Chaotic Itinerancy

in

Coupled Dynamical Recognizers

Takashi Ikegami and Gentaro Morimoto
Department of General Systems Sciences,
The Graduate School of Arts and Sciences,
University of Tokyo, 3-8-1 Komaba, Tokyo 153-8902, Japan

Abstract

We argue that chaotic itinerancy in interaction between humans originates in the fluctuation of predictions provided by the non-convergent nature of learning dynamics. A simple simulation model called the coupled dynamical recognizer is proposed to study this phenomenon. Daily cognitive phenomena provide many examples of chaotic itinerancy, such as turn taking in conversation. It is therefore an interesting problem to bridge two chaotic itinerant phenomena. A clue to solving this is the fluctuation of prediction, which can be translated as ‘hot prediction’ in the context of cognitive theory. Hot prediction is simply defined as a prediction based on an unstable model. If this approach is correct, the present simulation will reveal some dynamic characteristics of cognitive interactions.

PACS numbers: 05.45.Pq,84.35.+i,05.45.-a
In this paper, we re-discuss chaotic itinerant phenomena in terms of cognitive phenomena. We interpret chaotic itinerancy as a shifting dynamic process that continually creates and destroys a temporal structure in a deterministic manner.

To study these phenomena in relation to cognition, we propose a coupled dynamical recognizer (CDR) framework and introduce a co-learning concept. As an example here, we simulate a three-person coalition game. In the game, each player will show a red or blue card simultaneously. If two players have the same card, they form a coalition and the third person gets nothing. When everyone shows the same card, no one wins. By repeating this game, we observe an itinerant phenomenon of exchanging coalition pairs from time to time. Regarding this phenomenon as chaotic itinerancy, we analyse the behaviour. It is interesting to note that if one player is replaced with a finite machine, the itinerant phenomena will be suppressed. Coupled unstable learning dynamics provide the source of the itinerancy. Introducing other simulations by CDR, we consider the empirical studies of chaotic itinerancy. The source of the itinerant phenomena in simulations and psychological data is attributed to fluctuation of predictions, i.e., hot prediction. In the discussion, we insist that communications between co-learning systems can generate chaotic itinerancy with hot prediction.

I. INTRODUCTION

Vygotzky’s concept of ‘momentary child’, in contrast to Piaget’s ‘eternal child’, expresses well what kind of theory is required in cognitive science in general. A momentary child means an ever-changing rather than a static and stable nature of child developmental process. However, most of our theoretical tools can only deal with limiting attractor states. Chaotic itinerancy is one of the few ideas that can be used to characterize transient dynamics. We present here a new perspective of chaotic itinerancy in terms of cognitive behaviour.

Chaotic itinerant phenomena have been reported independently in several dynamic models [1–3]. Chaotic itinerancy (CI) has the following properties: 1) the existence of quasi-stable attractors; 2) jumping between quasi-stable states with some intrinsic dynamics; and 3) temporal changes in the effective dimensionality of a system.

The notion of CI has been claimed as important in a single brain system both theoretically
[4], and empirically [5]. It is also true in a coupled brain system. In this paper, we put stress on the importance of studying more than a single brain system. In other words, a multi-mind synthetic psychology is proposed. Because a brain can be assumed to be a device for communicating with other brains, we attempt here to reveal the importance of CI for communicative functions. CI in communication can thus be related to habituation and dis-habituation phenomena.

Habituation and dis-habituation are good descriptors of children’s play behaviour. Children tend to get habituated quickly for expected events but relatively slowly for unexpected ones. This is used as an objective index for testing the innate capability of infants (i.e., the expectation violation experiment [6]). When habituated, children change their rules of play or the kind of play itself. These habituation/dis-habituation cycles repeat themselves. What are the underlying dynamics? In this paper, we discuss the idea that the same mechanism is responsible for duality, which can be found in gaming communicative systems in general.

In §2, we introduce a coupled dynamical recognizer (CDR) model; in §3, a co-learning model with CDR is presented; and §4-6 describe the coalition model and simulation results. In §7, we briefly consider the simulation results of other CDR models; and in §8, we discuss some psychological phenomena in terms of chaotic itinerancy. The current approach is developed into a possible theory of developmental psychology and language. The paper concludes in §9 with a discussion of new perspectives in communication theory.

