Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007

Complementary Decision Making System for Aiding Medical Image Diagnosis

T. Z. Tan, G. S. Ng, and C. Quek

Abstract—**Decision making is important in problem solving. Human makes decisions everyday rather effectively. This suggests that functionally model the decision making process could potentially offer a competent decision support system. Hence, a *Complementary Decision Making System* (CDMS), which functionally models the neural decision making in human visual perception, is proposed. It is applied as a clinical decision support system for medical image diagnosis. Its feasibility as a diagnostician's adjunct is assessed in two experiments. The experimental results displayed CDMS as a promising tool for aiding medical image diagnosis.**

I. Introduction

EVERYDAY, human makes many decisions, from what to have for breakfast to which stock to purchase. There are consequences associated with each decision he or she made. Therefore, decision making is a crucial process as it can means lives and deaths. The decision making process is observed not only at the behavioral level, but also at the neural and cellular levels [1][2][3][4][5]!

Decision making is the process of generating, evaluating and selecting an option or course of action from a set of at least two alternatives [6]. It is the bridge between sensation and action [7]. Human interacts with the surrounding using senses, and then he/she decides what action to take. Unfortunately, decision making is nontrivial. Uncertainty, motives, past actions and outcomes, other's decision, rewards, emotions, preference, etc, are some, that influence the decision making [8][9][10]. Yet, human can perform it rather effectively. Thus, functionally model human decision making could offer promising approach to decision support system. Since the task of these systems is to assist user in decision making, the user will find the system more credible and trustable if the system have the decision making process that is similar to him or her.

Accordingly, *Complementary Decision Making System* (CDMS) is proposed as a *Clinical Decision Support System* (CDSS) to medical diagnosis. CDMS is a functional model of the neural decision making in visual perception. It functionally models the visual saccadic decision making in visual cortex [11]. The system is complementary as it uses positive and negative mechanisms towards decision making [9][12]. This exploitation of the relationship between positive and negative could be the mechanism to avoid confusion in inference process. Thus, integrate such complementary mechanism into the CDSS could result superior performance, and at the same time, enhance the system credibility and interpretability.

This paper is organized as follows: Section II describes the neural decision making system that CDMS modeled, the details of CDMS, as well as the experimental setup. Section II discusses the experimental results, and conclusion is given in Section IV.

Manuscript received January 31, 2007.

Authors are with the Centre for Computational Intelligence, School of Computer Engineering, Computer Science Department, Nanyang Technological University, N4 02a-32, Singapore 639798. (E-mail: asgsng@ntu.edu.sg).

II. Methodology

A. Visual Saccadic Decision Making

Visual saccadic decision making is a complementary decision making process which the CDMS modeled. Complementary decision making has three properties:

1. Learning from positive and negative stimuli.
2. Segregation of positive and negative knowledge.
3. Exploitation of the lateral inhibiting relationship between positive and negative.

Positive refers to decision that will bring forth rewards, whereas negative refers to the decision that results punishment. These properties are observed in how the brain responds to a visual motion discrimination task. In this task, the monkey judges the direction of random dot motion with saccadic eye movement. The pattern of complementary decision making can be observed here [11][13][14]:

The *Middle Temporal* visual area (MT, a.k.a V5) encodes the stimulus, which projected into the *Lateral Intraparietal* (LIP) area in *posterior parietal* cortex [13]. In LIP, there are two groups of neurons, whose firing pattern diverge markedly after stimulation. This configuration is for the ease of *lateral inhibition*. Neural populations associated with the preferred direction (high coherence) are activated, whereas the other are inhibited. After accumulation of the firing strength of each direction, the decision is made according to differences between the accumulated strengths correspond to the two directions (see Fig. 1).

In other words, the decision is based on the differences between the accumulated strengths of the opposing sensory pools (*opponent readout*). That is, whenever a rightward motion is detected, the neurons registered for rightward motion will accumulate greater strength than the leftward associated neurons, and these provide evidences for their firing and evidences against the opponent's firing.

