In Proceedings of the 16th International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 135-143, 2013.

TECHNICAL REPORT
R-407
First version: November 2012
Revised: February 2013

Meta-Transportability of Causal Effects: A Formal Approach

Elias Bareinboim
Cognitive Systems Laboratory
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA. 90095
eb@cs.ucla.edu

Judea Pearl
Cognitive Systems Laboratory
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA. 90095
judea@cs.ucla.edu

Abstract

This paper considers the problem of transferring experimental findings learned from multiple heterogeneous domains to a different environment, in which only passive observations can be collected. Pearl and Bareinboim (2011) established a complete characterization for such transfer between two domains, a source and a target, and this paper generalizes their results to multiple heterogeneous domains. It establishes a necessary and sufficient condition for deciding when effects in the target domain are estimable from both statistical and causal information transferred from the experiments in the source domains. The paper further provides a complete algorithm for computing the transport formula, that is, a way of fusing observational and experimental information to synthesize an unbiased estimate of the desired effects.

1 Motivation

The problem of transporting and synthesizing experimental knowledge from heterogeneous settings is pervasive in science. Conclusions that are obtained in a laboratory setting are transported and applied elsewhere, in an environment that differs in many aspects from that of the laboratory. In data-driven sciences, experiments are conducted on disparate domains, but the intention is almost invariably to fuse the acquired knowledge, and translate it into some meaningful claim about a target domain, which is usually different than any of the individual study domains.

Appearing in Proceedings of the 16^{th} International Conference on Artificial Intelligence and Statistics (AISTATS) 2013, Scottsdale, AZ, USA. Volume 31 of JMLR: W&CP 31.

However, 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), "overgeneralization" (Hofler, Gloster, and Hoyer (2010)), "quasi-experiments" (Shadish, Cook, and Campbell (2002), Ch. 3, Adelman (1991)), "heterogeneity" (Morgan and Winship (2007)), 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 to answer the practical challenge of generalizing causal knowledge across multiple domains posed in this paper.

Artificial Intelligence and Statistics provide the tools for tackling the meta-transportability problem. First, the distinction between statistical and causal knowledge has received syntactic representation through causal diagrams (Pearl, 1995, Spirtes, Glymour, and Scheines, 2000, Pearl, 2009). Second, the inferential machinery provided by the do-calculus (Pearl, 1995, 2009, Koller and Friedman, 2009) is particularly suitable for handling knowledge transfer across domains.

Armed with these techniques, Pearl and Bareinboim (2011) introduced a formal language for encoding differences and commonalities between domains and derived a complete set of conditions under which transportability of empirical findings is feasible between two domains, a source and a target. This paper generalizes their results to the case of multiple heterogeneous sources, a task that we call here "meta-transportability".

More formally, the meta-transportability problem concerns the transfer of causal knowledge from a heterogeneous collection of source domains $\Pi = \{\pi_1, ..., \pi_n\}$ to a target domain π^*. In each domain $\pi_i \in \Pi$, experiments can be performed and causal knowledge gathered. In π^*, potentially different from π_i, 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 trivial; no transfer can be justified. Yet the fact that all scientific 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 domains share common characteristics and that, owed to these commonalities, causal claims would be valid in new settings even where experiments cannot be conducted.

This paper generalizes the one dimensional theory introduced in (Pearl and Bareinboim, 2011) to cases where there exist multiple heterogeneous source domains. Remarkably, while the effects of interest might not be individually transportable to the target domain from anyone of the available sources, combining different pieces from the various sources can still enable us to estimate the effects on the target domain (to be shown later on).

The goal of this paper is to formally characterize the conditions under which the target quantity is (non-parametrically) estimable from the available data, and then to develop an effective procedure to decide when and how this transfer is possible.