II. COUPLED DYNAMICAL RECOGNIZER

Neural networks with recurrent interactions have been used to imitate human language performance [7], to mimic finite automaton behaviour [8], and to manipulate robot navigation [9]. Extending Pollack’s terminology, we term this a dynamical recognizer in the sense that a cognitive agent should be a dynamical agent, i.e., a momentary agent that is always changing its internal state. As Pollack first showed explicitly, a dynamical recognizer can imitate the behaviour of a finite automaton. To see how well a dynamical recognizer learns to imitate any given automaton, we examine geometrical patterns in the ‘context space’, i.e., a plot of output context node states against a series of random inputs. If a dynamical recognizer can successfully imitate the given finite automaton, the context space plotting will show finite islands of clusters, and a clear correspondence is observed between each cluster
and each node of the finite automaton. When the dynamical recognizer fails to imitate or
the opponent is not a finite automaton, the context space plotting shows a stretched and
folded, fractal-like structure.

We propose a coupled dynamical recognizer to simulate two agents communicating with
each other. Each agent generates an internal model of the other agent that best imitates
the other’s past behaviour pattern. Based on the best internal model, players choose their
next action by predicting their future behaviour outcomes.

We have applied our coupled dynamical recognizer model to two-person games. In this
paper, we introduce a three-person coalition game to discuss CI in a co-learning game. We
describe the common model framework below.

III. CO-LEARNING SYSTEM

Each agent tries to make a model of the other; however, the model of the first agent
generated by the second one is dependent on the model of the second one being generated
by the first one. Such a nested structure produces the complexity of the co-learning system.
Usually, a model generated by an agent cannot last and is taken over by other models. Yet
there exists no true solution in such a co-learning system, it often happens that agents are
trapped by a fixed pair of false models.

The model is generated as follows. The dynamical recognizer is used to mimic the op-
ponent’s behaviour. In each turn (game round), both players 1) generate a model of the
opponent that best mimics the opponent’s past moves; 2) anticipate the opponent’s future
moves, and 3) choose as their next move an action that is expected to induce the best future
outcomes for the game situation. These three steps are applied repeatedly in each game
round. We will explain those three steps in detail.

1) The past behaviour of the opponent will be mimicked by a dynamical recognizer.

We have input neurons \( u_0, u_1 \), output neurons \( z_0 \) and recurrent neurons \( z_1, z_2 \). A
function network is defined as a mapping process from the input to the output neurons. The
function network is parameterized via recurrent neural states, and its temporal dynamics at
each time step \( n \) are described as follows.


\[
x^{0}(t) = \begin{cases} 
u_{i}(t) & (i = 0, 1) \\ z_{i-1}(t-1) & (i = 2, 3) \end{cases} \\
x^{1}(t) = g\left(\sum_{j=0}^{3} w_{ij}^{0}x_{j}^{0} + b_{i}^{0}\right) (i = 0, 1, 2, 3) \\
z_{i}(t) = x^{2}(t) = g\left(\sum_{j=0}^{3} w_{ij}^{1}x_{j}^{1} + b_{i}^{1}\right) (i = 0, 1, 2) 
\]

where \(w_{ij}^{k}\) and \(b_{i}^{k}\) denote the weight and bias of the network and \(x_{i}^{k}\) are the activations of hidden neurons.

In the equations above, non-linearity exists only in the sigmoid function \(g(x) = (e^{-\beta x} + 1)^{-1}\).

We study a coalition game in this paper. This game has three players; that is, each player should have two models corresponding to the other two players. The game has two actions and the model takes two actions from each of the other players and has one action output. Therefore, a network structure with one output neuron and two input neurons is required.

The output is rounded off to 0 and 1, corresponding to two possible actions.

2) The degree to which the obtained dynamical recognizer can imitate the real opponent is measured by the sum of past behaviour patterns. In practice, the error \(E(n)\) after the \(n\)-th game is computed by:

\[
E(n) = \sum_{k=1}^{n} (z_{0}(k) - d(k))^{2}, 
\]

where \(d(k)\) is the actual opponent’s action in the \(k\)-th game, and \(z_{0}(k)\) is the action predicted by the network. We update the weight values in the direction of the gradient of the energy function in a learning phase:

\[
\Delta w_{ij}(n) = -\alpha \frac{\partial E(n)}{\partial w_{ij}} + \beta \Delta w_{ij}(n-1). 
\]

The updating method is called Back Propagation Through Time (BPTT), an extended back propagation technique for this recurrent neural architecture.

