{M_{effd}}$ by $\mathbf{M_{f11e}}$. The code in a core network composed of these five loops is shown in Table. 3.2. Both polysemy and synonymy are found. Polysemous tapes are $\mathbf{T_7}, \mathbf{T_{3f}}, \mathbf{T_d}, \mathbf{T_{2f}}$ and $\mathbf{T_{1b}}$. These tapes produce distinct machines through interpretations by distinct machines. The tape $\mathbf{T_7}$, for example, is interpreted as either $(\mathbf{T_7}, \mathbf{M_{effd}}) \rightarrow \mathbf{M_{3226}}$, $(\mathbf{T_7}, \mathbf{M_{700e}}) \rightarrow \mathbf{M_{3226}}$ or $(\mathbf{T_7}, \mathbf{M_{f11e}}) \rightarrow \mathbf{M_{effd}}$. Synonymous tapes produce one machine through interpretations by distinct machines. Tapes for $\mathbf{M_{effd}}$ is synonymous. It is produced from three different tapes, $\mathbf{T_7}, \mathbf{T_{2f}}$ and $\mathbf{T_5}$. We can also find that a tape is interpreted as one machine by distinct machines. For example, a tape $\mathbf{T_{15}}$ is interpreted as the same machine $\mathbf{M_{700e}}$ by the machines $\mathbf{M_{522a}}$ and $\mathbf{M_{566a}}$.

In minimal and fixed point core nets, the code cannot change even with large external noise. But for the oscillatory core network, the code can change temporally without noise. Table. 3.3 represents an example of temporal changes of coding system in an oscillatory core network. We depict interpretation ways only for tape $\mathbf{T_5}$ in autocatalytic loops of the oscillatory core net described in Fig. 3.11 in this table. There are 13 interpretation rules, 9 are distinct, in the period of time $t = 2408 \sim 2418$. A code at one time can not serve replication of the same core network. Constituents of a core net increase or decrease their concentrations by reactions according to the code at that time. But after a period, the same code reconstructed. For example, a machine $\mathbf{M_{4448}}$ exist at time steps $t = 2408 \sim 2412$. But by following the code at that time, the machine is not produced efficient amount. At last, it disappears from a core

Table 3.2: Example of code in a core network in fixed point state are shown in this table. This is a code in Fig. 3.9. A machine $\mathbf{M}$ reads a tape $\mathbf{T}$ and interprets it as a machine $\mathbf{M'}$ There are polysemous tapes, $\mathbf{T_7}, \mathbf{T_{3f}}, \mathbf{T_d}, \mathbf{T_{2f}}$ and $\mathbf{T_{1b}}$. These tapes are respectively interpreted as two different machines by two or three different machines. Synonymous tapes are also exist. $\mathbf{M_{effd}}$ is produced from three tapes, $\mathbf{T_7}, \mathbf{T_{2f}}$ and $\mathbf{T_5}$. A tape $\mathbf{T_{15}}$ is interpreted as the same machine $\mathbf{M_{700e}}$ by the machines $\mathbf{M_{522a}}$ and $\mathbf{M_{566a}}$.

T	M	M'	T	M	M'
7	effd	3226	3f	dffb	dffb
	700e			bff7	
	f11e	effd		effd	566a
15	522a	700e	d	688d	e33c
	566a			566a	e99d
2f	dffb	6eed	1b	e33c	fbbf
	fbbf	effd		e99d	7ccf
5	522a	effd	13	3226	522a
1d	6eed	688d	2b	522a	f11e

network at $t = 2413$. Its revival is at $t = 2418$. And the same interpretations with $t = 2408$ appear at $t = 2431$ again.

In general, diversity of code of core network in oscillatory state is higher than that of fixed point. It is also true for the length of autocatalytic loops. The diversity of code and length of an autocatalytic loop may have relevance to the stability of network. But it is not proved yet. Stability of autocatalytic loops can vary according to their topologies even if they have the same length.