2 Previous Work and Our Contributions

Consider Fig. 1(a) in which the node S represents factors that produce differences between source and target populations. Assume that we 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 the United States (U.S.), but we find the distribution $P(x, y, z)$ in LA to be different from the one in the U.S. (call the latter $P^*(x, y, z)$). In particular, the average age in the U.S. is significantly higher than that in LA. How are we to estimate the causal effect of X on Y in U.S., denoted $R = P^*(y|do(x))$? [1] [2]

The selection diagram for this example (Fig. 1) conveys the assumption that the *only* difference between the two populations are factors determining age distributions, shown as $S \rightarrow Z$, 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

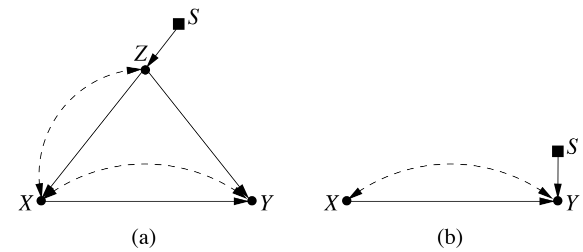


Figure 1: (a) Selection diagram illustrating when transportability among two domains is trivially solved through a simple recalibration. (b) Smallest selection diagram in which a causal relation is not transportable.

square nodes (■). From this assumption, the overall causal effect in the U.S. can be derived as follows:

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

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 the U.S. population, $P^*(z)$, to obtain an experimental claim $P^*(y|do(x))$ about the U.S.. In this trivial example, the transport formula amounts to a simple re-calibration (or re-weighting) of the age-specific effects to account for the new age distribution. In general, however, a more involved mixture of experimental and observational findings would be necessary to obtain a bias-free estimate of the target relation R, a full characterization of which is given in (Bareinboim and Pearl, 2012).

One might surmise that multiple pairwise transportability would be sufficient to solve the meta-transportability problem, but this is not the case. To witness, consider Fig. 2 which concerns the transference of experimental results from two sources ($\{\pi_a, \pi_b\}$) to infer the effect of X on Y in π^*, $R = P^*(y|do(x))$. In these graphs, X may represent the treatment (e.g., drug), Z represents an intermediate variable (e.g., biomarker), and Y represents the outcome (e.g., survival).

A simple analysis based on (Bareinboim and Pearl, 2012) shows that R cannot be transported from either source alone. However, we can decompose R into sub-relations such that each one is separately transportable from the source domains (in this case, π_a or π_b), as is demonstrated in Section 4 below.

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

[2] We use the structural interpretation of causal diagrams as described in (Pearl, 2000, pp. 205).

The goal of this paper is to go further and characterize the conditions for the existence of such decomposition, as well as its form, for **any** arbitrary collection of selection diagrams. Our contributions can be summarized as follows:

- We derive a general graphical condition for deciding meta-transportability of causal effects. We show that meta-transportability is feasible if and only if a certain graph structure does not appear as an edge subgraph of the inputted collection of selection diagrams.

- We construct a complete algorithm for deciding meta-transportability of joint causal effects and returning the correct transport formula whenever those effects are meta-transportable.

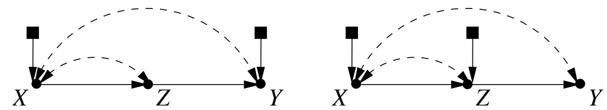
- We show that the do-calculus is complete for deciding meta-transportability.

3 Definitions and Preliminaries

The basic semantical framework in our analysis rests on *structural causal models* as defined in (Pearl, 2000, pp. 205), also called probabilistic causal models, or simply data-generating 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 $do(\mathbf{x})$.

We follow the conventions given in (Pearl, 2000). We will denote variables by capital letters and their realized values by small letters. Similarly, sets of variables will be denoted by bold capital letters, sets of realized values by bold letters. We will use the typical graph-theoretic terminology with the corresponding abbreviations $Pa(\mathbf{Y})_G$, $An(\mathbf{Y})_G$, and $De(\mathbf{Y})_G$, which will denote respectively the set of observable parents, ancestors, and descendants of the node set $\mathbf{Y}$ in G. By convention, these sets will include the arguments as well, for instance, the ancestral set $An(\mathbf{Y})_G$ will include $\mathbf{Y}$. We will usually omit the graph subscript whenever the graph in question is assumed or obvious. A graph $G_{\mathbf{Y}}$ will denote the induced subgraph G containing nodes in $\mathbf{Y}$ and all arrows between such nodes. Finally, $G_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}$ stands for the edge subgraph of G where all incoming arrows into $\mathbf{X}$ and all outgoing arrows from $\mathbf{Z}$ are removed.

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



(a) (b)

Figure 2: Selection diagrams illustrating impossibility of obtaining $P^*(y|do(x))$ through individual transportability from π_a and π_b to π^*, yet a more elaborated analysis yield the desired result combining different pieces from both domains.

