A Neural Model of Mentalization/Mindfulness based Psychotherapy

Abbas Edalat and Zheng Lin

Abstract—**We introduce and implement a neural model for mentalization/mindfulness based psychotherapy. It uses Dan Levine's neural model of pathways for emotional-cognitive decision making, which is integrated with a competitive Hopfield network built up from the new concept of strong patterns for the six basic emotions and for mentalization or mindfulness. We adopt a particular form of Q-learning to reinforce the mentalizing/mindful pattern in the network, which represents the process of psychotherapy. In a successful course of therapy, the mentalizing/mindful pattern becomes the more dominant pattern compared to negative emotions and the brain makes decisions that are more deliberate and thoughtful than heuristic and automatic.**

I. Introduction

We develop a neural model for psychotherapy based on mentalization using our current understanding of the human brain, psychopathology and the effect of psychotherapy on the individual. We adopt the fundamental viewpoint in attachment theory (see, for example, [1]): Psychopathology is, generally speaking, a reflection of suboptimal early neural development, integration and coordination which may be caused by a combination of adverse environmental and biological factors. Since early interactions with primary caregivers have a profound impact on the expression of genes, there is a period in an infant's life when nature and nurture actually coincide as experience is transformed into neural structure.

The quality and dynamics of the infant's relationships with the primary caregivers sculpt the neural circuits corresponding to the secure or insecure attachment types of the toddler which become the basis of the internal "working model" for the child with long term impact on the emotional well-being of the child and later the adult [2][3][4].

Symptoms of psychopathology such as depression and anxiety inflict psychic pain when the neural pathways required for healthy functioning are underdeveloped or under-regulated. The individual is then unable to modulate and regulate his or her arousal level and strong emotions. Psychotherapy helps the individual to take steps to integrate the cognitive and affective neural pathways so as to enable such modulation and regulation and thereby increase the capacity of the individual to withstand stress in problems of everyday life [6][7].

While there are currently many forms of psychotherapy available, an overarching goal in nearly all evidence based psychotherapies is to increase the capacity of the individual for mentalizing or mindfulness, i.e., to understand the emotional and mental state of others as well as the individual self [7]. Adults who had developed a secure attachment with their primary caregivers as children have a well-developed capacity for mentalizing or mindfulness, whereas those with an insecure attachment type lack such a strong capacity.

The three most widely used psychotherapies, namely Cognitive Behavioral Therapy (CBT), Mindfulness Based CBT and Psychodynamic Psychotherapy all have their main focus on nurturing and developing the capacity for mentalizing or mindfulness in the clients in one form or the other, an integrating process which would empower them to contain and come to terms with negative strong emotions. An increased capacity for mentalizing/mindfulness means that the individual is able to make more deliberate, thoughtful and conscious, as opposed to heuristic, automatic and subconscious decisions, which in turn provides a sense of agency and control in the individual [7]. We will from now refer to mentalizing/mindfulness as simply mentalizing.

Levine [8] has developed a neural model for pathways of emotional-cognitive decision making, which explains how deliberate as opposed to heuristic decisions are made in the brain pathways. It is comprised of a competitive model for basic and higher needs which are captured by attractors of a neural network [9] together with four main organs in the brain, namely the amygdala, the Orbitofrontal Cortex (OFC), the Dorsolateral Prefrontal Cortex (DLPFC) and the Anterior Cingulate Cortex (ACC), all of which are involved in decision making. An Adaptive Resonance Theory (ART) architecture [10] has been proposed to model the interactions of these four organs for decision making based on the input from the needs network.

We build on Levine's model to develop a model of mentalization based psychotherapy by integrating into his model the new concept of strong patterns in neural networks. Strong, in other words, multiply learned, patterns of the Hopfield network [11] have been recently introduced to model attachment types and behavioral and cognitive prototypes [12][13]. It has been shown that strong patterns are strongly stable, with an energy level decreasing proportionally to their degree (i.e., their multiplicity) and a retrieval capacity which, in the presence of random patterns, increases proportionally to the square of their degree. These results show that strong patterns are robust and persistent in the network memory as attachment types and behavioral prototypes are in the human memory system.

Strong patterns reflect cognitive and behavioral patterns or addictions that are deeply and repeatedly learned as in the process of Hebbian learning or long term potentiation and

Abbas Edalat and Zheng Lin are at the Department of Computing, Imperial College London. (emails: a.edalat@ic.ac.uk and zheng.lin12@ic.ac.uk)

therefore are suitable to model both the symptoms of psychopathology and the process of psychotherapy. They are incorporated in our framework to model a strong psychological need for cognitive closure, which can in particular be considered as a brain biased toward a combination of the basic six emotions, and in particular the negative emotions, namely, strong anger and fear. We also use them to model increased need for cognition and mentalizing that is developed in the individual as a result of psychotherapy. Facial expressions are employed to model the basic emotions as well as mentalizing; in fact for simplicity and as a first approximation in our implementation we employ animated smiley faces in the six basic emotional states and in a thinking mode to capture these seven states in our model. This means that in our implementation framework what matters is not the mental state itself but the relative strength of the mental states, i.e., the degree of the strong pattern representing each of them.

In this paper, reinforcement learning is used to model the process of behavioral change. Reinforcement learning, which originated from behavioral psychology and instrumental learning, is based on trial and error in order to maximize an agent's reward depending on the responses it receives from the environment [14]. Q-leaning is a particular form of reinforcement learning based on temporal difference where prediction of rewards of the agent's choices is carried out in immediate response to the environmental feedback. There is evidence from animal studies that the activity of the dopamine system, thought to be involved in decision making about reward, is actually consistent with and supports Q-leaning [15].

