Proceedings of the Twenty-Ninth International
Florida Artificial Intelligence Research Society Conference

Causal Probabilistic Graphical Models for Decoding Effective Connectivity in Functional Near InfraRed Spectroscopy

Samuel Antonio Montero-Hernández, Felipe Orihuela-Espina, Javier Herrera-Vega and **Luis Enrique Sucar**
Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Sta. Maria Tonantzintla, Puebla, Mexico
{samuel,f.orihuela-espina,vega,esucar}@ccc.inaoep.mx

Abstract

Uncovering effective relations from non-invasive functional neuroimaging data remains challenging because the physical truth does not match the modelling assumptions often made by causal models. Here, we explore the use of causal Probabilistic Graphical Models for decoding the effective connectivity from functional optical neuroimaging. Our hypothesis is that directions of arcs of the connectivity network left undecided by existing learning algorithms can be resolved by exploiting prior structural knowledge from the human connectome. A variant of the fast causal inference algorithm, seeded FCI, is proposed to handle prior information. For evaluation, we used an existing dataset from prefrontal cortical activity of a cohort of 62 surgeons of varying expertise whilst knot-tying was monitored using fNIRS. Seeded FCI is used to built the prefrontal effective networks across expertise groups to reveal expertise-dependent differences. As hypothesized, the incorporation of prior information from the connectome reduces the set of undecided links. Good nomological validity is achieved when data is retrospectively compared to the findings in the original publication of the dataset. We contribute to the analysis of effective connectivity in fNIRS with the incorportation of structural information, and contribute to the field of causal PGMs with a new structure learning algorithm capable of exploiting existing knowledge to reduce the number of links remaining undecided when only information from observations is used. This work has implications thus for both, the AI and the neuroscience communities.

Introduction

Modelling the effective (causal) connectivity is a necessity to understand the brain behaviour by interrogating the neural activity from a neuroimaging modality. Functional near infrared spectroscopy (fNIRS) (Jöbsis 1977) is an optical neuroimaging modality for monitoring the concentration of both oxygenated and reduced haemoglobin species allowing to obtain a set of observations of the neural activity. Notwithstanding, causal relations cannot be modelled from observations alone (Pearl 2009). It necessitates that either certain assumptions about the world are made, and/or external information is exploited.

Probabilistic graphical models (PGMs) encode statistical independence relations among random variables by decomposing the underlying joint probability distribution. Recently, the use of these models have been suggested as a suitable alternative to capture functional and effective connectivity from fMRI data. This opens an interesting new analytical line considering the wealth of variants available among PGMs. Surprisingly, the use of causal PGMs for effective connectivity analysis remains uncharted territory.

The exploration of effective connectivity has been attempted with analytical approaches of varied nature. For instance, coherence based methods such as directed transfer function (DTF) (Kaminski and Blinowska 1991) and partial directed coherence (PDC) (Baccalá and Sameshima 2001) were early attempts to retrieve the causal graph in neuroimages. Structural equation modelling (SEM) is a multivariate analysis technique in which the causal relationships are not inferred from the data but assumed *a priori* (McIntosh and Gonzalez Lima 1994) although is oblivious to temporal information. Dynamic causal modelling (DCM) (Friston, Harrison, and Penny 2003) proposes a double layered, neural and haemodynamic, bilinear model in terms of differential equations. DCM has become the *de facto* standard for revealing the effective connectivity. A closely related work was that of Rajapakse and Zhou (Rajapakse and Zhou 2007) and their use of DBNs.

In this research, we explore the suitability of a causal PGM to reveal the effective relations among brain regions as monitored with fNIRS. *A priori* knowledge was exploited to guide the learning of the structure of the causal PGM. In order to incorporate such background information, this work further contributes with an extension of the Fast Causal Inference (FCI) algorithm, the seeded FCI (sFCI), so that the remaining undecided relations can be resolved. Using real fNIRS dataset from an experiment for dexterity assessment in novices, trainees and expert surgeons (Leff et al. 2007) three connectivity networks were constructed. In all cases the number of undefined links decrease in contrast with those obtained when applying plain FCI to the dataset. The values in metrics of characteristic path length, global efficiency and network diameter show consistency with previous work reported by Ohuchida (Ohuchida et al. 2009).

