Causal Inference

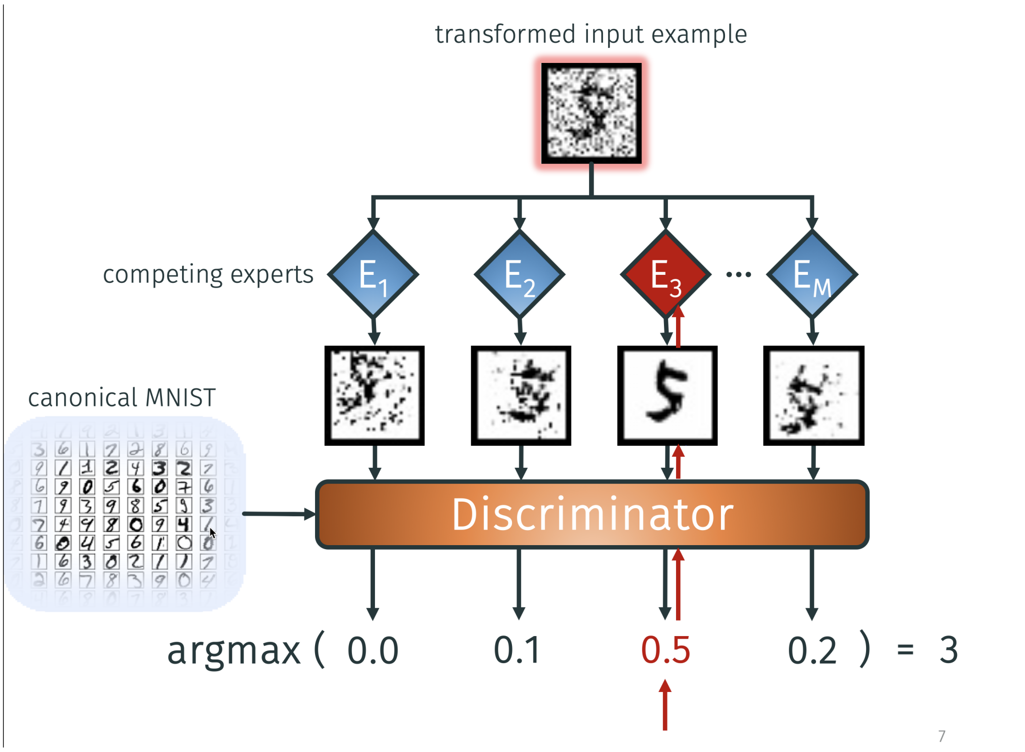

Figure 1.2: Through a competition of experts, inverse causal mechanisms are learned as independent modules in an image task. This results in generalisable and broadly applicable network modules, which can subsequently be combined and applied to new domains.

In the terminology of a book we recently published [2], the term causal inference comprises both *causal reasoning* and *causal discovery,* two somewhat inverse scenarios: While the former employs causal modelds for inferring about the expected observations (often, about their statistical properties), the latter is concerned with inferring causal models from empirical data. Both parts crucially depend on assumptions on the statistical properties entailed by hypothetical causal structures. During the past decade various assumptions have been proposed and assayed that go beyond traditional assumptions like the causal Markov condition and causal faithfulness. This has implications for both scenarios: it improves identifiability of causal structure, and it also entails additional statistical predictions if the causal structure is known.

Traditional causal discovery assumes that the units are connected by a causal directed acyclic graph a priori (mostly as random variables). In contrast, real-world observations are not necessarily a priori structured into those units (e.g. objects in images). The task of identifying reasonable units that admit *causal* models is challenging for both human and machine intelligence, but it aligns with the general goal of modern machine learning to learn *meaningful* representations for data, where 'meaningful' can for instance mean *interpretable*, *robust,* or *transferable.* The general idea that causal structure is a concept that often remains invariant across changing background conditions (first discussed by Herb Simon, and discussed in detail in our book) can be utilized for both causal reasoning and causal discovery. *Identifying causal units* will therefore be a third subsection below.

Causal reasoning Subject to sufficiently specific model assumptions (here: additive independent noise), causal knowledge admits novel techniques of noise removal such as our method of half-sibling regression published in PNAS [102], applied to astronomic data from NASA's Kepler space telescope [89].

Apart from entailing statistical properties for a *fixed* distribution, causal models also suggest how distributions *change* across data sets. To this end, one may assume, for instance, that structural equations remain constant across data sets and only the noise distributions change [2, 93], that some of the causal conditionals in a causal Bayesian network change, while others remain constant [30], or that they change *independently* [181], which results in new approaches to domain adaptation [242].

Causal discovery The toy problem of telling cause from effect in bivariate distributions, which we have earlier shown to be insightful also for more general causal inference problems, has been further explored [153, 251]. The performance of a broad variety of new approaches has been extensively studied in a long JMLR paper [101], suggesting that classification of cause and effect is indeed possible above chance level. New results for the multi-variable setting deal with, for instance, the problem of learning structural equation models in the presence of selection bias [239] and the idea of employing generalized score functions [145]. In [57], we introduce a kernel-based statistical test for joint independence of random variables which is a key component of multi-variate additive noise based causal inference.

Apart from progress on those 'classical' causal inference problems the domain of causal inference has been extended in several directions. Causal discovery for rare events has been further developed in terms of interacting Hawkes processes [240]. To study causal signals in images, the CVPR paper [197] infers whether the presence of an object on an image is the cause or the effect of the presence of another one, using additive noise based cause-effect inference.

In a study connecting principles of causal inference and foundations of physics [86], we relate asymmetries between cause and effect to asymmetries between past and future, deriving the thermodynamic arrow of time from the basic assumption of (algorithmic) independence of causal mechanisms. Within machine learning and time series modeling, new causal inference methods have revealed previously unknown aspects of the arrow of time [234].

Identifying causal units and causal learning Defining objects that are related by *causal models* typically amounts to appropriate coarse-graining of more detailed models of the world (e.g., physical models). Subject to appropriate conditions, causal models like structural equation models can arise from coarse-graining of 'microscopic' models including microscopic structural equation models [185], ordinary differential equations [156], temporally aggregated time series [186], or temporal abstractions of recurrent dynamical models [127]. Although every causal models in economics, medicine, or psychology uses variables that are abstractions of more elementary concepts, it is challenging to state general conditions under which coarse-grained variables admit causal models with well-defined interventions. Our work [185] provides some sufficient conditions.

Theoretical work in [152] shows that the independence of causal mechanisms can be formalized via group symmetry. The plausibility of a scene inferred from an image, for instance, can be formally assessed by testing whether some 'contrast function' attains values that are *typical* among those reached by symmetry transformations. This way, statistical independence (with permutations as corresponding symmetry group) is just a special case of a more general notion of independence.

In the context of a classical image recognition task, our recent ICML paper [139] shows that learning causal models that contain *invariant* mechanisms helps in transferring information across substantially different data sets (see figure). We plan to further pursue this direction, with the goal of moving from statistical representation learning towards causal world representations that should be more robust and support notions of intervention and planning.

More information: https://ei.is.mpg.de/project/causal-inference