The process of psychotherapy in our work is modeled by a form of Q-learning which reinforces mentalizing as well as happiness as opposed to negative emotions such as strong anger and fear. The reinforcement learning will result, on the one hand, in a modulation of the strong patterns reflecting negative emotions and, on the other hand, in an increase in the multiplicity of the strong patterns representing mentalizing and happiness. When the mentalizing pattern becomes the dominant pattern in the competitive neural network the psychotherapeutic process is considered to have been successful. The individual will now make far more deliberate, thoughtful and conscious decisions processed by the OFC and DLPFC than heuristic, automatic and subconscious ones that are amygdala driven and stem from suboptimal early neural development.

We essentially adopt the structure of the ART architecture as presented in [8] to model the relevant functions of amygdala, OFC, DLPFC and ACC in this work, but we use the Restricted Boltzmann Machine (RBM) [16] for classification of features in these brain regions rather than the ART classifier itself. This makes our implementation efficient and can be considered as a first approximation, even though the RBM has symmetric synaptic couplings and is therefore not as biologically plausible as the ART classifier.

We now briefly review neural models for psychotherapy in the literature. In [17] a neural model for personality has been proposed based on the character cube presented in [18], which, according to the authors, deals with individuals who are already functioning quite well but seek to improve their effectiveness. Similarly, the work in [19] uses Cloninger's character cube to develop a neural model for how human beings suppress or enhance certain types of behavior. In [20], a Hopfield network with two weakly connected layers has been used to model the concept of "working through" in psychoanalysis. Galatzer-Levy provides in [21] an expository account of how non-linear dynamical systems and attractors can be used qualitatively to model the psychoanalytical process; it also contains other related references in this subject. There are also several volumes on neural modeling of psychopathologies as in [22].

To our knowledge, however, this paper presents the first quantitative neural model of psychotherapy in general, which has been implemented, as well as the first neural model of any type of mentalization/mindfulness based psychotherapy, which in particular includes Cognitive Behavioral Therapy

II. DECISION MAKING NETWORK

A. The Network of Needs

In the human brain, multiple decision rules coexist and compete with each other. These rules are either irrationally heuristic or consciously deliberative. The decision rules at different levels of sophistication activate different regions of the brain [8].

Levine posits that the heuristic decisions are inherited from emotionally influenced decisions made by other mammals, and the deliberative decisions involve logic calculation and working memory manipulation [8]. He presents a decision-making system in the human brain encompassing a network of needs, a network of decision rules, and the communication between these two subsystems.

The network of needs involves physiological as well as psychological needs on different levels. These needs compete with each other, and the state of the needs network shifts within these needs when the individual experiences discontent.

The network of decision rules consists of four connected areas: the amygdala, the OFC, ACC and DLPFC, which account for various decision rules on specific tasks. These four regions comprise a three-layer network, in which the *vigilance threshold* of the individual determines the status of activation of each layer. The state of the needs network influences the vigilance threshold so that the winning needs become dominant to implement the corresponding decision rules.

Levine developed the needs module as the interpretation of Maslow's hierarchy of needs [23]. Maslow added psychological needs, such as love, esteem, and self-fulfillment, to the list of purely physiological needs.

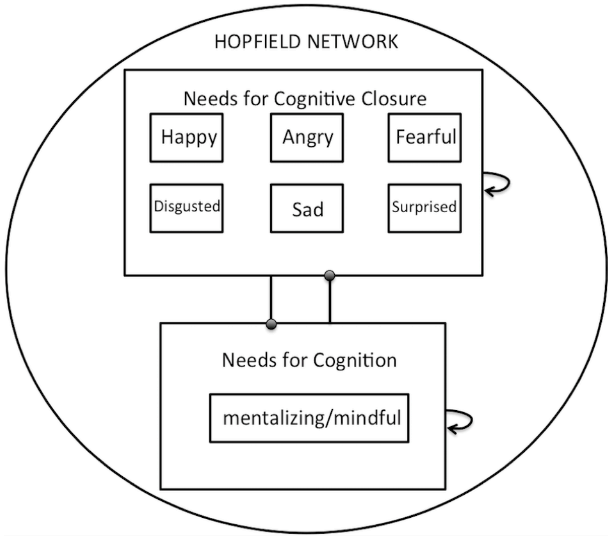

Fig. 1. The Hopfield network containing the needs for cognitive closure and needs for cognition. The needs for cognitive closure are correlated to the six basic emotional patterns, and the needs for cognition are correlated to the "mentalizing/mindful" pattern. Similar to the relationship of the needs in Maslow's hierarchy, each of these two kinds of needs receives inhibition signals from another and excitation signals from itself.

These needs inhibit others while sending excitatory signal to themselves; the physiological needs have normally the strongest self-excitation.

Maslow's hierarchy of needs can be considered as a two-layer hierarchy containing the need for cognition [24] and the need for cognitive closure [25]. The need for cognition is the motivation to analyze arguments deeply, and individuals with high need for cognition are more likely to make deliberative decisions. The need for cognitive closure is thought of as the motivation to make decisions without thinking about the relevant issues, and individuals with high need for cognitive closure are more likely to make heuristic decisions.