Causal Bayesian Networks

A Causal Bayesian Network (CBN) $\mathcal{C}$ consists of a $(\mathcal{G}(\mathbf{V},\mathbf{E}),\theta)$ pair, where $\mathcal{G}$ is a DAG and θ encodes the local

distribution such that the independence relations over variables can be obtained by the d-Separation (Sucar 2015) (as in the classical BN). However, unlike classical PGM each edge $(X, Y) \in \mathbf{E}$ means that X is a cause for Y. Let $\mathbf{V}$ be the set of variables on a causal system and $\mathbf{Y} \subset \mathbf{V}$, it is assumed that $\mathcal{C}$ is compatible with all the probability distribution over $\mathbf{V}$ resulting from interventions on $\mathbf{Y} \subseteq \mathbf{V}$.

Learning a causal model from data alone can be hard because: (i) the data set can be originated purely from observations without control over any variables, (ii) maybe there exist unobserved variables in the system, and/or (iii) the set of samples may be limited. Most existing algorithms attempt to go unaffected by these issues (Sucar 2015). In consequence, they rely on a set of assumptions. Common assumptions are *causal markov condition:* whereby a variable is independent of its non-descendants given its direct causes (parents in the graph), *faithfulness:* states that there are no additional independencies between the variables that are not implied by the causal Markov condition, and *causal sufficiency:* assumes that there are no common confounders of the observed variables in the model.

In many situations, the above are not met. For instance, although fNIRS measures the changes in concentration of oxygenated and reduced haemoglobin that are a proxy of brain activity, measures can be contaminated by the systemic contributions such as scalp blood flow, blood pressure or heart rate. Thus, assuming causal sufficiency in fNIRS is daring.

Current causal discovery algorithms are able to discover a equivalence class of the true causal structure. Those models share invariant features, such as same skeleton, v-structures and edge marks (tails or arrowheads) and can be represented via Ancestral Graphs (AG) (Richardson and Spirtes 2002). AG are an extension of DAGs in which bidirect arcs of the type $X \leftrightarrow Y$ and undirect arcs $X - Y$ are allowed.

The equivalence class for AGs is represented by a *Partial Ancestral Graph* (PAG) (Richardson and Spirtes 2002). In a PAG invariant marks are preserved, such that tails ($-$) and arrowheads ($>$) that appear in all members of the class are conserved and the non-invariant edge marks are represented by a circle ($\circ$). A special symbol "$*$" stands for any three kinds of marks, i.e. $X *\!\!\rightarrow Y$ stands for $X \circ\!\!\rightarrow Y$ or $X \leftrightarrow Y$ or $X \rightarrow Y$.

Seeded FCI: Incorporation of prior information to decide the direction of unresolved links

The FCI algorithm (Spirtes, Meek, and Richardson 1995) is a constraint-based algorithm for causal discovery. FCI allows the possibility of considering latent and selection variables. However the output of FCI may be more informative when prior information is encoded in the learning process.

The addition of *a priori* information may resolve some of the undecided directions present in the output of the FCI algorithm. This prior knowledge can be added in form of restrictions either by conditioning on fundamental or unnecessary relations. Algorithm 1 is our proposal, a seeded version of FCI (sFCI), for permitting incorporation of prior information to FCI.

The basic idea in sFCI is to start with a complete undirected graph $\mathcal{Q}(\mathbf{V}, \mathbf{E})$ and a set of invariant links $\mathbf{L}$ (prior information). Then iteratively, select a pair of adjacent variables X, Y in $\mathcal{Q}$ and select a subset of adjacent variables to both X and Y, remove the link between X and Y if they are independent conditioned on $\mathbf{L}$, otherwise hold it, and so on for the rest of adjacent pairs. Next, orient all edges as undefined ($\circ$) and using the result of d-separation test reorient the triplets of the form $A * - * B * - * C$.

The structural connections of the human brain establishes a set of constraints with respect to the possible paths in the causal graph. In neuroimaging, this set of constraints can be obtained from the so called *human connectome*, which establishes the expected physical links in the human brain.

