Thematic Role Assignment through a Biologically Plausible Symbolic-connectionist Hybrid System

João Luís Garcia Rosa
Centro de Ciências Exatas, Ambientais e de Tecnologias
PUC-Campinas
Campinas, SP, Brazil
E-mail: joaoluis@puc-campinas.edu.br

Alberione Braz da Silva
Centro de Ciências Exatas, Ambientais e de Tecnologias
PUC-Campinas
Campinas, SP, Brazil
E-mail: aberione.braz@dglnet.com.br

***Abstract*— There are several connectionist systems concerning natural language computational processing applications in artificial intelligence literature. Some try to predict the next word in a sentence context. Others attempt to discover grammatical symbolic rules from the neural network architecture, and others deal with semantic relationships between words. Unfortunately, very few of them employ psycholinguistic theories. Recently, some systems have made use of biologically plausible approaches to connectionist natural language processing, which are computationally as efficient as conventional methods. The system described here, in a way, presents all the properties and characteristics enumerated previously: it is a symbolic-connectionist hybrid system, designed to ”predict” thematic (semantic) roles assigned to words in a sentence context with a biologically inspired training algorithm and architecture, and adopting a psycholinguistic view of thematic theory. Its lexicon, which includes lexically ambiguous nouns and ”thematically ambiguous” verbs, can be increased, provided that semantic ”features” of words are entered. The system has proven to be even more computationally efficient than conventional connectionist systems.**

I. Introduction

Back propagation, which is largely employed nowadays as the most computationally efficient connectionist supervised learning algorithm, is argued to be biologically implausible [1], [2]. In addition, conventional recurrent architectures seem to be not supported by neuroscience. Current models are lacking in biology, mainly for the sake of mathematical simplicity [3]. In this paper, it is presented BIOPLA - *Biologically Plausible symbolic-connectionist hybrid system for thematic processing of natural Language sentences* - with a bi-directional connectionist architecture, which learns through a biologically plausible algorithm, adapted from the Generalized Recirculation algorithm [4], concerning the thematic role assignment in natural language sentence contexts. Thematic roles (AGENT, THEME, INSTRUMENT, etc.) assigned to nouns are operands of a semantic predicate, usually the main verb of a sentence [5]. Through the symbolic data regarding the expected thematic grids extracted from the connectionist architecture, it is possible to show that the biologically plausible system reflects successfully the thematic relationships learned.

II. Connectionist natural language applications

Several systems concerning natural language processing have made use of artificial neural networks. Some of them consider the prediction of next words in a sentence context [6]. Some try to discover and extract symbolic rules from the connectionist architecture [7], [8], and others deal with semantic relationships between words or phrases [7], [2], [9]. Recently, some approaches to connectionist natural language processing systems have employed biologically plausible models, which are computationally as efficient as conventional methods [6], [2], [9]. The system described here, in a way, presents all the properties and characteristics enumerated previously: it is a symbolic-connectionist hybrid system with a biologically inspired training algorithm and architecture, designed to ”predict” thematic roles assigned to words in a sentence context. Its lexicon, which includes lexically ambiguous nouns, lexically ambiguous verbs with same thematic grids, and ”thematically ambiguous” verbs with same senses but different thematic grids, can be increased, provided that their semantic ”features” are entered. It has already shown through several different sentences that earlier versions of this system are more computationally efficient than a conventional connectionist system with the well known back propagation learning procedure and a recurrent connectionist architecture [9].

III. Symbolic-connectionist hybrid systems

It is often argued that artificial neural networks lack transparency, that is, it is almost impossible to know how they develop internal representations. And it is well known that training often takes too long. A solution to such constraints is the symbolic-connectionist hybrid approach. In this method one can combine symbolic approach advantages, like expressive power of the general logical implications, ease of knowledge representation, and understanding through logical inference, with connectionism benefits, like learning, generalization, and fault tolerance [7].

