Extended version of paper in Proceedings of the 26th AAAI Conference, Toronto, Ontario, Canada, pp. 698-704, 2012.

TECHNICAL REPORT
R-390-L
First version: April 2012
Revised: April 2013

Transportability of Causal Effects: Completeness Results

Elias Bareinboim and Judea Pearl
Cognitive Systems Laboratory
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA. 90095
{eb,judea} at cs.ucla.edu

Abstract

The study of transportability aims to identify conditions under which causal information learned from experiments can be reused in a different environment where only passive observations can be collected. The theory introduced in [Pearl and Bareinboim, 2011] (henceforth [PB, 2011]) defines formal conditions for such transfer but falls short of providing an effective procedure for deciding whether transportability is feasible for a given set of assumptions about differences between the source and target domains. This paper provides such procedure. It establishes a necessary and sufficient condition for deciding when causal effects in the target domain are estimable from both the statistical information available and the causal information transferred from the experiments. The paper further provides a complete algorithm for computing the transport formula, that is, a way of fusing experimental and observational information to synthesize an estimate of the desired causal relation.

Motivation

The problem of transporting knowledge from one environment to another has been pervasive in many data-driven sciences. Invariably, when experiments are performed on a group of subjects, the issue arises whether the conclusions are applicable to a different but somehow related group. When a robot is trained in a simulated environment, the question arises whether it could put the acquired knowledge into use in a new environment where relationships among agents, objects and features are different.

Surprisingly, the conditions under which this extrapolation can be legitimized were not formally articulated. Although the problem has been discussed in many areas of statistics, economics, and the health sciences, under rubrics such as "external validity" [Campbell and Stanley, 1963; Manski, 2007], "meta-analysis" [Glass, 1976; Hedges and Olkin, 1985; Owen, 2009], "heterogeneity" [Hofler, Gloster, and Hoyer, 2010], "quasi-experiments" [Shadish, Cook, and Campbell, 2002, Ch. 3; Adelman, 1991], these discussions are limited to verbal narratives in the form of heuristic guidelines for experimental researchers – no formal treatment of the problem has been attempted.

This research was supported in parts by grants from NIH #1R01 LM009961-01, NSF #IIS-0914211 and #IIS-1018922, and ONR #N000-14-09-1-0665 and #N00014-10-1-0933.

AI is in a unique position to tackle this problem formally. First, the distinction between statistical and causal knowledge has received syntactic representation through causal diagrams [Pearl, 1995; Spirtes, Glymour, and Scheines, 2001; Pearl, 2009; Koller and Friedman, 2009]. Second, graphical models provide a language for representing differences and commonalities among domains, environments, and populations [PB, 2011]. Finally, the inferential machinery provided by the do-calculus [Pearl, 1995; 2009; Koller and Friedman, 2009] is particularly suitable for combining these two features into a coherent framework and developing effective algorithms for knowledge transfer.

Following [PB, 2011], we consider transferring causal knowledge between two environments Π and Π^*. In environment Π, experiments can be performed and causal knowledge gathered. In Π^*, potentially different from Π, only passive observations can be collected but no experiments conducted. The problem is to infer a causal relationship R in Π^* using knowledge obtained in Π. Clearly, if nothing is known about the relationship between Π and Π^*, the problem is unsolvable. Yet the fact that all experiments are conducted with the intent of being used elsewhere (e.g., outside the laboratory) implies that scientific progress relies on the assumption that certain environments share common characteristics and that, owed to these commonalities, causal claims would be valid even where experiments were never performed.

To formally articulate commonalities and differences between environments, a graphical representation named *selection diagrams* was devised in [PB, 2011], which represent differences in the form of unobserved factors capable of causing such differences. Given an arbitrary selection diagram, our challenge is to algorithmically decide whether commonalities override differences to permit the transfer of information across the two environments.

Previous Work and Our Contributions

Consider Fig. 1(a) which concerns the transfer of experimental results between two locations. We first conduct a randomized trial in Los Angeles (LA) and estimate the causal effect of treatment X on outcome Y for every age group $Z = z$, denoted $P(y|do(x), z)$. We now wish to generalize the results to the population of New York City (NYC), but we find the distribution $P(x, y, z)$ in LA to be different from the one in NYC (call the latter $P^*(x, y, z)$). In particular, the average age in NYC is significantly higher than that in LA. How are we to estimate the causal effect of X on Y in NYC, denoted $R = P^*(y|do(x))$? [1,2]

The selection diagram for this example (Fig. 1(a)) conveys the assumption that the *only* difference between the two population are factors determining age distributions, shown as $S \to Z$, while age-specific effects $P(y|do(x), Z = z)$ are invariant across cities. Difference-generating factors are represented by a special set of variables called *selection variables* S (or simply S-variables), which are graphically depicted as square nodes (■). From this assumption, the overall causal effect in NYC can be derived as follows [3]

$$
\begin{aligned}
R &= \sum_z P^*(y|do(x), z) P^*(z) \\
&= \sum_z P(y|do(x), z) P^*(z)
\end{aligned}
$$

The last line is the *transport formula* for R. It combines experimental results obtained in LA, $P(y|do(x), z)$, with observational aspects of NYC population, $P^*(z)$, to obtain an experimental claim $P^*(y|do(x))$ about NYC.

In this trivial example the transport formula amounts to a simple re-calibration of the age-specific effects to account for the new age distribution. In more elaborate examples, however, the full power of formal analysis would be required. For instance, [PB, 2011] showed that, in the problem depicted in Fig. 1(b), where both the Z-determining mechanism and the U-determining mechanism are suspect of being different, the transport formula for the relation $R = P^*(y|do(x))$ is given by

$$
\sum_z P(y|do(x), z) \sum_w P^*(z|w) \sum_t P(w|do(x), t) P^*(t)
$$

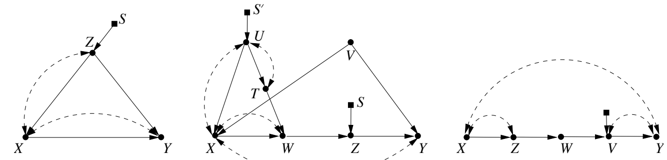
This formula instructs us to estimate $P(y|do(x), z)$ and $P(w|do(x), t)$ in the experimental domain, then combine them with the estimates of $P^*(z|w)$ and $P^*(t)$ in the target domain.

[PB, 2011] derived this formula using the following theorem, which translates the property of transportability to the existence of a syntactic reduction using a sequence of *do-calculus* operations.

[1]We will use $P_x(y)$ interchangeably with $P(y \mid do(x))$.

[2]We use the structural interpretation of causal diagrams. For example, Fig. 1(a) describes the following system of structural equations: $z \leftarrow f_1(s; u_z; u_{zx})$, $x \leftarrow f_2(z; u_x; u_{zx})$, $y \leftarrow f_3(y; z; u_y; u_{xy})$; each variable in the l.h.s. is assigned a value given by the respective deterministic function on the r.h.s. The exogenous (hidden) variables U are assigned a probability function which induces in turn, a probability distribution on all variables in the model. See Appendix 1 for a gentle introduction to the *do-calculus* and more details on this representation.

[3]This result can be derived by purely graphical operations if we write $P^*(y|do(x), z)$ as $P(y|do(x), z, s)$, thus attributing the difference between Π and Π^* to a fictitious event $S = s$. The invariance of the age-specific effect then follows from the conditional independence $(S \perp\!\!\!\perp Y|Z, X)_{G_{\overline{X}}}$, which implies $P(y|do(x), z, s) = P(y|do(x), z)$, and licenses the derivation of the transport formula.



(a) (b) (c)

Figure 1: Selection diagrams illustrating different facets of the transportability problem. (a) A selection diagram in which transportability is trivial. (b) A selection diagram in which the causal relation R is more involved and shown to be transportable using Theorem 1. (c) A selection diagram in which the procedure given in [PB, 2011] is unable to recognize a transportable relation R.

Theorem 1 (Do-calculus characterization [PB, 2011]). *Let D be the selection diagram characterizing Π and Π^*, and $\mathbf{S}$ a set of selection variables in D. The relation $R = P^*(\mathbf{y}|do(\mathbf{x}), \mathbf{z})$ is transportable from Π to Π^* if and only if the expression $P(\mathbf{y}|do(\mathbf{x}), \mathbf{z}, \mathbf{s})$ is reducible, using the rules of do-calculus, to an expression in which $\mathbf{S}$ appears only as a conditioning variable in do-free terms.*[4]

Theorem 1 is declarative but not computationally effective, for it does not specify the sequence of rules leading to the needed reduction, nor does it tell us if such a sequence exists.

