In: N. Biggs et al. (Eds.), Proceedings of Waves 2007. The 8th International Conference on Mathematical and Numerical Aspects of Waves, pp. 120 – 122, Dept. of Mathematics, University of Reading, 2007.

FOCK SPACE REPRESENTATIONS IN NEURAL FIELD THEORIES

Peter beim Graben[†,*], **Sabrina Gerth**[‡], **Douglas Saddy**[†] **and Roland Potthast**[§]
[†]**Dept. of Clinical Language Sciences, University of Reading, UK**
[‡]**Inst. for Linguistics, University of Potsdam, Germany**
[§]**Dept. of Mathematics, University of Reading, UK**
[*]**Email: p.r.beimgraben@reading.ac.uk**

Abstract

We outline a computational model of syntactic language processing based on Smolensky's Fock space representations of symbolic expressions using spherical wave functions. Symbolic computation, regarded as nonlinear operators acting upon these waves, provides a discrete sequence of training patterns that could be used to solve the inverse problem of neural field theories in order to determine the synaptic connectivity/weight kernels. The solutions of a neural field equation should then provide a model of event-related brain potentials that are elicited by syntactic processing problems.

Introduction

The difference between linguistic competence and performance becomes most obvious in the processing of *garden path sentences* such as "the lawyer charged the defendant was lying" [1], where an initially preferred interpretation of "the defendant" as an direct object of "charged" has to revised. *Psycholinguistics* investigates such effects that are due to *processing strategies* in performance rather than to syntactic competence e.g. by means of event-related brain potentials (ERP). Figure 1 illustrates the spatio-temporal ERP dynamics elicited by the processing of dispreferred object-verb-subject ($\mathtt{ovs}$) sentences (solid curves) compared to default subject-verb-object sentences ($\mathtt{svo}$) (dotted curves) in German [2], [3]. Evidently, the brain waves related to the disambiguating words diverge around 600 ms after stimulus presentation where the $\mathtt{ovs}$ condition exhibits a *syntactic positivity shift (P600)*.

Complementarily, *computational psycholinguistics* [4] is concerned with computational models of such experimental data that can provide insights into the nature of the cortical processes. Here, we shall outline a model of language-related brain waves in a neural field theory [5].

Fock Space Representations

The construction of our computational model comprises four steps: (1) a context-free toy grammar (CFG) and a processing automaton (the *parser*) [6] are (2) represented by activation vectors and a dynamics in a neural network using Smolensky's [7] *tensor product representations*. However, (3) these representations exhibit several disadvantages (cf. [8]) which could be remedied in an infinite setup, leading to (4) a representation by spherical wave functions.

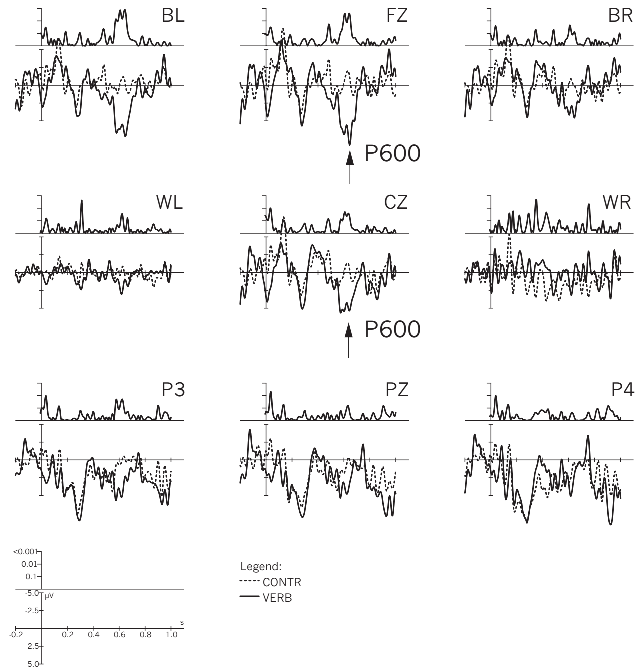


Figure 1: ERPs from a language processing experiment. "CONTR" refers to $\mathtt{svo}$, "VERB" to $\mathtt{ovs}$.