Data: set of variables $\mathbf{V}$, set of a priori links $\mathbf{L}$
Result: A partial ancestral graph $\mathcal{F}$
1) Start a complete undirected graph $\mathcal{Q}(\mathbf{V}, \mathbf{E})$ over the set of nodes $\mathbf{V}$;
2) $n = 0$;
repeat
  repeat
    select an ordered pair $X - Y \in \mathbf{E}$ such that
    $|\text{Adjacencies}(\mathcal{Q}, X) \backslash \{Y\}| \geq n$, and a subset
    $\mathbf{S} \subseteq \text{Adjacencies}(\mathcal{Q}, X) \backslash \{Y\}$ such that $|\mathbf{S}| = n$;
    if *d-separated*$(X, Y|\mathbf{S})$ *and* $X - Y \notin \mathbf{L}$ **then**
      delete the edge $X - Y$ from $\mathbf{E}$;
      record $\mathbf{S}$ in Sepset(X, Y) and Sepset(Y, X);
    end
  until $\forall \{X, Y\} : |$*Adjacencies*$(\mathcal{Q}, X) \backslash \{Y\}| \geq n$ *and*
  $\forall \mathbf{S} \subseteq$ *Adjacencies*$(\mathcal{Q}, X) \backslash \{Y\}$ *such that* $|\mathbf{S}| = n$ *have been tested for d-separation*;
  $n = n + 1$;
until $\forall \{X, Y\} : |$*Adjacencies*$(\mathcal{Q}, X) \backslash \{Y\}| < n$;
3) Let $\mathcal{F}(\mathbf{V}, \mathbf{E}')$ be the undirected graph resulting from step 2), then orient each edge as $\circ - \circ$;
$\forall A - B - C$ such that
$A - B$, $B - C \in \mathbf{E}'$ and $A - C \notin \mathbf{E}'$ and $A - B$, $B - C \notin \mathbf{L}$;
  if $B \notin$ *Sepset*(A, C) **then**
    orient $A * - * B * - * C$ as $A * - > B < - * C$
  end
4) $\forall A - B \in \mathbf{E}'$ and $A - B \notin \mathbf{L}$;
  if *d-separated*$(A, B|\mathbf{S})$ *such that* $\mathbf{S} \in$ *Possible-D-SEP*$(A, B) \backslash \{A, B\}$ *or Possible-D-SEP*$(B, A) \backslash \{A, B\}$ **then**
    remove $A - B$;
    record $\mathbf{S}$ in Sepset(A, B) and Sepset(B, A);
  end

Algorithm 1: Seeded Fast Causal Inference (sFCI) algorithm with a mechanism to consider prior information.

Decoding effective connectivity in fNIRS

Unlike the functional connectivity, effective connectivity is concerned with decoding cooperating brain regions, and most importantly, determining the direction of the flow of information.

The system of interest is a fNIRS neuroimaging which is able to take a snapshot of the cortical activity across brain regions by acquiring bivariate data at each channel while the subject is performing a certain task.

The differences in the physical processes of image formation, and the image reconstruction function across neu-

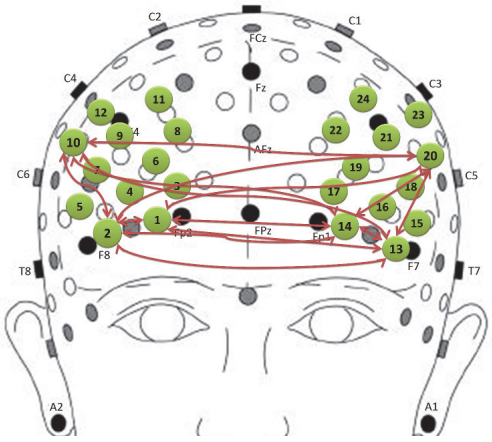

Figure 1: Channel distribution and structural information from the connectome used as prior knowledge (Hagmann et al. 2008; Joshi et al. 2010).

roimaging modalities (conferring different statistical properties to the observations) have to be considered for the modelling. For instance, a critical limitation in the fNIRS domain is that the set of observations only can incorporate a reduced number of samples due to the *habituation* effect in the brain, a response decrement due to stimulus repetition (Fischer et al. 2000).

In addition, the brain structural connectivity establishes a set of restrictions regarding to their anatomical paths, which is described by the *connectome*.

With this in mind, it is possible to state the problem of recovering the effective connectivity observed with fNIRS as the modelling of causal relations among a set of variables limited by a set of samples and integrating prior information from the connectome.

Experiments and results

The fNIRS neuroimaging dataset for this research was originally collected at Imperial College London back in 2007 to question about experience-dependent differences on prefrontal activity for a cohort of 62 surgeons (19 consultants, 21 trainees and 22 medical students) while repeating a knot-tying task 5 times at self-pace, allowing 30 seconds recovery between trials (Leff et al. 2007). All channels were set in the prefrontal cortex.