In a symbolic-connectionist hybrid approach, symbolic knowledge is inserted in a connectionist architecture as connection weights. The network is submitted to a training period, like conventional connectionist systems. After training, the symbolic theory, which gave initial knowledge to the network, is revised by the connectionist learning. This way, it is possible to overcome the drawbacks presented previously: since the system has initial knowledge, it takes less time to learn; and because of the nodes now naming concepts, the weights linked to them do make sense. The symbolic knowledge generated by the net can be extracted in a way analogous to the way initial knowledge was inserted. The symbolic data set and the network, from which it is extracted, are proved to be equivalent [10].

IV. Thematic roles

Verbs usually play a predicate function in a sentence context. The predicate represents the semantic relations between the verb itself and the nouns compounding the sentence. Thematic roles assigned to words in a sentence are the arguments of these semantic functions [5]. So, the verb *break*, for instance, in one possible reading of sentence (1), assigns the thematic roles AGENT and PATIENT, because *man* is supposed to be deliberately responsible for the action of breaking (the "agent"), and *window* is the "patient", that changes states, affected by the action.

The man broke the window (1)

But there are verbs that can assign more than one thematic grid, depending on the sentence they occur. For instance, in sentence (2), there is a different thematic grid ([CAUSE, PATIENT]) assigned by the same verb *break*, since *ball* causes the breaking, but in an involuntary way.

The ball broke the window (2)

Verbs with more than one thematic grid are called here *thematically ambiguous verbs*. Considering sentences (1) and (2) again, it is quite obvious that the nouns employed as subjects make the distinction between AGENT and CAUSE. In other words, thematic roles must be elements with semantic content [11].

A. Word representation

The representation chosen for words in the presented system is adapted from the classical semantic microfeature distributed representation [12]. Ten dimensions with two units each account for each verb and noun (see tables I and II). The values for each unit range from 0.0 to 1.0, in a *fuzzy*-like way. The more present the semantic feature is in the word, the closer the microfeature is to 1.0.

TABLE I

The semantic microfeature dimensions for verbs according to a thematic frame.

control of action	*no control of action*
direct process triggering	*indirect process triggering*
direction of action to source	*direction of action to goal*
impacting process	*no impacting process*
change of state	*no change of state*
psychological state	*no psychological state*
objective action	*no objective action*
effective action	*no effective action*
high intensity of action	*low intensity of action*
interest on process	*no interest on process*

Since the aim of the presented system is to deal with thematic relationships between words in a sentence, the microfeatures chosen for verbs attempt to contemplate the semantic issues considered relevant in a thematic frame. The microfeatures outside this context are meaningless [7].

TABLE II

The semantic microfeature dimensions for nouns.

human	*non-human*
soft	*hard*
small	*large*
pointed	*non-pointed*
fragile	*non-fragile*
food	*non-food*
tool/utensil	*non-tool/utensil*
value	*non-value*
furniture	*non-furniture*
animate	*inanimate*

It is important to know that, in BIOPLA lexicon, there are two-sense verbs with only one thematic grid (for instance, *love*: according to *WordNet*[1], there are four senses for verb *love* and here are employed two of them (1 and 2)), and there are "thematically ambiguous" verbs with only one sense, but two thematic grids (for instance, *break*, *frighten*, *hit*).

V. The biologically inspired proposed system

A. The system BIOPLA

BIOPLA is a symbolic-connectionist hybrid system designed to process the thematic roles (θ-roles) of natural language sentences. Symbolic data concerning the specific application are inserted as initial *hidden* knowledge (tables III and IV). These values reflect the expected features for verbs and nouns for each thematic role, according to linguistic theory. *Hidden* knowledge corresponds to connection weights between input and hidden layers, while *output* knowledge is related to connection weights between hidden and output layers. The inserted knowledge respects the *and neuron* notion: the summation of positive and negative weights, including bias, must be at most 1.0. Then the system begins to learn, in a supervised way, through presentations of semantically sound sentence-thematic grid pairs.