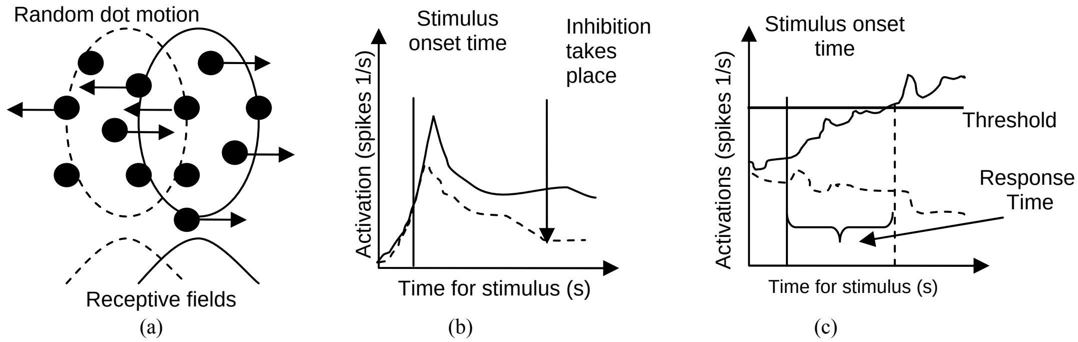

Fig. 1. Neural correlates of a perceptual decision. (a) When presented a set of random dot motion, the medial temporal neuron represents the stimulus direction using receptive fields (dotted line = preferred direction is left, solid line = preferred direction is right). (b) The preferred direction (in this case right) neurons in MT are activated, and at the same time the left-direction related neurons are inhibited. (c) No action is taken until the accumulation of activations reaches the threshold. The time taken from stimulus onset to action is termed response time. The overall strengths of right-direction are larger. Hence, the decision is made: right direction, direct eye movement to the right accordingly.

Subsequently, this decision activates the areas in *Superior Colliculus* (SC) and *Frontal Eye Field* (FEF), which produces the right *saccade* (rapid shift in the lines of sight [11]).

In an identical line of study, the selection of pursuit and saccades corroborate the complementary decision making process as well. The task pertains to the accurate selection of stimulus associated with pursuit target, after the saccade is completed. Neurons in the SC are responsible for the success of distinguishing the target. There are two groups of neurons in SC: the target-related and the nontarget-related neurons. Usually, target is reward or something that brings forth reward. Whenever a target is shown, the activity of target-related neurons is increased and those of nontarget related neurons are suppressed concurrently. This concurrent activation and deactivation leads to the correct saccade movement [14][15].

As shown, the visual saccadic decision making displayed the three characteristics of complementary decision making. Firstly, the brain learned from both positive and negative stimuli (left and right directions, target and nontarget). Secondly, the brain maintained two separate populations of neurons: target-related (positive) and nontarget-related (negative). Lastly, the correct decision is made by the concurrent activation and deactivation of target and nontarget-related neurons respectively. This concurrent activation and deactivation is known as lateral inhibition. The implementation of the three characteristics avoids confusion and ensures right decision is made. Thus, CDSS can be benefited by integrating these traits

The complementary decision making can be formalized as follows:

Let x_i be the sensory stimulus from i^{th} sensory modality, where $i \in [1, \mathrm{I}]$, I = total number of sensory modality, and A_{ij} be the j^{th} sensory neuron that represents x_i, where $j \in [1, \mathrm{J}]$, J = total number of sensory neurons. This A_{ij} is characterized by its receptive field $\mu_{A_{ij}}(\bullet)$, which subsequently projects its response to the neurons in prefrontal cortex, denoted as r_k, where $k = j, k \in [1, \mathrm{K}]$. The set of r_k that reaches the action threshold will activate its associated motor representation in motor cortical and subcortical structure, $B_{(lm)_k}$ (defined by $\mu_{B_{(lm)_k}}(\bullet)$), where $l \in [1, \mathrm{L}]$ is the number of motor representation, and $m \in [1, \mathrm{M}]$ corresponds to the number of actions, y. Superscript '+' and '-' are used to denote elements that are associated with positive and negative neurons respectively. Given a set of sensory stimuli $\mathbf{x} = (x_1, x_2, ..., x_\mathrm{I})$, a decision is made according to (1).

$$y_m = \begin{cases} \mu^{+}_{B_{(lm)_{k^+}}}(\mathbf{x}) \times \max_{k^+} \mu_{k^+}(\mathbf{x}), & \text{if } \mu_{r^+}(\mathbf{x}) > \mu_{r^-}(\mathbf{x}) \\ \mu^{-}_{B_{(lm)_{k^-}}}(\mathbf{x}) \times \max_{k^-} \mu_{k^-}(\mathbf{x}), & \text{otherwise} \end{cases} \tag{1}$$

where $\mu_{r^+}(\mathbf{x}) = \sum_{k=1}^{\mathrm{K}^+} \sum_{i=1}^{\mathrm{I}} \mu_{A_{ik}}(x_i)$, $\mu_{r^-}(\mathbf{x}) = \sum_{k=1}^{\mathrm{K}^-} \sum_{i=1}^{\mathrm{I}} \mu_{A_{ik}}(x_i)$

