NONLINEAR DECISION FEEDBACK EQUALIZATION USING RCPL NETWORK

Levent Demirekler, *Tülay Adalı*, and *Xiao Liu*
Information Technology Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD 21250, U.S.A.
Tel: +1 410 4551000/3521; fax:1 410 4553969
e-mail: adali@engr.umbc.edu

ABSTRACT

In this paper, we introduce a nonlinear decision feedback equalizer (DFE) structure which decomposes the nonlinear equalization problem into feedback and feedforward parts, thus allows flexibility in the choice of suitable structures for a given problem and improves on the performance of the conventional DFE and the nonlinear equalizer proposed in [1]. We show that the given general structure for this DFE can be represented by a recurrent canonical piecewise linear (RCPL) network [2] and that it satisfies the dynamic properties given in [3].

1 INTRODUCTION

Neural networks are good candidates for nonlinear signal processing tasks in communications due to their nonlinear computation capability, ability to learn from data, and potential for easy hardware implementation. In this paper we introduce a nonlinear DFE which utilizes a neural network structure to overcome the effects of nonlinearities in the channel.

Conventional adaptive equalizers such as linear equalizers and DFE use linear finite impulse response (FIR) filters. Linear filters perform poorly when there is nonlinearity in the channel. Although noncausal characteristic of DFE improves its performance over linear equalizers, both structures assume a linear channel model.

In [1], a multilayer perceptron based DFE is proposed which uses nonlinear FIR filters for the equalization of nonlinear channels. It is shown that the performance of the equalizer is improved with this equalizer structure compared to the equalizers using linear FIR filters.

In this paper, we propose a nonlinear DFE which uses a RCPL network at the feedforward and feedback part of a conventional DFE. RCPL network partitions the input signal space into finite disjoint regions and in each region, it can be represented by a FIR filter with infinite length. Thus it can approximate a large class of nonlinear functions when used as an equalizer.

The paper is organized as follows. In Section 2, we describe the structure of conventional DFE and the proposed nonlinear DFE. In Section 3, we give the definition of RCPL network and show that nonlinear DFE can be represented by a RCPL network and it satisfies its dynamic properties. Finally in section 4, we give the simulation results.

2 NONLINEAR DFE STRUCTURE

The structure of a conventional DFE is shown in Fig 1. The decision function is dependent on channel observations $\mathbf{y}$ and the previous decisions $\hat{\mathbf{x}}$. In conventional DFE, decision function is made up of linear combination of $\mathbf{y}$ and $\hat{\mathbf{x}}$. The advantage of conventional DFE over linear equalizers is the separation of the observation space into subset of observation centers for a particular decision for which the separation is achieved by the use of previous decisions $\hat{\mathbf{x}}$ [4]. Linear combination is not a good choice if there is nonlinearity in the channel. It is shown in [1] and [4]–[6] that the performance of a conventional DFE can be improved on by using nonlinear structures in the decision function.

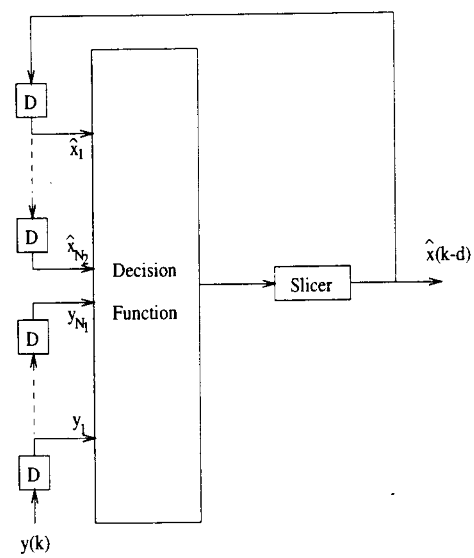

Figure 1: Decision Feedback Equalizer Structure

In this paper, we break the equalization problem into two parts, as does the conventional DFE, and introduce a general nonlinear decision function. The decision function with both the feedback and feedforward parts using recurrent nonlinear networks is shown in Figure 2. It is similar to the MLP structure proposed in [1] except that channel observations and the previous decisions are separately processed and there is full recurrency in the hidden nodes. It is observed that by breaking the problem into two parts, this structure resultes in improved error rates compared to the conventional MLP based DFE [1].

