IEEE Second International Conference on Data Stream Mining & Processing
August 21-25, 2018, Lviv, Ukraine

The Multidimensional Extended Neo-Fuzzy System and its Fast Learning for Emotions Online Recognition

Yevgeniy Bodyanskiy
Control Systems Research Laboratory
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
yevgeniy.bodyanskiy@nure.ua

Nonna Kulishova
Media Systems and Technologies Department
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
nokuliaux@gmail.com

Daria Malysheva
Control Systems Research Laboratory
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
darly.malysheva@gmail.com

***Abstract*—Many tasks require human facial expressions automatic recognition in real time. Recent solutions to this problem using machine learning methods have been based on the applying of training data sets that include hundreds of thousands of samples. The formation of these data is too costly. In this paper, the architecture of a system using extended neo-fuzzy neurons for online emotions recognition is examined. We propose the algorithm which is based on the entropy criterion for learning the system and reducing the amount of training data thousands of times.**

Keywords—extended neo-fuzzy system; online emotion recognition; entropy-information learning criterion

I. INTRODUCTION

Information technologies become an integral part of the lives of so many people; they are actively being introduced into education, business, healthcare, and entertainment. Many of these technologies are interactive, and they realize continuous two-way cooperation of a person with a computer or mobile device. One of the promising areas for the interfaces' development for such reciprocity is the approach that uses the recognition of people, their age, sex, state of health, emotional status on the real time video. This complex technical problem already finds its own solutions [1-10]. Usually, these decisions use the machine learning and neuro-fuzzy approach.

As a mathematical problem, the task of a user emotional status recognition by video is reduced to characteristic features detecting, and to the collected data clustering. The assumption of the modeled processes linearity is unreasonable. This leads to the need to select approaches that will be effective for nonlinear systems, especially in real-time conditions. Another problem is related to the fact that machine learning algorithms in this task require the training data sets in which the samples number can be tens or even hundreds of thousands. The creation of such sets is a serious, time-consuming task, significantly increasing the projects developing cost and implementation duration.

II. STANDARD AND EXTENDED NEO-FUZZY NEURON

The NFN could be very effective in the solving the problem of the person emotional state recognition from a video. The neo-fuzzy neuron (NFN) was firstly proposed in the early 1990s by Uchino and Yamakawa [11-13] to simplify the complex nonlinear systems modeling. The NFN is computationally simple, has high approximation accuracy and the ability to minimize the chosen criterion of learning in real time.

Recently, there have been publications about the NFN applying results in different tasks. In [14-16] various architectures of the NFN and the corresponding learning algorithms are proposed. Practical tasks related to the study of induction motors vibration, bearings functioning, bacteria colonies number optimization and classification problems were successfully solved using NFN [17-20].

The standard NFN is constructed on the so-called nonlinear synapses - the elements that realize the fuzzy zero-order Takagi-Sugeno inference [21, 22]:

IF x_i IS x_{li} THEN THE $f(x_i)$ IS $w_{li}, l = 1, 2, ..., h$.

This form corresponds to the transformation that the synapse performs:

$$f_i(x_i) = \sum_{l=1}^{h} w_{li}\mu_{li}(x_i) \tag{1}$$

where w_{li} is the synapse's weight, $\mu_{li}(x_i)$ - the membership function in the synapse, that fuzzify input component x_i l – weight number, $l = 1, 2, ..., h$, i – synapse number $i = 1, 2, ..., n$.

We improved the synapses possibilities - in one of the architectures [14] the so-called extended nonlinear synapse (ENS) was proposed, it is shown in Fig. 1.

