Logic, Linguistics and Connectionism

An exercise of translation of optimality theory constraints into logic

Teresita de Jesús Mijangos Martínez

Universidad de la Sierra Sur, Oaxaca, México
tmijangos@unsis.edu.mx

Abstract. Universal Grammar assumes that all human languages share a common structure with respect to their linguistic well-formedness conditions. In the linguistic theory called Optimality Theory (OT), this structure is represented by a set of linguistic candidate forms, a set of constraints and the definition of *being optimal*. From a nativist approach of OT is assumed that such structure is encoded in a Language Acquisition Device (LAD) that is provided to the learner of a language genetically. While OT is in a symbolic level, the LAD is in a biological one. These two levels have already been complemented with a connectionist intermediate level. In this paper our aim is to propose a complementary connection between the symbolic and the connectionist levels using *penalty logic*. We propose a translation from connectionism to logic expressed in the penalty knowledge base cv. This last translation try to simplify the encoding of the common structure $\mathfrak{U}$ suggested by CV_{net}.

Keywords: Non monotonic reasoning, Penalty logic, Universal Grammar, Optimality Theory, connectionism.

1 Introduction

In 1990 a new framework for generating linguistic theories arose, it was called "Optimality Theory" (OT). This Universal Grammar (UG) framework states —as every UG— that among human languages there are a set of universals, that is, there is a common structure that they share that constitutes the conditions of well-formedness of a language. Despite this heritage, OT differs of others traditional generative frameworks in the sense that it does not requiere inviolability in order to state universality. This OT aspect is quite interesting because the best candidate form will not be the perfect one (zero constraint violations) but the best one or *optimal*, i.e., that one whose violations of the constraints have the minimal cost. Besides in OT is assumed that the set of constraints is in principle inconsistent, so, every candidate of a linguistic form will violate at least one constraint. Which constraint or constraints are not fulfilled by a candidate form is important with respect to the position of those constraints in a hierarchy H that constitutes a specific language.[1]

OT was born in a context in which generative grammar was the main framework for researching works in linguistics. We can mention the name of "Chomsky" as an icon

[1] For an introduction to optimality theory *Cfr.*, [2].

of this framework. OT differs from generative grammar in many aspects, some of them are: (*Cfr.*[3])

1. In OT constraints a higher position in the hierarchy has more value, their violation has a bigger penalty or cost. But in order to select the *best* candidate, it is not enough to consider the position of its violated constraints. For selecting the *optimal* form it must be considered which constraints others competitors have violated. In other words, a grammatical or *optimal* form in OT depends always on other forms. This is a difference between OT and generative grammar, being the forms in the last one, singly evaluated with respect to certain rules.
2. In OT environment, constraints are universal and violable, in generative grammar they are universal but inviolable.
3. In OT markedness plays an important role in a grammar while in generative grammar is failthfulness the element that plays the central role.
4. OT does not conceive outputs in terms of duality, so, neither violability means inactivity nor satisfaction means activation. This dual way of thinking belongs to generative grammar.
5. For explaining language variation OT uses the re-ranking of constraints, each language has a special constraints ranking. Generative grammar explains language variation with elements like rules/parameters and the lexicon.

It is important to point out that in OT there are two kind of constraints: violable and inviolable constraints. In the CV Syllable Theory as is formulated in [8] the inviolable constraints are the *Gen* and the *Structural* constraints, these constraints are beyond any hierarchy, they have the highest cost of all constraints and are fulfilled in every candidate competition. The violable constraints are the *Con* constraints. For the examples presented here, we are going to consider only the set *Con*, i.e., the set of well-formedness conditions as is defined in [8, 972].

In *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, Paul Smolensky and Géraldine Legendre propose a connectionist model of OT called CV_{net}. This localist symmetric network assumes a nativist point of view of OT. CV_{net} is thought as the Language Acquisition Device(LAD) that encodes $\mathfrak{U}$, the common structure for all posible human languages. We are going to start from that work in order to give our translation of CV_{net} into a default logic system called *penalty logic*. The translation presented here enriches the CV_{net} proposal and simplifies it, besides it makes possible a link between OT and logic.