After the connectionist training, a symbolic data set is extracted (tables V, VI, and VII). For each input sentence, the system gives as output, its thematic grid. It learns through a training procedure based on the Recirculation [13] and GeneRec algorithms [4] in a bi-directional architecture representing a three-layer neural network with forty input units (twenty for the verb and twenty for the nouns), twenty hidden units, and ten output units, one for each of the ten thematic roles: AGENT, PATIENT, EXPERIENCER, THEME, SOURCE, GOAL, BENEFICIARY, CAUSE, INSTRUMENT, and VALUE.

[1] WordNet version 2.0: http://www.cogsci.princeton.edu/cgi-bin/webwn2.0

TABLE III

INITIAL HIDDEN SYMBOLIC KNOWLEDGE FOR VERBS INSERTED INTO THE NETWORK AS CONNECTION WEIGHTS. ABBREVIATIONS: y = YES (0.2); n = NO (-0.2).

θ-ROLE	*ca*	*dt*	*ds*	*im*	*cs*	*ps*	*oa*	*ea*	*hi*	*ip*
AGENT	y	y		y			y			y
PATIENT			n	y	y			y	y	
EXPERIENCER			y		n		n	n		n
THEME			n	n	n			y	n	
SOURCE		y	y		n			y		y
GOAL	y		n			n	y			y
BENEFICIARY	y	y	n		n			y		
CAUSE	n	n					n			n
INSTRUMENT	y			y			y	y		y
VALUE	y	y			n			y		y

Verb microfeatures legend: *ca = control of action, dt = direct process triggering, ds = direction of action to source, im = impacting process, cs = change of state, ps = psychological state, oa = objective action, ea = effective action, hi = high intensity of action, ip = interest on process.*

TABLE IV

INITIAL HIDDEN SYMBOLIC KNOWLEDGE FOR NOUNS INSERTED INTO THE NETWORK AS CONNECTION WEIGHTS. ABBREVIATIONS: y = YES (0.2); n = NO (-0.2).

θ-ROLE	*hu*	*so*	*sm*	*po*	*fr*	*fo*	*to*	*va*	*fu*	*an*
AGENT	y	y		n	n					y
PATIENT	n	n		n	y					n
EXPERIENCER	y	y		n			n			y
THEME	y	y		n			n			y
SOURCE	y	y		n	n					y
GOAL	y	y		n	n					y
BENEFICIARY	y	y		n	n					y
CAUSE	n				n	n		n		
INSTRUMENT	n	n		y	n		y			
VALUE	n		y			n	n	y		

Noun microfeatures legend: *hu = human, so = soft, sm = small, po = pointed, fr = fragile, fo = food, to = tool/utensil, va = value, fu = furniture, an = animate.*

Lateral inhibition occurs at the output layer. Figure 1 shows the architecture considering only one thematic role.

The words, represented by their semantic microfeatures, are presented at input layer, sequentially, one at a time, at their specific slots, depending on their syntactic categories (verb or noun), until the whole sentence is completely entered. This way, besides semantics, included as part of the distributed representation employed, syntactic constraints are also considered. At output layer, thematic roles are highlighted as soon as they are assigned. For instance, when the subject of a sentence is presented, no thematic role shows up, because it is unknown which will be the main verb, the predicate that assigns such role. When the verb appears, immediately the network displays the thematic role assigned to the subject

