A methodology for detection of causal relationships between discrete time series on systems

Rute Souza de Abreu

rute.s.abreu@gmail.com

Departamento de Computação e Automação

Universidade Federal do Rio Grande do Norte

January 25, 2019

Contents

Introduction

- Motivation of the study
  - Causal relationships importance
- Proposal
  - Define a methodology to identify causal relationships between discrete time series on systems.
    - Transfer Entropy
    - Structural Learning on Bayesian Networks
- Related Work
  - Implementation of transfer entropy for detecting causal relationships between industrial alarm series [1]
  - Use of transfer entropy to detect neuronal connections [2]
  - Use of transfer entropy, Granger causality and Dynamic Bayesian Networks for reconstruction of genetic networks [3]

[1] (SU et al., 2017) (HU et al., 2017), (YU; YANG, 2015)
[2] (VICENTE et al., 2011)
[3] (TUNG et al., 2007) (ZOU; FENG, 2009)

Theoretical Foundations

Information Theory

- Entropy
- Kullback-Leibler Divergence
- Mutual Information
- Transfer Entropy

Bayesian Networks

- Bayesian Networks - Concept
- Structural Learning in Bayesian Networks
- K2-Algorithm
- Medium Description Length (MDL)

Theoretical Foundations

Entropy

Definition: The amount of information produced by an information source.

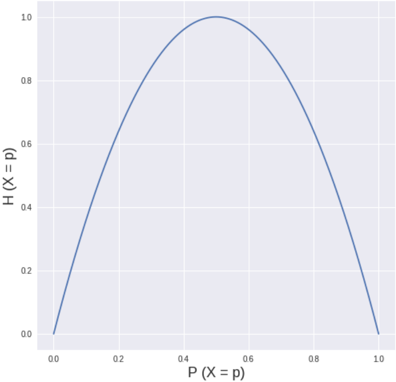
$$H = \sum_{i}^{n} p_i(i) log_2 \frac{1}{p(i)} \tag{1}$$

Properties:

- It is continuous in the domain of p_i, which is the probability mass function.
- It is monotonically increasing, in the n domain, when all events are likely equally.
- It is weighted additive when a choice is broken down into successive choices.

Theoretical Foundations

It is related to the frequency of the appearance of each value or the amount of surprise obtained given the appearance of a state or value.



Theoretical Foundations I

Kullback-Leibler Divergence

- Given a variable I, with probability distribution p. It measures the error or divergence when it is assumed that the probability distribution of I is q, instead of p.

Theoretical Foundations II

$$KL_I = \sum_{i}^{n} p(i) log \frac{p(i)}{q(i)} \tag{2}$$

$$K_{I,J} = \sum_{i,j}^{n} p(i,j) log \frac{p(i,j)}{q(i,j)} \tag{3}$$

$$K_{I|J} = \sum_{i,j}^{n} p(i,j) log \frac{p(i \mid j)}{q(i \mid j)} \tag{4}$$

Mutual Information

It is a measure derived from the Kullback-Leibler divergence. It is calculated between two processes I and J and quantifies the amount of information obtained about one process when observing the other.
It can be seen as the "information produced by erroneously assuming that the two processes are independent".[4]

$$MI_{I,J} = \sum_{i,j}^{n} p(i,j) log \frac{p(i,j)}{p(i) \cdot p(j)} \tag{5}$$

[4](SCHREIBER, 2000)

Transfer Entropy

It is a theoretical measure "that shares some of the desired properties of mutual information but takes the dynamics of information transport into account".[5]

$$TE_{J->I} = \sum_{i,i_{t+h},j} p(i_{t+h}, i_t^k, j_t^l) log \frac{p(i_{t+h} \mid i_t^k, j_t^l)}{p(i_{t+h} \mid i_t^k)} \tag{6}$$

[5](SCHREIBER, 2000)

$$TE_{J->I} = \sum_{i,i_{t+h},j} p(i_{t+h}, i_t^k, j_t^l) log \frac{p(i_{t+h} \mid i_t^k, j_t^l)}{p(i_{t+h} \mid i_t^k)} \tag{7}$$

$$i^k = [i_t, ..., i_{t-k+1}] \tag{8}$$

$$j^l = [j_t, ..., j_{t-l+1}] \tag{9}$$

- It infers the veracity of the equation:

$$p(i_{t+h} \mid i_t^k, j_t^l) = p(i_{t+h} \mid i_t^k) \tag{10}$$

Understanding of the parameters (Example)

