Proceedings of the International Symposium on Evolving Intelligent Systems, Plamen Angelov, Dimitar Filev, and Nikola Kasabov (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Dynamic Analysis of the Participatory Learning Algorithm

Elton Lima, Rosangela Ballini and Fernando Gomide

***Abstract*— This paper concerns the analysis of the dynamic behavior of the participatory learning process considering the compatibility and the arousal mechanisms simultaneously with the learning rate. The purpose is to study the convergence characteristics of participatory learning algorithm when learning data are a continuous and differentiable function of the learning variable. The main feature of the participatory learning model is variable compatibility and arousal values that entails variable learning rates. This feature involves the important problem of proving the stability and convergence of the learning system for this condition.**

I. Introduction

Participatory learning is an approach for learning that emphasizes the role of the current knowledge as a part of the learning process itself. Contrary to many learning paradigms, the participatory learning process includes a compatibility mechanism to let current knowledge influence how data affect learning. The impact of data in causing learning and knowledge revision depends on compatibility of data with the current knowldege [1].

One of the main applications of participatory learning is in incremental clustering, an essential element of information processing and machine learning systems [2]. Participatory learning also is part of many instances of data mining, evolving system modeling [3], control, forecasting [4] and classification systems.

Yager suggested in [1] an instance of participatory learning in the form of a basic exponential smoothing like model. The view of participatory learning as a combination of two regulatory mechanisms, participatory feedback mechanism and participatory arousal mechanism, was pursued in [5]. In particular, it was shown that the basic feedback participatory mechanism is stable whenever the compatibility of data with the current knowledge stays within the unit interval. In the present paper we analyse the dynamic behavior of the participatory learning process considering the compatibility and the arousal mechanisms simultaneously with the learning rate. More specifically, we look at the convergence of participatory learning algorithm when learning data are a function of the learning variable.

II. Participatory Learning

An essential feature of the participatory learning paradigm is that input data has the greatest impact in causing learning and knowledge revision when it is compatible with the current system knowledge. In particular, data too conflicting with the current knowledge are discounted. Let $v \in [0,1]^p$ be the object of learning, a variable encoding the knowledge of a system. We may say that our aim is to learn a value for v using data $x^k \in [0,1]^p$, $k = 1, 2, \ldots$. The current knowledge of the system consists of the vector v^k whose components, v_i^k, $i = 1, 2, \ldots, p$, are the system current knowledge about the values of the x_i^k. This is what has been learned after $k-1$ data are input.

Using the participatory learning mechanism [1], the current knowledge is updated using the following expression

$$v^{k+1} = v^k + \alpha G^k (x^k - v^k) \quad (1)$$

where $v^k \in [0,1]^p$, $\alpha \in [0,1]$ is the basic learning rate, and $G^k = G(\rho(x^k, v^k))$ is a function of the compatibility $\rho(x^k, v^k)$ of data x^k with current knowledge v^k at $k = 1, 2, \ldots$.

Compatibility can be viewed as an acceptance mechanism. The larger the $\rho(x^k, v^k)$, the more compatible the data is with the current knowlege. If most of data is compatible with the current knowledge, then we can use the portion of the data that deviates from the current knowledge to learn. A simple choice is $G^k = \rho^k$, that is, G^k is the identity function. This choice gives the basic participatory learning expression

$$v^{k+1} = v^k + \alpha \rho^k (x^k - v^k) \quad (2)$$

where

$$\rho^k = 1 - d(x^k, v^k) \quad (3)$$

and d is a distance measure between x^k and v^k.

Another important mechanism in participatory learning is the arousal mechanism. The arousal mechanism is an autonomous element, independent of the current knowledge, that observes the performance of the compatibility. The arousal mechanism controls compatibility to consider the possibility that the current knowledge may be erroneous. It essentially does this by noting how often incompatible data have been input. Its effect is not manifested by an individual incompatibility, but by an accumulation of a continued stream of low compatibilities. This should make the learning process more susceptible to learning. If too many observations are rejected as being incompatible with the current knowledge, then current knowledge should be revised.