The NFN extended nonlinear synapse realizes a fuzzy inference of an arbitrary order. For this, additional variables are used

$$y_{li}(x_i) = \mu_{li}(x_i)\left(w_{li}^0 + w_{li}^1 x_i + \right. \\ \left. + w_{li}^2 x_i^2 + ... + w_{li}^p x_i^p\right), \tag{2}$$

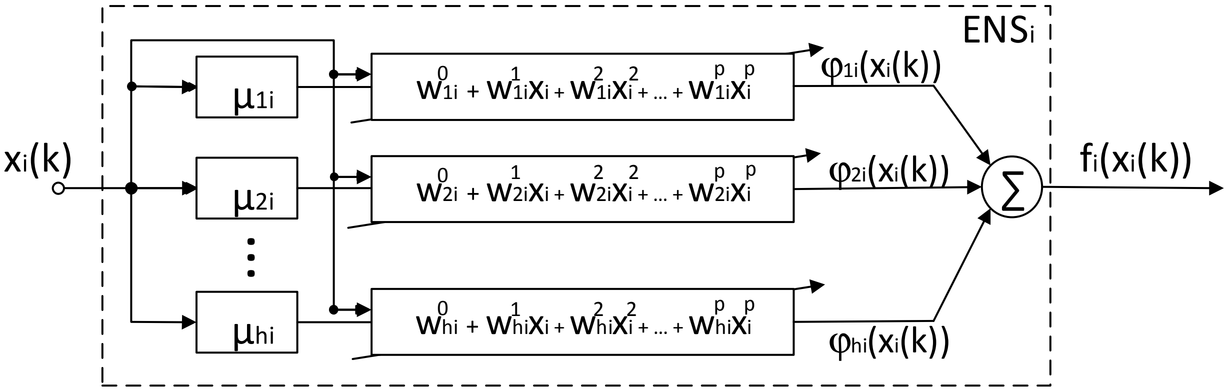

Fig. 1. Extended nonlinear synapse

$$f_i\left(x_i\right)=\sum_{l=1}^{h}\mu_{li}\left(x_i\right)\left(w_{li}^0+w_{li}^1x_i+\right.$$
$$\left.+w_{li}^2x_i^2+\ldots+w_{li}^px_i^p\right)=w_{1i}^0\mu_{1i}\left(x_i\right)+$$
$$+w_{1i}^1x_i\mu_{1i}\left(x_i\right)+\ldots+w_{1i}^px_i^p\mu_{1i}\left(x_i\right)+$$
$$+\ldots+w_{hi}^px_i^p\mu_{hi}\left(x_i\right), \tag{3}$$

$$w_i=\left(w_{1i}^0,w_{1i}^1,\ldots,w_{1i}^p,w_{2i}^0,\ldots,w_{2i}^p,\ldots,w_{hi}^p\right)^T, \tag{4}$$

therefore, we can write:

$$f_i\left(x_i\right)=w_i^T\tilde{\mu}_i\left(x_i\right), \tag{5}$$

$$y=\sum_{i=1}^{n}f_i\left(x_i\right)=\sum_{i=1}^{n}w_i^T\tilde{\mu}_i\left(x_i\right)=\tilde{w}^T\tilde{\mu}\left(x\right), \tag{6}$$

where $\tilde{w}^T=\left(w_1^T,\ldots,w_i^T,\ldots,w_n^T\right)^T$,

$$\tilde{\mu}\left(x\right)=\left(\tilde{\mu}_1^T\left(x_1\right),\ldots,\tilde{\mu}_i^T\left(x_i\right),\ldots,\tilde{\mu}_n^T\left(x_n\right)\right)^T. \tag{7}$$

Thus, ENS implements the output of the form:

IF x_i IS x_{li} THEN THE $f\left(x_i\right)$ IS $w_{li}^0+w_{li}^1x_i+\ldots+w_{li}^px_i^p, l=1,2,\ldots,h$,

that repeats with the formulation of Takagi-Sugeno r-th order inference.

Synapses are NFN structural blocks, which implements the mapping:

$$y=\sum_{i=1}^{n}f_i\left(x_i\right), \tag{8}$$

where x_i - the element of the vector of input data $x=\left(x_1,\ldots,x_i,\ldots,x_n\right)^T\in R^n$, *i* – component number, *n* – vector dimensionality, y - the scalar output of the NFN.

In the extended nonlinear synapse in Fig. 1 B-splines are used as membership function.