Here we employ a standard Hopfield network to model Levine's needs diagram by using strong patterns that represent the basic emotional and cognitive needs. The degrees of strong attractors in this network change dynamically according to the strength of the needs that the individual experiences at each point in time. The six basic emotions and a "mentalizing" pattern are stored as strong patterns represented by animated smiley faces in the system with various degrees. We assume that the six basic emotions are correlated to the need for cognitive closure and the mentalizing pattern is strongly correlated to the need for cognition. That is, increasing the degree of one of the emotional patterns leads to enhancing the satisfaction of the need for cognitive closure, and increasing the degree of the mentalizing pattern leads to enhancing the satisfaction of the need for cognition. As shown in Fig 1, these two needs compete with each other in the needs module, and the six emotional patterns compete with each other within the need for cognitive closure. In the Hopfield network, the initial

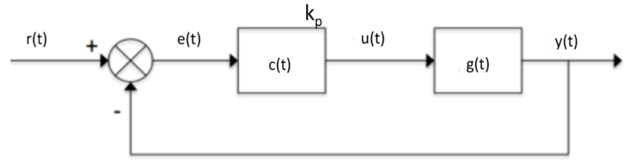

Fig. 2. A black box diagram of a P controller in a feedback loop. The dynamics of $c(t)$ and $y(t)$ are analogous to the control behavior of the OFC and the amygdala respectively.

configuration of the system will be updated during iterations, and it will, with high probability, eventually converge to the strongest pattern.

For example, if the mentalizing pattern is the dominant attractor stored in the Hopfield network, that is, the mentalizing pattern has the highest degree among all other strong patterns in the system, the brain model containing this Hopfield network is thought of having high need for cognition. Therefore, as in Levine's model, the energy of the mentalizing pattern will be stored in the working memory, and then any state of the network with higher energy will be disturbed by random noise so that the needs module can shift to the mentalizing attractor. Similar to this process, given any initial configuration, the dynamics of the Hopfield network will, with high probability, eventually converge to its strongest attractor, i.e., the attractor with highest degree and least energy.

B. *The Network of Decision Rules*

Three identical Restricted Boltzmann Machines (RBM) are introduced to simulate the behaviors of amygdala, OFC and DLPFC. These RBM's are able to categorize the seven different patterns (six emotional patterns and one mentalizing pattern) and represent them in the hidden units using seven distinct 17-unit vectors. Each of these RBM's responds to input signals in different ways (this will be discussed when combined with the network of needs in next subsection). And the outputs from them will be compared with respect to the value r of the vigilance threshold as a threshold in ACC. We say that mismatch occurs if the Hamming distance from the amygdala output to OFC output (or from OFC output to DLPFC output) is greater than r. For individuals with mentalizing pattern as the strongest attractor (high need for cognition), the vigilance threshold is always low, while for individuals with one of the emotional patterns as the strongest attractor (high need for cognitive closure), the vigilance threshold is high. In other words mismatch is easier to occur in a mentalizing person, while a person with high vigilance threshold most likely follows the heuristic signals.

The OFC is so densely connected to the limbic system that it has been sometimes characterized as the executive director of the limbic system [2]. Its optimal development in childhood allows it to regulate the stimuli from amygdala. We can therefore regard OFC as exerting a controlling mechanism, which we also incorporate in our model.

As far as the control from OFC to amygdala (or from

DLPFC to OFC) is concerned, we have designed two versions of our brain system in this project: one regards the OFC as a controller to adjust the outputs from amygdala, and the other disregards the control activity of OFC. However, before we can design such controller to simulate the control activity of OFC, we have to first consider the dynamics of the controlled system – the 17-neuron population in amygdala.

C. *Modeling Feedback Controlling*

To model the dynamics of the 17-neuron population in amygdala, we take these neurons as a neural network existing in a neurobiological system, even though the neurons used in this project are binary units. The dynamics of real neural population is very complex and highly nonlinear, but it is still a good approximation to consider the population as a linear system. Such linear approximation is still biologically plausible because it is relevant to the production of postsynaptic currents (PSC) in the postsynaptic cell [26]. For simplicity, a PSC model can be written as

$$h_{psc}(t) = \exp(-t/\tau_{psc}), \tag{1}$$

where τ_{psc} is the synaptic time constant. We can apply this model to the population system, so that the transfer function of the 17-neuron population is

$$h(t) = \frac{1}{\tau}\exp(-t/\tau), \tag{2}$$

where τ is the synaptic time constant. For convenience, we can rewrite Equation (2) in frequency domain using Laplace transformation $h(s)$ of $h(t)$ as given by Equation (2),

$$\bar{h}(s) = \frac{1}{1+s\tau}. \tag{3}$$

A proportional gain controller (P controller) is the simplest feedback controller used in industrial control system. Fig 2 shows a schema of a feedback control loop with P controller. The proportional controller $c(t)$, which is characterized by the proportional gain k_p, outputs a control signal $u(t)$ to the controlled system $g(t)$ according to the difference (error $e(t)$) between the measured system output $y(t)$ and the desired reference $r(t)$. This controller attempts to minimize the error and increase the speed of system response.

D. *Connection of Needs Network and Decision Network*

In [Levine 2009], it is posited that different needs in the module could result in different vigilance threshold. Here, we assign a high value of vigilance threshold to the need for cognitive closure, and a low value to the need for cognition. When the mentalizing pattern is dominant in this model, it has a strong ability to detect the mismatch between the OFC and amygdala (or the DLPFC and OFC). On the other hand, if the vigilance threshold is high, the OFC-amygdala loop is more activated than the DLPFC-OFC loop, and then the individual will undertake heuristic decisions; otherwise, the DLPFC-OFC loop is dominant in the decision rule network, and the person will make deliberative decisions.

In our DECIDER model, the input pattern will firstly stimulate the Hopfield network and the dynamics of this network will eventually evolve to an attractor with the lowest energy.

If the mentalizing pattern is recalled, then the Hopfield network will send a low-level vigilance threshold (a constant r) to the error detector (ACC), and the DLPFC-OFC loop is chosen to generate complex decision rules. As the secondary sensory device, the RBM in the DLPFC receives a mentalizing signal from the Hopfield network, and categorizes it into a 17-unit vector. And the RBM in the OFC, the primary sensory device, accounts for categorizing the input pattern into another 17-unit vector.