The connectome information was recovered from (Hagmann et al. 2008; Joshi et al. 2010) and adapted to the channel location described in (Leff et al. 2007), and it is shown in Fig 1.

A total of eight networks were built considering four groups (novices, trainees, experts and all subjects) and two variants, with and without connectome information using the sFCI in Algorithm 1 and the "classical" FCI respectively. Due to limitation of space Fig. 2 only presents novices, trainees and experts groups networks. Table 1 summarizes the number of undefined links and graph-theoretical measures (Rubinov and Sporns 2010) of functional integration using both FCI and our proposal sFCI. As expected, the number of undefined links decreases with the utilization of prior information.

	Novices		Trainees		Experts		All	
	FCI	*sFCI*	*FCI*	*sFCI*	*FCI*	*sFCI*	*FCI*	*sFCI*
UL	11	8	19	16	18	14	26	21
CPL	7.17	6.15	6.13	7.40	7.43	6.02	4.08	7.96
GE	0.20	0.21	0.25	0.20	0.20	0.22	0.35	0.20
ND	17	13	17	18	17	13	9	19

Table 1: Measures of brain connectivity (Rubinov and Sporns 2010). UL: undefined links in the networks, CPL: characteristic path length represents the average shortest path length between each pair of nodes, GE: global efficiency represents the brain capacity of parallel information flow and ND: network diameter is the largest number of nodes to travel between two nodes without loops.

In the original work of the dataset (Leff et al. 2007) the greater activity of the novices' prefrontal cortex leaded to a more lateralised response; an effect which can also be appreciated in the effective networks in Fig. 2. Interestingly, the network metrics in Tab. 1 tell a story very much alike that found in (Ohuchida et al. 2009) in which trainees evoked higher activity than novices or experts. These two observations provides strong nomological evidence about the working of the model. When considering the prior information, the network features tell a different story regarding the configuration of the network e.g. different pattern of path length and diameter, but shows a more intuitive higher efficiency of the network of experts, suggesting that the incorporation of the expected structural information can perhaps reveal more realistic effective information. However, further evidence is need before we make such claim.

Conclusions

Motivated by the neuroscientific demand of better modelling tools for the analysis of effective connectivity, we have presented a solution that is (i) innovative from the computational side with the proposal of the algorithm sFCI, and (ii) innovative from the fNIRS side with the exploitation of structural information to resolve effective links which would have otherwise been imposible to recover from data alone. Good nomological validity is achieved with previous findings. We plan to carry out additional experiments to afford other types of validity for the methodology proposed.

Acknowledgements. We thank Prof. Ara W. Darzi and Dr. Daniel Leff from Imperial College London for kindly permitting us to use their dataset in this research.

This work was supported by the CONACYT projects CB-2014-01-237251. Authors SAMH and JHV further received a PhD scholarship from the CONACYT.

References

Baccalá, L. A., and Sameshima, K. 2001. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics* 84:463–474.