Grammars

The sentence material of the ERP experiment [2], [3], $L = \{\mathtt{svo}, \mathtt{ovs}\}$, can be generated by the CFG [6] $G = G_1 \cup G_2$, where $G_1 = (\{\mathtt{s}, \mathtt{v}, \mathtt{o}\}, \{S, VP\}, \{S \to \mathtt{s}VP, VP \to \mathtt{vo}\}, S)$ and $G_2 = (\{\mathtt{s}, \mathtt{v}, \mathtt{o}\}, \{S, VP\}, \{S \to VP\mathtt{s}, VP \to \mathtt{ov}\}, S)$ where G_1, G_2 are locally unambiguous CFGs [9] reflecting the different processing strategies: G_1 generates the string $\mathtt{svo}$ while G_2 supplies $\mathtt{ovs}$.

Following [7], structured symbolic expressions such as lists or trees can be represented as activation vectors in a neural network by tensor products

$$\mathbf{v} = \bigoplus_n \mathbf{f}_n \otimes \mathbf{r}_n \tag{1}$$

of "filler", $\mathbf{f}_n$, and "role" vectors, $\mathbf{r}_n$, respectively. For the given case, elementary "fillers" are the variables $\mathtt{s}, \mathtt{v}, \mathtt{o}, S, VP \in \mathbb{R}^5$ of G, regarded as "spins", and "roles" are the positions ROOT $= |0\rangle$, LEFT DAUGHTER $= |1\rangle$, and RIGHT DAUGHTER $= |2\rangle \in \mathbb{R}^3$ of a labeled binary tree.

Equation (1) allows for recursion. In our example, the phrase structure tree for the sentence svo, $[_S\mathtt{s}[_{VP}\mathtt{vo}]]$, assumes the representation $S|0\rangle \oplus \mathtt{s}|1\rangle \oplus VP|02\rangle \oplus \mathtt{v}|12\rangle \oplus \mathtt{o}|22\rangle$, where $|k_1k_2\rangle$ is the "two-particle state" $|k_1\rangle \otimes |k_2\rangle$. In general, these vectors are elements of the "many particle" Fock space

$$\mathcal{F} = \bigoplus_{n=1}^{\infty} \mathcal{V}_F \otimes \bigotimes_{k=1}^{n} \mathcal{V}_R \tag{2}$$

where $\mathcal{V}_\mathcal{F}$ is a finite vector space spanned by the elementary fillers (analogous to particle spins) and $\mathcal{V}_R$ is a vector space of elementary roles.

Parsers

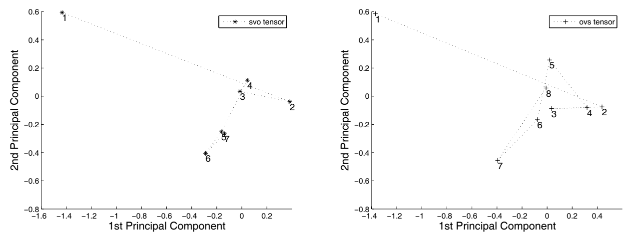
Since the grammars G_1, G_2 are locally unambiguous, the strings svo and ovs can be deterministically processed, e.g. by *top-down parsers* which successively construct the phrase structure trees by predicting the input [6]. Let us here consider the grammar G_1. The stack of a top-down parser is initialized with the start symbol of a grammar, $S|0\rangle$. As S is a nonterminal symbol, the content of the stack is replaced by the tensor product representation of the unique rule expanding S, i.e., by $S|0\rangle \oplus \mathtt{s}|1\rangle \oplus VP|2\rangle$. Next, the parser finds the nonterminal filler VP which is replaced by the representation $VP|0\rangle \oplus \mathtt{v}|1\rangle \oplus \mathtt{o}|2\rangle$, yielding the complete tree given above. Therefore, the automata are described by maps $\alpha : \mathcal{F} \to \mathcal{F}$ acting in the following way: each nonterminal filler $\mathbf{f}_n$ occurring in Eq. (1) is recursively replaced by a complex filler corresponding to the rule with left-hand-side $\mathbf{f}_n$.