3) Using the model of the opponent, each player anticipates any future actions. In practice, by giving all possible combinations of actions against the model over the next \(T\) rounds (i.e., \(x(t+1), x(t+2) \cdots x(t+T)\)), we can compute the expected reward \(U\) averaged over each future combination thus:
FIG. 1: In a coalition game, each player has to construct two models. Each player must evaluate each future action against both models to optimize future outcomes.

\[ < U(\hat{x}, \hat{y}) > = \sum_{n=1}^{T} U(\hat{x}(t + n), \hat{y}(t + n)) \omega^n, \]  

where the possible combinations are denoted by \( \hat{x} \) and the expected opponent's actions associated with \( \hat{x} \) are denoted by \( \hat{y} \). The outputs \( \hat{y}(t) \) are computed for each time step \( n \) by the model obtained in the time step \( t \).

4) Each player chooses the action combination that is expected to bring the highest reward to it and takes the first element of that action pattern as the next action.

By sequentially repeating this cycle, both agents continually update their internal models of the other players and proceed thereby. A schematic figure of the above cycle is depicted in Fig. 1.

IV. THE COALITION GAME

A minimal form of society begins when we have at least three persons. In a two-person game, co-operation can only be present or absent between the two persons. However, in a three-person situation, co-operation between two persons may develop into a coalition
against the third person. In many-person game situations, this is known as the ‘tragedy of the commons’ or the Elfarol bar problem [10]. A minimal aspect of the three-person game has been proposed by Akiyama and Kaneko [11] as follows.

Assume there are three players. Each player has two actions: to show a black or a white card. If three players show the same colour (i.e., all black or all white), everybody loses. Otherwise, two players of the same colour represent an alliance against the third. Formally, when everyone has the same colour, every player gets nothing. Otherwise, players in an alliance receive a reward of 1 and the third player $\epsilon$. The outcome of the game depends on how we chose the $\epsilon$ value.

In the case of $\epsilon = 0$, Akiyama and Kaneko simulated an evolution of game strategies. Those strategies use past memory patterns to decide the next action. New strategies are introduced into the system through random mutation. If a newly generated strategy can acquire better rewards from the rest of the society, it increases its population number in proportion to the amount of the reward. As a result of simulation experiments, they found that players’ behaviours change qualitatively as evolution proceeds. In the first stages, two players form an alliance so that an asymmetric solution is established. However, interestingly, temporary solutions emerge to replace the asymmetric solution. Thus, players change the card every third period and the average reward will then be equally distributed. This period-3 state is successively replaced by those players with longer periodicity.

This whole scenario has to be modified when each player makes models of the others to predict their future behaviours [12]. That is, co-learning dynamics provides a novel source for instability. In the following, we study the coalition games of players who use a dynamical recognizer to imitate other players’ behaviours. As in the second experiment, we discuss how game outcomes change when one player is replaced by a finite automaton.

V. SIMULATION RESULTS

Any game state can be either in a coalition or non-coalition state. In a coalition state, a single player is always excluded. There is no coalition state of three players at one time. The stability of the states is dependent on the reward ($\epsilon$) of the excluded player. In the case of $\epsilon = 0$, to remain in a state or to escape from a state does not change the reward. Thus, the player has no incentive to change its action. On the other hand, when $\epsilon$ takes a negative
value, it is beneficial for escaping from the exclusion state. Because each player anticipates future moves, when the absolute value of $\epsilon$ is small enough it may destabilize a no-coalition state. This is the basic view of the game, and the results support it as we see below.

1) $\epsilon = 0$ case (a weak incentive to escape from an excluded state):

Out of 24 simulations, 14 initial states ended up with a fixed coalition state within 500 game rounds and the other 10 initial states lasted for more than 500 rounds. In Fig.2, a typical set of dynamics is depicted. A system is finally captured by the coalition between A and C, excluding player B. The other dynamics from different initial states look similar.

During the transient state, players tend to make coalitions with other players, but each coalition is perturbed by another. Such transient coalitions are observed (e.g., around the 100th game round). In the lowest column of Fig.2, a temporary coalition between A and B, or A and C, appears at the later stages ($\geq 100$ game rounds). These transient coalitions should be correlated with the properties of internal models made by the players (see Fig.3-5). In these figures, we depict the Lyapunov exponents, RMS errors and the amount of information used for learning.