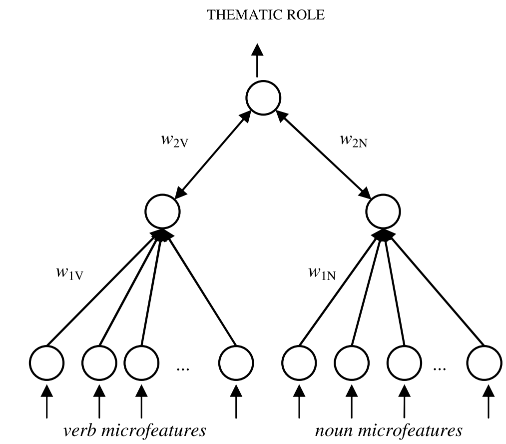

Fig. 1. The three-layer bi-directional connectionist architecture of BIOPLA for one thematic role (adapted from [9]). To the input layer the words, represented by their distributed microfeatures, are entered sequentially at their specific slots according to their syntactic category (verb or noun). These same inputs are fed into similar structure for the other thematic roles. The connections from the input layer to the hidden layer have weights w_{1V} and w_{1N}, representing the *hidden* weights of the verb input and the noun input respectively, and the bi-directional connections from the hidden layer to the output layer have *output* weights w_{2V} and w_{2N}. These weights can be extracted from the connectionist architecture as final symbolic data - see tables V, VI, and VII

TABLE V

FINAL HIDDEN SYMBOLIC KNOWLEDGE FOR VERBS EXTRACTED FROM THE NETWORK (CONNECTION WEIGHTS BETWEEN INPUT AND HIDDEN LAYERS - w_{1V} - SEE FIGURE 1)

θ	*ca*	*dt*	*ds*	*im*	*cs*	*ps*	*oa*	*ea*	*hi*	*ip*
A	**1.0**	**2.1**	**-4.8**	**3.0**	0.0	**-2.6**	**1.7**	**6.2**	**4.1**	**2.1**
P	**-3.3**	**-2.9**	**-8.7**	**1.9**	**1.5**	**-10.**	**-3.3**	**5.5**	**1.1**	**-2.9**
E	-0.8	**-2.0**	**4.2**	**-2.6**	0.0	**3.0**	**-2.0**	**-6.0**	**-3.3**	**-2.3**
T	-0.3	0.2	-0.1	**-1.3**	0.9	0.1	0.5	-0.3	-0.1	0.2
S	**3.9**	**5.3**	-0.5	**1.3**	-0.6	**-1.3**	**3.1**	**1.9**	**-1.2**	**5.3**
G	**4.1**	**4.9**	**-1.1**	**1.5**	-0.5	**-1.5**	**3.5**	**1.7**	**-1.3**	**5.3**
B	0.9	**2.1**	**-4.9**	**2.8**	-0.3	**-2.9**	**1.5**	**6.6**	**4.0**	**1.7**
C	**-4.2**	**-5.7**	0.7	**-1.2**	0.4	**1.4**	**-3.9**	**-1.4**	**1.4**	**-5.7**
I	**4.3**	**4.9**	-0.9	**1.7**	-0.6	**-1.3**	**3.4**	**2.1**	**-1.3**	**5.3**
V	**4.2**	**5.3**	-0.8	**1.3**	-0.6	**-1.3**	**3.1**	**2.1**	**-1.2**	**5.3**

Thematic role (θ) legend: A = AGENT, P = PATIENT, E = EXPERIENCER, T = THEME, S = SOURCE, G = GOAL, B = BENEFICIARY, C = CAUSE, I = INSTRUMENT, V = VALUE. Verb microfeatures legend: *ca = control of action, dt = direct process triggering, ds = direction of action to source, im = impacting process, cs = change of state, ps = psychological state, oa = objective action, ea = effective action, hi = high intensity of action, ip = interest on process.*

presented previously. For the other words, the correspondent thematic roles are displayed at the output, one at a time, for every input word [6].