2 The Basic CV Syllable Theory

In this section we will introduce one of the most important issues in OT, the Basic CV Syllable Theory. The translation into connectionism and then, the translation into logic will be based on the CV Syllable Theory.

In the definition of the nativist hypothesis of OT introduced in [8], OT is integrated by a set of linguistic candidate forms *S*, a set of constraints *Con* and the definition of *being optimal*. In that definition *Gen* and *Structural* constraints are not mentioned, given their necessity they are assumed fulfilled for every candidate competition. We are going to take that assumption and to work only with the violable constraints, the set of *Con* constraints.

In the Basic CV Syllable Theory, $\mathfrak{U}$ is constituted —in parallelism with OT— by a set of candidate forms *S*, constraints and a definition of *being optimal*. In particular, in the CV Theory the universe U_S of structural descriptions is formed by strings of consonants and vowels, i.e., C's and V's. Interpretations are strings of C's and V's that are not parsed into syllables like /VCVC/, /CVCVC/ while expressions are strings of C's and V's parsed into syllables. The parsing is indicated by a period. For example, from the interpretation /VCVC/ = /$V^1C^2V^3C^4$/ the following candidates could be generated:[2]

1. $\square V^1.C^2V^3.\langle C\rangle$
2. $\langle V\rangle.C^2V^3.\langle C\rangle$
3. $\langle V\rangle.C^2V^3.C^4\square.$
4. $V^1.C^2V^3C^4.$

From these candidates one could be selected as the *optimal*, but in order to do that selection we need a hierarchy of constraints and a definition of *being optimal*. Here we introduce the definition:

Definition 1. *A candidate w is considered to be optimal iff for each competitor w', the constraints that are violated by w must be ranked lower than at least one constraint violated by w'.* [1, 20]

The set of constraints *Con* for the CV Theory is formed by: [8],[5]

– *Onset*. Every syllable nucleus has a preceding onset.
– *NoCoda*. Codas are not permitted.
– *Parse*. For every element in the input there is a corresponding element in the output.
– *Fill*^{V}. Every syllable nucleus in the output has a corresponding element in the input.
– *Fill*^{C}. Every syllable onset or coda has a corresponding C in the input.

In OT a candidate competition is usually showed in a table where the candidate competitors are shown under a certain hierarchy H over the constraints of the set *Con*. Suppose that *Con* constraints are under the following hierarchy: $Fill^C$ » *Parse* » $Fill^V$ » NoCoda » *Onset* and for the input /$V^1C^2V^3C^4$/ the same candidates expressions mentioned above are generated as outputs. The OT table would look like:

Asterisks indicate violations on constraints. Candidate c_2 for example, has two violations on *Parse* constraint while candidate c_4 one violation over *NoCoda* and one over *Onset*. Observe that all the candidates have at least one constraint violation. The *optimal* is c_4 because for each competitor c' the constraints that are violated by c_4, i.e., *No Coda* and *Onset* are ranked lower than at least one constraint violated by any c'.

It is important to point out that under another hierarchy, the *optimal* output could not be the same. For example, if we consider the hierarchy H' where the constraints are ordered as follows: *Onset* » *No Coda* » $Fill^V$ » *Parse* » $Fill^C$, the *optimal* candidate is c_1.

[2] The symbol '□' indicates that an onset, nucleus or coda is empty; <*x*> is read as 'the element *x*' is unparsed

input: / $V^1C^2V^3C^4$/ →	*Fill* ^{C} >>	*Parse* >>	*Fill* ^{V} >>	*NoCoda* >>	*Onset*
c_1: .□V^1. C^2V^3.⟨C⟩	*	*			
c_2: ⟨V⟩ C^2V^3.⟨C⟩		**			
c_3: ⟨V⟩. C^2V^3.C^4□.		*	*		
c_4:. V^1. $C^2V^3C^4$.				*	*

Fig. 1. Candidate comparison in OT

3 A connectionist model of OT

In this section we are going to introduce CV_{net}, the model proposed by Paul Smolensky and Géraldine Legendre in [8]. In order to prevent confusions we will distiguish between two kind of constraints: constraints at the OT level and at the connectionist level. The first set of constraints we will call them OT-constraints, the second one, *net*-constraints.