A similar approach using minimalist grammars [10] gave the state space trajectories in a two-dimensional PCA projection of a 275,562-dimensional Fock space, shown in Fig. 2 [8].

Problems

The trajectories plotted in Fig. 2 were obtained by using linearly dependent role vectors because the dimension of the activation space became 1,879,048,192 using linearly independent ones [8]. Unfortunately, the treatment of these representations was numerically not feasible.

However, the very high space dimensions required even for finite language models is only one problem of Smolensky's proposal. Others are



(a) (b)

Figure 2: PCA projections of a "minimalistic" Fock space representation for language processing. (a) Trajectory of the svo parse, (b) that of the ovs parse.

- the processing of lists of unbounded length or of trees with arbitrary recursion depth, respectively, requires *infinite* dimensional spaces,
- but, only vertices have symbolic representations (brain state in-the-box model [7]); therefore
- most states occupy a meaningless vacuum; meaningful states are rather improbable.

Smolensky's solution of the dimensionality problem to use always finite-dimensional representations with linearly dependent role and/or filler vectors does not seem to be appropriate as it leads to cross-talk [7], thereby preventing a faithful mapping between the symbolic and the connectionist dynamics. Thus, Smolensky stresses that there is no *implementation* of symbolic processes in neural dynamics [11]. Since we are looking for such implementations, yet, we shall allow for infinite-dimensional Fock spaces to encounter the dimensionality problem in the following.

Wave functions

Our starting point are the linearly independent elementary role vectors of our CFG model. These can be identified with a triplet of angular momentum states: $|0\rangle \cong |1,0\rangle, |1\rangle \cong |1,-1\rangle, |2\rangle \cong |1,1\rangle$, which, in turn, have an $L_2(S)$ representation by spherical harmonics $|\ell m\rangle \cong Y_{\ell m}(\vartheta, \varphi)$ at the unit sphere S. This appears as a suitable choice as these functions are often used as basis functions of EEG waves and neural fields.

In the next step, tensor products of the role vectors are to be computed. In our notation, $|12\rangle$, e.g., denotes the left daughter of the right daughter of the tree's root, and is given as $|12\rangle = |1\rangle \otimes |2\rangle = |1,-1\rangle \otimes |1,1\rangle \cong Y_{1,-1}(\vartheta_1, \varphi_1)Y_{1,1}(\vartheta_2, \varphi_2)$.

In quantum theory, the product states $|\ell_1 m_1\rangle \otimes |\ell_2 m_2\rangle$ belong to different multipletts with total angular momentum ℓ given by the Clebsch-Gordan sums $|\ell, m, \ell_1, \ell_2\rangle = \sum_{m_1, m_2 = m - m_1} \langle \ell_1 m_1 \ell_2 m_2 | \ell m \ell_1 \ell_2 \rangle |\ell_1 m_1 \ell_2 m_2\rangle$ [12]. Thus, we can map the roles described by the tensor products $|\ell_1 m_1\rangle \otimes |\ell_2 m_2\rangle$ onto a family of spherical wave functions $Y_{\ell m}(\vartheta, \varphi)$ obeying the triangle relation $|\ell_1 - \ell_2| \leq \ell \leq \ell_1 + \ell_2$. Finally, a phrase structure tree will be represented by "spinors"

$$\Psi(\vartheta, \varphi) = \sum_{\ell, m} \mathbf{f}_{\ell m} Y_{\ell m}(\vartheta, \varphi) \tag{3}$$

with the elementary fillers $\mathbf{f}_{\ell m}$ denoting the tree labels.