Table 3.3: As an example of temporal changes of code, variable usages of tape $\mathbf{T_5}$ in the oscillatory core networks described in Fig. 3.11 are depicted in the period time step from 2408 to 2418. When a machine $\mathbf{M}$ reads a tape $\mathbf{T_5}$, it generate a machine $\mathbf{M'}$. In this example of core net, 13 interpretation mays of tape 5 appear. The tape is polysemously interpreted as 9 different machines. A machine can be produced by distinct machines. For example, a machine $\mathbf{M_{1883}}$ can be produced from machines $\mathbf{M_{1442}}, \mathbf{M_{4448}}$ or $\mathbf{M_{9993}}$.

time	2408 ~ 2410			2411			2412			2413		
	T	M	M’	T	M	M’	T	M	M’	T	M	M’
1)	5	1442	1883	5	1442	1883	5	1442	1883	5	1442	1883
2)	5	4448	1883	5	4448	1883	5	4448	1883			
3)				5	9993	1883						
4)												
5)				5	544a	9993						
6)	5	522a	600c	5	522a	600c						
7)				5	2224	600c						
8)				5	8550	c118						
9)				5	522a	e11c						
10)				5	8110	e99d						
11)				5	500a	b116						
12)												
13)												

time	2414			2415~2416			2417			2418		
	T	M	M’	T	M	M’	T	M	M’	T	M	M’
1)	5	1442	1883	5	1442	1883						
2)										5	4448	1883
3)												
4)										5	4448	9993
5)	5	544a	9993	5	544a	9993	5	544a	9993			
6)				5	522a	600c	5	522a	600c	5	522a	600c
7)												
8)										5	8550	c118
9)										5	522a	e11c
10)	5	8110	e99d	5	8110	e99d						
11)												
12)	5	9112	8770	5	9112	8770						
13)										5	c118	ebbd

3.8 Discussions

By introducing ensemble of machines and their description tapes, we have studied self-organization of a stable code for self-reproduction correlated with self-organization of an autocatalytic network. It is difficult to organize a stable code in general. An optimal coding of a machine may lead to the worse coding of the other. How to improve coding cannot be determined locally but should be determined in a network context. In a fixed core net, a stable code with polysemy and synonymy is organized. Alternation of coding driven by external noise has also been reported recently [44]. In their systems, a multi coding system, i.e. code with polysemous interpretations, is developed under a high noise environment. To make a stable code is to close logic consistently. It has succeeded in a core network in fixed point state. But for a oscillatory core network, its code dynamically changes. A code destabilizes itself and this results in oscillation of coding structure. We refer to such self-destabilizing nature of code as *openness of code*.

A minimal self-replicating loop, composed of one machine and one tape, emerges under influence of external noise. When external noise is elevated, a minimal self-replicating loop is exploited by parasitic networks. The populations of machines and tapes then begins to show unstable oscillation. In a realm of relatively high external noise, a system will evolve into a stable network against external noise spontaneously. When a network attains this state, it is stably sustained even after external noise is removed. We call the network acquiring such noise stability a "core network". The core network consists of many autocatalytic loops, being a fixed point or oscillatory state.

A transition from an unstable state to a core network state is similar to what has been seen in a host-parasite model [54, 58]. In the host parasite model, chaotic instability is shared by almost all species by sustaining a high mutation rate, leading to weak high-dimensional chaos, termed homeochaos. In the present model, roles of host and parasite emerge spontaneously in a network. Interaction between host and parasite loops causes dynamic instability as well. The core structure suppresses the instability and as the result a high active mutation rate is sustained in the present model.

We have introduced two different mutations – *passive and active mutation.* Taking the active mutation as a computation process, we propose a computational view of mutation and self-replication. A minimal self-replicating loop has no active mutation, that is no replication errors. A minimal self-reproduction loop can attain without any computation. If the computation process offers costs, the minimal loop may be the most economic self-replication system. Such autocatalytic network is stable only below some noise level. A large fixed core network, on the other hand, sustains a high active mutation rate in a high external noise regime. It absorbs external noise as active mutation of machine function. Namely, passive mutation caused by external noise is replaced as deterministic mutation by doing computation. The network carries and processes large amount of information in a core network.