If the Hamming distance of these two generated vectors is greater than the vigilance threshold, then we say that a mismatch occurs and the DECIDER would generate a deliberative rule. Otherwise, the DECIDER would make a heuristic decision. However, because of the low vigilance threshold, the DLPFC-OFC loop would most likely not make heuristic decisions. Note that there is no control from DLPFC to OFC in our current model.

On the other hand, if the recalled pattern is one of the six emotional patterns, the Hopfield network will send a high-level vigilance threshold to ACC, and the OFC-amygdala loop is chosen for making decisions. Similar to the loop above, as the secondary sensory device, RBM for OFC categorizes the emotional pattern from the Hopfield network, and amygdala categorizes the input pattern. The Hamming distance between the two output vectors is compared to the high-level vigilance threshold. A heuristic decision is generated if the Hamming distance is lower than the vigilance threshold; otherwise, a deliberative decision is made. Furthermore, the control from the OFC to amygdala influences the output of the RBM in amygdala. A well-trained OFC (with the proportional gain k_p tuned to be appropriate) can adjust the amygdala output to some value close to the output of the OFC.

III. MODELING PSYCHOTHERAPY

A. *Aims and Objectives*

The main purpose of this work is to apply a Q-learning algorithm to model the psychotherapeutic process. Two experiments are carried out to study how the discount factor of the Q-learning is related to the therapeutic process, and how the initial configuration of the attractor-competitive network influences therapy.

B. *Outline of Implementation*

As already pointed out, CBT aims at a more rational interpretation of events by the individual, which would ultimately enhance the control from the orbitofrontal cortex to amygdala. In this work, the interpretation of an event is modeled metaphorically by the recall pattern when the Hopfield network, with stored strong patterns for the six basic emotions and mentalization, is exposed to an initial

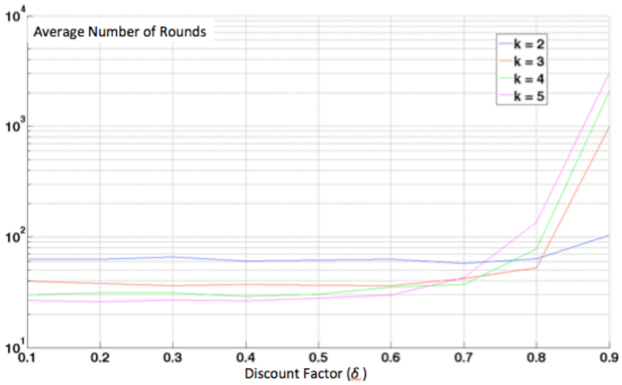

Fig. 3. The average number of rounds required for the mentalizing pattern become dominant when undergoing CBT treatment for different discount factor (δ). Each line in the plot corresponds to a particular exploration rate (k) in the action selection rule. The y-axis is presented in log scale because the values for $\delta = 0.9$ are unexpectedly high.

random pattern. These strong patterns are stored in the Hopfield network as smiley faces representing the six emotions and mentalization with different degrees. The regulating behavior from OFC to amygdala is realized by implementing the P-controller in which the parameter k_p is analogous to the control strength.

In addition now a Q-learning algorithm is designed to model how the modulating process from orbitofrontal cortex to amygdala is enhanced during therapy and how the relative strength of the strong patterns in the attractor-competitive network are altered. In other words, the degrees of patterns in the Hopfield network, and the parameter k_p keep changing during the operation of the Q-learning algorithm.

C. Hopfield Network

The Hopfield network is trained with eight types of patterns (Angry, Happy, Sad, Disgusted, Fearful, Surprised, Mentalizing and random patterns). The degrees of five of the basic seven patterns, namely Angry, Happy, Sad, Fearful and Mentalizing would change during the Q-learning process.

D. Feedback Control Loop

The initial value of k_p is set to 0.5. This value will increase in the course of treatment as the degree of the mentalizing pattern increases. The steady-state output of amygdala gets close to the output of OFC if we increase the value of parameter p.

E. Q-learning

We apply Q-learning to model how the choices that the patient makes will influence the state transition during the process of psychotherapy. In Q-learning it is assumed that there are a set S of Q-states and a set A of actions for an agent. The Q function $Q: S \times A \rightarrow \mathbb{R}$ calculates a Q value for each pair of Q-state $s \in S$ and action $a \in A$. These Q values are an estimate of the expected reward that the agent will receive from choosing this particular action when it comes to make an action decision. Following the choice of action a in Q-state s and an observed transition to Q-state s', the reward is given by $R_a(s, s')$ and the Q values are updated according to the conventional rule:

$$Q(s,a) \leftarrow Q(s,a) + \lambda[R_a(s,s') + \gamma]$$

where

$$\gamma = \delta \max_{a'} Q(s',a') - Q(s,a).$$

The learning rate $0 \leq \lambda \leq 1$ determines the extent to which newly acquired information overrides old information, and the discount factor $0 \leq \delta < 1$ signifies the relative value the agent places on immediate/future rewards.

TABLE I
REINFORCEMENT REWARDS FOR EACH ACTION

Action Name	Reward
Be_Angry	0
Be_Happy	0.5
Be_Sad	0.2
Be_Fearful	0.1
Be_Mentalizing	1

In our model, the state of the patient is given by a D-state and a Q-state:

- The D-state is the 5-vector, whose components are the degrees of the five chosen patterns in the following order: Angry, Happy, Sad, Fearful, Mentalizing.
- The Q-state is the ordinal 5-vector corresponding to the D-state, i.e., the relative rank of the degrees of the five patterns.