TABLE VI

FINAL HIDDEN SYMBOLIC KNOWLEDGE FOR NOUNS EXTRACTED FROM THE NETWORK (CONNECTION WEIGHTS BETWEEN INPUT AND HIDDEN LAYERS - w_{1N} - SEE FIGURE 1)

θ	*hu*	*so*	*sm*	*po*	*fr*	*fo*	*to*	*va*	*fu*	*an*
A	**1.2**	0.0	-0.6	**1.0**	0.6	**1.6**	0.4	**1.6**	**1.2**	0.8
P	**-13.**	**3.7**	**15.**	**-20.**	**11.**	0.6	**-6.7**	-0.8	**4.8**	0.1
E	**1.2**	0.1	-0.4	0.7	0.7	**1.1**	0.0	**1.1**	**1.1**	**1.0**
T	**-1.0**	0.3	0.7	**-1.7**	-0.7	**-1.6**	-0.5	**-1.9**	**-1.1**	-0.7
S	**7.7**	-0.4	-0.2	**1.1**	0.8	**1.5**	0.8	**1.4**	**1.2**	**1.6**
G	**4.0**	-0.2	-0.4	0.8	0.6	**1.3**	0.6	**1.1**	**1.0**	**1.4**
B	**4.8**	**2.1**	**-1.8**	-0.5	-0.5	**-1.8**	**-1.8**	**-1.1**	**-1.6**	**19.2**
C	**-16.**	**-1.0**	0.0	-0.2	0.0	-0.5	0.9	**-1.2**	0.4	**-3.6**
I	**-14.**	**-1.3**	0.3	0.2	-0.1	-0.3	**1.3**	-0.5	0.2	**-3.7**
V	-0.3	**2.4**	**4.5**	**2.8**	**-8.2**	**-1.6**	**-3.8**	-0.8	0.3	0.8

Thematic role (θ) legend: A = AGENT, P = PATIENT, E = EXPERIENCER, T = THEME, S = SOURCE, G = GOAL, B = BENEFICIARY, C = CAUSE, I = INSTRUMENT, V = VALUE. Noun microfeatures legend: *hu = human, so = soft, sm = small, po = pointed, fr = fragile, fo = food, to = tool/utensil, va = value, fu = furniture, an = animate.*

B. The learning procedure

The learning algorithm used in BIOPLA is inspired by the Recirculation [13] and GeneRec algorithms [4], and consists of two phases (figure 2).

TABLE VII

FINAL OUTPUT SYMBOLIC KNOWLEDGE FOR VERBS AND NOUNS EXTRACTED FROM THE NETWORK (CONNECTIONI WEIGHTS NETWEEN HIDDEN AND OUTPUT LAYERS - w_{2V} AND w_{2N} - SEE FIGURE 1)

THEMATIC ROLE	*verb* - w_{2V}	*noun* - w_{2N}
AGENT	**2.0**	0.7
PATIENT	**3.4**	**7.3**
EXPERIENCER	0.9	0.8
THEME	**1.1**	**1.0**
SOURCE	-0.6	**-2.1**
GOAL	-0.6	**-1.6**
BENEFICIARY	-0.6	**5.3**
CAUSE	**1.9**	**1.9**
INSTRUMENT	-0.7	**3.3**
VALUE	-0.6	**2.2**

In the *minus* phase, the semantic microfeature representation of the first word of a sentence is presented to the input layer A. Then, there is a propagation of these stimuli x to the output through the hidden layer B (*bottom-up propagation*). There is also a propagation of the previous actual output $o(t-1)$, which is initially empty, from output layer C back to the hidden layer B (*top-down propagation*). Then, a hidden minus activation (h^-) is generated (sum of the bottom-up and top-down propagations - through the sigmoid logistic activation function σ (equation 3)). Finally, the current actual output $o(t)$ is generated through the propagation of the hidden minus activation to the output layer (equation 4) [2].