As we consider the OT level, constraints are universal as they belong to the structure $\mathfrak{U}$ while constraints in the connectionist level are considered universal when their weights are identical or bounded. For OT-constraints we have the distinction between violable and inviolable constraints, this distinction is introduced in the connectionist level as constraints with variable weights —*net-Con* constraints— and constraints with fixed weights —*net-Gen* and *net-Structural* constraints.

CV_{net} is a localist symmetric network that represent the mapping given in CV Syllable Theory from underlying strings of C's and V's to surface forms of C's and V's, in other words, CV_{net} represents interpretations and expressions. In order to do that, this net has three layers: an input, an output and a corresponding layer. In the first one we will find the activation of an interpretation,[3] for example, the activation of the string /$V^1C^2V^3C^4$/. In the output layer we could find activated, one of the four candidate we have introduced above. The corresponding layer introduce the link between the input and the output.

CV_{net} has three kinds of units: vowels, consonants and coda consonants represented by an inverted triangle, a circle and a crescent respectively. In the input and the corresponding layer there is no distinction between onset or coda consonants, in these layers, the circle is interpreted only as a consonant. In the output layer the circle represents an onset consonant while a crescent is in place of a coda consonant. In figure 2 is showed the activation of candidate c_4 where the input is the interpretation /$V^1C^2V^3C^4$/ and the expression generated in the output is $V^1.C^2V^3C^4$. In this net input layer goes from the cell a_{12} to the cell a_{15}; output layer from a_{21} to a_{61}.[4]

In CV_{net} each constraint has assigned a bias and a connection coefficient. *net-Onset* has a negative bias -1 on every vowel unit in the output and a connection coefficient $+1$

[3] In this paper we are working only with the speakert's perspective for that reason we will find as inputs only interpretations, and as outputs, expressions.
[4] In a cell a_{ij}, *i* represents the row and *j* the column.

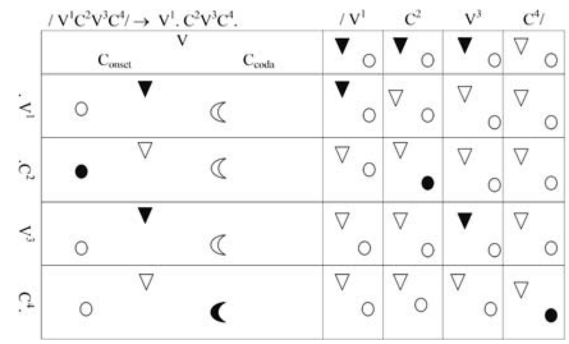

Fig. 2. Network activity for c_4 in CV_{net}

between every vowel and onset consonant units in the output. The connection coefficient will help to equilibrate the negative value of the bias. For *net-No Coda* a bias of -1 is in every coda unit at the output, but instead of a positive connection coefficient a negative coefficient -1 is introduced. The reason for this negative coefficient is that in the CV Syllable Theory codas are not permitted.

net-Parse has a bias of -1 in each input unit. A bias of -3 is assigned to each unit in the corresponding layer and $+2$ for all connections coefficients. For *net-Fill*^{V} and *net-Fill*^{C}, in the corresponding layer, the bias will be -3 and connection coefficients $+2$. Bias coefficient in the output will be -1. Observe that in *net-Parse* there was values with respect to the input and corresponding layers while in *net-Fill* the values are in the corresponding and output layers.

3.1 The energy function E of net-constraints

In OT the candidate with the less ranked violations in relation with other candidates is the *optimal*. In the connectionist environment the *optimal* candidate will have the minimal energy E with respect to the rest of candidates. In order to calculate the value E in relation to different activation states, we will use the Ljapunov or *energy* function:

$$E(s) = -\sum_{i<j} w_{ij} sisj \ . \tag{1}$$

In order to simplify the calculation of the energy function E, we will use *microconstraints*. It is possible to do this in CV_{net} because is a localist network. We identify the minimal significance part of the network –that constitutes the microconstraints— and then, we spread the result to a more complex net built with those elements.[5]

For describing the states of the microconstraints networks, we use instead of s_1, s_2, ... s_n the labels *v* for a vowel, *co* for an onset consonant, *cc* for coda consonants and *b* for biases, this will help us for an easier understanding. Graphical representations of