For example, if the D-state is [10,1,5,1,3], which gives the degrees of the five above patterns, then the Q-state would be [4,1,3,1,2]. The idea of using ordinal ranks in Q-learning was first introduced in [27]. There are in our case 5! Q-states for the permutations of all strict ordinal vectors, i.e. for the case where all five degrees are different, 5!/2! Q-states for the non-strict ordinal vector with two equal degrees, 5!/3! Q-states for the non-strict ordinal vector with three equal degrees, 5!/4! Q-states for the non-strict ordinal vector with four equal degrees, and 5!/5! states for the case where all degrees are equal. This thus gives a Q-state space S of size 120+60+20+5+1=206.

For each Q-state $s \in S$, there is a corresponding finite set of valid actions A to choose from. For simplicity it is assumed that all Q-states share the same action set

$$A = \{\text{Be\_Angry}, \text{Be\_Happy}, \text{Be\_Sad}, \text{Be\_Fearful}, \text{Be\_Mentalizing}\}$$

resulting in 206*5 (Q-state, action) combinations.

A probabilistic 'Boltzmann' action selection rule [28] is introduced to prevent the Q-values getting stuck in a local minimum. This rule allows state-action combinations with low Q-value to be chosen with some positive probability. According to the rule, the probability that action a is selected is:

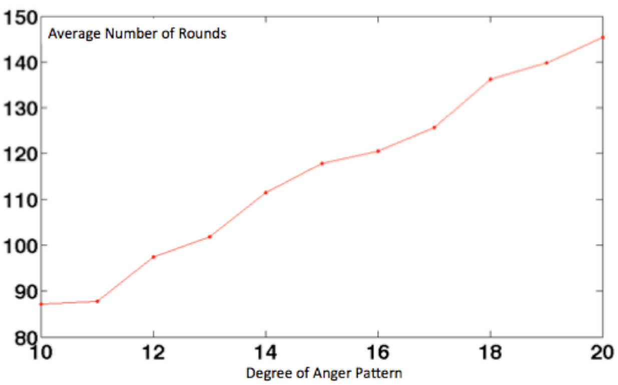

Fig. 4. The average number of rounds required for a "mentalizing/mindful" pattern when a patient received CBT treatment for initial degree of anger pattern.

$$P(a|s) = \frac{k^{Q(s,a)}}{\sum_{a \epsilon A} k^{Q(s,a)}}$$

where k is an exploration parameter, whose increase results in reducing the probability of selecting those actions with low Q value. Therefore, the action selection rule with high exploration parameter can be thought of deterministic selection rule (i.e. always choosing the action associated with the highest Q value).

The discount factor varies from 0.1 to 0.9, and for simplicity the learning rate is set to 1. The goal of CBT is to learn to deal with current problems in an appropriate way by mentalizing and therefore the "Be_Mentalizing" action receives the highest immediate reward. Be-Happy as a positive affect receives the next highest reward followed by Be-Sad to allow for grief and loss. Since a certain minimal amount of anxiety is also required for any effective performance, Be-Fearful is also slightly rewarded. The reward for each action as shown in Table I.

The initial Q value of each (Q-state, action) combination equals the value of the corresponding element in the Q-state vector. For example, if initially $s = [1, 4, 2, 3, 5]$, then we set $Q(s, \text{Be\_Angry})$ to 1 and $Q(s, \text{Be\_Happy})$ to 4.

The algorithm for changing the D-state and Q-state works as follows. At any point in time, we have a given D-state and its corresponding Q-state. According to the 'Boltzmann' rule, an action Be_X from the set A is selected, where X is one of the five chosen patterns: Angry, Happy, Sad, Fearful and Mentalizing. Subsequently, the Q-values are updated according to the Q-learning rule for Q-states. Concurrently, the degree of the pattern X is incremented by one, resulting in a new D-state that has the components of the old D-state except for the increment in the degree of the component of X. Note that the Q-state will now change if and only if the relative rank of the degrees of the five patterns changes.

The algorithm terminates once the mentalizing pattern becomes the dominant pattern in the Hopfield network. Otherwise, it will continue running until the maximum number of iterations. In the full version of this paper we will show mathematically that this algorithm eventually converges to a Q-state with a dominant Mentalizing pattern.

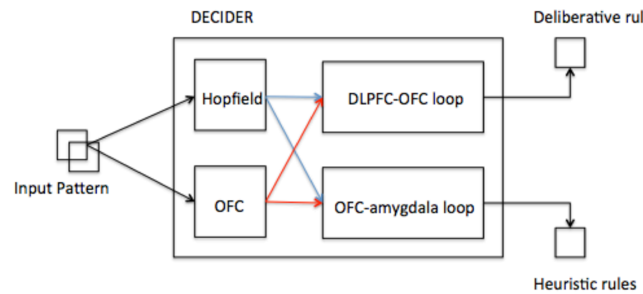

Fig. 5. An illustration of DECIDER, responding to the outside world.

F. Experiment 1

In this experiment we study how the discount factor is related to the number of iterations required to change the dominant pattern from one of the emotional patterns to the mentalizing pattern. Also, the relation between the exploration rate k in the probabilistic 'Boltzmann' action selection rule and the required rounds of iteration is examined.

The experiment goes as follows:

- The dominant pattern is initially set to the Angry pattern.
- The exploration rate k varies from 1 to 4.
- Discount factor varies from 0.1 to 0.9.
- An individual experiment is run up to 10,000 rounds.
- The number of rounds required to make the mentalizing pattern dominant is recorded.
- Each individual experiment is repeated for 20 times.