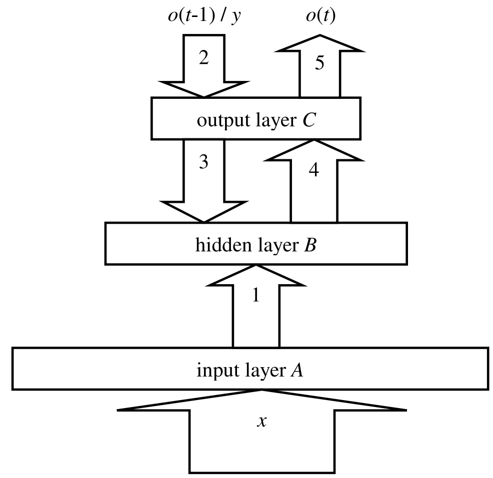

Fig. 2. The two phases of the GeneRec algorithm [4]. In the *minus* phase, when input x is presented to input layer A, there is propagation of these stimuli to the hidden layer (1). Then, a hidden minus signal is generated based on inputs and previous output stimuli $o(t-1)$ (2 and 3). Then, these hidden signals propagate to the output layer C (4), and an actual output $o(t)$ is obtained (5). In the *plus* phase, inputs x are presented to layer A again; there is propagation to hidden layer (1). After this, expected outputs y (2) are presented to the output layer and propagated back to the hidden layer (3), and a hidden plus signal is generated, based on inputs and on expected outputs. Recall that the architecture is bi-directional, so it is possible for the stimuli to propagate either forwardly or backwardly [9]

$$h_j^- = \sigma(\Sigma_{i=0}^{A} w_{ij}.x_i + \Sigma_{k=1}^{C} w_{jk}.o_k(t-1)) \ (3)$$

$$o_k(t) = \sigma(\Sigma_{j=1}^{B} w_{jk}.h_j^-) \ (4)$$

In the *plus* phase, there is a propagation of x from input layer A to the hidden layer B (bottom-up). After this, there is the propagation of the expected output y to the hidden layer (top-down). Then a hidden plus activation (h^+) is generated, summing these two propagations (equation 5). For the other words, presented one at a time, the same procedure (*minus* phase first, then *plus* phase) is repeated [2].

$$h_j^+ = \sigma(\Sigma_{i=0}^{A} w_{ij}.x_i + \Sigma_{k=1}^{C} w_{jk}.y_k) \ (5)$$

In order to make learning possible the synaptic weights are updated (equations 6 and 7), considering only the local information made available by the synapse. The learning rate η used in the algorithm is considered an important variable during the experiments [14].

$$\Delta w_{jk} = \eta.(y_k - o_k(t)).h_j^- \ (6)$$

$$\Delta w_{ij} = \eta.(h_j^+ - h_j^-).x_i \ (7)$$

VI. MAIN RESULTS

Researchers have shown that biologically plausible systems can be as efficient as conventional models, even better [15], [9]. This paper shows that a connectionist system, with biologically plausible architecture and learning procedures, is also computationally efficient, at least regarding a particular natural language processing application. Through the revealing results, one can see that BIOPLA learned the thematic relationships expected for a great number of syntactic constructions.

Firstly, symbolic data concerning thematic roles are inserted as initial *hidden* connection weights into BIOPLA architecture (tables III and IV). After training, a new data set can be extracted from the network, revising the initial thematic symbolic theory (tables V, VI, and VII).

A. Training

A sentence generator, which employs a given limited lexicon, generates syntactic and semantic sound sentences in order to train the system. After about $100,000$ training cycles, reaching an average output error of 10^{-2} (according to the *average squared error energy formula* [14]), with learning rate of 0.25, the system is able to display the learned thematic grid for an input sentence. It is important to mention here that for smaller learning rates, the performance decreases (several learning rates were tested; a better performance was achieved with 0.25).

B. Extracted symbolic data

For symbolic data extraction, real numbers are obtained from the network connections, corresponding to connection weights. The symbolic knowledge thus extracted from the connectionist architecture is related to the network learning and generalization capacities. As a consequence, the network is able to "revise" the initial symbolic theory [7].