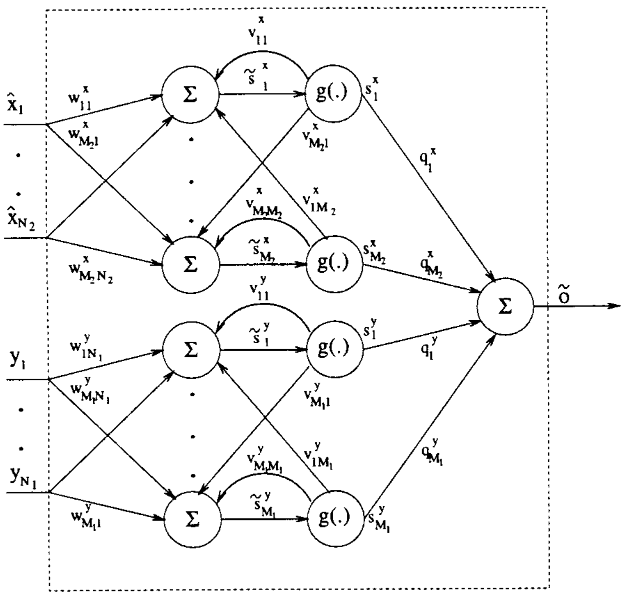

Figure 2: Decision Function Structure

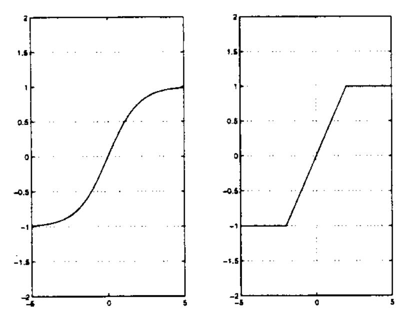

Figure 3: Sigmoidal and Piecewise Linear Activation Function $g(\cdot)$

3 RCPL Network

For the structure shown in Figure 2, we can replace the sigmoidal nonlinearity with a piecewise linear function as shown in Figure 3 and develop a RCPL formulation for the given DFE. RCPL allows for easy implementation because of its *piecewise linear* nature and savings in the number of total parameters required for representation because of its *canonical* structure, i.e., its efficient global representation capability [2]. We can show that the network structure with piecewise linear activation function is a recurrent canonical piecewise-linear (RCPL) network [3]. The recurrent canonical piecewise-linear function is defined as:

Definition 1.(*Recurrent Canonical Piecewise-Linear Function*): A function f: $D \to Q$ with sample space $D \subset R^N$ and compact subset $Q \subset R^M$ is said to be a RCPL function if it can be expressed by the global representation:

$$f(\mathbf{x}(n)) = \mathbf{a} + \mathbf{B}_0\mathbf{x}(n) + \mathbf{B}_1 f(\mathbf{x}(n-1)) + \sum_{i=1}^{\tau} c_i \left|\langle \boldsymbol{\alpha}_{0_i}, \mathbf{x}(n)\rangle + \langle \boldsymbol{\alpha}_{1_i}, f(\mathbf{x}(n-1))\rangle + \beta_i\right| \quad (1)$$

where $\mathbf{x} \in R^N$, $\mathbf{a}$, $\boldsymbol{\alpha}_{1_i} \in R^M$, $\mathbf{c}_i \in R^M$, $\boldsymbol{\alpha}_{0_i} \in R^N$. $\mathbf{B_0} \in R^{M\times N}$, $\mathbf{B_1} \in R^{M\times M}$, and $\tau, \beta_i \in R$. The elements in the vector $\mathbf{x}(n)$ are predicted by the following NARMA model:

$$x_i(n) = h_i(\mathbf{x}(n-1), \cdots, \mathbf{x}(n-p_1), f(\mathbf{x}(n-1)), \cdots, f(\mathbf{x}(n-p_2)) \quad (2)$$

A RCPL representation which is applicable to the general nonlinear DFE structure given in this paper can be obtained as follows: Let $p_1 = 1$, $p_2 = P$, $\mathbf{x}(n)$ is the state vector of the dynamic system and $\mathbf{u}(n)$ is the input vector of the system. The RCPL function is then defined as:

Definition 2 (*Recurrent Canonical Piecewise-Linear Network*): A function f: $D_1 \times D_2 \to Q$ with sample space $D_1 \subset R^N$, $D_2 \subset R^r$, and compact subset $Q \subset R^M$ is said to be a RCPL function if it can be expressed by the global representation:

$$f(\mathbf{x}(n), \mathbf{u}(n)) = \mathbf{a} + \mathbf{B}_0\mathbf{x}(n) + \mathbf{B}_1 f(\mathbf{x}(n-1), \mathbf{u}(n-1)) + \mathbf{B}_2\mathbf{u}(n) \quad (3)$$

$$x_k(n) = a_k + \langle \mathbf{h}_{0_k}, \mathbf{x}(n-1)\rangle + \langle \mathbf{h}_{1_k}, f(\mathbf{x}(n-1), \mathbf{u}(n-1))\rangle + \langle \mathbf{h}_{2_k}, \mathbf{u}(n)\rangle + \sum_{i=1}^{\tau} c_{k_i} \Big| \langle \boldsymbol{\alpha}_{0_i}, \mathbf{x}(n-1)\rangle + \langle \boldsymbol{\alpha}_{1_i}, \mathbf{u}(n)\rangle + \sum_{j=1}^{P} \langle \boldsymbol{\gamma}_{j_i}, f(\mathbf{x}(n-j), \mathbf{u}(n-j))\rangle + \beta_i \Big| \quad (4)$$

where $\mathbf{x}, \mathbf{h}_{0_k}, \boldsymbol{\alpha}_{0_i} \in R^N$, $f, \mathbf{a}, \mathbf{h}_{1_k}, \boldsymbol{\gamma}_{j_i} \in R^M$, $\mathbf{u}, \mathbf{h}_{2_k}, \boldsymbol{\alpha}_{1_i} \in R^r$, $\mathbf{B}_0 \in R^{M\times N}$, $\mathbf{B}_1 \in R^{M\times M}$, $\mathbf{B}_2 \in R^{M\times r}$, $a_k, c_{k_i}, \beta_{k_i}, \tau \in R$, $k = 1, 2, \cdots, N$ and x_k is the kth element in $\mathbf{x}$.

Based on the above definition we can express the new DFE structure as an RCPL network as follows: Let $M = 1$, $N = M_1 + M_2$, $r = N_1$,

$$\mathbf{u}(n) \equiv \mathbf{y}(n) = [y_1(n), \cdots, y_{N_1}(n)]$$

$$\mathbf{x}(n) \equiv \mathbf{s}(n) = [s_1^x(n), \cdots, s_{M_2}^x, s_1^y(n), \cdots, s_{M_1}^y(n)]$$

$$\tilde{o}(n) = f(\mathbf{s}(n), \mathbf{y}(n))$$

where $\mathbf{y}(n)$ is the input vector and $\tilde{o}(n)$ is the output of the network and $s_k^x(n), s_k^y(n)$ are the outputs of hidden nodes.

For the parameters in (3), let

$\mathbf{a}=0$, $\mathbf{B}_0=[q_1^x(n), \cdots, q_{M_2}^x, q_1^y(n), \cdots, q_{M_1}^y(n)]$

$\mathbf{B}_1=0$, $\mathbf{B}_2=\mathbf{0}$

and the parameters in (4), we choose

$a_k = 0$, $\mathbf{h}_{0_k}=0$, $\mathbf{h}_{1_k}=0$, $\mathbf{h}_{2_k}=\mathbf{0}$

$\tau = 2M_2$, $c_{k_i}=1/8$, $\beta_i=4$, $i = 1, \cdots, M_2$

$c_{k_i}=-1/8$, $\beta_i=-4$, $i=M_2+1, \cdots, 2M_2$

$\boldsymbol{\alpha}_{1_{i_k}}=(v_{k1}^x, \cdots, v_{kM_2}^x)^T$, $\boldsymbol{\alpha}_{2_{i_k}}=0$

$\boldsymbol{\gamma}_{j_{i_k}}=(\omega_{k1}^x, \cdots, \omega_{kN_2}^x)^T$, $i=1, \cdots, 2M_2$ $k=1, \cdots, M_2$

for states $[s_1^x(n), \cdots, s_{M_2}^x]$, and

$a_k = 0$, $\mathbf{h}_{0_k}=\mathbf{0}$, $\mathbf{h}_{1_k}=\mathbf{0}$, $\mathbf{h}_{2_k}=\mathbf{0}$