B. *Complementary Decision Making System*

The complementary decision making system models the visual saccadic decision making functionally. It is modeled using a five layer *Fuzzy Neural Network* (FNN), for a number of reasons. FNN exhibits characteristics that fit well with the features of decision making. For example, FNN interacts with the environment through taking in examples, whereas the decision making process starts by taking in stimuli through senses. Moreover, FNN is capable of capturing uncertainty by fuzzy sets. This is important, because factors that affect one's decision always comprise certain level of fuzziness. Apart from that, the "decision making" of FNN is closely akin to the human decision making process. Besides, the 5-layer architecture made the one-to-one mapping of the neural decision making possible (see Fig. 2). With these features, FNN allows explanation from the viewpoint of psychology and neurobiology, which adds credibility to the system.

As shown in Fig. 2, CDMS deducts its action based on the neurobiology of decision making. CDMS functionally models the three characteristics of complementary decision making:

1. CDMS constructs its knowledge base from both positive and negative examples. These knowledge bases are formed of fuzzy rules. The positive and negative fuzzy rules are data-driven, and are generated separately.

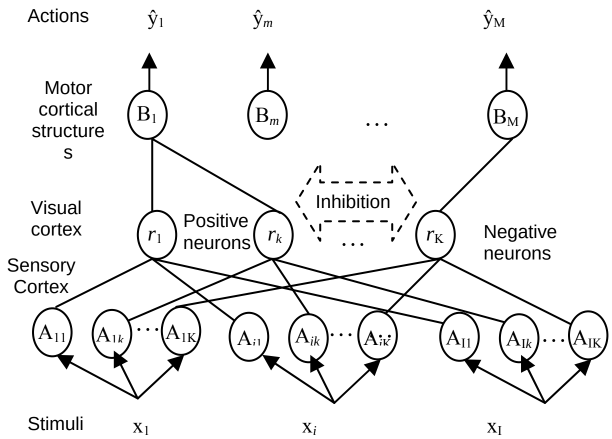

Fig. 2. The architecture of CDMS. Each layer corresponds to the steps in human decision making. The first layer is analogous to the sensory cortex, where the representation of the input stimuli is formed using fuzzy sets. The neuron in second layer resembles the visual cortical neuron. These neurons determine how the system behaves. Finally, the third layer corresponds to the motor cortical structures, where the motor representation is stored.

2. CDMS segregates the positive and negative knowledge in its architecture. This can be seen in Fig. 2. This configuration is to ease the exploitation of lateral inhibition.
3. During inference, CDMS exploits the lateral inhibition between the positive and negative knowledge. If the evidence favors the positive, CDMS will promote positive and suppress the negative.

However, the architecture in Fig. 2 is not sufficient for modeling the decision making. What is lacking from the architecture is the system that learns from the outcome of the decision. Human being has the capacity to learn from the outcome of their decision. If it is a positive outcome, i.e. reward, the decision is strengthened. If the outcome is undesirable or is a punishment, then change the decision so as to avoid the same mistake. Hence, in order to model this, a reward system is incorporated into the CDMS, as shown in Fig. 3. This reward system learns from the CDM and then reinforces good decision and change bad decision.

Although CDMS models the neurobiology of decision making, it is not as straightforward for user without biological background. Hence, a tractable system should be capable of mapping the human psychological decision making process. This is needed so that user can analyze the system using their familiar term. Since CDMS is to be realized as CDSS, hypothetical-deductive reasoning adopted by physicians [10] is mapped onto CDMS. A hypothetical-deductive reasoning named *Analogical Approximate Reasoning Schemes* (AARS [16]) is used. The mapping not only formalizes the operation of CDMS, but also equips CDMS with a decision making process that is analogous to the physician, as depicted in Table 1.

