Other titles in Foundations and Trends® in Signal Processing

*Energy-Based Models with Applications to Speech and Language Processing*
Zhijian Ou

*Model-Based Deep Learning*
Nir Shlezinger and Yonina C. Eldar
ISBN: 978-1-63828-264-8

*Generalizing Graph Signal Processing: High Dimensional Spaces, Models and Structures*
Xingchao Jian, Feng Ji and Wee Peng Tay
ISBN: 978-1-63828-150-4

*Learning with Limited Samples: Meta-Learning and Applications to Communication Systems*
Lisha Chen, Sharu Theresa Jose, Ivana Nikoloska, Sangwoo Park, Tianyi Chen and Osvaldo Simeone

Jeroen Berrevoets
University of Cambridge
jeroen.berrevoets@maths.cam.ac.uk

Krzysztof Kacprzyk
University of Cambridge
kk751@cam.ac.uk

Zhaozhi Qian
University of Cambridge
zhaozhi.qian@maths.cam.ac.uk

Mihaela van der Schaar
University of Cambridge, and
The Alan Turing Institute
mv472@cam.ac.uk
Editorial Scope

Foundations and Trends® in Signal Processing publishes survey and tutorial articles in the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multirate signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - Classification and detection
  - Estimation and regression
  - Tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2024, Volume 18, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.
# Contents

1 **Introduction** .......................................................... 3

2 **Background** .......................................................... 8
   2.1 Graphical Modelling .............................................. 8
   2.2 Causal Learning ................................................... 12

3 **CDL in 3 Dimensions: Structural, Parametric, and Temporal** .......... 17
   3.1 The Structural Scale and Rung 1.5 ............................ 17
   3.2 The Parametric Scale ............................................ 22
   3.3 The Temporal Scale .............................................. 25
   3.4 Characterizing CDL Models Using the Scale .................. 26

4 **Constructing and Testing Model Pipelines** ................................ 28
   4.1 Transition in the Knowledge .................................... 28
   4.2 Building CDL Model Pipelines .................................. 29
   4.3 Non-matching Assumptions ..................................... 31

5 **Treatment Effects: An Illustrative Example Using CDL** .................. 32
   5.1 Treatment Effects: The CDL Description ...................... 33
   5.2 Treatment Effects Over Time .................................... 36

Jeroen Berrevoets1, Krzysztof Kacprzyk1, Zhaozhi Qian1 and Mihaela van der Schaar1,2

1 University of Cambridge, UK; jeroen.berrevoets@maths.cam.ac.uk, kk751@cam.ac.uk, zhaozhi.qian@maths.cam.ac.uk, mv472@cam.ac.uk
2 The Alan Turing Institute, UK

ABSTRACT
Causality has the potential to truly transform the way we solve a large number of real-world problems. Yet, so far, its potential largely remains to be unlocked as causality often requires crucial assumptions which cannot be tested in practice. To address this challenge, we propose a new way of thinking about causality— we call this causal deep learning. Our causal deep learning framework spans three dimensions: (1) a structural dimension, which incorporates partial yet testable causal knowledge rather than assuming either complete or no causal knowledge among the variables of interest; (2) a parametric dimension, which encompasses parametric forms that capture the type of relationships among the variables of interest; and (3) a temporal dimension, which captures exposure times or how the variables of interest interact (possibly causally) over time. Our CDL framework enables us to precisely categorise and compare...
causal statistical learning methods. We use this categorisation to provide a comprehensive review of the CDL field. More importantly, CDL enables us to make progress on a variety of real-world problems by aiding us to leverage partial causal knowledge (including independencies among variables) and quantitatively characterising causal relationships among variables of interest (possibly over time). Our framework clearly identifies which assumptions are testable and which are not, so the resulting solutions can be judiciously adopted in practice. Our formulation helps us to combine or chain causal representations to solve specific problems without losing track of which assumptions are required to build these solutions, pushing real-world impact in healthcare, economics and business, environmental sciences and education, through causal deep learning.
Causality holds the promise to transform the way we solve a large number of real-world problems [158], [162]. Unlocking this promise amounts to adapting causality to the real world with its highly complex, unstructured, and abundant data. Beyond rich datasets, causal inference methods typically rely on diverse and extensive prior knowledge about the causal nature of the underlying system, many of which are not available or even testable in the application domain [83]. This has led to two consequences. Firstly, practitioners may shy away from adopting causal inference in general because the available causal knowledge is not sufficient, let alone complete. Secondly, the emphasis on causal knowledge may overshadow other important considerations such as the statistical, functional, or temporal properties. These consequences may lead the practitioner to under-utilize these methods which ultimately results in sub-optimal solutions.

We want to encourage researchers and practitioners to take causality to its next step: real-world impact. For this, we recognise the need to integrate causal inference with the sophisticated modelling capabilities of deep learning. This integration is particularly vital in real-world domains where understanding causal relationships can lead to better decision-making and predictions.
In this monograph, we introduce a new framework which enables pragmatically adopting causality ideas to solve real-world problems: Causal Deep Learning (CDL). Informally, CDL methods can leverage partial causal knowledge (among some and not necessarily all variables of interest), and quantitatively characterises the functional form (among variables of interest) and the time-evolution of the variables to provide significant insights to researchers and decision-makers. The reason why the above properties are important is two-fold: (i) We need a good way to match a model with prior knowledge in a complex system, this should allow information of any type (be it no information, partial information, or full information); and (ii) We need a good way to evaluate a solution. This latter point is important in any real-world setting, especially where we rely on modelling solutions to support impactful decisions. Inherently being built on a strong set of assumptions, model validation is a particularly tricky endeavour in causality and one which CDL can help us solve [160].