Lyapunov exponents can be calculated for given equations (1)-(3). Here the network can only take 0s and 1s as input values (i.e., $u_0$ and $u_1$), and the evolution of the state value $z(t)$ is given by the iterative application of four different mappings, $g_{00}$, $g_{01}$, $g_{10}$ and $g_{11}$, where the suffixes correspond to the possible input patterns into the network. For example, $n$ times applications of this mapping is given by:

$$ (z_1(t + n), z_2(t + n)) = g_{00}g_{01}g_{11}g_{10}g_{01}g_{00} \cdots (z_1(t), z_2(t)). $$

(7)

The context space plot provides an attractor of this iterated functional system. Assuming that each mapping is selected with the probability 1/4, we can compute the associated Lyapunov exponent. RMS errors are simply given by equation (4) with some re-normalization. At each round, the learning step is limited by 10000 steps. When it becomes smaller than a given threshold (here is given by 0.2), the learning process is truncated. By proceeding with the game rounds, the number of learning sequence increases one by one. In this simulation, the model imitates past behaviours by beginning with the last 16 rounds of each player. When the RMS error cannot become lower than the given upper threshold (0.4), a new model will be generated by forgetting all past steps except the last 16. In the figures, the
FIG. 2: Time evolution of action sequences for $\epsilon = 0.0$. Around 200 game rounds, the state is stabilized, with players’ continuously playing (0, 1, 0), respectively. The lowest column displays which player (A, B, C) is excluded at each game round. A non-coalition state is given by the bottom line.

FIG. 3: Time evolution of model player A by other players. Zigzag patterns of the memory lengths show that the players update their internal model for every tens of game rounds.

In this simulation, a no-coalition state has never appeared as a final attractor; nor has any bifurcation into a periodic state. Studying the details of the time evolution of those
characteristics, however, we noticed the following properties during the transients: 1) the Lyapunov exponent suddenly drops off when RMS errors increase at the longer memory lengths; 2) when resetting a memory, the Lyapunov exponent goes up and RMS error goes down; and 3) players’ changing actions do not always perturb model structures.

Those observations are confirmed by analysing the context space plots corresponding to the phases. In Fig.6, the model images get spread and abruptly collapse to a few points, which correspond to when the RMS error goes up and the Lyapunov exponent goes down. During the transient, the average rewards are distributed equally to each player. The complementary relationship between the Lyapunov exponent and the RMS rate can be found in Fig.7. On the other hand, when a system settles down to an attractor, the characteristic values also show relaxation dynamics. That is, memory length diverges and RMS error relaxes.
2) $\epsilon < 0$ case (stronger incentive to escape from the excluded state):

In this case, a fixed coalition state cannot last forever. A possible stable state is that all players take the same action, which merely appears when the absolute value of $\epsilon$ is small. Indeed, simulations with $\epsilon = -0.1$ show much longer transients than 500 game rounds, compared with the $\epsilon = 0$ case. By further simulating for longer than 500 rounds, three out of four runs have never settled down within 10000 rounds. In Fig. 8, a typical transient coalition established between player B and C is depicted. The coalition is not perfect and player A
FIG. 7: Model players with the Lyapunov exponent and RMS rates are overlaid on the same figure. This shows that the Lyapunov exponent is roughly inversely proportional to the RMS rate.

FIG. 8: $\epsilon = -0.1$ A transient coalition established between B and C.
intermittently perturbs the coalition. Here such a transient coalition starts near round 790 and collapses near round 820.

VI. A THIRD PLAYER AS A FINITE-STATE MACHINE

A player with a finite-state machine introduces a remarkable effect into the game. Two different FSAs are used in the following: FSA-1 changes the state when being excluded, while FSA-2 also changes the state, either when being excluded or in the non-coalition state. Both FSAs have only two internal states (nodes), so that the output behaviours become simpler compared with the dynamical recognizers. Thus, as we will see i) clearer transient coalitions evolve than when playing without FSA, and ii) a final attractor becomes periodic at periods of 2 or 3 instead of long transient (chaotic itinerant) behavior.

1) FSA1 with the $\epsilon = -0.1$ (Fig.9).

Out of 15 runs, seven establish a coalition between two cognitive players (dynamical recognizers inside) and excluding the FSA-1. This differs from the previous game with three dynamical recognizers. They show a period of two by playing 1 and 0, alternatively. The other seven runs end up with non-coalition states. Only one run shows an asymmetric coalition between two cognitive players with a period of three, excluding the FSA-2. Their time-averaged payoffs are 1 and 0.63, and the machine’s payoff is 0.27. Their actions are successive repetitions of (0,0,1), (1,1,0) and (1,0,1). Each cognitive player cannot stabilize the model of the other cognitive player, but the model of the machine player is fixed before long. Thus, the machine player is soon well-simulated by cognitive players.