- Series size = 6 samples
- k = 3 samples
- l = 2 samples
- h = 1 sample

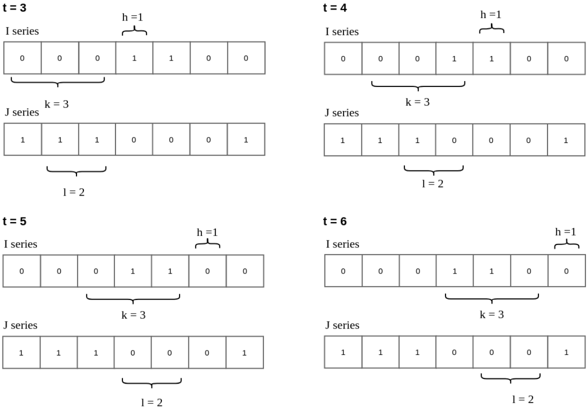


Figure 2: Choosing of the samples in TE.

Bayesian Networks

- Probabilistic models based on direct acyclic graphs.
- The nodes represent variables of interest, while the connections represent relationships of informational or causal dependencies.[6]
- Every dependence relationships is a conditional probability of a node, given its parents.

[6] (PEARL, 2011)

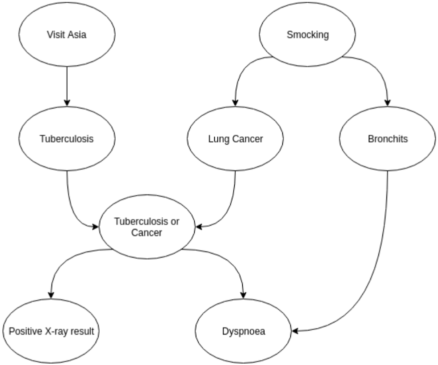


Figure 3: Bayesian Network Asia

Properties

- It factors the joint the distribution of the variable into local joint distributions.

$$P(A_n, \cap \ldots \cap A_1) = \prod_{i}^{n} P(A_i \mid \cap_{j=1}^{j-1} A_j) \tag{11}$$

$$P(A_n, \cap \ldots \cap A_1) = \prod_{i=1}^{n} P(A_i | pa_i) \tag{12}$$

$$P(Tuberc. \mid Cancer,\ Bronch.,\ Asia) = P(Tuberc. \mid Cancer) \tag{13}$$

Stuctural Learning in Bayesian Networks

K2 - Algorithm

- Bayesian method for estimating a probabilistic network from data.
- The algorithm searches for the network that has the highest posterior probability given a database of records
- Necessity of a prior ordering on nodes.
- Scores each local network with the following metric:

$$P(B_s, D) = c \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \tag{14}$$

Case	Variable values for each case		
	x_1	x_2	x_3
1	*present*	*absent*	*absent*
2	*present*	*present*	*present*
3	*absent*	*absent*	*present*
4	*present*	*present*	*present*
5	*absent*	*absent*	*absent*
6	*absent*	*present*	*present*
7	*present*	*present*	*present*
8	*absent*	*absent*	*absent*
9	*present*	*present*	*present*
10	*absent*	*absent*	*absent*

Figure 4: Database of cases

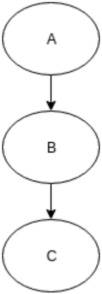


Figure 5: Possible structure.

Proposed Methodology

Objetivos

Objetivos

Objetivos

Aplicações

Resumo

References I

- HU, Wenkai et al. Cause-effect analysis of industrial alarm variables using transfer entropies. **Control Engineering Practice**, Elsevier, v. 64, p. 205–214, 2017.
- PEARL, Judea. Bayesian networks. 2011.
- SCHREIBER, Thomas. Measuring information transfer. **Physical review letters**, APS, v. 85, n. 2, p. 461, 2000.
- SU, Jianjun et al. Capturing Causality for Fault Diagnosis Based on Multi-Valued Alarm Series Using Transfer Entropy. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 19, n. 12, p. 663, 2017.

References II

- TUNG, Thai Quang et al. Inferring gene regulatory networks from microarray time series data using transfer entropy. In: IEEE. COMPUTER-BASED Medical Systems, 2007. CBMS'07. Twentieth IEEE International Symposium on Computer-Based Medical Systems. Maribor, Slovenia: IEEE, 2007. p. 383–388.
- VICENTE, Raul et al. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. **Journal of computational neuroscience**, Springer, v. 30, n. 1, p. 45–67, 2011.
- YU, Weijun; YANG, Fan. Detection of causality between process variables based on industrial alarm data using transfer entropy. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 17, n. 8, p. 5868–5887, 2015.

References III

ZOU, Cunlu; FENG, Jianfeng. Granger causality vs. dynamic Bayesian network inference: a comparative study. **BMC bioinformatics**, BioMed Central, v. 10, n. 1, p. 122, 2009.