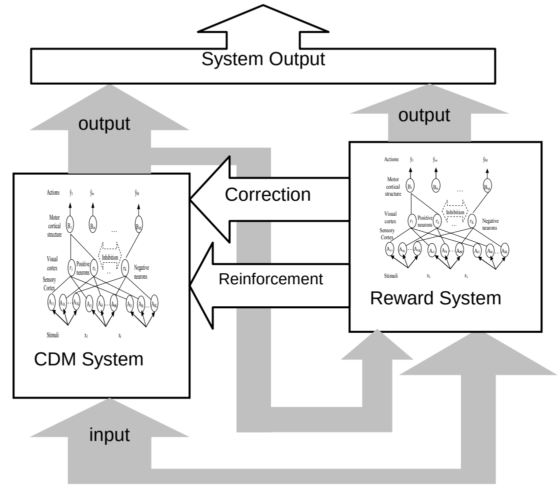

Fig. 3. The CDMS system. A CDMS consists of a CDM system and a reward system. The reward system serves as a teacher for the CDM system. If the CDM system makes a wrong decision, a correction signal is sent to the CDM system. If the CDM system makes a correct decision, then a reinforcement signal is sent instead. The final decision of CDMS is based on the outputs of CDM and reward system.

TABLE I: CDMS AND PHYSICIAN DECISION MAKING PROCESS

Steps	Physician	CDMS (AARS)
1	Observe the symptoms and ask the patient.	Take in the inputs from the data.
2	Assess similarity of present input with past knowledge and experience.	Perform the antecedent matching process, which compares existing knowledge with current input.
3	Select the best experience or knowledge that describes current situation.	Select the best rule that describes the current input.
4	Decides the appropriate response based on the selected knowledge.	Choose the appropriate output representation based on the selected rule.
5	Act based on the decisions.	Output the decision suggested by the output representation.

As shown, the decision making process is akin to that of the physician. Note that Step 3 is where the decision making takes place. When a sample is presented to CDMS, the CDMS contrasts the firing strengths of positive and negative rules. If it is a positive sample, then allow the positive rules to fire, and inhibit the negative rules, and vice versa, as defined in (2).

$$y_m = \begin{cases} \mu^{+}_{B_{(lm)_{k^+}}}(\mathbf{x}) \times \max\limits_{k^+} \mu_{k^+}(\mathbf{x}), & \text{if } \mu_{r^+}(\mathbf{x}) > \mu_{r^-}(\mathbf{x}) \\ \mu^{-}_{B_{(lm)_{k^-}}}(\mathbf{x}) \times \max\limits_{k^-} \mu_{k^-}(\mathbf{x}), & \text{otherwise} \end{cases} \quad (2)$$

where $\mu_{r^+}(\mathbf{x}) = \max\limits_{k} \sum_{k=1}^{K^+} \sum_{i=1}^{I} \mu_{A_{ik}}(x_i)$, and $\mu_{r^-}(\mathbf{x}) = \max\limits_{k} \sum_{k=1}^{K^-} \sum_{i=1}^{I} \mu_{A_{ik}}(x_i)$

As shown, (2) is very similar to the formalized neurobiological decision making listed in (1). These complementary sets of rules are constructed autonomously by CDMS based on the training data, and the construction is briefly, described below.

Given inputs **x** and desired actions **y**, the CDMS first forms the sensory representation of **x** and **y** using trapezoidal membership functions **A** and **B** respectively. Trapezoidal membership functions are used because of their lower computational requirement. A rule $r_k : \mathbf{A}_k \rightarrow \mathbf{B}_k$ is then formed to associate the input to the desired action.

Let $t \in [1,\mathrm{T}]$ = index for samples, $p \in [1,\mathrm{P}]$ = index for epochs. The learning of CDMS is as follows:

Given a set of sensory stimuli $\left(\mathbf{x}^t = \left(x_1^t, x_2^t, ..., x_{\mathrm{I}}^t\right), y^t\right)$:

Step 1: If there are existing rules, go to Step 2. Otherwise, construct the rule based on the following input-output pair. The rule will form input and output clusters as the antecedents and consequents respectively. The kernel of the membership function will center at the input data, and the support (max and min) will be preset to the maximum and minimum value of that dimension. i.e., $Max_i = \left(x_i^1, x_i^2, ..., x_i^{\mathrm{T}}\right)$, $Min_i = \left(x_i^1, x_i^2, ..., x_i^{\mathrm{T}}\right)$, and $\left(u_{ij}, v_{ij}\right) = \left(x_i^t, x_i^t\right)$. Then go to step 4.

