LIVRO DE RESUMOS

SIFSC 11

DÉCIMA PRIMEIRA SEMANA DA
GRADUAÇÃO E PÓS-GRADUAÇÃO DO
INSTITUTO DE FÍSICA DE SÃO CARLOS - USP

2021

Universidade de São Paulo
Instituto de Física de São Carlos

XI Semana Integrada do Instituto de Física de São Carlos

Livro de Resumos

São Carlos
2021

Semana Integrada do Instituto de Física de São Carlos

SIFSC 11

Coordenadores

Prof. Dr. Vanderlei Salvador Bagnato
Diretor do Instituto de Física de São Carlos – Universidade de São Paulo

Prof. Dr. Luiz Vitor de Souza Filho
Presidente da Comissão de Pós Graduação do Instituto de Física de São Carlos – Universidade de São Paulo

Prof. Dr. Luís Gustavo Marcassa
Presidente da Comissão de Graduação do Instituto de Física de São Carlos – Universidade de São Paulo

Comissão Organizadora

Arthur Deponte Zutião
Artur Barbedo
Beatriz Kimie de Souza Ito
Beatriz Souza Castro
Carolina Salgado do Nascimento
Edgard Macena Cabral
Fernando Camargo Soares
Gabriel dos Reis Trindade
Gabriel dos Santos Araujo Pinto
Gabriel Henrique Armando Jorge
Giovanna Costa Villefort
Inara Yasmin Donda Acosta
Humberto Ribeiro de Souza
João Hiroyuki de Melo Inagaki
Kelly Naomi Matsui
Leonardo da Cruz Rea
Letícia Cerqueira Vasconcelos
Natália Carvalho Santos
Nickolas Pietro Donato Cerioni
Vinícius Pereira Pinto

Normalização e revisão – SBI/IFSC

Ana Mara Marques da Cunha Prado
Maria Cristina Cavarette Dziabas
Maria Neusa de Aguiar Azevedo
Sabrina di Salvo Mastrantonio

Ficha catalográfica elaborada pelo Serviço de Informação do IFSC

Semana Integrada do Instituto de Física de São Carlos
(11: 06 set. - 10 set. : 2021: São Carlos, SP.)
Livro de resumos da XI Semana Integrada do Instituto de Física de São Carlos/ Organizado por João H. Melo Inagaki [et al.]. São Carlos: IFSC, 2021.
412 p.
Texto em português.
1. Física. I. Inagaki, João H. de Melo, org. II. Titulo

ISBN 978-65-993449-3-0 CDD 530

PG3

The deluge of spurious information in causal network inference

MARTINELLI, T.[1]; SOARES PINTO, D. O.[1]; RODRIGUES, F. A.[2]

tiago.martinelli@usp.br

[1]Instituto de Física de São Carlos - USP

[2]Instituto de Ciências Matemáticas e Computação - USP

Large databases are a prime opportunity for science, and data analytics is a remarkable new field of investigation in computer science and physics. Within this paradigm, causal inference (1) is an effective tool to reveal the mechanisms producing emergent phenomena in a multivariate data-generated complex system. One of the main questions regarding these mechanisms is how their organization, represented by large networks, influences the dynamical processes. Such a study is fundamental to control processes from changes in the network structure avoiding, for example, cascading failures. In practice, however, the only information available is data recorded from variables with unknown topology. Using causal algorithms and information theory, we can explore this scenario by quantifying individuals' influence on the system, inferring links among them, thus reconstructing the underlying network. In other words, we can make reverse engineering to obtain a causal graphical model via data. Built upon the concept of causal faithfulness, the so-called *causal discovery algorithms* have been proposing the breakdown of the contributions of the mutual information (MI) and conditional mutual information (CMI) among sets of variables to reveal influences. (2) These algorithms suffer from the lack of accounting *emergent causes* when connecting links resulting in a spuriously embellished view of the organization of complex systems. Using classical results from Ramsey's theory, we proved that CMI's with large conditioned datasets contain, necessarily, arbitrary correlations. These correlations appear only because of the size, not nature, of data. They can be found in "randomly"generated, large enough databases. (3) We also connected this result with the intrinsic violation of faithfulness to reveal causal emergency. The net result is an update of the well-known causa discovery algorithms, which can, in principle, detect and isolate emergent causal influences in the reconstructed network undetected so far.

Palavras-chave: Causal inference. Information theory. Data analysis.

Referências:

1 PEARL, J. **Causality**: models, reasoning and inference. Cambridge: Cambridge University Press, 2009.
2 RUNGE, J. Causal network reconstruction from time series: from theoretical assumptions to practical estimation. **Chaos**, v. 28, n. 7, p. 075310-1-075310-20, 2018. DOI 10.1063/1.5025050.
3 CALUDE, C. S.; LONGO, G. The deluge of spurious correlations in big data. **Foundations of Science**, v. 22, n. 3, p. 595-612, 2017. DOI 10.1007/s10699-016-9489-4.