Table V displays the *hidden* connection weights for verbs extracted from the network, regarding the ten thematic roles of BIOPLA, between input and hidden layers (w_{1V}). The numbers in **bold** show significant weights (values above 1.0). To arrive at the displayed numbers, it was considered the difference between the two features inside a dimension. The verb dimensions are composed of two real numbers. Inside each dimension (for instance, *control of action* and *no control of action* are values of the same dimension - see table I), what really matters is the difference between the values. For instance, concerning the thematic role PATIENT, *control of action* is -1.6 and *no control of action* 1.7, which gives a difference of -3.3, number shown on table V. This means that the verb that assigns the thematic role PATIENT does not present *control of action* as one of its semantic features.

Another interesting attribute about PATIENT is that it is the only thematic role that presents the feature *change of state*. It was assigned as initial *hidden* feature (table III) and it was confirmed as final hidden extracted data (table V). Another number that is worth of paying attention to is the very large negative value of the feature *psychological state* in the extracted data. This means that verbs that assign the thematic role PATIENT learned by the system should have *no psychological state* as one of its semantic features.

Regarding thematic role AGENT, initially the verb that assigns this role is expected to present *control of action*, the process should have *direct triggering*, be *impacting*, be *objective*, and the subject should have *interest on process* (initial *hidden* symbolic data inserted - see table III). One can notice that all these features are highlighted at the end of training (table V), and others appear also: *direction to goal*, *no psychological state*, *effective action*, and *high intensity of action*. This means that the system "discovered" new features to be associated to the verb that assigns the thematic role AGENT.

According to table IV, the noun to which is assigned the thematic role AGENT is initially considered *human*, *soft*, *non-pointed*, *non-fragile*, and *animate*. Neither of these features was strongly confirmed at the end of training, as shown on table VI. The reason is that the noun does not influence the thematic role AGENT as the verb does (see on table VII, the *output* weight between hidden and output layers, w_{2V}, relative to the verb, is relevant (2.0), while the output weight w_{2N} relative to noun is below 1.0).

The thematic role CAUSE is a good example to be examined because this role has similar final *output* weights between hidden and output layers for verb and noun (table VII). The verb that assigns the thematic role CAUSE is expected to present (table III): *no control of action*, *indirect process triggering*, *no objective*, and *no interest on process*. All of these were highly strengthened. In addition, *no impacting process*, *psychological state*, *no effective action*, and *high intensity of action* appear, but with lesser values.

The noun that is assigned the thematic role CAUSE is supposed to be (table IV): *non-human*, *non-fragile*, *non-food*, and *non-value*. *Non-human* was very highlighted. *Non-value* appears also, but with a small rate. *Non-fragile* and *non-food* have insignificant values. Maybe this is so because they are not relevant to this thematic role. Also, BIOPLA training showed that CAUSE should be *inanimate* and *hard* also.

For the other thematic roles, similar results can be observed. This outcome is very representative since it shows that a connectionist system with a bi-directional architecture and an algorithm which are supposed to be more biologically realistic reveals that the "symbolic" data extracted from the connectionist architecture confirm, in a more consistent way, the semantic features expected for each thematic role.

It is interesting to notice that the way the system learns allows improving itself. In other words, if a designer sets a wrong, or inadequate, value to a particular weight, it will be possible to correct this value after training, because of the inconsistency observed on the extracted data displayed.

VII. CONCLUSIONS

The proposed system BIOPLA is a symbolic-connectionist hybrid approach to natural language processing. In this approach, the advantages of symbolic systems are combined with the benefits of connectionism to yield a more discriminating thematic role processing.

BIOPLA deals with ambiguity, since its lexicon contains lexically ambiguous nouns and verbs. There are also the *thematically ambiguous* verbs, which, despite having only one sense, can assign two thematic grids for different sentences.

This paper aims to show that a biologically plausible symbolic-connectionist hybrid system, consisting of a bi-directional architecture and a learning algorithm that uses only local information to update its connection weights, is able not only to take care of a natural language processing problem, but also to be computationally efficient. This is confirmed by symbolic data extracted from the connectionist architecture, reflecting the semantic features expected for ten thematic roles taught to the system.