[5] For a graphical representation of *Net-constraints* look at [8]; for microconstraints [3].

microconstraints are given in a similar way as we have stated in CV$_{net}$, we only add to this representation one figure: a circle with a cross inside for representing biases. The representation of the *micro-Onset* is also a symmetric network. As we have mentioned above, in this microconstraint there is a bias -1 on every vowel, equilibrated by the connection coefficient $+1$. This is showed in the figure below.[6]

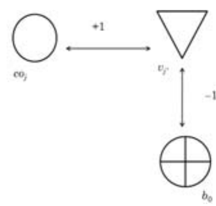

Fig. 3. *Micro-Onset*

Then, we calculate the energy function E for this system as: E(s) $= -co_j v_{j'} + v_{j'} b_0$. The optimal scenarios for the *micro-Onset* are two, with -2 as value: when the onset consonant and the vowel are both activated but the bias are not; or when consonant and the vowel are both inactivated but the bias is activated. In analogous ways we can get the calculation of the energy function E for the rest of constraints. The importance of energy function here is that it will permit us to calculate the *optimal* candidate of a network that represents a knowledge base. This function will be the link with the logical model introduced in the next section.

4 A logic model of OT

Our model is based on [4] but it gives to that proposal a specific orientation. Here we are going to expose only some elements of our model, those that can help you to understand the translation of the next section.

Penalty logic is an extended version of propositional calculus, that helps for modeling nonmonotonic reasoning, specifically cases with conflicting beliefs or inconsistencies that arise when there are no reliable knowledge sources. As its name indicates Penalty Logic associates a penality or cost to each formula of a knowledge base Ψ. This cost —expressed in a real positive number— will generate a ranking among the formulae in the knowledge base. Our strong beliefs will have a high price and will be at the top of the ranking, the penalty for losing these formulas will be more costly than our weaker beliefs.

Definition 2. *Let $\mathcal{L}$ be the basic language of a standard propositional calculus. A penalty knowledge base (PKB) Ψ is a finite set of pairs (S, k), being S a set of well*

[6] Other microconstraints are represented taking into account the global representation of the constraints. In order to have a graphical representations of the constraints *Cfr.*[8].

formed formulae of $\mathfrak{L}$ and k the penalty function mapping from S to the set (0, ∞). Δ_Ψ represents the set of propositional formulae of $\mathcal{L}$ that are in Ψ.[7]

Example of a PKB. An interesting aspect of Penalty Logic is that can deal with inconsistent knowledge bases. Penalties are used as a criterion for selecting the preferred consistent subsets. Here we are going to introduce an example of a PKB but instead of introducing the well known example of Tweety, the penguin, we have worked in a new example with Fifo, the platypus as the main actor.

Suppose that we believe that none mammal is oviparous, that all platypus are mammals and that they are oviparous; and the strength of our beliefs has the values 20, 5 and 10 respectively. Consider a set of propositional atoms $\mathcal{M} = \{m_0, m_1, \ldots, m_n\}$ where m_i stands for '*i* is a mammal'. A set $\mathcal{O} = \{o_0, o_1, \ldots, o_n\}$ and a set $\mathcal{P} = \{p_0, p_1, \ldots, p_n\}$ also of propositional atoms, where o_i and p_i are in place of '*i* is an oviparous' and '*i* is a platypus' respectively. Our PKB will be constituted by the following propositional formulae with their corresponding penalties to the right:

$m_i \to \neg o_i$	20
$p_i \to m_i$	5
$p_i \to o_i$	10

This PKB that we call *example* is inconsistent.[8] Suppose that the individual *i* is Fifo and Fifo is a platypus. With this assumption plus the PKB we can infer from the first two formulae that Fifo is not an oviparous, but from the third formula it follows that Fifo is an oviparous. Then, Fifo is and is not an oviparous. In order to deal with inconsistencies penalty logic works with scenarios, that is, consistent subsets formed from the PKB.