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 usually associated with one particular domain (also called setting, study, population, environment). In ordinary transportability, this representation was extended to capture properties of two domains simultaneously. This is made possible if we assume that the structural equations share the same set of arguments, though the functional forms of the equations may vary arbitrarily (Bareinboim and Pearl, 2012). [3]

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.[4] Alternatively, the absence of a selection node pointing to a variable represents the assumption that the mechanism responsible for assigning value to that variable is identical in both domains.

[3] As discussed in the reference, the assumption of no structural changes between domains can be relaxed.

[4] 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 might be used to infer the diagrams from data (Pearl and Verma, 1991, Pearl, 2000, Spirtes et al., 2000).

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

Definition 3 (μ-Transportability)**.** *Let $\mathcal{D} = \{D_1, ..., D_n\}$ be a collection of selection diagrams relative to source domains $\Pi = \{\pi_1, ..., \pi_n\}$, and target domain π^*, respectively. Let $\langle P^i, I^i \rangle$ be the pair of observational and interventional distributions of π_i, and P^* be the observational distribution of π^*. The causal effect $R = P^*_{\mathbf{x}}(\mathbf{y})$ is said to be μ-transportable from Π to π^* in $\mathcal{D}$ if $P^*_{\mathbf{x}}(\mathbf{y})$ is uniquely computable from $\bigcup_{i=1,...,n} \langle P^i, I^i \rangle \cup P^*$ in any model that induces $\mathcal{D}$.*

A powerful way of solving μ-transportability problems invokes the principle of decomposition, as formulated in the next theorem:

Theorem 1. *Given a set of domains $\Pi = \{\pi_1, \pi_2, ..., \pi_n\}$ characterized by selection diagrams $\mathcal{D} = \{D_1, D_2, ..., D_n\}$ relative to domain π^*, a relation $R(\pi^*)$ is μ-transportable if it can be decomposed into a set of subrelations of the form:*

$$R_k = P^*(\mathbf{V_k} | do(\mathbf{W_k}), \mathbf{Z_k}) \qquad k = 1, 2, \ldots, K$$

such that each R_k is uniquely computable from some $D_k \in \mathcal{D}$.

Proof. The proof is immediate, if the target relation is decomposable in such a way that each subrelation is computable from the available information, the quantity is μ-transportable by definition. □

Through successive decompositions, we can reduce the target relation to more elementary and easily recognizable μ-transportability problems as defined below:

Definition 4. *(Trivial μ-Transportability)*
A causal relation R is said to be trivially μ-transportable *from Π to π^*, if $R(\pi^*)$ is identifiable from the data available in π^* (i.e., (G^*, P^*)).*

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

Definition 5. *(Direct μ-Transportability)*
A causal relation R is said to be directly μ-transportable *from Π to π^*, if for some domain $\pi_i \in \Pi$, $R(\pi_i) = R(\pi^*)$.*

A graphical test for direct μ-transportability of $R = P^*(y|do(x), z)$ follows from do-calculus and reads: for some domain π_i, $(S \perp\!\!\!\perp Y | X, Z)_{D^{(i)}_{\overline{X}}}$; in words, X blocks all paths from S to Y once we remove all arrows pointing to X and condition on $\{X, Z\}$ in the selection diagram D_i. As an example, the Z-specific effects in Fig. 1 is the same in both domains, hence, it is directly μ-transportable. The same effects are directly μ-transportable in Fig. 2(b), but not in 2(a).

These two cases will act as a basis to decompose the problem of μ-transportability into smaller and more manageable subproblems based on Theorem 1.

So far we have been concerned with the positive cases of μ-transportability, but it is also important to understand when a certain quantity cannot be μ-transported. The following lemma provides an auxiliary tool to prove that a quantity is not μ-transportable based on the refutation of the uniqueness property required by Def. 3, which will show to be instrumental for proving completeness in the sequel.