To overcome this deficiency, [PB, 2011] proposed a recursive procedure (their Theorem 3) which can handle many examples, among them Fig. 1(b), but is not complete. We will show[5], for example, that their procedure fails to recognize R as "transportable" in the diagram of Fig. 1(c) whereas the procedures developed in this paper will recognize it as such and will support it with the transport formula:

$$
R = \sum_z P(z|do(x)) \sum_w P(w|do(x, z)) \sum_v P^*(v|w) P^*(y|v, w)
$$

We summarize our contributions as follows:

- We derive a general graphical condition for deciding transportability of causal effects. We show that transportability is feasible if and only if a certain graph structure does not appear as an edge subgraph of the inputted selection diagram.
- We provide necessary or sufficient graphical conditions for special cases of transportability, for instance, controlled direct effects.
- Finally, we construct a complete algorithm for deciding transportability of joint causal effects and returning the correct transport formula whenever those effects are transportable.

Preliminary Results

The basic semantical framework in our analysis rests on *probabilistic causal models* as defined in [Pearl, 2000, pp. 205], also called structural causal models or data-generating

[4]The only if part is proven in Corollary 4.

[5]See Corollary 5 in Appendix 2 .

models. In the structural causal framework [Pearl, 2000, Ch. 7], actions are modifications of functional relationships, and each action $do(\mathbf{x})$ on a causal model M produces a new model $M_{\mathbf{x}} = \langle \mathbf{U}, \mathbf{V}, \mathbf{F}_{\mathbf{x}}, P(\mathbf{U})\rangle$, where $F_{\mathbf{x}}$ is obtained after replacing $f_X \in \mathbf{F}$ for every $X \in \mathbf{X}$ with a new function that outputs a constant value x given by $do(\mathbf{x})$.

Key to the analysis of transportability is the notion of "identifiability," defined below, which expresses the requirement that causal effects be computable from a combination of data P and assumptions embodied in a causal graph G.

Definition 1 (Causal Effects Identifiability [Pearl, 2000, pp. 77]**).** *The causal effect of an action $do(\mathbf{x})$ on a set of variables $\mathbf{Y}$ such that $\mathbf{Y} \cap \mathbf{X} = \emptyset$ is said to be identifiable from P in G if $P_{\mathbf{x}}(\mathbf{y})$ is uniquely computable from $P(V)$ in any model that induces G.*

Causal models and their induced graphs are normally associated with one particular domain (also called setting, study, population, environment). In the transportability case, we extend this representation to capture properties of several domains simultaneously. This is made possible if we assume that there are no structural changes between the domains, that is, all structural equations share the same set of arguments, though the functional forms of the equations may vary arbitrarily. [6, 7]

Definition 2 (Selection Diagram)**.** *Let $\langle M, M^* \rangle$ be a pair of structural causal models [Pearl, 2000, pp. 205] relative to domains $\langle \Pi, \Pi^* \rangle$, sharing a causal diagram G. $\langle M, M^* \rangle$ is said to induce a selection diagram D if D is constructed as follows:*

1. *Every edge in G is also an edge in D;*
2. *D contains an extra edge $S_i \rightarrow V_i$ whenever there might exist a discrepancy $f_i \neq f_i^*$ or $P(U_i) \neq P^*(U_i)$ between M and M^*.*

In words, the S-variables locate the *mechanisms* where structural discrepancies between the two domains are suspected to take place.[8] Alternatively, one can see a selection diagram as a carrier of invariance claims between the mechanisms of both domains – the absence of a selection node pointing to a variable represents the assumption that the mechanism responsible for assigning value to that variable is the same in the two domains.

Armed with a selection diagram and the concept of identifiability, transportability of causal effects (or transportability, for short) can be defined as follows:

Definition 3 (Causal Effects Transportability)**.** *Let D be a selection diagram relative to domains $\langle \Pi, \Pi^* \rangle$. Let $\langle P, I \rangle$ be the pair of observational and interventional distributions of Π, and P^* be the observational distribution of Π^*. The causal effect $R = P_{\mathbf{x}}(\mathbf{y})$ is said to be transportable from Π to Π^* in D if $P_{\mathbf{x}}(\mathbf{y})$ is uniquely computable from P, P^*, I in any model that induces D.*

The problem of transportability generalizes the problem of identifiability, to witness note that all identifiable causal relations in (G^*, P^*) are also transportable, because they can be computed directly from Π^* and require no experimental information from Π. This observation engender the following definition of *trivial transportability*.

Definition 4. *(Trivial Transportability)*
A causal relation R is said to be trivially transportable *from Π to Π^*, if $R(\Pi^*)$ is identifiable from (G^*, P^*).*

The following observation establishes another connection between identifiability and transportability. For a given causal diagram G, one can produce a selection diagram D such that identifiability in G is equivalent to transportability in D. First set $D = G$, and then add selection nodes pointing to all variables in D, which represents that the target domain does not share any commonality with its pair – this is equivalent to the problem of identifiability because the only way to achieve transportability is to identify R from scratch in the target domain.

Another special case of transportability occurs when a causal relation has identical form in both domains – no recalibration is needed. This is captured by the following definition.

Definition 5. *(Direct Transportability)*
A causal relation R is said to be directly transportable *from Π to Π^*, if $R(\Pi^*) = R(\Pi)$.*

A graphical test for direct transportability of $R = P(y|do(x), z)$ follows from do-calculus and reads: $(S \perp\!\!\!\perp Y|X, Z)_{G_{\overline{X}}}$; in words, X blocks all paths from S to Y once we remove all arrows pointing to X and condition on Z. As a concrete example, the z-specific effects in Fig. 1(a) is the same in both domains, hence, it is directly transportable.

These two cases will act as a basis to decompose the problem of transportability into smaller and more manageable subproblems (to be shown later on).

The following lemma provides an auxiliary tool to prove non-transportability and is based on refuting the uniqueness property required by Definition 3.

Lemma 1. *Let $\mathbf{X}, \mathbf{Y}$ be two sets of disjoint variables, in population Π and Π^*, and let D be the selection diagram. $P_{\mathbf{x}}(\mathbf{y})$ is not transportable from Π to Π^* if there exist two causal models M^1 and M^2 compatible with D such that $P_1(\mathbf{V}) = P_2(\mathbf{V})$, $P_1^*(\mathbf{V}) = P_2^*(\mathbf{V})$, $P_1(\mathbf{V} \setminus \mathbf{W}|do(\mathbf{W})) = P_2(\mathbf{V} \setminus \mathbf{W}|do(\mathbf{W}))$, for any set $\mathbf{W}$, all families have positive distribution, and $P_1^*(\mathbf{y}|do(\mathbf{x})) \neq P_2^*(\mathbf{y}|do(\mathbf{x}))$.*

Proof. Let I be the set of interventional distributions $P(\mathbf{V} \setminus \mathbf{W}|do(\mathbf{W}))$, for any set $\mathbf{W}$. The latter inequality rules out the existence of a function from P, P^*, I to $P_{\mathbf{x}}(\mathbf{y})$. $\square$

While the problems of identifiability and transportability are related, Lemma 1 indicates that proofs of non-transportability are more involved than those of non-identifiability. Indeed, to prove non-transportability requires

[6] This definition was left implicit in [PB, 2011].

[7] The assumption that there are no structural changes between domains can be relaxed starting with $D = G^*$ and adding S-nodes following the same procedure as in Def. 2, while enforcing acyclicity.

[8] Transportability assumes that enough structural knowledge about both domains is known in order to substantiate the production of their respective causal diagrams. In the absence of such knowledge, *causal discovery* algorithms can be used to infer the diagrams from data [Pearl and Verma, 1991; Pearl, 2000; Spirtes, Glymour, and Scheines, 2001].

the construction of two models agreeing on $\langle P, I, P^* \rangle$, while non-identifiability requires the two models to agree solely on the observational distribution P.

The simplest non-transportable structure is an extension of the famous 'bow arc' graph named here 's-bow arc', see Fig. 2(a). The s-bow arc has two endogenous nodes: X, and its child Y, sharing a hidden exogenous parent U, and a S-node pointing to Y. This and similar structures that prevent transportability will be useful in our proof of completeness, which requires a demonstration that whenever the algorithm fails to transport a causal relation, the relation is indeed non-transportable.

Theorem 2. *$P_x^*(y)$ is not transportable in the s-bow arc graph.*

Proof. The proof will show a counter-example to the transportability of $P_x^*(Y)$ through two models M_1 and M_2 that agree in $\langle P, P^*, I \rangle$ and disagree in $P_x^*(y)$.

Assume that all variables are binary. Let the model M_1 be defined by the following system of structural equations: $X_1 = U, Y_1 = ((X \otimes U) \otimes S), P_1(U) = 1/2$, and M_2 by the following one: $X_2 = U, Y_2 = S \vee (X \otimes U), P_2(U) = 1/2$, where $\otimes$ represents the *exclusive or* function.