Definition 3. *Let φ be a formula of $\mathcal{L}$. A scenario φ in a PKB is a consistent subset Δ' of Δ_Ψ such that Δ' $\cup$ {φ} is consistent. The cost of a scenario Δ' in a PKB Ψ, written $\mathcal{K}_\Psi$(Δ'), is the sum of the penalties of the formulae of Ψ that do not belong to Δ'. That is,*

$$\mathcal{K}_\Psi = \sum_{\delta \epsilon (\Delta_\Psi - \Delta')} k(\delta) \ . \tag{2}$$

where $k(\delta)$ stands for the penalty or cost of an expression δ. [3, 54]

Example of scenarios
Take φ = p_i and assume the PKB *example*, the following scenarios could be formed:[9]
$\Delta_1 = \{m_i \to \neg o_i, p_i \to m_i\}$
$\Delta_2 = \{m_i \to \neg o_i, p_i \to o_i\}$
$\Delta_3 = \{m_i \to \neg o_i\}$

[7] Here we have made a little change in the Pinkas definition of a well formed formula (wff) in penalty logic. Instead of consider a wff as a set of pairs given by the union of a propositional formula plus a positive real number, we have called that wff a knowledge base, a wff of a penalty logic is for us, each pair formed by *S* and *k*.
[8] We will write Δ_{ex} in order to indicate the formulae that belong to this PKB.
[9] Many more escenarios can be formed from Δ_{ex}, here we have only indicated as examples.

Definition 4. *The optimal scenarios of φ are those that have the minimal cost $\mathcal{K}$ with respect to a PKB Ψ.*

Examples. According to definition 3 the cost of the scenarios mentioned above will be: $\mathcal{K}_{ex}(\Delta_1) = 10$, $\mathcal{K}_{ex}(\Delta_2) = 5$ and $\mathcal{K}_{ex}(\Delta_3) = 15$. According to definition 4, the *optimal* scenario of p_i will be Δ_2.

Definition 5. *Let v* an interpretation function that maps the set S of well formed formulae of $\mathcal{L}$ to the set $\{-1, 1\}$. The system energy of v is calculated by:*

$$\mathcal{E}_{\Psi} = \sum_{\delta \epsilon \Delta_{\Psi}, \|\delta\|_v = 1} k(\delta) \ . \tag{3}$$

The function *v** maps well formed formulae to the set -1, 1. Where $v(\varphi)= -1$ is interpreted as φ is unsatisfied and $v(\varphi)= 1$ as φ is satisfied. The *energy function* E as considered in CV_{net} introduce the values $[1, 0, -1]$ for describing the different states of the network. The mapping introduced from the connectionist environment to the logical one, is parcial and bivalent. We take only the cases 1 and -1 from the energy function E, excluding those that consider 0.

Example. This formula tells us that in order to calculate the *energy system* of an interpretation *v**, we have to take into account those formulae that belong to the PKB Ψ but that they are not satisfied under that interpretation, that is, under *v**. Suppose that *v** assigns the following valuation for the formulae of the PKB Ψ: $v(m_i \rightarrow \neg o_i)= 1$, $v(p_i \rightarrow m_i)= v(p_i \rightarrow o_i) = -1$. Given that the last two formulae are not satisfied under *v**, we should sum up the penalties of these formulae (5 + 10) according to definition 3. The *energy system* for p_i will be 15.

Definition 6. *The models of a PKB Ψ with the minimal energy $\mathcal{E}$ are the preferred models of Ψ.*

Example of preferred models. We are going to introduce analogous models to the scenarios presented above. To the right of each model its energy is calculated according to definition 5.
In $\mathcal{M}_1$, $v(m_i \rightarrow \neg o_i)= v(p_i \rightarrow m_i)= 1$, $v(p_i \rightarrow o_i) = -1$. $\mathcal{E}_{ex}(v^*)= 10$.
In $\mathcal{M}_2$, $v(m_i \rightarrow \neg o_i)= v(p_i \rightarrow o_i) = 1$, $v(p_i \rightarrow m_i)=-1$. $\mathcal{E}_{ex}(v^*)= 5$.
In $\mathcal{M}_3$, $v(m_i \rightarrow \neg o_i)= 1$, $v(p_i \rightarrow m_i)= v(p_i \rightarrow o_i) = -1$. $\mathcal{E}_{ex}(v^*)= 15$.

According to the definition 6, the preferred model of Ψ will be $\mathcal{M}_2$, because its has the minimal energy $\mathcal{E}$. This result matches with the obtained for the scenarios above. The matching is not by chance, equivalences between preferred models of penalty logic and minimal energy has been proven in [4].