The arousal mechanism can be viewed as a value $a^k \in [0,1]$ called the arousal rate. It is inversely related to the confidence in the model based upon its past performance. The smaller the a^k, the more credible the current knowledge is. The larger the arousal rate, the more doubtful is the current

Elton Lima and Fernando Gomide are with Faculty of Electrical and Computer Engineering, University of Campinas, Campinas, Brazil (gomide@dca.fee.unicamp.br).

Rosangela Ballini is with the Institute of Economics, University of Campinas, Campinas, Brazil (email: ballini@eco.unicamp.br).

knowledge. The expression to compute a^k suggested in [1] is the following

$$a^{k+1} = a^k + \beta(1 - \rho^{k+1} - a^k) \tag{4}$$

where $\beta \in [0, 1]$ is a learning rate, $k = 1, 2, \ldots$.

Figure 1 overviews the participatory learning process and the interplay between the compatibility and arousal mechanisms.

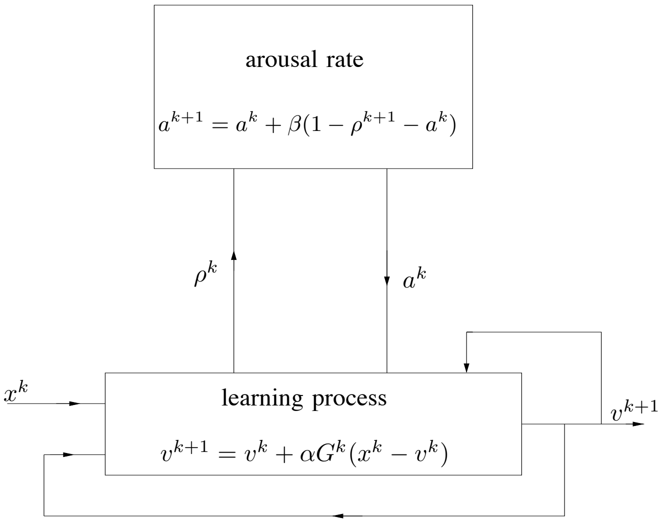

Fig. 1. Participatory learning process.

The complete participatory learning [6] chooses $G^k = (\rho^k)^{1-a^k}$ which gives the prototypical participatory learning process

$$v^{k+1} = v^k + \alpha(\rho^k)^{1-a^k}(x^k - v^k) \tag{5}$$

III. Dynamic Analysis of Participatory Learning

In what follows we assume, without loss of generality, that $p = 1$. Let us also suppose that the inputs $x^k = f(v^k)$, that is, values x^k, $k = 1, 2, \ldots$ are a continous function $f : [0, 1] \rightarrow [0, 1]$ of the learning variable. In this case, the dynamics of the participatory learning process can be expressed by

$$v^{k+1} = v^k + \alpha^k G^k(f(v^k) - v^k) \tag{6}$$

Since function f is continuous in $[0, 1]$, there exists at least a point $\overline{v} \in [0, 1]$ such that $f(\overline{v}) = \overline{v}$, that is, $\overline{v}$ is a fixed point of f [7]. Indeed, if we let $h(v) = f(v) - v$, then $h : [0, 1] \longrightarrow [-1, 1]$ is also continuous and, because the origin $0 \in [-1, 1]$, from the intermediate value theorem [7] we know there exists $\overline{v} \in [0, 1]$ such that $h(\overline{v}) = 0$.

Assume the participatory learning process expressed by (6).

Lemma 1: Let $f : [0, 1] \longrightarrow [0, 1]$ be a continuous function and $I \subseteq [0, 1]$ an interval in which $f'(v) = \dfrac{df(v)}{dv}$ also is continuous for $\overline{v} \in I$. If $f'(v) < 0$ in I, then $\lim_{k\to\infty} v^k = \overline{v}$ if $\alpha^k \in (0, 1]$, that is, the sequence $\{v^k\}$ produced by (6) converges to a fixed point of f.