Lemma 2. *The two models agree in the distributions $\langle P, P^*, I \rangle$.*

Proof. We show that the following equations must hold for M_1 and M_2:

$$\begin{cases} P_1(X|S) = P_2(X|S), \quad S = \{0,1\} \\ P_1(Y|X,S) = P_2(Y|X,S), \quad S = \{0,1\} \\ P_1(Y|do(X), S=0) = P_2(Y|do(X), S=0) \end{cases}$$

for all values of X, Y. The equality between $P_i(X|S)$ is obvious since $(S \perp\!\!\!\perp X)$ and X has the same structural form in both models. Second, let us construct the truth table for Y:

X	S	U	Y_1	Y_2
0	0	0	0	0
0	0	1	1	1
0	1	0	1	1
0	1	1	0	1
1	0	0	1	1
1	0	1	0	0
1	1	0	0	1
1	1	1	1	1

To show that the equality between $P_i(Y = 1|X, S = 0), X = \{0,1\}$ holds, we rewrite it as follows:

$$P_i(Y=1|X,S=0) = \frac{P_i(Y=1|X,S=0,U=1)P_i(X|U=1)P_i(U=1)}{P_i(X)} + \frac{P_i(Y=1|X,S=0,U=0)P_i(X|U=0)P_i(U=0)}{P_i(X)} \quad (1)$$

In eq. (1), the expressions for $X = \{0,1\}$ are functions of the tuples $\{(X=1,S=0,U=1),(X=0,S=0,U=0)\}$, which evaluate to the same value in both models. Similarly, the expressions $P_i(Y=1|X,S=1)$ for $X = \{0,1\}$ are functions of the tuples $\{(X=1,S=1,U=1),(X=0,S=1,U=0)\}$, which also evaluate to the same value in both models.

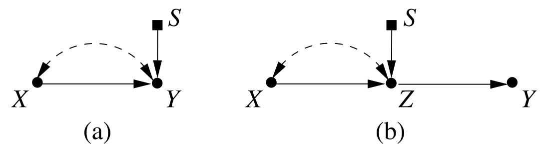


Figure 2: (a) Smallest selection diagram in which $P(y|do(x))$ is not transportable (s-bow graph). (b) A selection diagram in which even though there is no S-node pointing to Y, the effect of X on Y is still not-transportable due to the presence of a sC-tree (see Corollary 2).

We further assert the equality between the interventional distributions in Π, which can be written using the do-calculus as

$$\begin{aligned} &P_i(Y=1|do(X),S=0) = \\ &\sum_U P_i(Y|do(X),S=0,U)P_i(U|do(X),S=0) = \\ &P_i(Y=1|X,S=0,U=1)P_i(U=1) + \\ &P_i(Y=1|X,S=0,U=0)P_i(U=0), \quad X = \{0,1\} \end{aligned} \quad (2)$$

Evaluating this expression points to the tuples $\{(X=1,S=0,U=1),(X=1,S=0,U=0)$ and $(X=0,S=0,U=1),(X=0,S=0,U=0)\}$, which map to the same value in both models. □

Lemma 3. *There exist values of X, Y such that $P_1(Y|do(X), S=1) \neq P_2(Y|do(X), S=1)$.*

Proof. Fix $X = 1, Y = 1$, and let us rewrite the desired quantity $R_i = P_i(Y=1|do(X=1), S=1)$ in Π^* as

$$\begin{aligned} R_i = &P_i(Y=1|X=1,S=1,U=1)P_i(U=1) + \\ &P_i(Y=1|X=1,S=1,U=0)P_i(U=0) \end{aligned} \quad (3)$$

Since R_i is a function of the tuples $(X=1,S=1,U=1),(X=1,S=1,U=0)$, it evaluates in M_1 to $\{1,1\}$ and in M_2 to $\{1,0\}$.

Hence, together with the uniformity of $P(U)$, it follows that $R_1 = 1$ and $R_2 = 1/2$, which finishes the proof. □

Lemma 1, Lemmas 2 and 3 prove Theorem 2. □

Characterizing Transportable Relations

The concept of confounded components (or *C-components*) was introduced in [Tian and Pearl, 2002] to represent clusters of variables connected through bidirected edges, and was instrumental in establishing a number of conditions for ordinary identification (Def. 1). If G is not a C-component itself, it can be uniquely partitioned into a set $\mathcal{C}(G)$ of C-components. We now recast C-components in the context of transportability.[9]

Definition 6 (sC-component)**.** *Let G be a selection diagram such that a subset of its bidirected arcs forms a spanning tree over all vertices in G. Then G is a sC-component (selection confounded component).*

[9] C-components can itself be seen as an extension of the more elementary notion of *inducing path*, which was introduced much earlier in [Verma and Pearl, 1990].

A special subset of C-components that embraces the ancestral set of Y was noted by [Shpitser and Pearl, 2006b] to play an important role in deciding identifiability – this observation can also be applied to transportability, as formulated in the next definition.

Definition 7 (sC-tree)**.** *Let G be a selection diagram such that $\mathcal{C}(G) = \{G\}$, all observable nodes have at most one child, there is a node Y, which is a descendent of all nodes, and there is a selection node pointing to Y. Then G is called a Y-rooted sC-tree (selection confounded tree).*

The presence of this structure (and generalizations) will prove to be an obstacle to transportability of causal effects. For instance, the s-bow arc in Fig. 2(a) is a Y-rooted sC-tree where we know $P_x(y)$ is non-transportable.

In certain classes of problems, the absence of such structures will prove sufficient for transportability. One such class is explored below, and consists of models in which the set X coincides with the parents of Y.

Theorem 3. *Let G be a selection diagram. Then for any node Y, the causal effects $P^*_{Pa(Y)}(y)$ is transportable if there is no subgraph of G which forms a Y-rooted sC-tree.*

Proof. See Appendix 2. □

Theorem 3 provides a tractable transportability condition for the Controlled Direct Effect (CDE) – a key concept in modern mediation analysis, which permits the decomposition of effects into their direct and indirect components [Pearl, 2001; 2012]. CDE is defined as the effect of X on Y when all other parents of Y are held constant, and it is identifiable if and only if $P_{Pa(Y)}(y)$ is identifiable [Pearl, 2009, pp. 128].

The selection diagram in Fig. 1(a) does not contain any Y-rooted sC-trees as subgraphs, and therefore the direct effects (causal effects of Y's parents on Y) is indeed transportable. In fact, the transportability of CDE can be determined by a more visible criterion:

Corollary 1. *Let G be a selection diagram. Then for any node Y, the direct effect $P^*_{Pa(Y)}(y)$ is transportable if there is no S node pointing to Y.*

Proof. See Appendix 2. □

Generalizing to arbitrary effects, the following result provides a necessary condition for transportability whenever the whole graph is a sC-tree.

Theorem 4. *Let G be a Y-rooted sC-tree. Then the effects of any set of nodes in G on Y are not transportable.*

Proof. See Appendix 2. □

The next corollary demonstrates that sC-trees are obstacles to the transportability of $P_x(y)$ even when they do not involve Y, i.e., transportability is not a local problem – if there exists a node W that is an ancestor of Y but not necessarily "near" it, transportability is still prohibited (see Fig. 2(b)). This fact anticipates that transporting causal effects of singleton Y is not necessarily easier than the general problem of transportability.

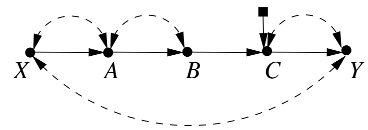


Figure 3: Example of a selection diagram in which $P(Y|do(X))$ is not transportable, there is no sC-tree but there is a s^*-tree.

Corollary 2. *Let G be a selection diagram, and $\mathbf{X}$ and $\mathbf{Y}$ a set of variables. If there exists a node W that is an ancestor of some node $Y \in \mathbf{Y}$ such that there exists a W-rooted sC-tree which contains any variables in $\mathbf{X}$, then $P_{\mathbf{x}}(\mathbf{y})$ is not transportable.*

Proof. See Appendix 2. □

We now generalize the definition of sC-trees (and Theorem 4) in two ways: first, Y is augmented and can be a set of variables; second, S-nodes can point to any variable within the sC-component, not necessarily to root nodes. For instance, consider the graph G in Fig. 3. Note that there is no Y-rooted sC-tree nor W-rooted sC-tree in G (where W is an ancestor of Y), and so the previous results cannot be applied even though the effect of X on Y is not transportable in G – still, there exists a Y-rooted s^*-tree in G, which will prevent the transportability of the causal effect.