5 Translation of CV_{net} into Penalty Logic

The PKB of the Basic Syllable Theory will be formed by all the constraints of OT. Here, we will introduce only the logic representation of the *Con* constraints. The translation is based on CV_{net}, so, the symmetric aspect of this network will be translated into logic as biconditionals. First, we introduce the formulae used in PKB *cv* and then, the symbolization of the *Con* constraints.

5.1 Formulae

Let INPUT, OUTPUT AND CORRESPOND be the sets of input, output and corresponding units respectively and let *i*, *j* be the indexes of the units that can be substituted by any natural number. We have the following sets of propositional atoms:[10]
$\{v^0, v^1, \dots v^n\} \in$ INPUT with v^i standing for: 'unit *i* is a vowel input unit'.
$\{c^0, c^1, \dots c^n\}$ with c^i standing for: 'unit *i* is a consonant input unit'.
$\{v_{ij}\} \in$ CORRESPOND where $i \geq 0$, $j \geq 0$ and $\{v_{ij}\}$ and stands for 'unit *i* from *j* is a corresponding vowel unit'.
$\{c_{ij}\} \in$ CORRESPOND where $i \geq 0$, $j \geq 0$ and $\{c_{ij}\}$ and stands for 'unit *i* from *j* is a corresponding consonant unit'.
$\{v_0, v_1, \dots v_n\} \in$ OUTPUT with v_j standing for: 'unit *j* is a vowel output unit'.
$\{co_0, co_1, \dots co_n\} \in$ OUTPUT with co_j standing for: 'unit *j* is an onset consonant (output) unit'.
$\{cc_0, cc_1, \dots cc_n\} \in$ OUTPUT with cc_j standing for: 'unit *j* is a coda consonant (output) unit'.
$\{b_0, b_1, \dots b_n\}$ where b_i stands for: 'unit *i* is a bias on'.

5.2 Translation of CV_{net}

In order to identify the OT constraints in the penalty logic environment, we will use the prefix *p* for calling each constraint. For example, *Onset* constraint will be called *p-Onset*. Below the formal representation of *Con* constraints.

p-Onset $\{j < j'$: $co_j \leftrightarrow v_j'\}$, $\{v_j' \leftrightarrow \neg\, b_0\}$
p-NoCoda $\{cc_j \leftrightarrow \neg b_0\}$
p-Parse $\{v^i \leftrightarrow v_{ij}\}$, $\{v_j \leftrightarrow v_{ij}\}$, $\{v^i \leftrightarrow \neg b_0\}$, $\{v_{ij} \leftrightarrow \neg b_1\}$, $\{c^i \leftrightarrow c_{ij}\}$, $\{co_j \leftrightarrow c_{ij}\}$, $\{cc_j \leftrightarrow c_{ij}\}$, $\{c^i \leftrightarrow \neg b_2\}$, $\{c_{ij} \leftrightarrow \neg b_3\}$.
p-Fill^V $\{v^i \leftrightarrow v_{ij}\}$, $\{v_j \leftrightarrow v_{ij}\}$, $\{v_{ij} \leftrightarrow \neg b_0\}$ $\{v_j \leftrightarrow \neg b_1\}$.
p-Fill^C $\{c^i \leftrightarrow c_{ij}\}$, $\{co_j \leftrightarrow c_{ij}\}$, $\{cc_j \leftrightarrow c_{ij}\}$, $\{c_{ij} \leftrightarrow \neg b_2\}$, $\{co_j \leftrightarrow \neg b_3\}$, $\{cc_j \leftrightarrow \neg b_4\}$.

These formulae represent the *preferred models* as it has shown in [3], then in a network environment they have the minimal energy $\mathcal{E}$.

Additional to the logical representation of constraints *Con*, we need to induce the hierarchy among these constraints. This we can get it with exponential penalties assigned to each constraint. For example, we can take the decimal position system, then *n* in $(\frac{v}{n})^k$ would be equal to 10, where n represents the number of violations that are possible under a single constraints;[11] v is instead of the number of violations (or asterisks in OT table) and k is a natural number that grows as the hierarchy goes down, i.e. $(\frac{v}{n})^1$, $(\frac{v}{n})^2, \dots (\frac{v}{n})^n$.