Initially the degrees of the seven patterns are set as follows:

Angry: 8	Fearful: 5	Sad: 3	Disgusted: 1
Happy: 1	Surprised: 1	Mentalizing: 2	

In addition 500 random patterns are stored in the network.

After, 200 rounds, we take the mean value of the number of rounds required to make the mentalizing pattern to become dominant. This average provides a more accurate estimate for the Q-learning behavior. The results are displayed in Fig 3.

As the exploration rate k increases, the average number of rounds required for obtaining a successful CBT treatment (i.e. the mentalizing pattern becomes dominant in the Hopfield network) decreases. The patient with a higher value of k is thought of as a patient who is more willing to comply with the treatment protocol, a major assumption of successful CBT. We also observed that, in general, the patient with higher discount factor, i.e., the patient waiting for long-term rewards, achieves a slower rate of success than those striving for more short term change and reward.

We also observed that the number of rounds required for the mentalizing pattern to become dominant is quite high

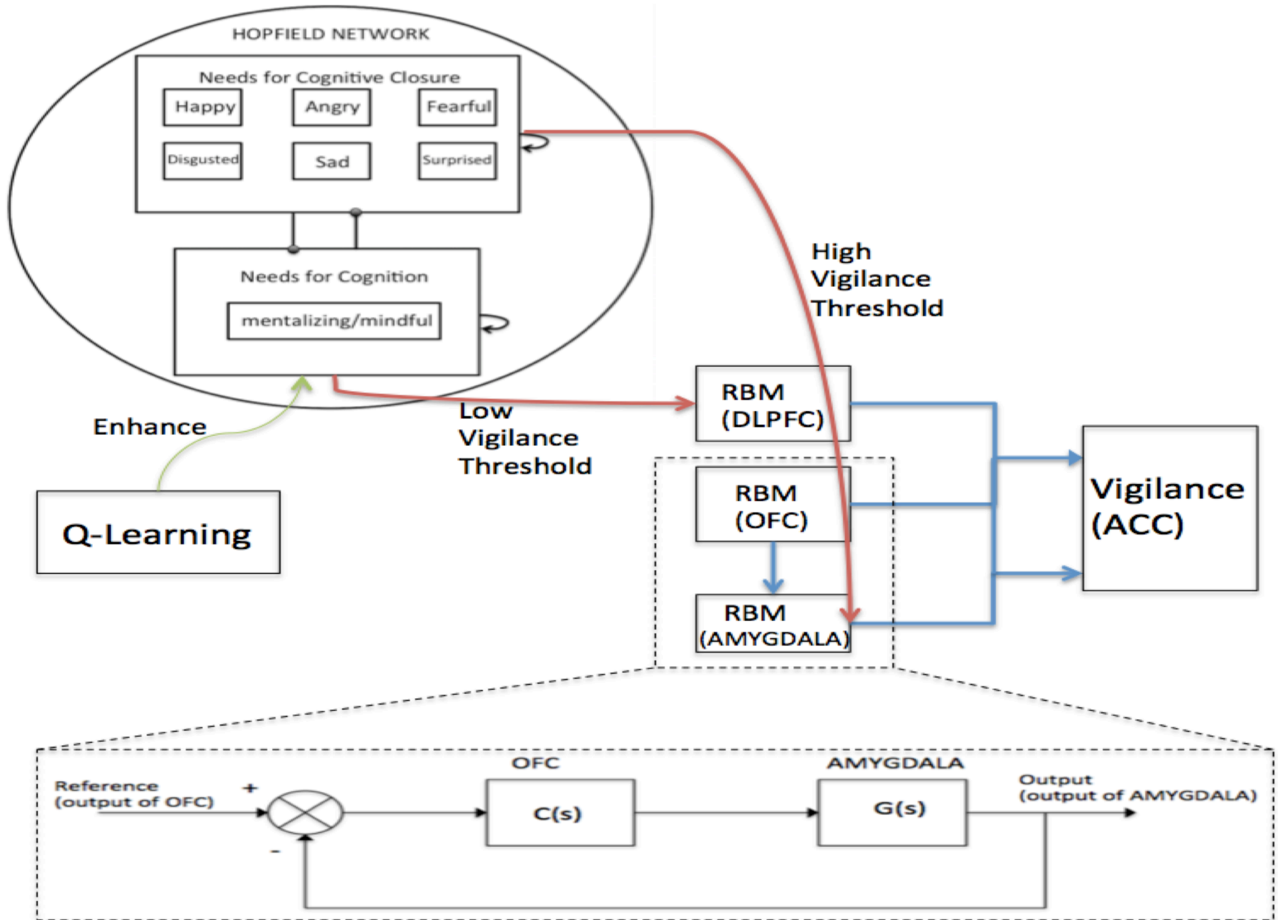

Fig. 6. DECIDER model combining Hopfield network as needs network, and RBMs as decision network. Q-learning influences the attractors in the Hopfield. Needs for cognitive closure outputs high-level vigilance threshold signals to ACC. Needs for cognition outputs low-level vigilance threshold signals to ACC. The control from OFC to amygdala is designed as a feedback control loop.

when the discount factor $\delta > 0.9$ (i.e. the patient who prefers to wait for future rewards). The patient receives reward every time a reinforced action is chosen, even though no state transition occurs, that is, $R_a(s, s') > 0$ for $s = s'$. Therefore, the value of $Q(s, a)$ is updated even if no state transition has occurred. For the situation that a patient has high discount factor $\delta > 0.9$, these non-reinforced actions will induce large Q and more state-action exploration.

G. *Experiment 2*

In this experiment we studied how the initial degrees of the stored patterns in the Hopfield network are related to the number of iterations required to change the dominant pattern from one of the emotional patterns to the mentalizing pattern.