Step 2: Compute the firing strength of each rule using (3).

$$\mu_{r_k}(\mathbf{x}) = \frac{1}{\mathrm{I}} \sum_{i=1}^{\mathrm{I}} \mu_{\mathbf{A}_{ij}}(x_i) \tag{3}$$

Step 3: Find the maximally fired rule from the positive and negative experiences using (4) and (5), respectively.

$$r^{*+} = \arg\max_{k^+} R^+ = \arg\max_{k^+} \bigcup_{k=1}^{\mathrm{K}^+} r_k^+ \tag{4}$$

$$r^{*-} = \arg\max_{k^-} R^- = \arg\max_{k^-} \bigcup_{k=1}^{\mathrm{K}^-} r_k^- \tag{5}$$

If the current desired output is positive, then look for the maximally fired positive rule. If this rule is sufficiently close to the input, then generalize this rule to incorporate the input, else create new rule for this input-output pair. The definition of "close" is according to the user-defined threshold, ρ, which can be ρ^+ or ρ^-. The positive and negative populations have their own threshold settings, i.e. $\rho^+ \neq \rho^-$. When the threshold is surpassed, the rule is adjusted according to the input. The learning is based on simple updating, and the learning rate is an exponentially decreasing function that simulates the aging process, as described by (6) and (7). If there is no similar experience, the rule will not be updated, and create a new rule based on the present input-output pair.

$$\left(u_{ij}^{p+1}, v_{ij}^{p+1}\right) = \begin{cases} \left[u_{ij}^p - \beta^t * \left(u_{ij}^p - x_i^t\right), v_{ij}^p + \beta^t * \left(x_i^t - v_{ij}^p\right)\right], & \text{if } \mu_{r^*}(x) > \rho \\ \left(u_{ij}^p, v_{ij}^p\right), & \text{Otherwise} \end{cases} \tag{6}$$

where

$$\beta^p = \eta * \exp\left(-\frac{(p-1)^2}{Age}\right) \tag{7}$$

and η = initial plasticity, and β^p = learning constant at epoch *p*. *Age* is a preset parameter that determines the plasticity of the system, and hence, it is the aging rate of the system. Together, the η and *Age* monitor the plasticity of system. In general, larger η and *Age* mean larger plasticity. The gradually decreasing of the system plasticity, aside from modeling the aging process, is to avoid the instability that may happen when the data is noisy.

Step 4: Check if terminating condition is met. This can be done by setting the maximum iteration or based on some reduction of cost function. In this study, the training is terminated if the overall accuracy achieved by the system is greater than some preset threshold (8). Unfortunately, there is no systematic way of determining this variable.

$$Learning = \begin{cases} t \rightarrow t+1, \text{Go to Step 1}, & \text{if } Accuracy > \rho_{threshold} \\ & \quad || \ p < Max\_iteration \\ \text{Terminate Learning}, & \text{Otherwise} \end{cases} \tag{8}$$

While the training of CDM is in progress, the reward system learns to predict the CDM system behavior, as well as the potential reward it will receive. The learning is exactly the same as the CDMS, except that it has two additional inputs: the prediction of CDM system, and the fitness value of the rule. That is, the input-output pair fed into the reward system is $\left[\left(x_1^t, x_2^t, ..., x_{\mathrm{I}}^t, y_1^t, y_2^t, ..., y_{\mathrm{M}}^t, fitness_{r*}\right)\left(d_{reward,1}, d_{reward,2}, ..., d_{reward,\mathrm{M}}\right)\right]$. d_m^t is the desired output of the m^{th} dimension, $d_{reward,m}$ is the m^{th} desired output of the reward system for the given input, and is delineated in (9).

$$d_{reward,m} = \begin{cases} 1, & \text{if } y_m^t = d_m^t \\ 0, & \text{Otherwise} \end{cases} \tag{9}$$

Whenever there is a real reward (CDM elicits desired response, i.e. $d_{reward,m} = 1$), the reward system will strengthen the rule responsible for the correct behavior, and punished it otherwise. In this work, this behavior is done by giving each rule a fitness value based on the reward or punishment they received, as shown in (10).

$$fitness_{r*}^{p+1} = \begin{cases} fitness_{r*}^p + 1, & \text{if } d_{reward,m} = 1 \\ fitness_{r*}^p, & \text{Otherwise} \end{cases} \tag{10}$$

This fitness value can be viewed as how credible the rule is, in performing the task. Thus, when the CDMS decides, it considers the output of the CDM network, the output of the reward system, and the fitness value of the rule responsible for that decision. Given a positive sample $\mathbf{x}^t = \left(x_1^t, x_2^t, ..., x_{\mathrm{I}}^t\right) \in X^+$, CDMS derives its decision:

1. Presentation of the sample to the system,

$\mathbf{x} = (x_1, x_2, ..., x_I)$.