Definition 8 (s^*-tree)**.** *Let G be a selection diagram, where $\mathbf{Y}$ is the maximal root set. Then G is a $\mathbf{Y}$-rooted s^*-tree if G is a sC-component, all observable nodes have at most one child, and there is a selection node pointing to some vertex of G (not necessarily in $\mathbf{Y}$).*

We next conveniently introduce a structure that witnesses non-transportability characterized by a pair of s^*-trees. Transportability will be shown impossible whenever such structure exists as an edge subgraph of the given selection diagram.

Definition 9 (s-hedge)**.** *Let $\mathbf{X}, \mathbf{Y}$ be set of variables in G. Let F, F' be $\mathbf{R}$-rooted s^*-trees such that $F \cap X \neq 0$, $F' \cap X = 0$, $F' \subseteq F$, $\mathbf{R} \subset An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$. Then F and F' form a s-hedge for $P_{\mathbf{x}}(\mathbf{Y})$ in G.*

For instance, in Fig. 3, the s^*-trees $F' = \{C, Y\}$, and $F = F' \cup \{X, A, B\}$ form a s-hedge to $P_x(y)$.

We state below the formal connection between s-edges and non-transportability.

Theorem 5. *Assume there exist F, F' that form a s-hedge for $P_x(y)$ in Π and Π^*. Then $P_x(y)$ is not transportable from Π to Π^*.*

Proof. See Appendix 2. □

To prove that the s-hedges characterize non-transportability in selection diagrams, we construct in the next section an algorithm which transport any causal effects that do not contain a s-hedge.

A Complete Algorithm For Transportability of Joint Effects

The algorithm proposed to solve transportability is called **sID** (see Fig. 4) and extends previous analysis and algorithms of identifiability given in [Pearl, 1995; Kuroki and Miyakawa, 1999; Tian and Pearl, 2002; Shpitser and Pearl, 2006b; Huang and Valtorta, 2006]. We build on two observations developed along the paper:

(i) Transportability: Causal relations can can be partitioned into trivially and directly transportable.

(ii) Non-transportability: The existence of a s-hedge as an edge subgraph of the inputted selection diagram can be used to prove non-transportability.

The algorithm **sID** first applies the typical c-component decomposition on top of the inputted selection diagram D, partitioning the original problem into smaller blocks (call these blocks sc-factors) until either the entire expression is transportable, or it runs into the problematic s-hedge structure.

More specifically, for each sc-factor Q, **sID** tries to directly transport Q. If it fails, **sID** tries to trivially transport Q, which is equivalent to solving an ordinary identification problem. **sID** alternates between these two types of transportability, and whenever it exhausts the possibility of applying these operations, it exits with failure with a counterexample for transportability – that is, the graph local to the faulty call witnesses the non-transportability of the causal query since it contains a s-hedge as edge subgraph.

Before showing the more formal properties of **sID**, we demonstrate how **sID** works through the transportability of $Q = P(y|do(x))$ in the graph in Fig. 1(c).

Since $D = An(Y)$ and $\mathcal{C}(D \setminus \{X\}) = (C_0, C_1, C_2)$, where $C_0 = D(\{Z\})$, $C_1 = D(\{W\})$, and $C_2 = D(\{V, Y\})$, we invoke line 4 and try to transport respectively $Q_0 = P^*_{x,w,v,y}(z)$, $Q_1 = P^*_{x,z,v,y}(w)$, and $Q_2 = P^*_{x,z,w}(v, y)$. Thus the original problem reduces to transporting $\sum_{z,w,v} P^*_{x,w,v,y}(z) P^*_{x,z,v,y}(w) P^*_{x,z,w}(v, y)$.

Evaluating the first expression, we trigger line 2, noting that nodes that are not ancestors of Z can be ignored. This implies that $P^*_{x,w,v,y}(z) = P^*_x(z)$ with induced subgraph $G_0 = \{X \rightarrow Z, X \leftarrow U_{xz} \rightarrow Z\}$, where U_{xz} stands for the hidden variable between X and Z. **sID** goes to line 5, in which in the local call $\mathcal{C}(D \setminus \{X\}) = \{G_0\}$. Note that in the ordinary identifiability problem the procedure would fail at this point, but **sID** proceeds to line 6 testing whether $(\mathbf{S} \perp\!\!\!\perp Z|X)_{D_{\overline{X}}}$. The test comes true, which makes **sID** directly transport Q_0 with data from the experimental domain Π, i.e., $P^*_x(z) = P_x(z)$.

Evaluating the second expression, we again trigger line 2, which implies that $P^*_{x,z,v,y}(w) = P^*_{x,z}(w)$ with induced subgraph $G_1 = \{X \rightarrow Z, Z \rightarrow W, X \leftarrow U_{xz} \rightarrow Z\}$. **sID** goes to line 5, in which in the local call $\mathcal{C}(D \setminus \{X\}) = \{G_1\}$. Thus it proceeds to line 6 testing whether $(\mathbf{S} \perp\!\!\!\perp W|X, Z)_{D_{\overline{X,Z}}}$. The test comes true again, which makes **sID** directly transport Q_1 with data from the experimental domain Π, i.e., $P^*_{x,z}(w) = P_{x,z}(w)$.

Evaluating the third expression, **sID** goes to line 5 in which $\mathcal{C}(D \setminus \{X, Z, W\}) = \{G_2\}$, where $G_2 = \{V \rightarrow Y, S \rightarrow V, V \leftarrow U_{vy} \rightarrow Y\}$. It proceeds to line 6 testing whether $(\mathbf{S} \perp\!\!\!\perp W|X, Z)_{D_{\overline{X,Z}}}$, which is false in this case. It tests the other conditions until it reaches line 9, in which $C' = G_0 \cup G_2 \cup \{X \leftarrow U_{xy} \rightarrow Y\}$. Thus it tries to transport $Q'_2 = P^*_{x,z}(v, y)$ over the induced graph C', which stands for ordinary identification, and trivially yields (after simplification) $\sum_v P^*(v|w) P^*(y|v, w)$. The return of these calls composed indeed coincide with the expression provided in the first section.

function $\mathbf{sID}(\mathbf{y}, \mathbf{x}, P^*, I, D)$

INPUT: $\mathbf{x}, \mathbf{y}$ value assignments, P^* observational distribution in Π^*, I set of interventional distributions in Π, D a selection diagram, $\mathbf{S}$ set of selection nodes.

OUTPUT: Expression for $P^*_{\mathbf{x}}(\mathbf{y})$ in terms of P^*, I or $FAIL(F, F')$.

1 **if** $\mathbf{x} = \emptyset$, **return** $\sum_{\mathbf{V} \setminus \mathbf{Y}} P^*(\mathbf{V})$

2 **if** $\mathbf{V} \setminus An(\mathbf{Y})_D \neq \emptyset$,
  return $\mathrm{sID}(\mathbf{y}, \mathbf{x} \cap An(\mathbf{Y})_D, \sum_{\mathbf{V} \setminus An(\mathbf{Y})_D} P^*, An(\mathbf{Y})_D)$

3 Set $\mathbf{W} = (\mathbf{V} \setminus \mathbf{X}) \setminus An(\mathbf{Y})_{D_{\overline{\mathbf{X}}}}$.
  if $\mathbf{W} \neq \emptyset$, **return** $\mathbf{sID}(\mathbf{y}, \mathbf{x} \cup \mathbf{w}, P^*, D)$

4 **if** $\mathcal{C}(D \setminus \mathbf{X}) = \{C_0, C_1, ..., C_k\}$,
  return $\sum_{\mathbf{V} \setminus \{\mathbf{Y}, \mathbf{X}\}} \prod_i \mathbf{sID}(c_i, \mathbf{V} \setminus c_i, P^*, D)$

5 **if** $\mathcal{C}(D \setminus \mathbf{X}) = \{C_0\}$

6   **if** $(\mathbf{S} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{X})_{D_{\overline{\mathbf{X}}}}$, **return** $P(\mathbf{y}|do(\mathbf{x}))$

7   **if** $\mathcal{C}(D) = \{D\}$, **FAIL**(D, C_0)

8   **if** $C_0 \in \mathcal{C}(D)$, **return** $\sum_{s \setminus \mathbf{Y}} \prod_{i|V_i \in S} P^*(v_i|V_D^{(i-1)})$

9   **if** $(\exists C')C_0 \subset C' \in \mathcal{C}(D)$, **return** $\mathbf{sID}(\mathbf{y}, \mathbf{x} \cap C',$
  $\prod_{i|V_i \in C'} P^*(V_i|V_D^{(i-1)} \cap C', v_D^{(i-1)} \setminus C'), C')$.

Figure 4: Modified version of identification algorithm capable of recognizing transportable relations.