Thus, an extended neo-fuzzy neuron (ENFN), receiving an input vector $x\left(k\right)=\left(x_1\left(k\right),\ldots,x_i\left(k\right),\ldots,x_n\left(k\right)\right)^T$ ($k=1,2,\ldots$ - the current count of discrete time), generates a resulting scalar value

$$y\left(k\right)=\sum_{i=1}^{n}\sum_{l=1}^{h}w_{li}\left(k-1\right)\mu_{li}\left(x_i\left(k\right)\right) \tag{9}$$

where $w_{li}\left(k-1\right)$ are the values of the synaptic weights, that obtained as a result of training based on the previous k-1 observations. Fig. 2 shows how the elements of the ENFN [14] combine.

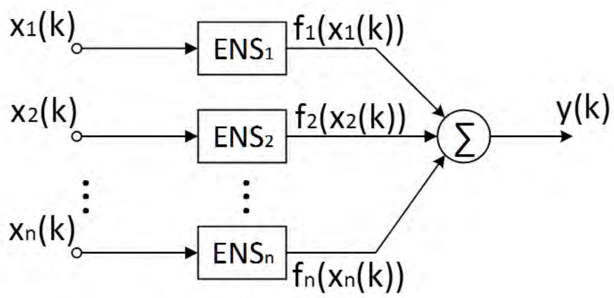

Fig. 2. Extended neo-fuzzy neuron

III. MULTIDIMENSIONAL EXTENDED NEO-FUZZY NEURON

The considered architecture of the ENFN was further developed in [23], where a multidimensional extended neo-fuzzy neuron (MENFN) was considered. In MENFN the input vector signal $x=\left(x_1,\ldots,x_i,\ldots,x_n\right)^T\in R^n$ generates the output vector response. This structure contains several layers. The input layer consists of ENFN; an intermediate layer of elements rejecting negative values, the output layer normalizes the output values and combines them into the resulting vector. For learning of the developed architecture, an algorithm based on a gradient procedure was used. Writing the learning criterion in the form:

$$E(k)=\frac{1}{2}\left(d(k)-y(k)\right)^2=\frac{1}{2}e^2(k)= \\ =\frac{1}{2}\left(d(k)-\sum_{i=1}^{n}\sum_{l=1}^{h}w_{li}\mu_{li}\left(x_i(k)\right)\right)^2, \tag{10}$$

we obtain the learning algorithm:

$$\begin{cases} w(k)=w(k-1)+r^{-1}(k)e(k)\mu\left(x(k)\right), \\ r(k)=\alpha r(k-1)+\left\|\mu\left(x(k)\right)\right\|^2, 0\le\alpha\le 1. \end{cases} \tag{11}$$

where $d(k)$ - external data for training, $e(k)$ - error of learning, η - parameter of learning rate. Depending on the value of α , the (12) is converted to a Goodwin-Ramage-Caines algorithm [24] or an one-step Kaczmarz-Widrow-Hoff algorithm [25].

Despite its universal properties, this learning algorithm does not provide the fulfillment of strict requirements for the system learning in real time on a small training data samples number, which is present in the formulation of the user emotional status recognition task in a video sequence.

To accelerate the MENFN's learning, it was suggested to use the entropy-information learning criterion [26]:

$$E_j(t)=\frac{1}{2}\left(1+d_j(t)\right)\ln\frac{1+d_j(t)}{1+y_j(t)}+ \\ +\frac{1}{2}\left(1-d_j(t)\right)\ln\frac{1-d_j(t)}{1-y_j(t)}. \tag{12}$$

It was noted in [26] that this criterion becomes essentially effective if the hyperbolic tangent is chosen as the activation functions for $y_j(t)$:

$$y_j(x)=\tanh\left(\tilde{w}^T x\right) \tag{13}$$

Then differentiating (13) respecting to w_{ij} , accounting (14), gives a simple learning algorithm of the form:

$$\frac{dw_{ij}}{dt}=\eta e_j(t)x_i, \tag{14}$$

where $e_j(t)=d_j(t)-y_j(t)$ is the local learning error. In the discrete case this simple expression takes the form:

$$w_j(k+1)=w_j(k)+\eta(k)e_j(k)x(k) \tag{15}$$

This recording simplicity ensures both the computational simplicity of the algorithm, as well as the high learning rate required for on-line applications.

As a result, the architecture of the MENFN acquired the following form (Fig. 3).