The experiment goes as follows:

- The dominant pattern is initially set to the Angry pattern.
- The initial degree of the Angry pattern varies from 10 to 20.
- The initial degrees of other patterns are constant.
- Discount factor $\delta = 0.2$.
- Exploration rate $k = 2$.
- Each experiment is repeated 10,000 rounds.
- The number of rounds required to make the mentalizing pattern dominant is recorded.
- Each experiment is repeated1000 times.

Initially 500 random patterns are stored in the network together with the seven basic patterns with the following degrees:

Angry: 10 to 20	Fearful: 5,	Sad: 3	Disgusted: 1
Happy: 1	Surprised: 1		Mentalizing: 2

We again run the experiment 1000 times and take the average of the number of rounds required to make the mentalizing pattern dominant and depict the result in Fig 4.

Except for the Angry pattern, the degree of any other pattern is viewed as constant. In general, as the degree of the Angry pattern is increased, the average number of rounds required for obtaining a successful CBT treatment (i.e. the mentalizing pattern becoming dominant in the Hopfield network) increases.

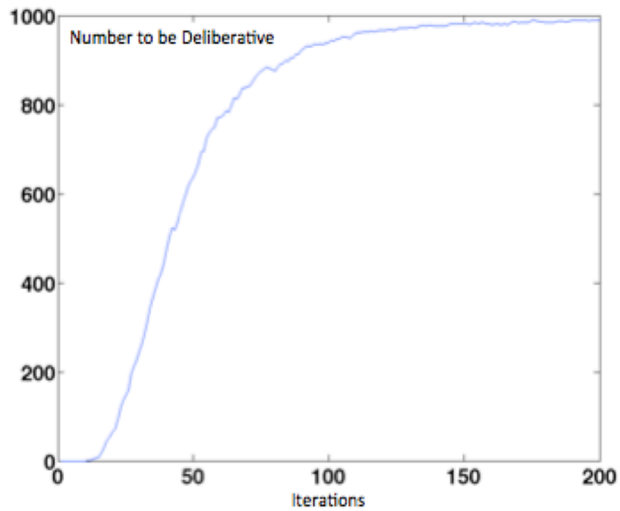

Fig. 7. Simulation of DECIDER model: average numbers to make deliberative decisions versus iterations during Q-learning.

IV. Bringing it all together

Integrating the above components, we now obtain a super-model containing a super-network DECIDER, which simulates decision making and reinforcement learning, mimicking the effect of CBT treatment. Fig 5 illustrates how DECIDER tackles the input pattern from outside world and makes decisions as output. Fig 6 illustrates the structure of DECIDER and how its subsystems communicate with each other.

As described previously, the seven basic patterns are stored in the Hopfield network as strong attractors, and one of them is recalled once the network receives a stimulus. If the recalled attractor is one of the six emotional patterns, it will be treated as an input signal to amygdala; otherwise, it will be viewed as an input signal to DLPFC.

Three identical well-trained Restricted Boltzmann machines (RBMs) represent three regions accounting for various decision rules in the brain: DLPFC, OFC and amygdala. This trained RBM has the capability of categorizing the seven basic patterns, the six basic emotions and the mentalizing pattern. It receives binary images with 25*25 pixels as input, and reduces the dimension of the input to 17-unit vectors. As mentioned above, the DLPFC receives the mentalizing pattern as input and the amygdala receives the emotional patterns as input. In addition, the OFC categorizes the input patterns from outside world directly. The DLPFC-OFC loop works if the recalled pattern of the Hopfield network is the mentalizing pattern, and it accounts for generating deliberative rules. On the other hand, the OFC-amygdala works if the recalled pattern is one of the emotional patterns, and it deals with heuristic rules and some deliberative rules.

The ACC has the function to detect the potential conflict between the outputs of DLPFC and OFC (or OFC and amygdala). It is viewed as a selective constant vigilance threshold: the vigilance threshold is low if the recalled pattern is the mentalizing pattern, and it is high if the recalled pattern is an emotional pattern.

The decision-making loop outputs heuristic rules if the Hamming distance between the two RBM outputs is lower than the vigilance threshold; otherwise it makes deliberative decisions. Since the vigilance threshold for the DLPFC-OFC loop is low, it is unlikely that the DLPFC-OFC loop makes heuristic decisions. For the OFC-amygdala loop, if a mismatch does not occur (i.e. the Hamming distance is less than the vigilance threshold), the decisions are made based on the emotion output from amygdala; otherwise, decisions are made based on sophisticated rules.

The first task of the reinforcement learning is to reinforce the positive patterns stored in the Hopfield network. It works with a variety of adjustable parameters, such as exploration rate, discount factor, and learning rate. Different combinations of the values of these parameters represent the patient in various situations. Secondly, the reinforcement learning aims to reinforce the control from the OFC to amygdala by simply increasing the parameter of the controller in the feedback control system. Such feedback system consists of the OFC as controller, and the amygdala as the plant to be controlled. It aims to adjust the output of amygdala to get close to the output of OFC, which is treated as a reference signal in the control loop.

A. *Experiment*

Initially we store 500 random patterns and the seven basic patterns with the following degrees:

Angry: 10 Fearful: 5 Sad: 3 Disgusted: 1
Surprised: 1 Happy: 1 Mentalizing: 2

The value of high vigilance threshold is set to 16 and the value of low vigilance threshold is set to 2. As described above, if the recalled pattern of the Hopfield network is “mentalizing”, the low value is assigned to the vigilance threshold; otherwise, the high value is assigned to it. Then, the output vectors of the two sensory devices in the loop will be compared, and if the Hamming distance (the number of different bits) is greater than the vigilance threshold value, the object will make deliberative decisions; otherwise, it will make heuristic decisions.