We prove next soundness and completeness of **sID**.

Theorem 6 (soundness)**.** *Whenever **sID** returns an expression for $P_{\mathbf{x}}(\mathbf{y})$, it is correct.*

Proof. See Appendix 2. □

Theorem 7. *Assume **sID** fails to transport $P_{\mathbf{x}}(\mathbf{y})$ (executes line 7). Then there exists $\mathbf{X}' \subseteq \mathbf{X}$, $\mathbf{Y}' \subseteq \mathbf{Y}$, such that the graph pair D, C_0 returned by the fail condition of **sID** contain as edge subgraphs s^*-trees F, F' that form a s-hedge for $P_{\mathbf{x}'}(\mathbf{y}')$.*

Proof. See Appendix 2. □

Corollary 3 (completeness)**.** ***sID** is complete.*

Proof. See Appendix 2. □

Corollary 4 (do-calculus characterization)**.** *The rules of do-calculus, together with standard probability manipulations are complete for establishing transportability of all effects of the form $P^*_{\mathbf{x}}(\mathbf{y})$.*

Proof. See Appendix 2. □

Conclusions

We provide a complete (necessary and sufficient) graphical condition for deciding when the causal effect of one set of variables on another can be transported from experimental to non-experimental environment. We further provide a complete algorithm for computing the correct transport formula whenever this graphical condition holds.

Appendix 1

The do-calculus [Pearl, 1995] consists of three rules that permit us to transform expressions involving do-operators into other expressions of this type, whenever certain conditions hold in the causal diagram G. (See footnote 1 for semantics.)

We consider a DAG G in which each child-parent family represents a deterministic function $x_i = f_i(pa_i, \epsilon_i), i = 1, \ldots, n$, where pa_i are the parents of variables X_i in G; and $\epsilon_i, i = 1, \ldots, n$ are arbitrarily distributed random disturbances, representing background factors that the investigator chooses not to include in the analysis.

Let X, Y, and Z be arbitrary disjoint sets of nodes in a causal DAG G. An expression of the type $E = P(y|do(x), z)$ is said to be compatible with G if the interventional distribution described by E can be generated by parameterizing the graph with a set of functions f_i and a set of distributions of $\epsilon_i, i = 1, \ldots, n$

We denote by $G_{\overline{X}}$ the graph obtained by deleting from G all arrows pointing to nodes in X. Likewise, we denote by $G_{\underline{X}}$ the graph obtained by deleting from G all arrows emerging from nodes in X. To represent the deletion of both incoming and outgoing arrows, we use the notation $G_{\overline{X}\underline{Z}}$.

The following three rules are valid for every interventional distribution compatible with G.

Rule 1 (Insertion/deletion of observations):

$$P(y|do(x), z, w) = P(y|do(x), w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}$$

Rule 2 (Action/observation exchange):

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$$

Rule 3 (Insertion/deletion of actions):

$$P(y|do(x), do(z), w) = P(y|do(x), w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\,\overline{Z(W)}}},$$

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

The do-calculus was proven to be complete [Shpitser and Pearl, 2006a; Huang and Valtorta, 2006], in the sense that if an equality cannot be established by repeated application of these three rules, it is not valid.

Appendix 2

Lemma 4. *Let $\mathbf{X}, \mathbf{Y}$ be sets of variables. Let M, M^* be a pair of causal models and G the respective selection diagram. Then $Q = P^*_{\mathbf{x}}(\mathbf{Y})$ is transportable in G if and only if Q is transportable in $G_{An(\mathbf{Y})}$.*

Proof. See [Tian, 2002, Chapter 5] that provides analogous construction. □

Theorem 3. *Let G be a selection diagram. Then for any node Y, the direct effect $P^*_{Pa(Y)}(y)$ is transportable if there is no subgraph of G which forms a Y-rooted sC-tree.*

Proof. We known from [Tian, 2002, Theorem 22] that whenever there exists no subgraph G_T of G satisfying all of the following: (i) $Y \in T$; (ii) G_T has only one c-component, T itself; (iii) All variables in T are ancestors of Y in G_T, the direct effect on Y is identifiable, as sC-trees are structures of this type. Further [Shpitser and Pearl, 2006b, Theorem 2] showed that the same holds for C-trees, which also implies the inexistence of a sC-trees. Since such structure does not show up in G, the target quantity is identifiable, and hence transportable.

It remains to show that the same holds whenever there exists a subgraph that is a C-tree and in which no S node points to Y, i.e., there is no Y-rooted sC-tree at all. It is true that $(S \perp\!\!\!\perp Y|Pa(Y))_{G_{\overline{Pa(Y)}}}$, given that all paths from S to Y are closed. This follows from the following facts: 1) all paths from S passing through Y's ancestors were cut in $G_{\overline{Pa(Y)}}$; 2) all bidirected paths were also closed given that the conditioning set contains only root nodes, and a connection from S must pass through at least one collider; 3) by Lemma 4, transportability does not depend on descendants of Y. Thus, it follows that we can write $P^*_{Pa(Y)}(Y) = P_{Pa(Y)}(Y|S) = P_{Pa(Y)}(Y)$, concluding the proof. □

Corollary 1. *Let G be a selection diagram. Then for any node Y, the direct effect $P^*_{Pa(Y)}(y)$ is transportable if there is no S node pointing to Y.*

Proof. Follows directly from Theorem 3. □

Lemma 5. *The exclusive OR (XOR) function is commutative and associative.*

Proof. Follows directly from the definition of the XOR function. □

Remark 1. Despite the fact of being a strict generalization of Theorem 2, the construction provided below is still worth to make for two reasons: first, it will provide a simplified construction of the one given in Theorem 2; second, it will set the tone for proofs of generic graph structures which will in the sequel show to be instrumental in proving non-transportability in arbitrary structures.

Theorem 4. *Let G be a Y-rooted sC-tree. Then the effects of any set of nodes in G on Y are not transportable.*

Proof. The proof will proceed by constructing a family of counterexamples. For any such G and any set $\mathbf{X}$, we will construct two causal models M_1 and M_2 that will agree on $\langle P, P^*, I \rangle$, but disagree on the interventional distribution $P_{\mathbf{x}}(y)$.

Let the two models M_1, M_2 agree on the following features. All variables in $\mathbf{U} \cup \mathbf{V}$ are binary. All exogenous variables are distributed uniformly. All endogenous variables except Y are set to the bit parity (xor) of the values of their parents. The two models differ is respect to Y's definition. Consider the function for Y, $f_Y : U, Pa(Y) \rightarrow Y$ to be defined as follows:

$$\begin{cases} M_1 : Y = \big((pa(Y) \otimes u) \otimes s\big) \\ M_2 : Y = \big((pa(Y) \otimes u) \vee s\big) \end{cases}$$

Lemma 6. *The two models agree in the distributions* $\langle P, P^*, I\rangle$.

Proof. Since the two models agree on $P(\mathbf{U})$ and all functions except f_Y, it suffices to show that f_Y maintains the same input/output behavior in both models for each domains.
Subclaim 1: Let us show that both models agree in the observational and interventional distributions relative to domain Π, i.e., the pair $\langle P, I\rangle$. The index variable S is set to 0 in Π, and f_Y evaluates to $(pa(Y) \otimes u)$ in both models, which proves the subclaim.
Subclaim 2: Let us show that both models agree in the observational distribution relative to Π^*, i.e., P^*. The index variable S is set 1 in Π^*, and f_Y evaluates to $((pa(Y) \otimes u) \otimes 1)$ in M_1, and 1 in M_2. Since the evaluation in M_1 can be rewritten as $\neg((pa(Y) \otimes u)$, it remains to show that $(pa(Y) \otimes u)$ always evaluates to 0.

This fact is certainly true, consider the following observations: a) each variable in $\mathbf{U}$ has exactly two endogenous children; b) the given tree has Y as the root; c) all functions are *XOR* – these imply that Y is computing the bit parity of the sum of all U nodes, which turns out to be even, and so evaluates to 0 and proves the subclaim. $\square$

Lemma 7. *For any set* $\mathbf{X}$, $P_1(Y|do(\mathbf{X}), S = 1) \neq P_2(Y|do(\mathbf{X}), S = 1)$.

Proof. Given the functional description and the discussion in the previous Lemma, the function f_Y evaluates always to 1 in M_2.