2. Antecedent matching of the positive and negative rules, i.e. compute $\mu_{\mathbf{A}_{ij}}(x_i)$ for every rules.
3. Rule fulfillment is computed according to $\mu_{r_k}(\mathbf{x}) = \frac{1}{I}\sum_{i=1}^{I}\mu_{\mathbf{A}_{ij}}(x_i)$, and rule with maximum firing strength is selected, $r_k^* = \begin{cases} r_k^{*+}, & \text{if } \arg\max_k \{R^+\} > \arg\max_k \{R^-\} \\ r_k^{*-}, & \text{Otherwise} \end{cases}$.
4. Consequent derivation based on the selected rules. In this phase, the predictions of the CDM network and the reward system are compared. In this work, the derivation of conclusion is based on the output of the CDM network, the output of reward system, and the fitness of the rule (see (11)).

$$y_{cdms,m} = \begin{cases} +, & \text{if } y_m = + \,\&\, d_{reward,m} = 1 \\ +, & \text{if } y_m = - \,\&\, d_{reward,m} = 0 \,\&\, fitness_{r^{*-}} < \rho \\ +, & \text{if } y_m = + \,\&\, d_{reward,m} = 0 \,\&\, fitness_{r^{*+}} > \rho \\ -, & \text{if } y_m = - \,\&\, d_{reward,m} = 1 \\ -, & \text{if } y_m = + \,\&\, d_{reward,m} = 0 \,\&\, fitness_{r^{*+}} < \rho \\ -, & \text{if } y_m = - \,\&\, d_{reward,m} = 0 \,\&\, fitness_{r^{*-}} > \rho \end{cases} \quad (11)$$

5. Deduct conclusion and output the conclusion to the user; in this case, $y_1 = +$.

Thus, CDMS has a reasoning process highly akin to that of human. This facilitates the system validation and justification. By making use of CDM, the lateral inhibition of positive and negative classes is exploited, and hence, may offer superior performance. Furthermore, the employment of reward system enables CDMS to take into account the probability when making the decision. As such, CDMS is believed to be a competent CDSS.

C. Experiment Setup

Two datasets are used: Wisconsin diagnostic breast cancer [17]. The dataset has 569 instances (357 benign and 212 malignant). Each instance has 30 features extracted from the *Fine needle aspiration* (FNA) biopsy images, such as radius of the cell nucleus, texture, perimeter, etc.

The second dataset is a FNA breast lesions dataset. It is obtained from the Department of pathology, Royal Hallamshire Hospital, Sheffield [18]. The dataset comprise 692 (235 malignant and 457 benign) consecutive specimens of FNA breast lumps. Each specimen is described by 11 features such as the cellular dyshesion, intracytoplasmic lumina, etc.

The datasets are divided into three stratified cross-validation sets. The training set contains 1/3 of the dataset, whereas the testing set contains the remaining unseen 2/3. All features are employed to perform the diagnosis.

The performance of the system is benchmarked against popular decision support systems [19] such as *Naïve Bayesian* (NB), *Multilayer Perceptron* (MLP), *Logistic Regression* (LR), *Support Vector Machine* (SVM), *Linear Discriminant Analysis* (LDA), and C4.5. Note that the same experimental settings are used for all the methods presented. All simulations are run on Intel Pentium 4, 2 GHz machine with 1 GB RAM. For SVM, the kernels used are linear, polynomial, sigmoid, and radial basis function. The presented performance is the best SVM result using one of these kernels. The metrics used are

- Recall – accuracy on training set;
- Predict – accuracy on testing set;
- Sensitivity and specificity, as described by (12) and (13), respectively.

In diagnosis, the positive concerns the malignant case, and the negative concerns the benign case.

$$\text{Sensitivity} = \frac{\text{Positive samples correctly classified}}{\text{Total number of positive samples}}. \quad (12)$$

$$\text{Specificity} = \frac{\text{Negative samples correctly classified}}{\text{Total number of negative samples}}. \quad (13)$$

III. Experimental Results

The averaged results are summarized in Tables 2 and 3. From the tables, CDMS illustrates its superior performance compared against other common tools for CDSS. In the two experiments, CDMS shows superior recall and generalization ability, hitting close to 100% accuracy even when only 1/3 data is used for training. When benchmarked against results reported in past usage, CDMS outperform some of the system in spite of it uses less samples for learning. For example, WDBC was achieved 93.6% by SVM built upon 1/3 data [20], 96%-97% by Fuzzy Coco [[21] and NeuroRule [22] built upon 50% of the data. As for the second dataset, the accuracy obtained by experts and LR is 94% and 95% respectively [18]. This suggests that CDMS outperforms the conventional CDSS in generalization.