$\tau = 2M_1$, $c_{k_i}=1/8$, $\beta_i=4$ $i=1, \cdots, M_1$

$c_{k_i}=-1/8$, $\beta_i=-4$, $i=M_1+1, \cdots, 2M_1$

$\boldsymbol{\alpha}_{1_{i_k}}=(v_{k1}^y, \cdots v_{kM_1}^y)^T$, $\boldsymbol{\alpha}_{2_{i_k}}=(\omega_{k1}^y, \omega_{k2}^y, \cdots, \omega_{kN_1}^y)^T$

$\boldsymbol{\gamma}_{j_{i_k}}=0$, $i=1, \cdots, 2M_1$ $k=1, \cdots, M_1$

for states $[s_1^y(n), \cdots, s_{M_1}^y]$.

The RCPL structure is then expressed by the following equations:

$$\tilde{o}(n) = \sum_{i=1}^{M_2} q_i^x s_i^x(n) + \sum_{i=1}^{M_1} q_i^y s_i^y(n) \tag{5}$$

$$s_k^x(n) = g(\tilde{s}_k^x(n)) \text{ and } s_k^y(n) = g(\tilde{s}_k^y(n))$$

$$\tilde{s}_k^x(n) = \sum_{i}^{M_2} v_{ki}^x s_k^x(n-1) + \sum_{i=1}^{N_2} \omega_{ki}^x \hat{x}_i(n), \quad k = 1, \cdots, M_2 \tag{6}$$

$$\tilde{s}_k^y(n) = \sum_{i}^{M_1} v_{ki}^y s_k^y(n-1) + \sum_{i=1}^{N_1} \omega_{ki}^y y_i(n), \quad k = 1, \cdots, M_1 \tag{7}$$

The network defined by (5)-(7) is shown in Figure 2. Note that the output of the network is fed to a slicer to find $\hat{x}(n-d)$, and then fedback into the network. This modification introduced into the structure is observed to improve the performance of the DFE.

3.1 Dynamics of RCPL

To study the dynamics of the RCPL function described by (3) and (4), we first rewrite the function in the following form:

$$\begin{aligned}\bar{\mathbf{x}}(n) &= \bar{\mathbf{a}} + \bar{\mathbf{B}}_1\bar{\mathbf{x}}(n-1) + \bar{\mathbf{B}}_2\mathbf{u}(n) \\ &+ \sum_{i=1}^{\tau} \bar{\mathbf{c}}_i | \langle \bar{\boldsymbol{\alpha}}_{0i}, \bar{\mathbf{x}}(n-1)\rangle + \langle \bar{\boldsymbol{\alpha}}_{1i}, \mathbf{u}(n)\rangle \\ &+ \sum_{j=1}^{P} \langle \bar{\boldsymbol{\gamma}}_{j_i}, \bar{\mathbf{x}}(n-j)\rangle + \bar{\beta}_i |\end{aligned} \tag{8}$$

where

$$\bar{\mathbf{x}}(n) = \begin{bmatrix} \mathbf{x}(n) \\ f(\mathbf{x}(n), \mathbf{u}(n)) \end{bmatrix} \bar{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ \mathbf{a} + \mathbf{B}_0\mathbf{a} \end{bmatrix} \bar{\mathbf{c}}_i = \begin{bmatrix} \mathbf{c}_i \\ \mathbf{B}_0\mathbf{c}_i \end{bmatrix}$$

$$\bar{\mathbf{B}}_1 = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 \\ \mathbf{B}_0\mathbf{H}_0 & \mathbf{B}_1 + \mathbf{B}_0\mathbf{H}_1 \end{bmatrix} \bar{\mathbf{B}}_2 = \begin{bmatrix} \mathbf{H}_2 \\ \mathbf{B}_2 + \mathbf{B}_0\mathbf{H}_2 \end{bmatrix}$$