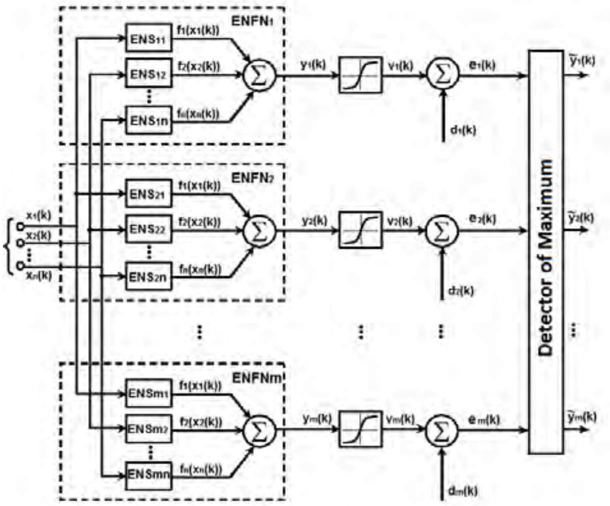

Fig. 3. Multidimensional extended neo-fuzzy neuron with activation functions $y_j(t)=\tanh\left(\tilde{w}^T x\right)$

The first layer consists of the ENFN, the number of which corresponds to the output vector $y_m(k)$ dimension. Nonlinear synapses amount, that forms each of the ENFNs, corresponds to the input feature vector $x_n(k)$ dimension. The next layer implements the function

$$v_j(k)=\psi\left(y_j(k)\right)=\tanh\left(y_j(k)\right) \tag{16}$$

The last layer of MENFN detects the maximums in the calculated learning algorithms values $v_j(k)$:

$$\tilde{y}(k)=\sup_{j=1}^{m}\left\{v_j(k)\right\}, \tag{17}$$

is necessary, if the learning vector set in the range [0,1].

IV. Experiment

The proposed architecture ability to recognize individual emotions was investigated using photographs from two open bases - Psychological Image Collection at Stirling (PICS) [27], partly from the Extended Cohn-Kanade (CK +) database [28]. Some images are in public use as objects for recognition.

In the proposed MENFN architecture, 11 membership functions were used in each non-linear synaptic ENS. The amount of terms in the fuzzy inference rules is assumed to be 3, so that the network realizes the Takagi-Sugeno output of the second order in such form:

I IF x_i IS x_{li} THEN THE $f(x_i)$ IS $w_{li}^0+w_{li}^1x_i+...+w_{li}^p x_i^p, l=1,2,...,h$.

The NFNs amount m corresponds to the dimensionality of the output data vector. Seven basic emotions are selected for recognition: anger, disgust, fear, surprise, happiness, sadness, neutral expression. Therefore, m = 7. The character features vector contains the two-dimensional coordinates of 35 feature points position (Fig. 4).

Fig. 4. Examples of training images and position of characteristic points

To increase the network learning rate we decided to use variable value, that decreases as the number of the learning epoch and the photo position in the training sample increases:

$$\eta = \frac{1}{N_{EP} \cdot n_{InSet}} \tag{18}$$

where N_{EP} is the number of the current network learning epoch; n_{InSet} - current photo number in the set.

In this task, special attention was paid to a learning data set small size. To examine how the proposed architecture and learning algorithm will recognize facial expressions, small photo sets are used. Their dimensions are given in Table I.

TABLE I. DIMENSIONS OF TRAINING SETS OF PHOTOS FOR INDIVIDUAL EMOTIONS

Emotion	***Anger***	***Disgust***	***Fear***	***Happiness***	***Sorrow***	***Surprise***	***Neutral***
Data set size	49	66	35	45	19	50	80

The network learns to recognize each emotion separately in several algorithm steps; the resulting learning error is very small. The plots of the error depending on the learning epoch are shown in Fig. 5.