Proof: Let $v^1 \in I$; two cases are of interest (except the obvious case $v^1 = \overline{v}$):

(i) $f(v^1) - v^1 > 0$
(ii) $f(v^1) - v^1 < 0$

(i): In the first case $f(v^1) - v^1 > 0$ and for $\alpha^k G^k \in [0, 1]$ we have:

$$\begin{array}{ll} k = 1: & v^2 = v^1 + \alpha^1 G^1(f(v^1) - v^1) > v^1; \\ k = 2: & v^3 = v^2 + \alpha^2 G^2(f(v^2) - v^2) > v^2; \\ \vdots & v^{k+1} > v^k \end{array}$$

which suggests, as we show below (see also Figure 2), that $(f(v^k) - v^k)$ does not changes sign. Thus, $\{v^k\}$ is a limited, monotone increasing sequence and hence there is a $\overline{v}$ such that $\lim_{k\to\infty} v^k = \overline{v}$ [7].

Knowing that the sequence $\{v^k\}$ has a limit, to show that $f(\overline{v}) = \overline{v}$, namely, that $\overline{v}$ is a fixed point, it suffices to note from (6) that

$$\lim_{k\to\infty} v^{k+1} = \lim_{k\to\infty} v^k + \lim_{k\to\infty} \alpha^k G^k(f(v^k) - v^k)$$

$$\lim_{k\to\infty} \alpha^k G^k(f(v^k) - v^k) = 0$$

and, because $\alpha^k G^k$ is by definition finite, $\lim_{k\to\infty}(f(v^k) - v^k) = 0$.

Therefore, once f is continuous we have

$$f(\lim_{k\to\infty}(v^k)) = \lim_{k\to\infty} v^k \quad \Rightarrow \quad f(\overline{v}) = \overline{v}$$

which means that $\overline{v} = \sup_k\{v^k\}$. In other words, the fixed point is the upper boud of the sequence $\{v^k\}$ [7].

Next, we need to show that $(f(v^k) - v^k)$ indeed does not change sign in I. Suppose there is a j for which $(f(v^j) - v^j) > 0$ and $(f(v^{j+1}) - v^{j+1}) < 0$. In this case

$$f(v^j) - f(v^{j+1}) - v^j + v^{j+1} > 0$$

$$\frac{f(v^j) - f(v^{j+1})}{v^j - v^{j+1}} > 1 > 0$$

From the mean value theorem [7], there is a $v^* \in (v^j, v^{j+1})$ with $f'(v^*) > 1 > 0$, which contradicts the assumption that $f'(\overline{v}) < 0$ because the fixed point $\overline{v} \in (v^j, v^{j+1})$. Therefore, we conclude that $(f(v^k) - v^k)$ does not change sign for $k = 1, 2, \ldots$.

(ii): The proof when $f(v^1) - v^1 < 0$ is similar. Note, however, that in this case $\{v^k\}$ is a limited, monotone decreasing sequence with fixed point $\overline{v} = \inf_k\{v^k\}$, that is, the fixed point is the lower bound of $\{v^k\}$. ■

When $\overline{v}$ is a fixed point, but $f'(\overline{v}) > 0$, the conditions for convergence of the sequence $\{v^k\}$ are as given by the following lemma.

Proceedings of the International Symposium on Evolving Intelligent Systems, Plamen Angelov, Dimitar Filev, and Nikola Kasabov (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

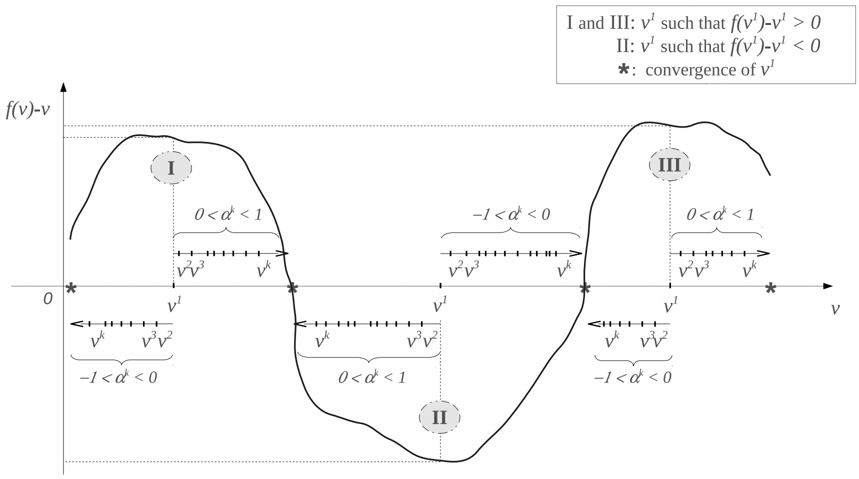

Fig. 2. Behavior of the sequence $\{f(v^k) - v^k\}$ for different initial values v^1.