Several autocatalytic loops are embedded in a core network. These loops are divided into two types. One has no active mutation. Tapes are replicated locally. It corresponds to the Eigen-Schuster's hypercycle. We call this type of network Eigen-Schuster (ES) type autocatalytic network. The other has a high active mutation rate. It consists of double loops of machines and tapes. Namely, each tape as well as machine produces a successive tape. Both entities are replicated as a whole, not individual. We call this type of network double loop type autocatalytic network. We also have mentioned that this type of network can be regarded as a editing system. Since tapes in the network are rewritten to templates of different machines, in contrast to tapes in ES type network, which are replicated without errors.

It is thought that DNA is replicated individually and RNA is a copy of DNA in real biological systems. However, an example [38] may be related to our double loop autocatalytic network. The authors of [38] has shown that DNA of macronucleus has been generated from DNA of micronuclear in Oxytricha nova. DNA of micronuclear is transcribed once into RNA. Then the exons are completely rearranged and reversely transcribed into the DNA of macronucleus. If editing and reversely transcribing RNA is more stable than replicating DNA itself, an active mutation as editing will be favored in early genetic systems.

It is said that biological system has evolved from RNA world to DNA world through RNP world [36, 35, 109]. RNA world consists of RNA molecules having catalytic functions as well as replicating function. Then it has evolved to the world including RNAs and proteins (RNP world). Even in such RNP world, self-reproduction is still unstable. Editing of RNA must repeatedly occur. By using more stable polynucleotide, DNA world has evolved. RNA, protein and DNA cooperatively work in the DNA world as in the present biological systems.

Our modeling can correspond to the second level of real evolution (RNP world). Differentiation of ES type and double loop type in autocatalytic networks can be taken as a transition from RNP to DNA world.

Autocatalytic loops in our model overcome the difficulties asserted by Niesert et al. [76]. They have discussed that an autocatalytic network become unstable either by selfish replicator, short circuit or zero concentrations (see §3.1.3). And the network is destabilized by inhomogeneous couplings. The inhomogeneous coupling is that a tape catalyzes more (less) tapes than being catalyzed and the similar asymmetry exists for machine reactions. In our core network, autocatalytic loops with different length, even minimal loop, can coexist. Oscillatory core nets does not break down by the zero concentrations of component, which is Niesert's third catastrophe. A similar topology is recovered after a period of time. The inhomogeneous couplings are also permitted. Polysemy in a tape, encoding several machines, are important to resolve there difficulties. As Kauffman has discussed [61], redundancy caused by non-specific replication makes an autocatalytic network robust.

The state where no subcomponents of the system are themselves self-replicating has been referred to as a collective reproducing state by Kauffman [61]. The emergence of such ordered state in his autocatalytic polymer model could also escape Niesert's catastrophes. But the inhomogeneous couplings were not able to be managed. There was not a gene like substance in his system. However, reproduction and evolution of unstable entity requires a stable description of itself.

Fontana's "algorithmic chemistry" [30,31] has shown evolution of replicating programs similar to our work. His chemical world consists of strings of variables and operators called functions. Reactions among functions are governed by λ-calculus. A system is dominated by single self-copying strings in a level-0 stage, which corresponds to our minimal self-reproduction loop. By forbidding self-copying function, a mutually catalytic set emerges (level-1 stage). This level-1 stage corresponds to the fixed point core network in our model. In spite of no explicit prohibition of self copying, our system spontaneously attains mutual catalyticy under noisy environment.

In Fontana's next level (level-2 stage), interconverting autocatalytic sets can be combined. This level corresponds to our advanced model [51], where tapes and machines are wrapped within a cell. In an ensemble of the cellular systems, evolution to complex self-reproduction is caused not by external noise buy by influx machines from other cells. Several translations are performed by external machines from other cells. This cooperative behavior brings about diversification of core networks.