# Learning the States: A Brain Inspired Neural Model

András Lőrincz

Eötvös Loránd University, Budapest H-1117, Hungary
andras.lorincz@elte.hu

**Abstract.** AGI relies on Markov Decision Processes, which assume deterministic states. However, such states must be learned. We propose that states are deterministic spatio-temporal chunks of observations and notice that learning of such episodic memory is attributed to the entorhinal hippocampal complex in the brain. EHC receives information from the neocortex and encodes learned episodes into neocortical memory traces thus it changes its input without changing its emerged representations. Motivated by recent results in exact matrix completion we argue that step-wise decomposition of observations into 'typical' (deterministic) and 'atypical' (stochastic) constituents is EHC's trick of learning episodic memory.

**Keywords:** sparse coding, exact matrix completion, hippocampus, MDP.

## 1 Introduction

We think that learning of states is the focal problem of Artificial General Intelligence (AGI) in many respects. For example, Markov Decision Process (MDP) model is the key components of AGI [8,12] and MDP starts from the concept of state. In MDP, state has a deterministic flavor since it has no hidden component, it is not spoiled by noise, and is valid during a finite time window. Learning of 'states' matching the MDP framework is challenging since in real world problems there are many variables giving rise to combinatorial explosion. MDP becomes tractable if close-to-deterministic (CtD) processes can be identified. Then a state is the list of ongoing processes, supporting the Markovian assumption. Thus, the separation, memorizing, and recognition of CtD processes or episodes seem to be the key problem. Intriguingly, mammalian species have a special learning architecture, the entorhinal-hippocampal complex (EHC see, e.g., [2]) for this. EHC has puzzling properties, like (a) EHC learns episodic instances and encodes those into the neocortex where information have come from without influencing its own representation and (b) lesion to EHC spoils episodic learning but a large portion of learned episodes is spared. Motivated by recent results in Exact Matrix Completion (EMC) we argue that step-wise decomposition of observations into 'typical' (deterministic) and 'atypical' (stochastic) constituents is EHC's trick and it suits MDPs.

## 2 Two Stages of the Architecture

**Neocortical models.** typically start from overcomplete sparse code (OSC). Assume that $x^i \in \mathbb{R}^n\,(i = 1, \ldots, I)$ is the $i^{th}$ input to be reconstructed, or matched in $\ell_2$

J. Schmidhuber, K.R. Thórisson, and M. Looks (Eds.): AGI 2011, LNAI 6830, pp. 315–320, 2011.

norm, $I$ is the number of training inputs, $h^i \in \mathbb{R}^m$ denotes the coefficient vector of the sparse decomposition, and $D = [d_1, \ldots, d_m]$ ($d_j \in \mathbb{R}^n$, $m \geq n$, and $d_j^T d_j \leq 1$, $j = 1, \ldots, m$) denotes the so called the dictionary or reconstruction matrix consisting of unit norm basis features. OSC task is to optimize both the code and the dictionary [14]:

$$\min_{D\in\mathbb{R}^{n\times m},\ h\in R^m} \sum_{i=1}^{I} \frac{1}{2}\|x^i - Dh^i\|_2^2 + \kappa\|h^i\|_0 \tag{1}$$

where $\|\cdot\|_0$ denotes the $\ell_0$-norm, the number of nonzero components of the argument.

Although the $\ell_0$ minimization problem (1) is NP-hard, under certain conditions exact polynomial solutions can be found by replacing the $\ell_0$ norm with $\ell_1$ norm [5,6].

Our concepts are based on recent revolutionary findings of signal processing about recovering low-dimensional data from high dimensional observations [4,3]. Let us assume that observation matrix $X = [x^1, \ldots, x^I] \in \mathbb{R}^{n\times I}$ ($x^i \in \mathbb{R}^n$, $i = 1, \ldots, I$) is composed of a low-rank component $L = [l^1, \ldots, l^I]$ and a sparse matrix $S = [s^1, \ldots, s^I]$ with few but arbitrarily large components and that $X = L + S$. Under mild conditions (e.g., on the rank of $L$ and the sparsity of $S$) *both* matrices can be *exactly* recovered [4] via, e.g., Robust Principal Component Analysis (RPCA) having the following objective:

$$\text{minimize} \ \ \|L\|_* + \lambda\|S\|_1 \tag{2}$$

subject to $X = L + S$, where $\|L\|_*$ denotes the nuclear norm of matrix $L$, i.e. the *sum* of the singular values of $L$, $\|S\|_1$ denotes the $\ell_1$ norm of matrix $S$, i.e., $\|S\|_1 = \sum_{j=1}^{n}\sum_{i=1}^{I}|S_{ji}|$, and $\lambda$ is the so-called trade-off parameter, which governs the dimension of matrix $L$. On the other hand, matrix $S$ may assume maximal rank, independent of $\lambda$. We use the normalized parameter $\lambda^* = \lambda/\sqrt{\max(n, I)}$ [3].