Now let us consider M_1. Note that performing the intervention and cutting the edges going toward $\mathbf{X}$ creates an asymmetry on the sum of the bidirected edges departing from U, and consequently in the sum performed by Y. It will be the case that some $\mathbf{U}'$ will appear only once in the expression of Y. Therefore, depending on the assignment $\mathbf{X} = \mathbf{x}$, we will need to evaluate the sum (mod 2) over $\mathbf{U}'$ in Y or its negation, which given the uniformity of the distribution of U will yield $P_1(Y|do(\mathbf{X}), S = 1) = 1/2$ in both cases. $\square$

By Lemma 1, Lemmas 6 and 7 together prove Theorem 4. $\square$

Corollary 2. *Let G be a selection diagram, let $\mathbf{X}$ and $\mathbf{Y}$ be set of variables. If there exists a node W which is an ancestor of some node $Y \in \mathbf{Y}$ and such that there exists a W-rooted sC-tree which contains any variables in $\mathbf{X}$, then $P_{\mathbf{x}}(\mathbf{y})$ is not transportable.*

Proof. Fix a W-rooted sC-tree T, and a path p from W to Y. Consider the graph $p \cup T$. Note that in this graph $P^*_x(Y) = \sum_w P^*_x(w)P^*(Y|w)$. From the last Theorem $P^*_x(w)$ is not transportable, it is now easy to construct $P^*(Y|W)$ in such a way that the mapping from $P_{\mathbf{x}}(W)$ to $P_{\mathbf{x}}(\mathbf{Y})$ is one to one, while making sure all distributions are positive. $\square$

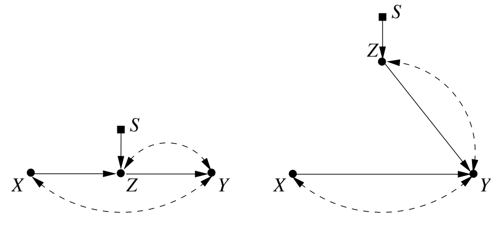


Figure 5: Selection diagrams in which $P(y|do(x))$ is not transportable, there is no sC-tree but there is a s^*-tree. These diagrams will be used as basis for the general case; the first diagram is named sp-graph and the second one sb-graph.

Remark 2. The previous results comprised cases in which there exist sC-trees involved in the non-transportability of Y – i.e., Y or some of its ancestors were roots of a given sC-tree. In the problem of identifiability, the counterpart of sC-tree (i.e., C-tree) suffices to characterize non-identifiability for singleton Y. But transportability is more subtle and this is not the case here – it depends not only on X and Y "locations" in the graph, but also the relative position of S. Consider Figures 3 and 5(a) (sp-graph). In these graphs there is no sC-tree but the effect of X on Y is still non-transportable.

The main technical subtlety here is that in sC-trees, a S-node combines its effect with a X-node intersecting in the root node (considering only the bidirected edges), which is not the case for non-transportability in general. Note that in the graphs in Figures 3 and the sp-graph, the nodes S and X intersect first through ordinary edges and meet through bidirected edges only on the Y node. This implies a certain "asynchrony" because in the structural sense when we have a S-node this implies a difference in the structural equations between domains. But only a difference in the structural sense does not imply non-transportability, for instance, $P^*_x(z)$ is transportable in the sp-graph even though the equations of Z being different in both models.

The key idea to produce a proof for non-transportability in these cases is to keep the effect of S-nodes after intersecting with X "dormant" until they reach the target Y and then manifest. We implement this idea in the next proof, which is one base case and should pavement the way for the most general problem.

Theorem 8. *$P^*_x(Y)$ is not transportable in the sp-graph (Fig. 5(a)).*

Proof. We will construct two causal models M_1 and M_2 compatible with the sp-graph that will agree on $\langle P, P^*, I\rangle$, but disagree on the interventional distribution $P_x(Y)$.

Let us assume that all variables in $\mathbf{U} \cup \mathbf{V}$ are binary, and let U_1 be the common cause of X and Y, U_2 be the common cause of Z and Y, and U_3 be the random disturbance exclusive to Z. Let M_1 and M_2 be defined as follows:

$$M_1 = \begin{cases} X = U_1 \\ Z = \Big(((X \otimes U_2 \otimes 1) \otimes U_3) \vee S \Big) \otimes \Big(S \wedge (X \otimes U_2) \Big) \\ Y = Z \otimes U_1 \otimes U_2 \end{cases}$$

and:

$$M_2 = \begin{cases} X = U_1 \\ Z = \Big(((U_2 \otimes 1) \otimes U_3) \vee S \Big) \otimes (S \wedge U_2) \\ Y = Z \otimes U_2 \end{cases}$$

Both models agree in respect to $P(\mathbf{U})$, which is defined as follows: $P(U_1) = P(U_2) = P(U_3) = 1/2$.

Lemma 8. *The two models agree in the distributions $\langle P, P^*, I \rangle$.*

Proof. **Subclaim 1**: Let us show that both models agree in the observational and interventional distributions relative to domain Π, i.e., the pair $\langle P, I \rangle$. In both models X has the same expression, which entails the same probabilistic behavior in both cases. The index variable S is set to 0 in Π, and Z evaluates to $(X \otimes U_2 \otimes 1 \otimes U_3)$ in M_1 and $(U_2 \otimes 1 \otimes U_3)$ in M_2. Since $\mathbf{U}$ is equal and uniformly distributed in both models, we obtain the same input/output probabilistic behavior in M_1 and M_2. In similar way, Y evaluates to $(1+R)$ in both models, which entails the same input/output probabilistic behavior in both models.
Subclaim 2: Let us show that both models agree in the observational distribution P^* relative to Π^*. The index variable S is set 1 in Π^*, f_Z evaluates to $(X \otimes U_2 \otimes 1)$ in M_1, and $(U_2 \otimes 1)$ in M_2. Together with the uniformty of $\mathbf{U}$, $P(Z = 1 | X = x) = 1/2$ in both models. Further, f_Y evaluates to 1 in both models, which yields the same input/output behavior in both models. □

Lemma 9. *There exist values of X, Y such that $P_1(Y|do(X), S = 1) \neq P_2(Y|do(X), S = 1)$.*

Proof. Fix $X = 1, Y = 1$. First notice that f_Z evaluates to U_2 in M_1 and $\overline{U_2}$ in M_2. Given that U_2 is distributed uniformly, both quantities coincide (and they represent the effect of X on Z, which is transportable in G). Now the evaluation of f_Y in M_1 reduces to U_1, while it reduces to 1 in M_2, which show disagreement and finishes the proof of this Lemma. □

By Lemma 1, Lemmas 8 and 9 together prove Theorem 8. □

Remark 3. We have a different sort of asymmetry in the case of Fig. 5(b) (called sb-graph). In this case, the nodes X and S do not intersect before meeting Y – i.e., they have disjoint paths and Y lies precisely in their intersection.

Still, this case is not the same of having a sC-tree because in sb-graphs we need to keep the equality from the S nodes to Y until S intersects X on Y. Employing a similar construct as in the sp-graph, we keep the effect of S dormant until it reaches Y and then emerges.

Theorem 9. *$P^*_x(Y)$ is not transportable in the sb-graph (Fig. 5(b)).*

Proof. We construct two causal models M_1 and M_2 compatible with the sb-graph that will agree on $\langle P, P^*, I \rangle$, but disagree on the interventional distribution $P_x(Y)$.

Let us assume that all variables in $\mathbf{U} \cup \mathbf{V}$ are binary, and let U_1 be the common cause of X and Y, U_2 be the common cause of Z and Y, and U_3 be the random disturbance exclusive to X. Let M_1 and M_2 agree with the following definitions:

$$M_1, M_2 = \begin{cases} X = U_1 \otimes U_3 \\ Y = X \otimes Z \otimes U_1 \otimes U_2 \end{cases}$$

and disagree in respect to Z as follows:

$$\begin{cases} M_1 : Z = U_2 \otimes S \\ M_2 : Z = \big((U_2 \vee S) \otimes (S \wedge (\overline{U_2}) \big) \end{cases}$$

Both models also agree in respect to $P(\mathbf{U})$, which is defined as follows: $P(U_1) \neq 1/2$, $P(U_2) = P(U_3) = 1/2$.