TABLE II: PERFORMANCE ON WDBC DIAGNOSIS

Method	Recall	Sens.	Spec.	Predict	T. Time
NB	93.7±0.9	95.5±2.1	90.0±4.6	93.8±1.1	0.1
MLP	99.8±0.3	96.8±2.5	92.7±4.4	95.7±1.3	10.24
LR	100	95.0±1.6	91.9±1.1	94.0±1.1	0.17
SVM	97.5±0.3	99.2±1.1	91.7±6.6	96.8±1.1	0.68
LDA	98.0±0.6	89.3±5.8	98.5±0.9	95.8±0.9	0.04
C4.5	98.6±0.8	96.4±2.3	90.5±4.5	94.5±1.7	0.09
CDMS	97.5±0.4	97.8±0.3	99.6±0.7	98.9±1.0	2.35

TABLE III: PERFORMANCE ON FNA BREAST LESIONS

Method	Recall	Sens.	Spec.	Predict	T. Time
NB	94.5±2.1	89.8±5.1	96.9±1.4	94.5±1.1	0.42
MLP	98.5±0.9	88.8±3.9	97.5±0.7	94.5±1.6	2.43
LR	97.5±2.2	90.2±5.9	97.0±1.2	94.7±2.7	0.11
SVM	96.9±1.3	92.6±1.7	97.4±1.1	95.7±0.5	0.09
LDA	96.1±2.2	86.6±3.8	99.0±0.3	94.8±1.1	0.02
C4.5	96.5±0.4	86.9±6.5	97.5±1.3	93.8±1.4	0.01
CDMS	98.9±0.1	98.8±0.2	99.0±0.1	98.9±0.1	1.28

Abbreviation: Sens – Sensitivity, Spec – Specificity, T. Time – training time. Recall, Sens, Spec. and Predict are in percentage. T. Time is in second.

CDMS has an acceptable training time despite its architecture is relatively more complicated. The training time of CDMS is also shorter than the MLP system. For system response time, all systems responded in less than a second.

A significant strength of CDMS is the ability to generate fuzzy rules to explain its computation. These rules capture the uncertainties in the data, and may give insights into the problem. In contrast to the crisp rule generated by decision tree, fuzzy rules allows the use of linguistic hedges such as "very", "rather", and hence is more expressive. Furthermore, fuzzy rule employs linguistic terms to encapsulate unnecessary details from the user, which ease the interpretation of the rule. Examples of the rules generated in FNA diagnosis are listed in Table 4. CDMS has two sets of rules: one from the CDM system and one from the reward system. The rules of reward system describe the behavior of the CDM system. Note that the negative rule of CDMS reward system, when the CDM system is not confident, the reward system may output the opposite diagnosis. In such cases, the CDMS output will follow the output of the system with higher fitness or confidence level.

TABLE 4: FUZZY RULES FOR BREAST CANCER DIAGNOSIS

Rules	CDMS (CDM)	CDMS (Reward)	Decision Tree
+	**IF** x1 is very low, x2 is very high, x3 is high, x4 is…. **THEN** Malignant.	**IF** x1 is very low, x2 is very high, x3 is high, x4 is…, **AND** CDM output is Malignant, **AND** fitness value is high **THEN** Malignant.	**IF** x7 ≤ 0, x3 ≤ 0, x1 > 54, x4 ≤ 0, x1 > 75, **THEN** Malignant
-	**IF** x1 is high, x2 is medium, x3 is marginal high, …, **THEN** Benign.	**IF** x1 is high, x2 is medium, x3 is marginal high, …, **AND** CDM output is Malignant, **AND** fitness value is very low, **THEN** Benign.	**IF** x7 > 0, x9 ≤ 0, x1 ≤ 59, **THEN** Benign

The rules of CDMS use all the features for its antecedents. The one shown here is simplified version.

However, CDMS constructed more rules than decision tree algorithm. The fact that CDMS has two rulebase systems exacerbates the situation. The numbers of rules generated by the CDMS and reward systems are 34 and 27 respectively. Whereas decision tree C4.5 generates only 6 rules. In this aspect, CDMS is less efficient than C4.5. Nevertheless, the performance of CDMS is higher than C4.5 in terms of prediction accuracy. There is a trade-off between the number of rules and prediction accuracy. One can according to his or her need, adjust the number of rules, with some sacrifices on the accuracy.