Sparse components are then expanded into OSC via the double optimization of the code and the dictionary giving rise to our new model

$$\text{minimize} \ \ \sum_{i=1}^{I} \frac{1}{2}\|s^i - Dh^i\|_2^2 + \kappa\|h^i\|_1 \tag{3}$$

$$\tag{4}$$

subject to $x^i = l^i + s^i$ $(\forall\, i)$ as sketched in Fig. 1(a).

**EHC.** is at top of the sensory processing hierarchy and we assume a hierarchical model built from PCA and ICA [13]. The model suits autoregressive (AR) worlds. The interpretation is that at the top of the hierarchy typical and atypical parts have identical dimensions and OSC relaxes to independent component analysis (ICA) [15]. Feedback from EHC targets OSC representations leaving typical channels intact.

## 3 Illustrative Simulations on Natural Images

For the model of the neocortex we tested the impact of RPCA preprocessing on sparse coding. Normalized natural image patches were first decomposed with RPCA at different $\lambda^*$ values and then the resulting sparse parts were further encoded by OSC: 16-fold

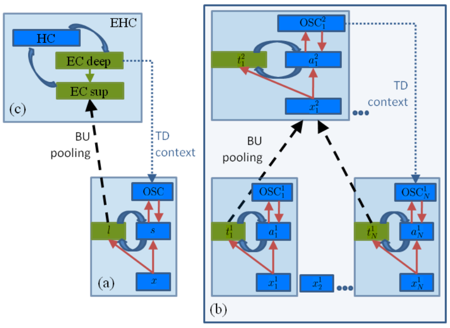


**Fig. 1. Sketch of the hierarchy.** (a): flow diagram. Input $x$ is decomposed into 'typical' ('atypical') part $l$ ($s$) and $x = l+s$. Atypical part is expanded into an OSC. Curved arrows: RPCA based pre-filtering between $l$ and $s$. (b): functional interpretation and hierarchical embedding: step-wise separation of typical ($t$) and atypical ($a$) components. Typical representations: $t_1^q, t_2^q, \ldots, t_p^q$, $p$ and $q$: number of partitions and corresponding layers, $x_1^q, x_2^q, \ldots, x_p^q$: small input parts, dashed arrows: pooling. Dotted arrows: higher order atypical representations provide contextual information for upstream layers, influence atypical representations and leave the bottom-up hierarchy of typical information flow intact. Typical part encodes increasingly invariant features downstream. Sparse part specifies the borders of typical regions. (c): the top of the hierarchy is EHC (HC: hippocampus, EC sup/deep: entorhinal superficial/deep layers). OSC relaxes to ICA [13].

overcompleteness with input dimension $n = 16 \times 16 = 256$ and OSC dimension $m = 4096$. OSC was optimized by a combination of $\ell_0$ and $\ell_1$ methods [11], while the overcomplete dictionary ($D$) was tuned online via stochastic gradient learning.

For natural images, subspace of low-rank matrix $L$ is basically the same for PCA and RPCA algorithms. It is known that PCA produces *global orthogonal grid-like filters* on natural images [9].

The PCA algorithm, however, is insufficient for developing noise free local and overcomplete dictionary: PCA prefiltering can increase the spatio-temporal frequencies of the filters, but these filters remain global (Fig 2(a)). However, when the RPCA algorithm is used for preprocessing, elements of the sparse dictionary become structurally sparse; i.e., localized with decreasing noise content (Fig 2(b)). Locality depends on parameter $\lambda^*$: for large $\lambda^*$ values the Gabor-filter like characteristics vanishes.