REFERENCES

[1] F. H. C. Crick, "The recent excitement about neural networks," *Nature*, vol. 337, pp. 129–132, 1989.

[2] J. L. G. Rosa, "A biologically inspired connectionist system for natural language processing," in *Proceedings of the 2002 VII Brazilian Symposium on Neural Networks (SBRN 2002)*. Recife, Brazil: IEEE Computer Society Press, Nov. 11–14, 2002, pp. 243–248.

[3] ——, "An artificial neural network model based on neuroscience: Looking closely at the brain," in *Artificial Neural Nets and Genetic Algorithms - Proceedings of the International Conference in Prague, Czech Republic - ICANNGA-2001*, V. Kurková, N. C. Steele, R. Neruda, and M. Kárný, Eds. Prague, Czech Republic: Springer-Verlag, Apr. 22–25, 2001, pp. 138–141.

[4] R. C. O'Reilly, "Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm," *Neural Computation*, vol. 8:5, pp. 895–938, 1996.

[5] L. Haegeman, *Introduction to Government and Binding Theory*. Blackwell, 1991.

[6] J. L. G. Rosa, "A biologically motivated connectionist system for predicting the next word in natural language sentences," in *Proceedings of the 2002 IEEE International Conference on Systems, Man, and Cybernetics - IEEE-SMC'02*, vol. 4, Hammamet, Tunisia, Oct. 6–9, 2002.

[7] J. L. G. Rosa and E. Françozo, "Hybrid thematic role processor: Symbolic linguistic relations revised by connectionist learning," in *Proceedings of IJCAI'99 - Sixteenth International Joint Conference on Artificial Intelligence*, vol. 2. Stockholm, Sweden: Morgan Kauffman, July 31–6 Aug. 1999, pp. 852–857.

[8] ——, "Linguistic relations encoding in a symbolic-connectionist hybrid natural language processor," in *Lecture Notes in Computer Science - Advances in Artificial Intelligence, Proceedings of the International Joint Conference 7th. Ibero-American Conference on AI - 15th. Brazilian Symposium on AI - IBERAMIA-SBIA 2000*, M. C. Monard and J. S. Sichman, Eds., vol. 1952. São Paulo, Brazil: Springer-Verlag Heidelberg, Nov. 19–22, 2000, pp. 259–268.

[9] J. L. G. Rosa, "A biologically plausible and computationally efficient architecture and algorithm for a connectionist natural language processor," in *Proceedings of the 2003 IEEE International Conference on Systems, Man, and Cybernetics - IEEE-SMC'03*, Washington, D.C., USA, Oct. 5–8, 2003, pp. 2845–2850.

[10] A. S. d'Avila Garcez, K. Broda, and D. M. Gabbay, "Symbolic knowledge extraction from trained neural networks: A sound approach," *Artificial Intelligence*, vol. 125, pp. 155–207, 2001.

[11] D. Dowty, "On the semantic content of the notion of 'thematic role'," in *Properties, Types and Meaning*, G. Chierchia, B. H. Partee, and R. Turner, Eds. Dordrecht, Kluwer, 1989.

[12] J. L. McClelland and A. H. Kawamoto, "Mechanisms of sentence processing: Assigning roles to constituents of sentences," in *Parallel Distributed Processing*, J. L. McClelland and D. E. Rumelhart, Eds. A Bradford Book, MIT Press, 1986, vol. 2.

[13] G. E. Hinton and J. L. McClelland, "Learning representations by recirculation," in *Neural Information Processing Systems*, D. Z. Anderson, Ed. New York: American Institute of Physics, 1988, pp. 358–366.

[14] S. Haykin, *Neural Networks - A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[15] R. C. O'Reilly and Y. Munakata, *Computational Explorations in Cognitive Neuroscience - Understanding the Mind by Simulating the Brain*. Cambridge, Massachusetts, USA: A Bradford Book, The MIT Press, 2000.