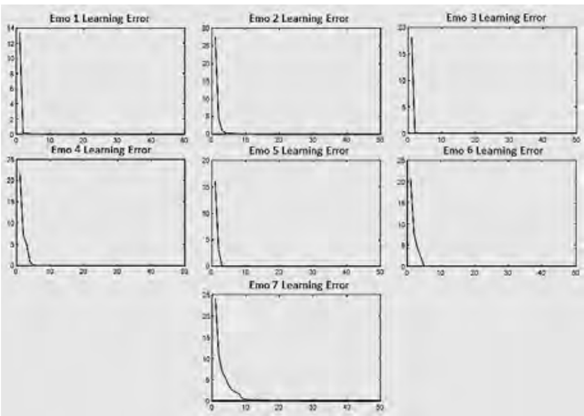

Fig. 5. The plots of the MENFN learning error on small sets for each of the seven emotions separately

Obviously, the learning rate varies from 2 to 15 epochs for different emotions. The sizes of learning sets do not exceed 100 samples each.

Then the architecture ability to learn from a mixed set was examined, and sets total size was 344 photos. The number of unrecognized emotions is given in Table II.

TABLE II. THE NUMBER OF UNRECOGNIZED EMOTIONS AS A RESULT OF MENFN LEARNING FROM A MIXED SET

	Primary emotions						
	Anger	***Disgust***	***Fear***	***Happiness***	***Sorrow***	***Surprise***	***Neutral***
The number of images in the learning set	1	0	2	2	0	0	2
The percentage of unrecognized images,%	2.04	0	5.71	4.44	0	0	2.5

A MENFN learning error change is shown in Fig. 6.

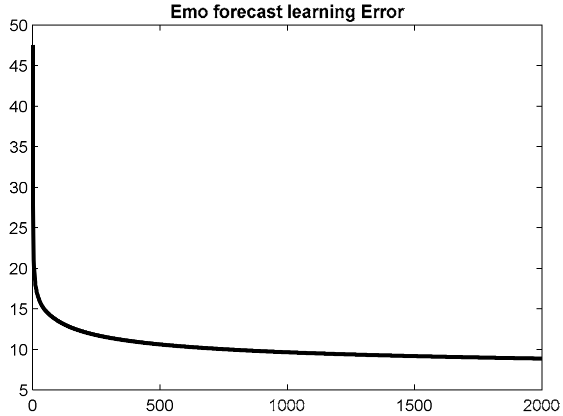

Fig. 6. MENFN learning error for the mixed data set

V. CONCLUSIONS

The paper proposes architecture of the MENFN. Its structure generalizes the standard NFN for the case of arbitrary order fuzzy inference procedure and multidimensional input and output data. The proposed learning algorithm allows effectively distribute the data between the different clusters. Considered MENFN has high learning rate, provided by the learning algorithm based on the entropy criterion; improved clustering properties, easy to numerical implementation.

REFERENCES

[1] A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch, M. R. Wrobel, "Human-Computer Systems Interaction: Backgrounds and Applications," ch. 3, Emotion Recognition and Its Applications. Cham: Springer International Publishing, 2014, pp. 51 – 62.

[2] Kaggle. Challenges in representation learning: Facial recognition challenge, 2013.

[3] G.U. Kharat, S.V. Dudul, "Emotion Recognition from Facial Expression Using Neural Networks," in Human-Computer Systems Interaction. Advances in Intelligent and Soft Computing, vol 60, Z.S. Hippe, J.L. Kulikowski, Eds. Berlin, Heidelberg: Springer, 2009.

[4] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, no. 6, 2009, pp. 803 – 816.

[5] B. Fazel, J. Luettin, "Automatic facial expression analysis: a survey", Pattern Recognition, 36(1), 2003, pp. 259 – 275.

[6] Ch.-Yi Lee, Li-Ch. Liao, "Recognition of Facial Expression by Using Neural-Network System with Fuzzified Characteristic Distances Weights," IEEE Int. Conf. Fuzzy Systems FUZZ-IEEE 2008. [IEEE World Congress on Computational Intelligence, pp. 1694 – 1699, 2008].

[7] N. Kulishova, "Emotion Recognition Using Sigma-Pi Neural Network," Proc. of 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2016, pp. 327 –331.

[8] A. Graves, J. Schmidhuber, C. Mayer, M. Wimmer, B. Radig, “Facial Expression Recognition with Recurrent Neural Networks,” International Workshop on Cognition for Technical Systems, Munich, Germany, October 2008.