Lemma 10. *The two models agree in the distributions $\langle P, P^*, I \rangle$.*

Proof. **Subclaim 1**: Let us show that both models agree in the observational and interventional distributions relative to domain Π, i.e., the pair $\langle P, I \rangle$. The index variable S is set to 0 in Π, and Z evaluates to $(X \otimes U_2)$ in both models. Since the two models agree on $P(\mathbf{U})$ and all other other functions, the two models generate the same distributions for Π.
Subclaim 2: Let us show that both models agree in the observational distribution P^* relative to Π^*. The index variable S is set 1 in Π^*, f_Z evaluates to $\big((U_2 \otimes 1\big)$ in M_1, and U_2 in M_2. Given that these variables are uniformly distributed, both models agree in $P(Z)$. Now let us consider the behavior of f_Y, it evaluates to $\overline{U_3}$ in M_1, and U_3 in M_2, and since $P(U_3)$ is uniformly distributed, it is the same in both models. □

Lemma 11. *There exist values of X, Y such that $P_1(Y|do(X), S = 1) \neq P_2(Y|do(X), S = 1)$.*

Proof. Fix $X = 1, Y = 1$. First notice that f_Z evaluates to $\overline{U_2}$ in M_1 and U_2 in M_2. The evaluation of f_Y in M_1 reduces to U_1, while it reduces to $\overline{U_1}$ in M_2. It follows that in M_1, f_Y evaluates to 1 with probability $P(U_1 = 1)$, while in M_2 it evaluates with probability $P(U_1 = 0)$, which disagree by construction, finishing the proof of this Lemma. □

By Lemma 1, Lemmas 10 and 11 together prove Theorem 9. □

Remark 4. We have two complementary components to forge a general scheme to prove arbitrary non-transportability. First, the construct of Theorem 4 shows how to prove non-transportability for general structures such as sC-trees. In the sequel, the specific proofs of non-transportability for the sp-graph (Theorem 8) and sb-graph (Theorem 9) partition the possible interactions between X, S and Y. In the former, X and S intersect before meeting with Y, while in the latter they have disjoint paths and Y

lies in their intersection. Not surprisingly, the proof for the general case basically combines these analyses, which we show below.

Theorem 5. *Assume there exist F, F' that form a s-hedge for $P_x(y)$ in Π and Π^*. Then $P_x(y)$ is not transportable from Π to Π^*.*

Proof. We first consider counterexamples with the induced graph $H = De(F)_G \cap An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$, and assume, without loss of generality, that H is a forest. We construct two causal models M_1 and M_2 that will agree on $\langle P, P^*, I\rangle$, but disagree on the interventional distribution $P^*_{\mathbf{x}}(\mathbf{y})$.

Let F be an $\mathbf{R}$-rooted sC-forest, let $\mathbf{V}'$ be the set of observable variables and $\mathbf{U}'$ be the set of unobservable variables in F. Let us assume that all variables in $\mathbf{U}' \cup \mathbf{V}'$ are binary. Call $\mathbf{W}$ the set of variables pointed by S-nodes in F', which by the definition of sC-forest is guaranteed to be non-empty.

In model 1, let each $V_i \in \mathbf{V}' \setminus \mathbf{W}$ compute the bit parity of all its observable and unobservable parents (i.e., $f_i^{(1)} = \otimes(\bigcup_{V_j \in \mathbf{Pa_i}} V_j)$, where the xor is applied for each element of the set and the result computed so far), while in model 2, let V_i compute the bit parity of all its parents except that any node in F' disregards the parents values if the parent is in F (i.e., $f_i^{(2)} = \otimes(\bigcup_{V_j \in \mathbf{Pa_i} \cap F'} V_j)$ if V_i is in F', and $f_i^{(2)} = f_i^{(1)}$, otherwise).

Define $W \in \mathbf{W}$ as follows:

$$\begin{cases} M_1 : W = \left(\left(f_w^{(1)} \otimes U_w^*\right) \vee S\right) \otimes \left(S \wedge \left(1 \otimes f_w^{(1)}\right)\right) \\ M_2 : W = \left(\left(f_w^{(2)} \otimes U_w^*\right) \vee S\right) \otimes \left(S \wedge \left(1 \otimes f_w^{(2)}\right)\right) \end{cases}$$

where f_w is constructed in similar way as f_i in M_1 and M_2 above, and U_w^* is an additional fair coin exclusively pointing to W. Let us call $\mathbf{U_w}$ the collection of such coins. Furthermore, let us assume that each $U_i \in \{\mathbf{U}' \setminus \mathbf{U_w}\}$ is also a fair coin (i.e., $P(U_i) = 1/2$).

Lemma 12. *The two models agree in the distribution of P^* and there exists a value assignment $\mathbf{x}$ for $\mathbf{X}$ such that $P_1(\mathbf{Y}|do(\mathbf{x}), S = 1) \neq P_2(\mathbf{Y}|do(\mathbf{x}), S = 1)$.*

Proof. For $S = 1$, the result follows directly since the systems of equations in both models reduce to the construction given in Theorem 4 at [Shpitser and Pearl, 2006b]. □

Lemma 13. *The two models agree in the distributions $\langle P, I\rangle$.*

Proof. Let us show that both models agree in the observational distribution P relative to domain Π. The selection variable S is set to 0 in Π, and note that both systems are the same as in Π^* except that now each variable $W \in \mathbf{W}$ has an extra variable U_w^* pointing to it that should be taken into account in W's evaluation, and in turn in the whole system.

We have a forest over the endogenous nodes and all functions compute the bit parity of the value of their parents, and so we can view each node as computing the sum mod 2 of its exogenous ancestors in H. We want to show that the distribution of each family is equally likely for each possible assignment (i.e., $P(v_i|\mathbf{pa_i}) = 1/2$, for all $v_i, \mathbf{pa_i}$).

Let us partition the analysis in two cases. First consider the case of $V_i \in \mathbf{R}$ in which there exists a S-node in the respective sC-tree. Note that that the evaluation of V_i relies only on the value of $U_w^* \in \mathbf{U_w}$ in its respective tree since $U \in \{\mathbf{U}' \setminus \mathbf{U_w}\}$ has an even number of endogenous children in F, and it is counted twice, so evaluates to zero (i.e., it does not affect V_i's evaluation). For now, let us assume that there is only one U_w^* that affects the evaluation of V_i. Given the uniformity of U_w^*, it suffices to show that U_w^* can vary independently for any configuration of the parents of V_i.

For any configuration of $\mathbf{U}' = (U_1 = u_1, ..., U_w^* = u_w^*, ...)$, consider the corresponding evaluation of $\mathbf{Pa_i} = \mathbf{pa_i}$, and also $V_i = u_w^*$. We want to show that it is possible to flip the current value of U_w^* from u_w^* to $\neg u_w^*$ while preserving the parents' evaluation $\mathbf{pa_i}$. Assume this is not so. This implies that the evaluation of $\mathbf{Pa_i}$ and V_i count the same $\mathbf{U}$'s, contradiction.

To see why, consider $\mathbf{Pa_i^*} \subseteq \mathbf{Pa_i}$ the set of parents of V_i that are descendents of U_w^*. Now, for each of these parents flip the minimum number of variables from $\mathbf{U} \setminus \mathbf{U_w}$, and call this set $\mathbf{U}^*$. (Note that this is always possible since we need at most one U for each parent, which should exist by construction of sC-forest.) Now, make $U_w^* = \neg u_w^*$, and note that $\mathbf{Pa_i} = \mathbf{pa_i}$ since flipping the values of $\mathbf{U}^*$ compensates the flip of U_w^*. But it is also true now that V_i evaluates to $\neg u_w^*$ since, in the same way as before, all other variables in $\{\mathbf{U} \setminus \mathbf{U_w}\}$ are cancelled out in V_i's evaluation, including the ones in $\mathbf{U}^*$. This proves the claim.

Consider the following two facts: **Subclaim 1**: Let X and Y be two binary variables such that $P(X = x) = p \neq 1/2$ and $P(Y = y) = q = 1/2$. Then the probabilistic input/output behavior of $Z = XOR(X, Y)$ is the same of Y. The variable $Z = 1$ whenever $\{(X = 1, Y = 0), (X = 0, Y = 1)\}$, which happens with probability $pq + (1 - p)(1 - q)$. Since $q = 1/2$, the expression reduces to $p * 1/2 + (1 - p) * 1/2 = 1/2$. **Subclaim 2**: Let X and Y be two binary variables such that $P(X = x) = P(Y = y) = p = 1/2$. Then the probabilistic input/output behavior of $Z = XOR(X, Y)$ is the same of X (or Y). This follows directly from Subclaim 1. It is clear that if there are multiple nodes from $\mathbf{U_w}$ in the evaluation of V_i, the same construction is also valid given the subclaim above. It is also not difficult to generalize this argument to consider root set that are not singleton, including roots in which there are not S-nodes as ancestors.

Finally, let us consider the case of $V_i \in \{F \setminus \mathbf{R}\}$. It suffices to show that the function from $\mathbf{U}' \setminus \mathbf{U_w}$ to $\mathbf{V}' \setminus \mathbf{R}$ is 1-1 when we fix $\mathbf{U_w} = \mathbf{u_w}$. We use the same argument as Shpitser. Assume this is not so, and fix two instantiations of $\mathbf{U}' \setminus \mathbf{U_w}$ that map to the same value of $\mathbf{V}' \setminus \mathbf{R}$, and differ by the set $\mathbf{U}^* = \{U_1, ..., U_k\}$. Since the bidirected edges form a spanning tree, there exists $\mathbf{V}^*$ with an odd number of parents in $\mathbf{U}^*$ (and were not in $\mathbf{R}$, by construction). Order them topologically and let the topmost be called X. Note that if we flip all values in $\mathbf{U}^*$, the value of X will also flip, contradiction. Given the uniformity of $\mathbf{U}'$, the claim follows. We can put this together with the previous claim, and the result follows. We can add fair coins as the input to all other variables outside F, which will imply the claim for the whole graph G.