Lemma 2: Consider the hypothesis of Lemma 1 and $f'(\overline{v}) > 0$ in I. Then, the sequence $\{v^k\}$ produced by the participatory learning process (6) is such that $\lim_{k\to\infty} v^k = \overline{v}$ when $\alpha^k \in [-1, 0)$.

Proof: The proof is similar to the proof of Lemma 1. Just note that, in this case, $\{v^k\}$ is monotone increasing if $f(v^1) - v^1 < 0$ and decreasing if $f(v^1) - v^1 > 0$. ■

The next result is a characterizion of the convergence of the sequence $\{v^k\}$ induced by (6), whenever a fixed point exists.

Theorem 1: Let $f : [0, 1] \longrightarrow [0, 1]$ and $I \subseteq [0, 1]$ an interval in which $\dfrac{df(v)}{dv}$ exists and the fixed point $\overline{v} \in I$. Then the sequence $\{v^k\}$ produced by (6) converges if $\alpha^k \neq 0$ and $|\alpha^k| \leqslant 1$.

Proof: Immediate from Lemmas 1 and 2 because either $\alpha^k \in [-1, 0)$ or $\alpha^k \in (0, 1]$ which means $|\alpha^k| \leqslant 1$. ■

The signal of the sequence $\{f(v^k) - v^k\}$ is the same as of the initial value $f(v^1) - v^1$, as Lemma 1 shows. See Figure 2 for an illustration. The fixed point depends on the sign of α^k. Lemmas 1 and 2 show that $\{v^k\}$ is a monotone and convergent sequence. A simple example is given in Figure 3 for $f(v) = sin(v) + v$.

As suggested by Figure 2 and summarized by Theorem 1, if $0 < \alpha^k < 1$, then the fixed point when the initial value v^1 is in region I is the same when v^1 is in II. Similarly, when $-1 < \alpha^k < 0$, the fixed point $\{v^k\}$ is the same if v^1 is in region II or III. Figure 2 also stress that $f(v^1) - v^1$ and $\{f(v^k) - v^k\}$ always have the same sign.

Summing up, the results above give conditions for which the the sequence of values $\{v^k\}$ produced by participatory learning expressed by (6) converges when inputs are a continuous and differentiable function of the learning variable. Continuity and differentiability of f are reasonable requirements in many practical applications. It remains to explain how the learning algorithm converges. Next we characterize the convergence rate of $\{v^k\}$, namely, the speed with which $v^k \to \overline{v}$.

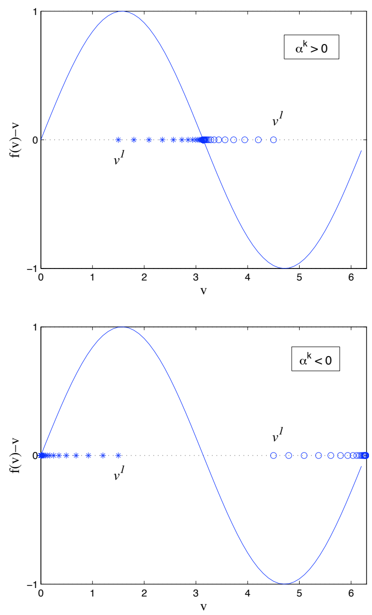

Fig. 3. Convergence of the participatory learning process.

Let us recall that a sequence $\{v^k\}$ converges to avalue $\overline{v}$ in order $q > 0$ if there exists a positive, real-valued constant K such that [8]

$$|v^{k+1} - \overline{v}| \leqslant K|v^k - \overline{v}|^q$$

Theorem 2: Consider the same hypothesis of Theorem 1. If $\lim_{k\to\infty} v^k = \overline{v}$ exists, then there exists a neighborhood of $\overline{v}$ in which $\{v^k\}$ converges in order $q = 1$ (linear convergence).