[9] S. Ouelett, “Real-time emotion recognition for gaming using deep convolutional network features,” CoRR, vol. abs./1408.3750, 2014.

[10] B. Kim, J. Roh, S. Dong, and S. Lee, “Hierarchical committee of deep convolutional neural networks for robust facial expression recognition,” Journal on Multimodal User Interfaces, 2016, pp. 1–17.

[11] J. Miki, J. Yamakawa, “Analog implementation of neo-fuzzy neuron and its on-board learning,” in Computational Intelligence and Applications, Ed. N.E. Mastorakis, Piraeus: WSES Press, 1999, pp. 144 – 149.

[12] J. Yamakawa, E. Uchino, J. Miki, H. Kusanagi, “A neo-fuzzy neuron and its application to system identification and prediction of the system behavior,” Proc. 2-nd Int. Conf. on Fuzzy Logic and Neural Networks “IIZUKA-92”, Iizuka, Japan, 1992, pp. 477 – 483.

[13] E. Uchino, J. Yamakawa, “Soft computing based signal prediction, restoration and filtering,” in Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algoritms, Ed. Da Ruan, Boston: Kluwer Academic Publishers, 1997, pp. 331 – 349.

[14] Ye.V. Bodyanskiy, N.Ye. Kulishova, “Extended neo-fuzzy neuron in the task of images filtering,” Radioelectronics. Computer Science. Control, № 1(32), 2014, pp. 112 – 119.

[15] Ye. Bodyanskiy, Y. Victorov, “The cascade of neo-fuzzy architecture and its online learning algorithm,” Int. Book Series Inf. Sci. Comput., 17(1), 2010, pp. 110 – 116.

[16] Ye. Bodyanskiy, I. Kokshenev, V. Kolodyazhniy, “An adaptive learning algorithm for a neo-fuzzy neuron,” Proc. of the 3rd Conference of the European Society for Fuzzy Logic and Technology, pp. 375 – 379, 2005.

[17] D. Zurita, M. Delgado, J.A. Carino, J.A. Ortega, G. Clerc, “Industrial Time Series Modelling by Means of the Neo-Fuzzy Neuron,” IEEE Access, vol. 4, 2016, pp. 6151 – 6160.

[18] M. Pandit, L. Srivastava, V. Singh, “On-line voltage security assessment using modified neo-fuzzy neuron based classifier,” IEEE Int. Conf. Ind. Technol., 2006, pp. 899 – 904.

[19] H.D. Kim, “Optimal learning of neo-fuzzy structure using bacteria foraging optimisation,” Proceedings of the ICCA, 2005.

[20] A.M. Silva, W. Caminhas, A. Lemos, F. Gomide, “A fast learning algorithm for evolving neo-fuzzy neuron,” Applied Soft Computing, vol. 14, Part B, January 2014, pp. 194 – 209.

[21] T. Takagi, M. Sugeno, “Fuzzy identification of systems and its application to modeling and control,” IEEE Trans. On System, Man and Cybernetics, 15, 1985, pp. 116 – 132.

[22] J.-S. Jang, C.-T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Upper Saddle River: Prentice Hall, 1997.

[23] Z. Hu, Ye.V. Bodyanskiy, N.Ye. Kulishova, O.K. Tyshchenko, “A Multidimensional Extended Neo-Fuzzy Neuron for Facial Expression Recognition,” International Journal of Intelligent Systems and Applications (IJISA), vol.9, No.9, 2017, pp.29 – 36.

[24] G.C. Goodwin, P.J. Ramage, P.E. Caines, “Discrete time stochastic adaptive control,” SIAM J. Control and Optimisation, 19, 1981, pp. 829 – 853.

[25] S. Haykin, Neural Networks. A Comprehensive Foundation. Upper Saddle River: Prentice Hall, 1999.

[26] A. Cichocki, R. Unbehauen, Neural Networks for Optimization and Signal Processing. Stuttgart: Teubner, 1993.

[27] http://pics.psych.stir.ac.uk/2D_face_sets.htm

[28] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," Proceedings of IEEE workshop on CVPR for Human Communicative Behavior Analysis, San Francisco, USA, 2010.