In a single experiment, the patient chooses an action subject to the reinforcement rules, and then the Hopfield network is updated according to the chosen action. Subsequently, a random pattern is used to stimulate the Hopfield network and the output is recorded. This is considered as one iteration. Such iteration will be repeated 200 times in the single experiment.

This experiment is run 1000 times and each time we have recorded the decisions that are made during the Q-learning procedure (200 iterations). As shown in Fig.7, the average numbers to make deliberate decisions over the 1000 runs are essentially monotonically increasing.

V. Conclusion and Future Work

This work can be regarded as a first attempt to develop and implement a quantitative neural model for mentalization/ mindfulness based psychotherapy which includes CBT and psychodynamic therapy.

Here we list a number of themes for future research:

- Improving the performance of the Hopfield network so that it is robust to correlated patterns.
- Using Deep Learning networks to be able to categorise emotions on real human faces, so that the system is able to cope with the real world.
- Using artificial neural network to represent the control from the OFC to amygdala, so that the model is more biologically plausible.
- Modeling the connections between the DLPFC to OFC.
- Designing a learning scheme such that it reinforces the control in terms of time (or iteration in discrete domain).
- Allowing for Long Term Depression of strong patterns by reducing their degrees if they are not recalled for a long time.
- Redesigning the super model using spiking neurons.

Acknowledgment

We would like to thank an anonymous referee for useful comments and additional pointers to the literature, which have improved the presentation of this paper.

References

[1] J. Bowlby, The making and breaking of affectional bonds. Taylor & Francis, 2005.

[2] A. N. Schore, Affect Regulation and the Repair of the Self (Norton Series on Interpersonal Neurobiology). Vol. 2. WW Norton & Company, 2003.

[3] A. N. Schore, Affect Dysregulation and Disorders of the Self (Norton Series on Interpersonal Neurobiology). Vol. 1. WW Norton & Company, 2003, pp. 178-234.

[4] L. Cozolino, The neuroscience of human relationships: Attachment and the developing social brain. WW Norton & Co, 2006.

[5] L. Cozolino, The Neuroscience of Psychotherapy: Healing the Social Brain (Norton Series on Interpersonal Neurobiology). WW Norton & Company, 2010.

[6] A. N. Schore, The Science of the Art of Psychotherapy (Norton Series on Interpersonal Neurobiology). WW Norton & Company, 2012.

[7] J. G. Allen, P. Fonagy, and A. W. Bateman, *Mentalizing in clinical practice*. American Psychiatric Pub, 2008.

[8] D. S. Levine, "Brain pathways for cognitive-emotional decision making in the human animal." *Neural Networks* 22.3 (2009): 286-293.

[9] M. A. Cohen, and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks."*Systems, Man and Cybernetics, IEEE Transactions on* 5 (1983): 815-826.

[10] G. A. Carpenter, and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine." *Computer vision, graphics, and image processing* 37.1 (1987): 54-115.

[11] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the national academy of sciences* 79.8 (1982): 2554-2558.

[12] A. Edalat, and F. Mancinelli, "Strong attractors of Hopfield neural networks to model attachment types and behavioural patterns." *Neural Networks (IJCNN), The 2013 International Joint Conference on*. IEEE, 2013.

[13] A. Edalat, "Capacity of strong attractor patterns to model behavioural and cognitive prototypes." *Advances in Neural Information Processing Systems*. 2013.

[14] R. S. Sutton, and A. G. Barto, Reinforcement learning: An introduction. MIT press, 1998.

[15] M. R. Roesch, D. J. Calu, and G. Schoenbaum, "Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards." *Nature neuroscience* 10.12 (2007): 1615-1624.

[16] G. E. Hinton, and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507.

[17] A. M. C. Aleksandrowicz, and D. S. Levine, "Neural dynamics of psychotherapy: what modeling might tell us about us." *Neural networks* 18.5 (2005): 639-645.

[18] C. R. Cloninger, "A new conceptual paradigm from genetics and psychobiology for the science of mental health." *Australian and New Zealand Journal of Psychiatry* 33.2 (1999): 174-186.

[19] D. S. Levine, "Angels, devils, and censors in the brain." *ComPlexus* 2.1 (2005): 35-59.

[20] R. S. Wedemann, R. Donangelo, and L. A. de Carvalho, "Generalized memory associativity in a network model for the neuroses." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 19.1 (2009): 015116.

[21] R. M. Galatzer-Levy, "Good vibrations: Analytic process as coupled oscillations." *The International Journal of Psychoanalysis* 90.5 (2009): 983-1007.

[22] R. W. Parks, D. S. Levine, and D. L. Long, eds, Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience. Mit Press, 1998.

[23] A. H. Maslow, "Toward a psychology of being." (1968).

[24] J. T. Cacioppo, and R. E. Petty, "The need for cognition." *Journal of personality and social psychology* 42.1 (1982): 116.

[25] D. M. Webster, and A. W. Kruglanski, "Individual differences in need for cognitive closure." *Journal of personality and social psychology* 67.6 (1994): 1049.

[26] C. Eliasmith, and C. H. Anderson, Neural engineering: Computation, representation, and dynamics in neurobiological systems. MIT Press, 2004.

[27] P. Dayan, The role of value systems in decision making. In: C. Engel and W. Singer (eds), Better Than Conscious? Implications for Performance and Institutional Analysis. Cambridge, MA: MIT Press, Strungmann Forum Report, 2008.

[28] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey." *arXiv preprint cs/9605103* (1996).