The learner of a language has to identify violations on the output candidates, for this task he can help himself with a logical reasoning using the logical representation of

[10] Remember that in CV_{net} we had input, output and corresponding units, some of them represented vowels, coda or onset consonants and the bias.
[11] n might be infinite.

constraints *Con* plus the inference mechanism of penalty logic. A second task that must do is to select the *optimal* output, for this task the constraint hierarchy of a language will support the selection in a inconsistent set of constraints.[12]

In this paper it has been assumed a nativist reading of OT, but it important to clarify that OT is not necessary linked to nativism. From the nativist approach, the structure $\mathfrak{U}$ integrated by a set of linguistic candidate forms, a set of well formedness conditions *Con* and the definition of best satisfying a hierarchy of violable constraints (*being optimal*) is encoded in the learner since his birth by a Language Acquisition Device (LAD). In this route, we could think in the logical representation proposed above, as a logic implementation of the human system. For example, if the system believes that $v(b_0)= 1$ for *p-NoCoda*, that means that he believes that the bias associated to *p-NoCoda* is turn on.[13] From this valuation, he could infer for example, $\neg\, cc_j$, that is, there is no a unit *j* that is a coda consonant. The truth of the statement that unit *i* is a bias on, will help to the learner to infer that he should look for no codas.

6 Conclusions and Future Work

We have presented a translation of CV_{net} network into Penalty Logic. The translation has been done using a bivalent mapping. For further work we consider important to extend it to a *n*-valent mapping in order to model situations in which we cannot assign values consistently for calculating the energy function of the system.[14]

Acknowledgments

I am grateful to José Alberto Cruz Tolentino for his suggestions for possible applications of the translation here presented and to Samuel Sánchez Hernández for his bibliography recommendations on phonological stuff. Besides I want to thank to Víctor Alberto Gómez Pérez and Alejandro Jarillo Silva for their invaluable support in technical business with LaTeX.

[12] In the execution of the first task the learner could realize that even in an accepted syllabification of a word, there are constraints violated. So, he could ask himself how should I select the *optimal* output if there is no output without mistake? The ranking induced by the hierarchy will help him to order candidates by priority. The set of constraint is inconsistent so, there is no possibility that all of them were fulfilled or not violated.

[13] The learner should believe that valuation as product of the neural encoding that he also has of the constraints, in the connectionist level the *net-Coda* shows a connection coefficient negative. In the symbolic environment, this value of the coefficient multiplied by the negative bias will be positive. For that reason it is said that the valuation for the sentence b_0 will be equal to 1.

[14] We are thinking in scenarios like $\Delta = \{p_i \rightarrow o_i\}$.

References

1. Blutner, R..: Language and Optimality. Seminar at Humboldt University/University of Oslo/ ESSLLI, Helsinki (2001)
2. McCarthy, J.: A Thematic Guide to Optimality Theory. Cambridge University Press, Cambridge (2002)
3. Mijangos Martínez, T.: Penalty Logic and Genomic Enconding. UVA, Amsterdam (2006)
4. Pinkas, G.: Reasoning, Nonmonotonicity and Learning in Connectionist Networks that Capture Propositional Knowledge. Artificial Intelligence Vol. 77, pp. 203-247. Elsevier (2006)
5. Prince, A. and Smolensky, P.: Optimality Theory, Constraint Interaction in Generative Grammar. Blackwell, Malden, MA. (1993/2004)
6. Prince, A. and Smolensky, P.: Optimality Theory, Constraint Interaction in Generative Grammar. Technical Report, Linguistic Department, Rutgers University, New Brunswick, NJ and Department of Computer Science, University of Colorado at Boulder. Rutgers Optimality Archive 537 (1993/2002)
7. Rumelhart and McClelland: Nativism vs. Empiriscism. In: Rumelhart, McClelland, and the PDP Research Group, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Vol. 1: Foundations; pp. 139-142. The MIT Press, Cambridge, MA (1986)
8. Smolensky, P. and Legendre, G.: The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. Vol. 1: Cognitive Architecture. The MIT Press, Cambridge, MA (2006)