2) FSA2 with $\epsilon = -0.1$ (Fig.10).

The game dynamics appear similar to that for three cognitive players. It generally takes a longer time before it settles down to attracting behaviours. The final attractors are the same as those found in the game with FSA1. However, cognitive players cannot generate the model of FSA2 easily. During the transient phase, a local coalition takes place, where two cognitive players tend to make a period-two synchrony in Fig.10. However, the synchronization is destabilized by the FSA2. In Fig.11, we show each internal model of the machine player for each of the cognitive players. As is clearly seen in these context space plots, the generated models show continuous linear images. Sometimes the image of FSA-2 becomes much more complex, as in Fig.12.
FIG. 9: Dynamics go to the period-2 (above) and the period-3 (below) attractors.

VII. ON OTHER SIMULATED GAMES WITH CDR

We summarize other simulated games with coupled dynamical recognizers below. In each game, a dynamical recognizer is used for imitating other players’ behaviours to decide which action to play next based on the prediction.
FIG. 10: A game against FSA-2. It finally goes to a period-2 attractor.

FIG. 11: Imitated images of FSA2 by th player A and B, respectively.

A. the IPD game

Here, we study the iterated prisoner’s dilemma (IPD) game as an example. This has been studied extensively in the past two decades. In the two-person IPD game, each player has to play to ‘defect’ or to ‘co-operate’ repeatedly. If both play to co-operate in a given round, the players get three points each. However, when one plays to defect against the other’s co-operation, the former gets five points but the latter gets nothing. If both play to
defect against each other, the players only get one point each. Knowing that iterations of
this game can bring about mutual co-operation, R. Axelrod has opened computer program
tournaments of the IPD game [13]. The program Tit-for-Tat, which simply mimics what the
opposing player did in the previous round, has won the tournament twice. In the presence of
noise, more complex strategies than Tit-for-Tat emerge [? ]. Depending on some parameter
values, it is believed to show open-ended evolution. In below, we study the IPD game
without noise but with co-learning players.

As well as in the coalition game, each player tends to make a model of the other by
using DR [17, 18]. This time, the structure of the DR is different from the present coalition
game simulation. It has only two layers, but the connections from the context nodes are
added in a non-linear way to the connections from the input neurons. By imitating the past
behaviour of the opponent with DR using BPTT, a player predicts future action sequences
to optimize the associated outcomes. Elaborating all possible action sequences of length 10
to search for the best action is supposed to bring the best outcome.

In games with fixed strategies, the system can easily reach equilibria. However, in games
between learning strategies, complicated transients were again observed. The length of the

FIG. 12: An example of a complicated image of FSA2. $\lambda = -0.827$. 
transients, i.e., the length of time it took to reach the fixed point of complete defection, depends on the size of the networks. When the number of the recurrent output nodes was two, equilibrium was reached after 90 games. When the network size is small, the player tends to over-simplify the opponent and easily recognizes the opponent’s strategy as all-defection. As the network size becomes larger, the ability to express complex models increases. Thus, the transients become longer. Sudden changes in the RMS error are often observed, which corresponds to the large changes in the weights of the network. This means that many possible models can express the actions with similar precisions. Actually, when we give the same learning set different initial weights, the learning often produces different results. In the model space, there are many models at local minima that minimize the error nearly equally. To capture this aspect, we computed a possible fitness landscape for each time step. In the initial stage, when each player successfully believes that the opponent is playing Tit-for-Tat, the landscape is smooth and has only one local minimum. However, after a few to 10 steps, it becomes quite complex and reveals a lot of local minima. At the end of the game when each player assumes that the opponent always defects, it gradually returns to a simple flat landscape. These complex landscapes reflect this complex behaviour. The error increases after each new game, then the model changes to a new local minimum when the error exceeds the barrier. Thus, such an itinerancy between local minima has been observed, and it constitutes an origin of complex behaviour. Note that these landscapes do not indicate that the evolution of the behaviour is considered the optimization process in the landscape. Each player assumes that the landscape reflects the reality of the opponent, but actually it is almost always an illusion. Because no unique model has an ability to express behaviour, the landscapes have many local minima with similar errors.