$\bar{\boldsymbol{\alpha}}_{0_i} = (\boldsymbol{\alpha}_{0_i}, 0)^T$, $\bar{\boldsymbol{\alpha}}_{1_i} = \boldsymbol{\alpha}_{1_i}$, $\bar{\boldsymbol{\gamma}}_{j_i} = (0, \boldsymbol{\gamma}_{j_i})^T$, $\bar{\beta}_i = \beta_i$

$\mathbf{a} = (a_1, a_2, \cdots, a_N)^T$, $\mathbf{c}_i = (c_{1_i}, c_{2_i}, \cdots, c_{N_i})^T$

$\mathbf{H}_j = (\mathbf{h}_{j_1}, \cdots, \mathbf{h}_{j_N})^T$, $i = 1, 2 \cdots, \tau$, $j = 0, 1, 2$

Then by using the definition given above, we can show that the RCPL function is bounded for bounded inputs.

Theorem 1: For the RCPL function defined by (3) and (4), assume that the input vector $\mathbf{u}(n)$ is bounded and the parameters satisfy the following condition: If there exists an $\varepsilon_0 \in (0,1)$ such that

$$||\bar{\mathbf{B}}_1|| + \sum_{i=1}^{\tau} ||\bar{\mathbf{c}}_i|| \, (||\, \bar{\boldsymbol{\alpha}}_{0_i} + \sum_{j=1}^{P} \bar{\boldsymbol{\gamma}}_{j_i} ||) \leq 1 - \varepsilon_0 \tag{9}$$

then, there is a real number d, such that for all $K \geq d$, the ball $D(K) = \{\mathbf{x} : ||\mathbf{x}|| \leq K\}$ is invariant under (3) and (4).

Proof of the theorem is given in [3]. Note that the structure used here is a more general form. The proof of the theorem however, follows the same procedure with this new definition. The definition given in (8) provides a convenient framework to study dynamics of RCPL function which we also use to prove the following:

Theorem 2: The map that defines the RCPL function (3) and (4) is a contractive mapping if the condition given in (9) is satisfied.

Proof: Let

$$\begin{aligned}k(\mathbf{x}) &= \bar{\mathbf{a}} + \bar{\mathbf{B}}_1\bar{\mathbf{x}} + \bar{\mathbf{B}}_2\mathbf{u}(n) + \sum_{i=1}^{\tau} \bar{\mathbf{c}}_i | \langle \bar{\alpha}_{0_i}, \mathbf{x}\rangle + \langle \bar{\boldsymbol{\alpha}}_{1_i}, \mathbf{u}(n)\rangle \\ &+ \sum_{j=1}^{P} \langle \bar{\boldsymbol{\gamma}}_{j_i}, \bar{\mathbf{x}}(n-j)\rangle + \bar{\beta}_i |\end{aligned} \tag{10}$$

then,

$$\mathrm{k}(\mathbf{x}_1) - \mathrm{k}(\mathbf{x}_2) = \bar{\mathbf{B}}_1(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

$$\begin{aligned}&+ \sum_{i=1}^{\tau} \bar{\mathbf{c}}_i (| \langle \bar{\alpha}_{0_i}, \bar{\mathbf{x}}_1\rangle + \langle \bar{\boldsymbol{\alpha}}_{1_i}, \mathbf{u}(n)\rangle + \sum_{j=1}^{P} \langle \bar{\boldsymbol{\gamma}}_{j_i}, \bar{\mathbf{x}}_1(n-j)\rangle + \bar{\beta}_i | \\ &- | \langle \bar{\alpha}_{0_i}, \bar{\mathbf{x}}_2\rangle + \langle \bar{\boldsymbol{\alpha}}_{1_i}, \mathbf{u}(n)\rangle + \sum_{j=1}^{P} \langle \bar{\boldsymbol{\gamma}}_{j_i}, \bar{\mathbf{x}}_2(n-j)\rangle + \bar{\beta}_i |)\end{aligned} \tag{11}$$

and by using (9), and defining $\bar{\mathbf{d}}_i = \bar{\alpha}_{0_i} + \sum_{j=1}^{P} \bar{\boldsymbol{\gamma}}_{j_i}$ we get

$$||k(\mathbf{x}_1) - k(\mathbf{x}_2)|| \leq (||\bar{\mathbf{B}}_1|| + \sum_{i=1}^{\tau} ||\bar{\mathbf{c}}_i|| \, ||\bar{\mathbf{d}}_i||) \, ||\mathbf{x}_1 - \mathbf{x}_2||$$

where $\varepsilon_0 \in (0,1)$.