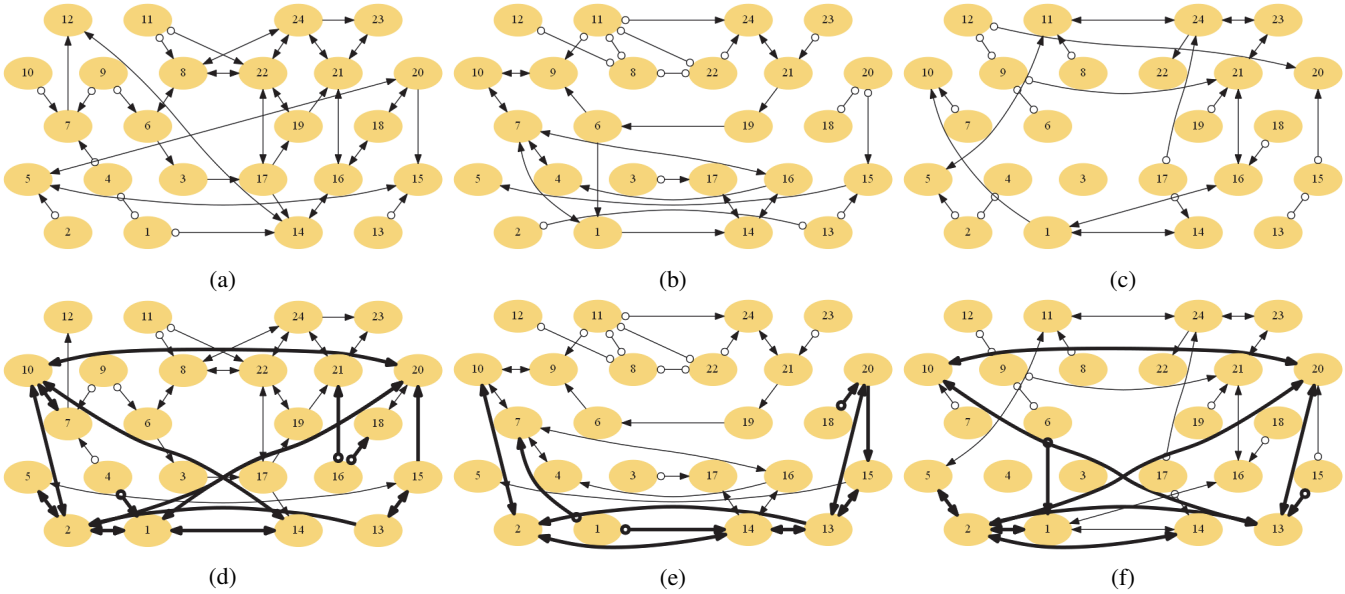

Figure 2: Effective connectivity networks. Top row show the connectivity revealed with plain FCI whereas bottom row show the connectivity networks exploiting prior information from the connectome with sFCI. Columns represent the different expertise groups; from top to bottom; novices (a and d), trainees (b and e), and consultants (c and f) networks. Resolved features (initial $-$ or final $>$) are represented by bold links.

Fischer, H.; Furmark, T.; Wik, G.; and Fredrikson, M. 2000. Brain representation of habituation to repeated complex visual stimulation studied with pet. *Neuroreport* 11(1):123–126.

Friston, K. J.; Harrison, L.; and Penny, W. 2003. Dynamic causal modelling. *NeuroImage* 19:1273–1302.

Hagmann, P.; Cammoun, L.; Gigandet, X.; Meuli, R.; Honey, C. J.; Wedeen, V. J.; and Sporns, O. 2008. Mapping the structural core of human cerebral cortex. *PLoS Biol* 6(7):e159.

Jöbsis, F. F. 1977. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science, New Series* 198(4323):1264–1267.

Joshi, A.; Joshi, S.; Dinov, I.; Shattuck, D.; Leahy, R.; and Toga, A. 2010. Anatomical structural network analysis of human brain using partial correlations of gray matter volumes. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, 844–847.

Kaminski, M., and Blinowska, K. J. 1991. A new method of the description of the information flow in the brain structures. *Biological Cybernetics* 65:203–210.

Leff, D. R.; Orihuela-Espina, F.; Atallah, L.; Darzi, A. W.; and Yang, G.-Z. 2007. Functional near infrared spectroscopy in novice and expert surgeons: a manifold embedding approach. In Ayache, N.; Ourselin, S.; and Maeder, A., eds., *Medical Image Computing and Computer-Assisted Intervention (MICCAI'07)*, volume 4792, 270–277. Australia: Lecture Notes in Computer Science.

McIntosh, A. R., and Gonzalez Lima, F. 1994. Structural equation modelling and its application to network analysis in functional brain imaging. *Human Brain Mapping* 2:2–22.

Ohuchida, K.; Kenmotsu, H.; Yamamoto, A.; Sawada, K.; Hayami, T.; Morooka, K.; Takasugi, S.; Konishi, K.; Ieri, S.; Tanoue, K.; Iwamoto, Y.; Tanaka, M.; and Hashizume, M. 2009. The frontal cortex is activated during learning of endoscopic procedures. *Surgical Endoscopy* In press.

Pearl, J. 2009. *Causality: models, reasoning and inference*. Cambridge Univ Press, second edition.

Rajapakse, J. C., and Zhou, J. 2007. Learning effective brain connectivity with dynamic bayesian networks. *Neuroimage* 37:749–760.

Richardson, T., and Spirtes, P. 2002. Ancestral graph Markov models. *Annals of Statistics* 30(4):962–1030.

Rubinov, M., and Sporns, O. 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52(3):1059–1069.

Spirtes, P.; Meek, C.; and Richardson, T. 1995. Causal inference in the presence of latent variables and selection bias. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, 499–506. Morgan Kaufmann Publishers Inc.

Sucar, L. E. 2015. *Probabilistic Graphical Models: Principles and applications*. Advances in Computer Vision and Pattern Recognition. Springer.