Now we consider the three-person IPD simulations. Some effective combinations of FSA strategies in the three-person IPD game have been reported [19], that differ from Tit-for-Tat in the two-person IPD game. Although their simulation is a spatial game, effective strategies are reported as three-node Tit-for-Tat-like strategies. As in the two-person IPD game, a mutual defection state is a strong attractor. Players use DR to imitate other players in this three-person IPD game [12]; indeed, it has been shown that almost all runs end with the mutual defection state. However, unexpectedly long transient states are observed. In particular, five out of ten runs do not settle down to the mutual defection state before 500 rounds. However, we have found some exceptional cases for some learning parameters. One
example shows a simultaneous co-operation of three players. One player becomes tolerant against twice-successive defection and the second player becomes tolerant against a single defection. It is interesting to note that all players simulate eight forward rounds, planning six successive co-operations and two successive defections in the last two rounds.

B. Dubey game

The Dubey game [20] has two players, where each player has an optimal function with two parts; one given by the function of its own position $x$ and one given by the function of other player’s position $y$. Given a home position at $(a, b)$ for a player 1, the optimal functional ($U_1$) of this player is given by, $U_1(x, y) = -[(x - a)^2 + (y - b)^2]$. The egocentric player 1 is most rewarded when $x = a$ and $y = b$. Similarly, given the other player’s optimal function as $U_2(x, y) = -[(x - c)^2 + (y - d)^2]$, the egocentric player 2 wants to have the player 1 at $x = c$ and $y = d$. If $a \neq c$ and $c \neq d$, classical game theory indicates that there is a unique Nash equilibrium pair of this game, i.e., $(x = a, y = c)$, assuming that the strategy set of each player is a point on each axis[34].

Each player moves along a separate spatial axis to take an advantageous position over the other player. Although the players are egocentric in principle, it is shown that some altruistic behaviour will be performed as a dynamical attractor phase. DR is used to model the other player as well. A possible action is to step forward or backward every time step [21]. The altruistic behaviour is no longer attainable by continually modelling the opponent player merely as a Tit-for-Tat player. Rather, players have to change their model of imitation dynamically to achieve mutual co-operation, otherwise they go to a static non-co-operative Nash solution. Instead of using the BPTT learning algorithm, players are restricted to use DR of a finite-sized network of quantized weight values. At each game round, each player chooses the image that best mimics the other player’s behaviour. It is worth comparing the above co-operation states with the iterated Prisoner’s Dilemma (IPD) game situations. In the IPD game, playing C immediately implies co-operation. Therefore, the ‘Tit-for-Tat’ image becomes the fixed point of this coupled recognizer system. In the present case, this is not true. One has to deceive others by stepping back to one’s home position, which paradoxically will sustain mutual co-operation. To do so, players must temporarily hold complex, non-automaton images while changing between two distinct finite
automaton images, Tit-for-Tat and negative Tit-for-two-Tats. This model switching is a temporarily periodic dynamic behaviour, however the underlying model images include the ‘strange attractor’ as in Fig.12.

Inability to have a static and unique model of the other player is due to the context dependency of the internal models. It is worth noting that anti-phase synchronization in action sequences between players emerges as a precursory phenomenon of this mutual co-operation phase, but without any intentions. This is also observed before players come to have Tit-for-Tat images in the two-person IPD game with the same learning algorithm here [18].

C. Language game

We simulate a discourse process between two persons [22]. Each person alternatively makes an utterance to the other person by giving a sequence of words. To make an utterance at each turn, each person generates an internal model of their conversation, not the other speaker’s behaviour. Namely, each person induces a particular rule on what utterance follows the other from the past conversation pattern. The internal model of the conversation is given as a DR. In practice, the DR returns successive utterances against input utterances, where the utterances are represented by bit strings.

The core of this model is the smoothness principle. That is, choosing one utterance out of many is governed by the principle; i.e., choose the utterance that most preserves the context flow of the conversation. Utterances are composed of words, which are composed of alphabets. These alphabets have a bit string representation. Context space plots are regarded as an internal representation of the corresponding alphabets; that is, how a sequence of alphabets is mapped onto the context space. We measure the smoothness of successive utterances by quantifying the continuity of the internal representations. Depending on the situation (i.e., part-conversation pattern), the same alphabets can have different internal representations. Thus, the smoothness is preserved but with different utterances. As an example of the smoothness measure, we took an inner product of the successive context state vector, whose components are the states of context nodes. The alphabet that least changes the value of the inner product will be selected to speak next. Considering this fact, we observe what conversation model persons generate and how their utterances are going to
be formed.