With new methods being proposed at various machine learning and statistics venues, tracking solved problems across research efforts is hard. Our CDL framework allows this necessary comparison across these fields. Allowing: (i) accurate comparison of existing methods, (ii) identifying gaps in contemporary research, ultimately driving research forward, (iii) communicate methods to a practical audience, encouraging the adoption of causal deep learning in practice. Throughout this text, we will use “Causal deep learning” and its acronym “CDL” to refer to methods that leverage causal ideas in their models as presented in this work.

**Illustrative examples.** Let us explore four diverse domains where causal deep learning, characterised by its focus on causal structure, functional relationships among variables of interest, and time, to understand the need for this new way of thinking.

First, consider the medical and healthcare domain, where determining the effectiveness of medications is paramount. Traditional models from pharmacology or physiology often struggle to capture the complexity of drug interactions and patient responses. To make progress in studying the effects of medications on health outcomes, CDL might focus on the causal relationship between the drug dosage and patient...
recovery rate while controlling for other variables like age, gender, and pre-existing conditions. In such a case, a CDL method would not assert complete causality of overall health factors but rather isolate the drug’s impact. By employing only a partial causal structure, CDL can model the nuanced relationship between drug dosage and patient recovery, considering both linear and non-linear effects. Moreover, by incorporating temporal dynamics, CDL may unravel how medication effectiveness evolves in time, providing critical insights for personalized medicine. Naturally, without first identifying methods that allow partial causal knowledge as input, we may make mistakes and employ a method requiring full (and correct) causal knowledge. Using our framework may avoid such critical mistakes.

In economics and business, the relationship between interest rate changes and consumer spending is a classic example of a complex causal interaction. Traditional econometric models might fail to capture the intricacy of this relationship. CDL, with its ability to handle non-linear and high-dimensional data, can provide a more robust understanding. By adopting a pragmatic approach, CDL could enable the investigation of partial causality between interest rate changes and consumer spending. While acknowledging that other factors like education, employment rates, inflation, and economic policies also impact spending, the focus here is on understanding how variations in interest rates specifically influence consumer behaviour. By considering the temporal dimension, CDL may also uncover lag effects, where changes in interest rates take time to manifest in consumer behaviour, a crucial insight for policymakers.

Environmental science, especially the study of climate change, presents another compelling case for causal deep learning. An interesting research aspect might be to explore the causal relationship between carbon dioxide emissions and global temperature increase, acknowledging that other factors like deforestation and solar radiation also play roles in climate change. Using such a pragmatic CDL approach which requires only partial causal structure would allow researchers to isolate and understand the specific impact of CO2 emissions. In addition, the relationship between CO2 emissions and global temperature increase is not straightforward and likely non-linear. Here, causal deep learning framework can capture the complexity of this relationship beyond tradi-
tional linear models. Additionally, understanding how this relationship evolves over time is critical for predicting future climate patterns and for formulating effective environmental policies.

Lastly, in the education sector, the impact of classroom size on student performance is a topic of ongoing debate. While simple linear models might suggest a straightforward inverse relationship, the reality is likely more nuanced. Causal deep learning can help in identifying not just whether, but how class size impacts student performance, including potential threshold effects or non-linear dynamics. Furthermore, by examining how this relationship changes over an academic year or across different educational stages, deeper insights can be gained into effective educational planning and resource allocation. Solving this problem requires detailed descriptions of these dynamics are typically only possible with a set of fully known equations describing the data-generating process. These methods exist but often require a large set of modelling assumptions which are best charted with our framework.

In these and other real-world domains, causal deep learning can be used as a powerful framework to make progress. Unlike current causal models in machine learning, CDL allows for a nuanced understanding of causal relationships without requiring complete causal knowledge, learns in a data-driven manner the parametric form of the relationships among the variables of interest using powerful deep-learning models, and importantly, considers the dimension of time. This approach can uncover insights that traditional models might miss, leading to more effective interventions and policies across various sectors. The exploration of causal deep learning in these diverse real-world domains not only underscores its versatility but also highlights its potential to revolutionize how we understand and interact with the world around us.

With our detailed definitions, we can better align a problem with the (causal) solution. While research in causality has a long history of listing out the various assumptions necessary to identify causal effects, these assumptions often do not map easily into practice. Furthermore, deep learning has no such history, typically because most deep learning architectures are assumed to be non-parametric. However, not all problems require the enormous flexibility of deep neural networks, but may
still be non-linear. Our proposal encompasses all the above in hopes of aiding the correct and useful adoption of causal deep learning methods.

Please refer to https://nowpublishers.com/article/Details/SIG-123 for the online version of this monograph.
References


References


References


References


References


References


<table>
<thead>
<tr>
<th>Reference</th>
<th>Author(s)</th>
<th>Title/Description</th>
</tr>
</thead>
</table>
References