In regard to the equality between I, note that given that the equality of both models holds for P, and removing edges due to interventions will just make some nodes from $\mathbf{U}' \setminus \mathbf{U_w}$ to have an odd number of children, it it not difficult to see based on the previous argument that this just creates more variables that are free to vary, which will entail the same probabilistic uniform behavior in both models. Another way to see this fact is to consider the new exogenous variables from $\{\mathbf{U} \setminus \mathbf{U_w}\}$ that have only one children after the intervention as analogous to U^*_w, and so the same argument follows. □

Finally, Lemma 1 together with Lemmas 12 and 13 prove Theorem 5. □

Theorem 6 (soundness)**.** *Whenever **sID** returns an expression for $P_{\mathbf{x}}(\mathbf{y})$, it is correct.*

Proof. The correctness of the identifiability calls were already established elsewhere [Huang and Valtorta, 2006; Shpitser and Pearl, 2006b], which are performed by **sID** over Π^* and called trivial transportability.

It remains to show the correctness of the test in line 6 of **sID**. First note that, by construction, $\mathbf{X}$ is always a set of pre-treatment covariates. But now the correctness follows directly by S-admissibility of $\mathbf{X}$ together with Corollary 1 in [PB, 2011]. □

Remark 5. The next results are similar to the analogous ones for identification given in [Tian and Pearl, 2002] and [Shpitser and Pearl, 2006a].

Theorem 7. *Assume **sID** fails to transport $P_{\mathbf{x}}(\mathbf{y})$ (executes line 7). Then there exists $\mathbf{X}' \subseteq \mathbf{X}$, $\mathbf{Y}' \subseteq \mathbf{Y}$, such that the graph pair D, C_0 returned by the fail condition of **sID** contain as edge subgraphs s^*-trees F, F' that form a s-hedge for $P_{\mathbf{x}'}(\mathbf{y}')$.*

Proof. Before failure **sID** evaluated false consecutively at line 5 and 6, so D local to this call is a sC-component, and let $\mathbf{R}$ be its root set. We can remove some directed arrows from D while preserving $\mathbf{R}$ as root, yielding a $\mathbf{R}$-rooted s^*-tree F. Since by construction $F' = F \cap C_0$ is closed under descendants and only directed arrows were removed, both F, F' are s^*-trees. Also by construction, $\mathbf{R} \subset An(\mathbf{Y})_{D_{\overline{\mathbf{X}}}}$ together with the fact that $\mathbf{X}$ and $\mathbf{Y}$ from the recursive call are clearly subsets of the original input, finish the proof. □

Lemma 14. *Let $\mathbf{X}, \mathbf{Y}$ be sets of variables. If $Q = P^*_{\mathbf{x}}(\mathbf{Y})$ is not transportable in G, then Q is not transportable in the graph resulted from adding a directed or bidirected edge to G. Equivalently, if Q is transportable in G, then it is also transportable in graph resulted from removing a directed or bidirected edge from G.*

Proof. This result is obvious, see [Tian, 2002, Chapter 5] that provides an analogous construction. □

Lemma 15. *Let $\mathbf{X}, \mathbf{Y}$ be sets of variables. If $Q = P^*_{\mathbf{x}}(\mathbf{Y})$ is not transportable in respect to the selection diagram G, then Q is not transportable in the selection diagram resulted from adding selection nodes to G. Equivalently, if Q is transportable in G, then it is also transportable in graph resulted from removing selection nodes from G.*

Proof. This proof is also obvious and follows the same structure of Lemmas 4 and 14. □

Corollary 3 (completeness)**.** ***sID** is complete.*

Proof. The result follows from Theorem 5 together with Lemmas 4, 14, and 15. □

Corollary 4 (do-calculus characterization)**.** *The rules of do-calculus, together with standard probability manipulations are complete for establishing transportability of all effects of the form $P^*_{\mathbf{x}}(\mathbf{y})$.*

Proof. It was shown elsewhere [Shpitser and Pearl, 2006a] that the steps of **sID** but line 6, 7 correspond to sequences of standard probability manipulations and applications of the rules of do-calculus. The line 6 is constituted by a conditional independence judgement, and standard probability operations for the replacement of the functions based on the invariance allowed by the S-admissibility of the local $\mathbf{X}'$ in each recursive call. Note that by Theorem 7, for all graphs in which the failure conditions is triggered (line 7), it is possible to construct a counterexample for transportability. □

Corollary 5. *Theorem 3 in [PB, 2011] is incomplete.*

Proof. Figure 1(c) demonstrates a selection diagram in which the relation $R = P^*(y|do(x))$ is transportable, but Theorem 3 is not capable of recognizing it.

Let us test the applicability of each of its conditions:

- Step 1. R is not trivially transportable due to the confounding arc $X \leftrightarrow Z$ due to Tian's identifiability [Tian and Pearl, 2002];
- Step 2. There is no S-admissible set because the confounding arc $V \leftrightarrow Y$ and Verma's inducing path condition [Verma and Pearl, 1990];
- Step 3. There is no set W which makes $(X \perp\!\!\!\perp Y|W)$ to hold, this is due to the confounding arc $X \leftrightarrow Y$;

Since there is no remaining actions to be taken, the algorithm exits without returning any expression. □

References

Adelman, L. 1991. Experiments, quasi-experiments, and case studies: A review of empirical methods for evaluating decision support systems. *Systems, Man and Cybernetics, IEEE Transactions on* 21(2):293 –301.

Campbell, D., and Stanley, J. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Wadsworth Publishing.

Glass, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educational Researcher* 5(10):pp. 3–8.

Hedges, L. V., and Olkin, I. 1985. *Statistical Methods for Meta-Analysis*. Academic Press.

Huang, Y., and Valtorta, M. 2006. Identifiability in causal bayesian networks: A sound and complete algorithm. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press. 1149–1156.

Koller, D., and Friedman, N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Kuroki, M., and Miyakawa, M. 1999. Identifiability criteria for causal effects of joint interventions. *Journal of the Royal Statistical Society* 29:105–117.

Manski, C. 2007. *Identification for Prediction and Decision*. Cambridge, Massachusetts: Harvard University Press.

Owen, A. B. 2009. Karl pearsons meta-analysis revisited. *Annals of Statistics* 37(6B):3867–3892.

Pearl, J., and Bareinboim, E. 2011. Transportability of causal and statistical relations: A formal approach. In *Proceedings of the Twenty-Fifth National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press. 247–254.

Pearl, J., and Verma, T. 1991. A theory of inferred causation. In Allen, J.; Fikes, R.; and Sandewall, E., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*. San Mateo, CA: Morgan Kaufmann. 441–452.

Pearl, J. 1995. Causal diagrams for empirical research. *Biometrika* 82(4):669–710.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press. Second ed., 2009.

Pearl, J. 2001. Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*. San Francisco, CA: Morgan Kaufmann. 411–420.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press, second edition.

Pearl, J. 2012. The mediation formula: A guide to the assessment of causal pathways in nonlinear models. In Berzuini, C.; Dawid, P.; and Bernardinell, L., eds., *Causality: Statistical Perspectives and Applications*. New York: Wiley. Chapter 12.

Shadish, W.; Cook, T.; and Campbell, D. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin, second edition.

Shpitser, I., and Pearl, J. 2006a. Identification of conditional interventional distributions. In Dechter, R., and Richardson, T., eds., *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*. Corvallis, OR: AUAI Press. 437–444.

Shpitser, I., and Pearl, J. 2006b. Identification of joint interventional distributions in recursive semi-Markovian causal models. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press. 1219–1226.

Spirtes, P.; Glymour, C.; and Scheines, R. 2001. *Causation, Prediction, and Search*. Cambridge, MA: MIT Press, 2nd edition.

Tian, J., and Pearl, J. 2002. A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press/The MIT Press. 567–573.

Tian, J. 2002. *Studies in Causal Reasoning and Learning*. Ph.D. Dissertation, Computer Science Department, University of California, Los Angeles, CA.

Verma, T., and Pearl, J. 1990. Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*, 220–227. Also in P. Bonissone, M. Henrion, L.N. Kanal and J.F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 6*, Elsevier Science Publishers, B.V., 255–268, 1991.