Sometimes, a conversation rule takes a finite automaton-like representation, but sometimes it has a cloud-like appearance (See Fig.13). When this occurs, it means that one alphabet can possess different meanings (i.e., different context space representations). Overlapping of those clouds produces divergence of utterance patterns. When one agent has a cloudy image of the conversation, the other has a fixed point-like image. However, as the conversation proceeds, those images change alternatively and sometimes last long enough to settle down to a final attractor. Discourse thus essentially becomes open-ended, and the conversation model never settles into stable attractors. This study shows that the speech-utterance pattern changes dynamically when we adopt the smoothness principle. Contrary to intuition, the smoothness principle can enhance a variety of utterance patterns.

VIII. CHAOTIC ITINERANCY

It should be repeated here that jumping out from one local attractor to the other before reaching a final attractor is a generic feature of co-learning game systems. As a final attractor cannot be reachable within a finite time, it is natural to define a local (or momentary) attractor in a practical sense. In everyday life, our mind and brain continually change their states as well as their structures. CI is a normal process in a single brain system [4] and especially in the coupled brain system concept introduced in this paper. What causes CI to occur in coupled brain systems? CI is different from the static landscape plus noise system (i.e. it is composed of several attractors separated by a (rugged) fitness landscape with random jumping between them caused by random forces). In the case of the co-learning system, the fitness landscape itself varies from time to time. It is revealed from the simulation of a coalition game playing with a finite-state machine, where two players can finally imitate the machine’s behaviour and the CI is suppressed. Instead of jumping by random noise, the jumping mechanism of the co-learning system has a deterministic origin. When players imitate other players’ behaviours, their images become either clusters of points or cloud-like in a context space. Players with fixed-point images tend to behave like finite machines (e.g., determining their actions from a finite history). On the other hand, players with cloud-like images behave like pseudo machines. The cloud-like images contain continuous state structures. In the coalition game simulation, the players’ behaviours change abruptly
FIG. 13: Changing inner representations of agents. Each utterance consists of 4 combinations of three basic words made out of 6 letters (e.g. $ba, di_1i_2, and gu_1u_2u_3$). (a) shows a part of conversation in the simulation, where $U_A(t)$ and $U_B(t)$ indicate agent A’s and agent B’s utterances at time $t$. The utterances enclosed with broken and dash-dotted lines are used for making models by agents A ($t=70$) and B ($t=69$), respectively. (b)-(e) show the associated context space plots at time step 69 and 70. (d) shows a typical scattered points representation and the alphabet i occupies a small volume in the context space, although at the next step ($t = 70$), its volume seen in (e) is spread out and has an overlap with other alphabets. Compared with agent B, (b) and (c) show that agent A revises his model independently of B’s.
because of the cloud-like patterns. However, when players have these cloudy images, they can imitate their opponents successfully in a coalition game. This is revealed from the correlation between the RMS errors and the Lyapunov exponent.

There are some empirical examples, which we will try to reinterpret from our observation.

1) Experimental IPD games by human players. Since Axelrod’s computer program tournament, the evolution of mutual co-operation without mutual belief/commitment has been most discussed. However, it is known from Yamagishi’s review on experimental gaming [23] that human players’ co-operative actions are much more dependent on how they develop trust between themselves. Therefore, pre-gaming conversation may affect future co-operation. It is more difficult to cope with strangers than those whom one knows well. Recently, the functional image study of co-operation in the IPD game shows that the prefrontal cortex is more active in co-operative states than the non-co-operative ones [24]. The experiment also showed that prefrontal regions are more active for human–human playing than for human–computer playing. A prefrontal region is said to be responsible for planning, prediction, and decision making, · · ·. This reminds us of the coalition game with dynamical recognizers and with a finite machine player. We interpret the experiment as each human tries to predict the other’s behaviour by detecting the intention, but this is not as significant when playing against a computer. McCabe et al., the authors of the experiment, also try to understand their results by saying that the difference between human–human and human–computer gaming is related to each’s inference of the other’s mental states [25], i.e., a theory of mind function. A human playing with a human predicts the other’s behaviour knowing that their mental states may fluctuate. However, when a human plays with a computer, the prediction becomes more stable as it converges (in the experiment, the computer uses a fixed, known probabilistic strategy). On the other hand, players with DR cannot sustain mutual co-operation, owing to the fluctuation of their prediction based on unstable internal models. This prediction we named ‘hot prediction’ [26]. We therefore expect to see CI in prefrontal activities when a human plays against another human.