Neural Field Theory

In the preceding sections we outlined a way how to represent symbolic content by tensor products in an infinite-dimensional Fock space of spherical wave functions $Y_{\ell m}(\vartheta, \varphi)$. Correspondingly, symbolic computation, such as language processing α, will be represented by (presumably nonlinear) operators $\mathbf{A}_\alpha : \Psi(\vartheta, \varphi) \to \Phi(\vartheta, \varphi)$. A time-discrete trajectory of these wave functions $\{\Psi_t | t \in \mathbb{N}_0\}$ describes a cognitive process with initial condition Ψ_0.

These functions can be regarded as stationary solutions of a neural field equation [5]

$$\begin{aligned} & \tau(x) \frac{\partial \Psi(x,t)}{\partial t} + \Psi(x,t) = \\ = \quad & \int_{-\infty}^{t} \mathrm{d}t' \int_S \mathrm{d}x' \, w(x, x') \; \times \\ \times \quad & G(t - t') \, f \left[\Psi \left(x', t' - \frac{||x - x'||}{c} \right) \right] \end{aligned} \tag{4}$$

at the unit sphere, $x = (\vartheta, \varphi)$. Here, $\tau(x)$ describes the neural time constants, the kernel $w(x, x')$ the synaptic connectivity and weights, $G(t - t')$ the postsynaptic impulse response, f the activation function and c the propagation velocity of neural activity.

Regarding Ψ_t as *training patterns* for the field equation, the *inverse problem* of determining the kernel $w(x, x')$ from the solutions Ψ_t, entails a cognitive implementation by a neural field theory. Such a model would be able to predict the dynamics of ERPs evoked by processing problems such as garden path sentences [1], [2], [3].

References

[1] L. Osterhout, P. J. Holcomb, and D. A. Swinney. Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *J. Exp. Psych. LMC*, 20(4):786 – 803, 1994.

[2] P. beim Graben, D. Saddy, M. Schlesewsky, and J. Kurths. Symbolic dynamics of event-related brain potentials. *Phys. Rev. E*, 62(4):5518 – 5541, 2000.

[3] D. Saddy and P. beim Graben. Measuring the neural dynamics of language comprehension processes. In T. Witruk, A. D. Friederici, and T. Lachmann, editors, *Basic Fuctions of Language, Reading and Reading Disorder*, pages 41 – 60. Kluwer Academic Publishers, Boston, 2002.

[4] R. L. Lewis. Computational psycholinguistics. In *Encyclopedia of Cognitive Science*. Macmillan Reference Ltd., 2000.

[5] C. J. Rennie, P. A. Robinson, and J. J. Wright. Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybern.*, 86:457 – 471, 2002.

[6] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison–Wesley, Menlo Park, California, 1979.

[7] P. Smolensky and G. Legendre. *The Harmonic Mind. From Neural Computation to Optimality-Theoretic Grammar*, volume 1: Cognitive Architecture. MIT Press, Cambridge (MA), 2006.

[8] S. Gerth. Parsing mit minimalistischen, gewichteten Grammatiken und deren Zustandsraumdarstellung. Master's thesis, Universität Potsdam, 2006.

[9] P. beim Graben, B. Jurish, D. Saddy, and S. Frisch. Language processing by dynamical systems. *Int. J. Bif. Chaos*, 14(2):599 – 621, 2004.

[10] E. P. Stabler. Derivational minimalism. *SLNCS 1328*, pages 68 – 95. Springer, New York, 1997.

[11] P. beim Graben. Incompatible implementations of physical symbol systems. *Mind and Matter*, 2(2):29 – 51, 2004.

[12] A. R. Edmonds. *Angular Momentum in Quantum Mechanics*. Princeton University Press, 1957.