Theorem 2 shows that $k(\cdot)$ is a contractive mapping whenever (9) is satisfied. Thus, after receiving input vector $\mathbf{u}(n)$, which is assumed to be bounded, the function will always reach a unique equilibrium regardless of its initial state $\mathbf{x}_0$.

4 Simulation Results

As an example, consider a channel with transfer function $H(z)=1+0.5z^{-1}+0.2z^{-2}$. The output of the channel is passed through a memoryless nonlinearity $f(\cdot)=(\cdot)-0.03(\cdot)^2$ and then corrupted by additive white Gaussian noise. The input to the channel is drawn from a 4-level alphabet {-3,-1,1,3} for which each symbol is equally probable. The number of channel observations N_1 is selected as 12, and the number of previous decisions N_2, used in the decision function is selected as 4. The number of hidden nodes for the channel observations M_1 and previous decisions M_2 are selected as 11. The network is trained based on minimum mean square error criteria. 1000 training samples are used and the learning parameter η is chosen as 0.01 for modified DFE, MLP-DFE and 0.001 for conventional DFE. Several learning parameters are tried and the ones that yield best performance are used in the simulation. 10 independent realizations are performed for which the system is tested for a total of 2×10^6 samples.

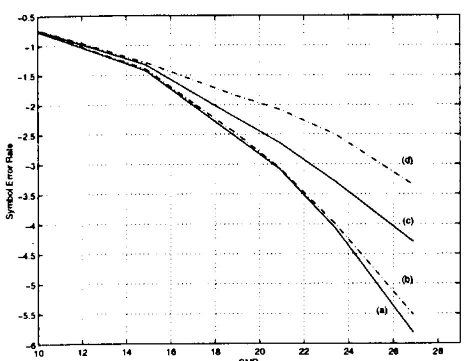

Figure 4: Symbol Error Rates for (a) sigmoidal (b) piecewise linear activation function by using the modified DFE, (c) MLP-DFE (d) DFE

From the figure we see that symbol error rates for the modified DFE is better than MLP-DFE and conventional one. We also note that using sigmoidal or piecewise linear activation function result in similar performance.

In this paper, we introduced a nonlinear DFE which uses a neural network structure in the decision function. We showed that the structure can be represented as a RCPL network if we use a piecewise linear activation function in the hidden nodes. We discussed the dynamic properties of a RCPL network and observed that the proposed nonlinear DFE structure exhibits similar performance for both activation functions. The piecewise linear equalizer, on the other hand, provides access to a variety of analysis and development tools that are linear and allows development of efficient learning algorithms while effectively approximating functions that are highly nonlinear. We compared our structure with MLP-DFE and conventional DFE and gave our simulation results.

References

[1] M. Meyer and G. Pfeiffer, "Multilayer perceptron based decision feedback equalizers for channels with intersymbol interference," *IEE PROCEEDINGS-1*, vol. 140, no.6, pp. 420-424, December, 1993.

[2] T. Adalı and X. Liu, "Canonical piecewise linear network for nonlinear adaptive filtering and its application to blind equalization," *Signal Processing*, vol. 61, no. 2, pp. 145-155, 1997.

[3] X. Liu and T. Adalı, "Recurrent canonical piecewise linear network: theory and application in blind equalization," in *Proc. IEEE Workshop on Neural Networks for Signal Processing (NNSP)*, Amelia Island, FL, Sep. 1997, pp. 446-455.

[4] B. Mulgrew, "Applying Radial Basis Functions," *IEEE Signal Processing Magazine*, pp. 50-65, March 1996.

[5] S. Theodoridis, C.F.N. Cowan, and C.M.S. See, "Schemes for equalisation of communication channels with nonlinear impairments," *IEE Proc. Comm*, vol. 142, no. 3, Jun. 1995, pp. 1350-2425.

[6] S. Ong, C. You, S. Choi and D. Hong, "A Decision Feedback Recurrent Neural Equalizer as an Infinite Impulse Response Filter," *IEEE Transactions on Signal Processing*, vol. 45, no.11, pp. 2851-2858, November, 1997.