2) Turn taking. Turn taking is not caused simply by periodic entrainment. Thus, autistic children tend to repeat the same phrases but fail to take turns with their caretakers [27, 28]. Some developmental psychologists argue that autism is characterized by the lack of a ‘theory of mind’ module [28, 29]. If we assume that this ‘theory of mind’ is related to ‘hot prediction’, a failure to take turns can be explained as in the case of the IPD experiments.
In our terminology, some incomplete learning is necessary to take turns. This may lead to new perspectives on turn taking that are not caused by mere entrainment. For example, sensitivity to the ongoing-ness of interaction and/or subtle differences of body-movements. We also expect to see chaotic itinerancy here. In a recent simulation of turn taking with CDRs [30, 31], it was suggested that two agents should actively interact with each other to maintain turn taking. Non-reactive, thus finite machine-like, players fail to take turns. But, paradoxically, we argue that the unstable dynamics underlying the interaction can be used to hold the interaction itself.

IX. DISCUSSION

The notion of chaotic itinerancy in coupled dynamical recognizers accounts for some important aspects of communication dynamics. Usually, communication is assumed to be a process in which information is exchanged between agents. Linguistic communication represents one such example. On the other hand, we should discuss the different aspects of communication, or in other words, proto-conversational or meta-communicative aspects of communication. It is widely observed in communication systems ranging from play behaviour to turn taking that temporary transitions occur between each salient phase and the next. Non-recursivity and an open-ended nature is expected in lively communication, and indeed this characterizes CI.

In simulated gaming systems with coupled dynamical recognizers, players mutually predict future behaviour by imitating their past behaviours. However, as we discussed above, those predictions are usually unreliable as other players are also changing their behaviour according to their predictions. Such nested prediction provides for the complexity of gaming dynamics. For example, it quite often happens that a subtle difference in model structures will be amplified in future games. This fluctuation in prediction, as stated in §9, is ‘hot prediction’, as opposed to ‘cold prediction’. The difference between a physical cause–effect relationship and a motive–action relationship is a matter of reliability, i.e., the strength of the linkage between cause and effect. When A is followed by B, we call A the motive, if we notice any fluctuation in the linkage or the inference from A to B. To detect the fluctuation in the inference, we have to have a model that computes $A \rightarrow B$ beforehand. If the linkage is between a cause and effect, the more we experience the event, the more we will be con-
vinced that the cause is followed by the effect. This we term ‘cold prediction’. On the other hand, the relationship usually depends on the hidden context being internally or externally provided. Knowing the context dependency, people dare to predict the effect from the cause. We think that imitating the behaviour of others is an example of such prediction. Then, we call the cause a ‘motive’ and the prediction a ‘hot prediction’. We predict the behaviour of others even though we know that the motive is unseen and unreliable. As we argued in [26], this theory of mind means to detect autonomy and to make a prediction based on that at the same time. In terms of our simulation, ‘hot prediction’ can be related to the prediction with cloudy context images. That the same input pattern does not give the same output pattern is the typical behaviour associated with cloudy images.

To conclude, we have discussed that human communication can be imitated by coupled dynamical recognizers as co-learning dynamic processes. In §3, games with and without a machine player revealed that chaotic itinerancy was suppressed in the latter case. A continually changing structure, that is, the change of topics in a language game and the generation of coalition pairs in a coalition game, is a key feature of human ‘hot prediction’. Human learning is very different from mere neural net learning with fast convergence. Such human learning has a duality of convergence and divergence. The divergence part is important as it gives a potential for exploration activity. The non-convergent nature of learning, and hot prediction as its result, are the foundations of our learning behaviour. Chaotic itinerancy in communication is based on such convergence and divergence nature of dynamics. Empirically, the initiation of such live communication has been observed between an infant and its mother, and named the ‘development of inter-subjectivity’ by Trevarthen [33]. Using double monitor experiments, Trevarthen showed that a child and its mother are sensitive to the liveliness and continuity of their mutual images. An infant immediately distinguishes between live and recorded images. A coupled dynamical recognizers model gives a possible interpretation of such a dynamic nature of human communication, that amplifies the subtle differences between sense and action patterns.
This work was partially supported by Grants-in-aid nos. 13831003 and 12-09422 from the Ministry of Education, Science, Sports and Culture.


[27] Rosenberg, S. and Abbeduto, L. *Language and Communication in Mental Retardation* (LEA


[34] As long as one player sticks to a point, there is no reason for the other player to deviate from the point either. Furthermore, if it is impossible for both players to take simultaneous advantage from a deviation from the point, it is called a Pareto-efficient point. In this game, such a Pareto-efficient point $t$ forms a line segment joining $(a, b)$ and $(c, d)$ [20]