## 4 RPCA Based Hierarchical Architecture

RPCA (as opposed to PCA) enables the formation of local OSC dictionary. Note that the philosophy behind RPCA differs from that of PCA: RPCA is motivated by EMC: find the subspace where a small portion of the input is sufficient to fill in the rest *exactly*) even if large (but sparse) outliers are present. RPCA works for spatio-temporal inputs (Fig. 2(c)) and filling in can be extended to the temporal domain that corresponds to the idea of CtD processes. However, EMC conditions are not fulfilled for natural signals with heavy tailed distribution: the 'outliers' are not sparse. In turn, we conjecture that RPCA involves a hierarchy.

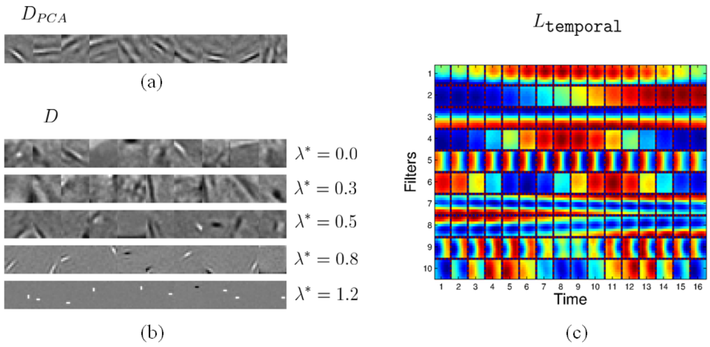


**Fig. 2. Basis types of (R)PCA preprocessing and Sparse Coding.** (a): samples of learned sparse filters after projecting out from PCA subspace: wavy, mostly global and noisy structure. (b): samples of learned sparse filters after RPCA for different $\lambda^*$ values. With increasing $\lambda^*$ the filters get smaller and *cleaner*. (c): first 10 spatio-temporal RPCA bases learned on temporally concatenated input sequences (shown as sequences of 16 frames of size $8 \times 8$ pixels): low-frequency spatio-temporally separable and non-separable filters appear.

Low-dimensional representation corresponds to large (smooth and slowly varying grid-like) structures, whereas OSC receptive fields represent edges. In our interpretation the low-dimensional part corresponds to the predictable slowly varying smooth part of the signal, whereas OSC represents the spatio-temporal borders of large domains. In turn, RPCA gives rise to CtD spatio-temporal chunks with sparse OSC delimiters. Spatio-temporal chunks make the hierarchical and compositional representation of the episodes at different time scales from locally moving edges to autobiographical events.

The EHC loop is similar: typical part represents global hexagonal grids, sparse part corresponds to local places, see e.g. [17]. In addition, HC outputs events of about 1s duration compressed into about 50ms time windows (see [7] and the cited references). This time compression is ideal for learning at the level of synapses and seems relevant for the encoding of episodic memory or *specific sequences* into the neocortex in a top-down fashion. According to our (verbal) model, this top-down prediction influences OSC representations upstream that can hold the details of the individual episodes. Then the typical part of the representation is left intact during top-down encoding.

Key features of our proposed hierarchical architecture are depicted in Fig. 1.

## 5 Discussion

We started by saying that learning of discrete states is crucial for AGI, especially for decision making, and argued that modular spatio-temporal chunks could serve as combinatorial (factored) state-descriptors. We argued that RPCA and OSC can separate predictable chunks and unpredictable markers representing the borders of these chunks. We showed simulations about the effect of RPCA prefiltering that promoted the learning of edges (atypical components) and represented typical low-frequency components also when time was involved.

The interplay between low-rank and sparse representations is of high importance. Consider the categorization of an object as a face. It can be based on the typical properties. In contrast, recognition of a somebody's face requires the encoding of the atypical properties. However, generalization across the actual hair style, etc. is crucial for robust recognition.

**Factors and modules.** We note that OSC and ICA are combinatorial representations and can represent factors of decision making. The EHC loop, indeed holds other representations beyond grids and place cells, e.g., head direction cells [19] as well as the conjunctive representations [17] and decision making can select, e.g., egocentric or allocentric representations, whichever is better in a given situation. Factored reinforcement learning [1,10], which is known to be polynomial [18], captures such ideas.