Proof: Consider $h(v) = (f(v) - v)$, $\theta^k = (v^k - \overline{v})$ and

$\alpha^k G^k \in [0, 1]$. From expression (6) for $\{v^k\}$ we have

$$v^{k+1} = v^k + \alpha^k G^k h(v^k)$$

$$\theta^{k+1} = \theta^k + \alpha^k G^k h(v^k) \leqslant \theta^k + h(v^k)$$

We must show that there exists a positive real-valued constant K such that $|\theta^{k+1}| \leqslant K|\theta^k|$. For all v^k close to $\overline{v}$, there exists a ξ between v^k and $\overline{v}$ such that (mean value theorem)

$$h'(\xi) = \frac{h(v^k) - h(\overline{v})}{v^k - \overline{v}} = \frac{h(v^k) - 0}{\theta^k}$$

$$h(v^k) = h'(\xi)\theta^k$$

Therefore $\theta^k + h(v^k) = [1 + h'(\xi)]\theta^k$ and, assuming $h'(\xi)$ bounded, there exists a positive constant K with $|h'(\xi)| < K$. Hence

$$|\theta^{k+1}| \leqslant |[1 + h'(\xi)]\theta^k| \leqslant (1 + K)|\theta^k|.$$

■

We see that $\{v^k\}$ converges in order $q = 1$, or equivalently, the sequence $\{v^k\}$ converges linearly to fixed point $\overline{v}$ whenever it exists. Generally speaking, learning processes such as (6) converge linearly because they use zero-order information of f only, that is, they need just the value of f at each step.

Notice that Theorems 1 and 2 provide necessary conditions only. There may exist situations in which f is neither differentiable nor continous, but the sequence $\{v^k\}$ may still converge.

IV. Conclusions

In this paper we have studied the dynamic behavior of the participatory learning model. The study assumes that learning data is a function of the object of learning. A characteristic of this learning model is the variable compatibility and arousal index that involves variable learning rates. This feature raises the important problem of proving the convergence of the learning system for this condition. Necessary conditions have been developed considering the joint effect of compatibility, arousal and learning rates. Future work shall address the general case, the situation when no assumption is made on how the learning variable is related with the input data.

V. Acknowledgments

The first author acknowledges FAPESP, the Research Foundation of the State of São Paulo for fellowship 05/58879-1. The last authors thank the Brazilian National Research Council, CNPq, for grants 302407/2008-1 and 304857/2006-8, respectively.

References

[1] R. Yager, "A model of participatory learning," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 20, no. 05, pp. 1229–1234, September/October 1990.

[2] L. Silva, F. Gomide, and R. Yager, "Fuzzy clustering with participatory learning and applications," in *Advances in Fuzzy Clustering and its Applications*, J. Oliveira and W. Pedrycz, Eds., pp. 139–153. Wiley, N.J., USA, 2007.

[3] P. Angelov, D. P. Filev, and N. Kasabov, *Evolving Intelligent Systems: Methodology and Applications*, Wiley-IEEE Press, March 2010.

[4] E. Lima, M. Hell, R. Ballini, and F. Gomide, "Evolving fuzzy modeling using participatory learning," in *Evolving Intelligent Systems*, P. Angelov, D. Filev, and N. Kasabov, Eds., chapter 4, pp. 67–86. Wiley, N.J., USA, 2010.

[5] R. Yager and D. Filev, "Modeling participatory learning as a control mechanism," *International Journal of Intelligent Systems*, , no. 08, pp. 431–450, 1993.

[6] R. Yager, "Participatory learning: A paradigm for more human like learning," in *2004 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE*, Budapest, Hungary, July 2004, pp. 79–84.

[7] R. Bartle, *The Elements of Real Analysis*, Wiley, NY, EUA, 2nd edition, 1976.

[8] S. Conte and C. de Boor, *Elementary Numerical Analysis - An Algorithmic Approach*, McGraw-Hill, NY, EUA, 3nd edition, 1981.