IV. CONCLUSION

In this work, a functional model of human visual saccadic decision making is proposed. The CDMS employs the way brain ensures rewarding decision making. Experimental results demonstrated the superior performance of CDMS in diagnosis, as well as competent generalization. On top of that, CDMS autonomously constructs its knowledge to describe and to monitor its operation. The inference process of CDMS is highly akin to human as well. Unfortunately, CDMS has its limitation. Since CDMS requires explicit generation and segregation of positive and negative knowledge, CDMS has larger rulebase. To complicate the matters, CDMS has two rule systems. This situation can be alleviated by introducing more efficient rule construction or rule pruning algorithm to CDMS.

REFERENCES

[1] K. L. Briggman, H. D. I. Abarbanel, W. B. Kristan Jr, "Optical imaging of neuronal populations during decision-making", *Science*, vol. 307, 896-901, 2005.

[2] G. S. Berns, T. J. Sejnowski, "How the basal ganglia make decisions", in A. R. Damasio, H. Damasio, Y. Crhsiten (Eds.), *Neurobiology of Decision Making*, Heidelberg: Springer-Verlag, 101-114, 1996,

[3] J. I. Gold, M. N. Shadlen, "Banburismus and the brain: decoding the relationship between sensory, stimuli, decisions, and reward", *Neuron*, vol. 36, 299-308, 2002.

[4] M. J. Frank, E. D. Claus, "Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making and reversal", *Psychological Review*, vol. 113, 300-326, 2006.

[5] H. Korn, D. S. Faber, "The Mauthner cell half a century later: a neurobiological model for decision making? *Neuron*, vol. 47, 13-28, 2005.

[6] A. Marschner, T. Mell, A. Villringer, F. M. Reischies, H. R. Heekeren, "Reward-based decision-making and aging", *Brain Research Bulletin*, vol. 67, 382-390, 2005.

[7] D. Lee, B. P. McGreevy, D. J. Barraclough, "Learning and decision making in monkeys during a rock-paper-scissors game", *Cognitive Brain Research*, vol. 25(2), 416-430, 2005.

[8] M. P. Paulus, "Neurobiology of decision making: quo vadis?" *Cognitive Brain Research*, vol. 23, 2-10, 2005.

[9] P. Slovic, E. Peters, M. L. Finucane, D. G. MacGregor, "Affect, risk, and decision making", *Health Psychology*, vol. 24, S35-S40, 2005.

[10] C. Evans, "Clinical decision making theories: patient assessment in A&E", *Emergency Nurse*, vol. 13(5), 16-19, 2005.

[11] P. W. Glimcher, "The neurobiology of visual-saccadic decision making", *Annual Review of Nueroscience*, vol. 26, 133-179, 2003.

[12] A. Bechara, H. Damasio, A. R. Damasio, "Emotion, decision making and the orbitofrontal cortex", *Cerebral Cortex*, vol. 10, 295-307, 2000.

[13] J. D. Shall, "Neural correlates of decision processes: neural and mental chronometry", *Current Opinion in Neurobiology*, vol. 13, 182-186, 2001.

[14] X. –J. Wang, "Probabilistic decision making by slow reverberation in cortical circuits", *Neuron*, vol. 36, 955-968, 2002.

[15] M. L. Platt, "Neural correlates of decisions", *Current Opinion in Neurobiology*, vol. 12, 141-148, 2002.

[16] I. B. Turksen, Z. Zhong, "An approximate analogical reasoning schema based on similarity measures and interval-valued fuzzy sets", *Fuzzy Sets and System*, vol. 34, 323-346, 1990.

[17] C. L. Blake, C. J. Merz, UCI Repository of machine learning database. Irvine CA: University of California, Department of Information and Computer Science. Available: http://www.ics.uci.edu/~mlearn/MLR-epository.html.

[18] S. Cross, R. F. Harrison, "Fine needle aspirate of breast lesions dataset", Artificial Neural Networks in Medicine World Map, 2006. Available: http://www.phil.gu.se/ann/annworld.html

[19] I. H. Witten, E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, San Francisco: Morgan Kaufmann, 2005.

[20] J. T. –Y. Kwok, "The evidence framework applied to support vector machines", *IEEE Transactions on Neural Networks*, vol. 11(5), 1162-1173, 2000.

[21] C. A. Penä-Reyes, *Coevolutionary Fuzzy Modeling*, Lecture Notes in Computer Sciences, vol. 3204, Heidelberg: Springer-Verlag, 2004.

[22] R. Setiono, "Generating concise and accurate classification rules for breast cancer diagnosis", *Artificial Intelligence in Medicine*, vol. 18(3), 205-219, 2000.