### 5.1 Searching for States and Top-Down Encoding.

Efforts to learn the MDP states have a long history, see, e.g., [16] and references therein. We conjectured that reinforcement learning in complex environments can be efficient if learned states are composed of combinations of spatio-temporal processes. We suggested that episodes are made of close-to-deterministic parts of spatio-temporal processes together with the sparse delimiters of those that make the low-dimensional and the sparse components of the representation, respectively.

We have argued that the hierarchical separation of typical and atypical processes enable the EHC to keep its internal representation, while encoding the details into the neocortex: novel temporal processes and their markers can be learned in sparse representations upstream. Although the convergence of this procedure remains to be shown, we expect that convergence may be warranted by means of subtle constraints: the sparse dictionaries of different layers have to be matched. Such matching is possible and has been demonstrated for high and low resolution portions of sparse representation [20].

## 6 Conclusions and Outlook

We suggested a model of the neocortex and made and attempt to connect it to an EHC model [13] in order to learn discrete states for decision making and planning in the MDP framework with discrete Markovian 'states'. We argued that such states correspond to deterministic processes that may start or halt. Recent advances in EMC offer a novel possibility since one may separate spatio-temporal inputs to typical and to specific parts. The former can be seen as an approximation of deterministic process, whereas specific features can represent the spatio-temporal boundaries of the processes. Interestingly, RPCA delivers meaningful decomposition of signal for which the conditions can not be validated [3]. We conjectured that learning at different spatio-temporal scales may require a hierarchical architecture and model matching for all levels might be guided by the top, the EHC.

**Acknowledgments.** I am grateful for enlightening discussions with Zsolt Palotai and Gábor Szirtes. Research was supported by the European Union and co-financed by the European Social Fund (grant agreement no. TÁMOP 4.2.1./B-09/1/KMR-2010-0003).

## References


1. Boutilier, C., Dearden, R., Goldszmidt, M.: Stochastic dynamic programming with factored representations. Artificial Intelligence 121(1-2), 49–107 (2000)
2. Buzsáki, G.: Theta rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. Hippocampus 15, 827–840 (2005)
3. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? (2009), http://arxiv.org/abs/0912.3599
4. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. of Comput. Math. 9, 717–772 (2008)
5. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory 52, 489–509 (2006)
6. Donoho, D.: Compressed sensing. IEEE Trans Inf. Theory 52, 1289–1306 (2006)
7. Geisler, C., Diba, K., Pastalkova, E., Mizuseki, K., Royer, S., Buzsáki, G.: Temporal delays among place cells determine the frequency of population theta oscillations in the hippocampus. Proc. Nat. Acad. Sci. 107, 7957–7962 (2010)
8. Hutter, M.: Feature reinforcement learning: Part I. Unstructured MDPs. J. Artif. Gen. Intell. 1, 3–24 (2009)
9. Hancock, P., Baddeley, R., Smith, L.: principal components of natural images, T. Network: Comp. Neural Syst. 3, 61–70 (1992)
10. Kearns, M.J., Koller, D.: Efficient reinforcement learning in factored MDPs. In: Proc. 16th Int. J. Conf. on Artif. Intel., pp. 740–747 (1999)
11. Lőrincz, A., Palotai, Z., Szirtes, G.: Spike-based cross-entropy method for reconstruction. Neurocomputing 71, 3635–3639 (2008)
12. Lőrincz, A., Bárdosi, Z., Takács, D.: Sketch of an AGI architecture with illustration. In: Conference on Artificial General Intelligence (AGI 2010). Adv. Intell. Syst. Res. (2010)
13. Lőrincz, A., Szirtes, G.: Here and now: how time segments become events in the hippocampus. Neural Networks 22, 738–747 (2009)
14. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–609 (1996)
15. Olshausen, B.A., Field, D.J.: Sparse coding with an overcomplete basis set: A strategy employed by V1? Vision Res. 37, 3311–3325 (1997)
16. Samejima, K., Omori, T.: Adaptive internal state space construction method for reinforcement learning of a real-world agent. Neural Networks 12, 1143–1155 (1999)
17. Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.B., Moser, E.I.: Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science 312, 758–762 (2006)
18. Szita, I., Lőrincz, A.: Factored value iteration converges. Acta Cybern. 18, 615–635 (2008)
19. Taube, J.S.: The head direction signal: Origins and sensory-motor integration. Ann. Rev. Neurosci. 30, 181–207 (2007)
20. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. IEEE Trans. Image Proc. (to appear, 2011)