Lemma 1. *Let $\mathbf{X}, \mathbf{Y}$ be two sets of disjoint variables in populations Π and π^*, and let $\mathcal{D} = \{D_1, ..., D_n\}$. $P^*_{\mathbf{x}}(\mathbf{y})$ is not μ-transportable from Π to π^* if there exist two models M^1 and M^2 compatible with $\mathcal{D}$ such that $P^i_{M_1}(\mathbf{V}) = P^i_{M_2}(\mathbf{V})$, $P^*_{M_1}(\mathbf{V}) = P^*_{M_2}(\mathbf{V})$, $P^i_{M_1}(\mathbf{V} \setminus \mathbf{W} | do(\mathbf{W})) = P^i_{M_2}(\mathbf{V} \setminus \mathbf{W} | do(\mathbf{W}))$, for all $\mathbf{W} \subset \mathbf{V}$, all families have positive distribution, and $P^*_{M_1}(\mathbf{y}|do(\mathbf{x})) \neq P^*_{M_2}(\mathbf{y}|do(\mathbf{x}))$.*

Proof. Let $\mathcal{I} = \bigcup_i I^i$, the collection of all interventional distributions $P^i(\mathbf{V} \setminus \mathbf{W} | do(\mathbf{W})) \in I^i$, for all $\mathbf{W} \subset \mathbf{V}$ in domain D_i, and similarly $\mathcal{P} = \bigcup_i P^i$ for the observational distributions in each domain. The latter inequality of the Lemma rules out the existence of a function from $\langle \mathcal{P}, \mathcal{I}, P^* \rangle$ to $P^*_{\mathbf{x}}(\mathbf{y})$. □

While the problems of transportability and μ-transportability are related, Lemma 1 indicates that proofs of non-μ-transportability are more involved than those of non-transportability. Indeed, to prove non-μ-transportability requires the construction of two models agreeing on $\langle \mathcal{P}, \mathcal{I}, P^* \rangle$ ($\mathcal{P}, \mathcal{I}$ represent the collection of observational and interventional distributions in all source domains), while non-transportability requires the two models to agree solely on the distributions $\langle P^j, I^j, P^* \rangle$ of a domain π_j.

4 Deriving μ-Transportability

We provide examples where it is feasible and infeasible to μ-transport certain relations, and this will help to build our intuition to study more formal and general conditions for μ-transportability.

Let us consider again the example in Fig. 2.. Recall that our goal is to establish whether $R(\pi^*) = P^*(y|do(x))$ is μ-transportable from $\{\pi_a, \pi_b\}$ to π^*. Note that in this case, if we wish to directly transport R from each place this is not allowed since we can rewrite $P^*(y|do(x)) = P(y|do(x), S)$, and the condition for directly μ-transport this relation does not hold, i.e., it is not true that $(S \perp\!\!\!\perp Y|X)_{D^{i}_{\overline{X}}}$ in π_i (for π_a and π_b)[5]. Now we decompose the target relation as follows (possible implementation of Theorem 1):

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

where the first line we condition on Z, the second line follows by rule 2 of do-calculus since $(Z \perp\!\!\!\perp Y|X)_{D_{\overline{X}\underline{Z}}}$ holds, and the third line follows from rule 3 of the do-calculus since $(X \perp\!\!\!\perp Y|Z)_{D_{\overline{X,Z}}}$ holds, where D is the causal diagram of π^* (despite the S-nodes).

Finally, we can now directly transport each of these pieces individually from the source domains noticing that $P^*(y|do(z))$ is directly μ-transportable from π_b giving that $(S \perp\!\!\!\perp Y|Z)_{D^{(b)}_{\underline{Z}}}$, and $P^*(z|do(x))$ is directly μ-transportable from π_a given that $(S \perp\!\!\!\perp Z|X)_{D^{(a)}_{\overline{X}}}$. These yields, respectively, $P^*(y|do(z)) = P^{(b)}(y|do(z))$ and $P^*(z|do(x)) = P^{(a)}(z|do(x))$, and therefore the target relation can be rewritten as,

$$R(\pi^*) = \sum_z P^{(a)}(z|do(x)) P^{(b)}(y|do(z))$$

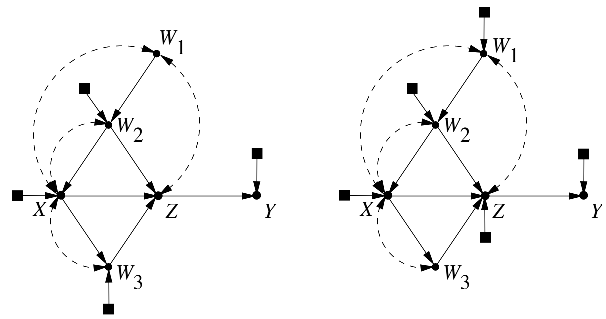
For a somewhat more challenging example, consider the selection diagrams in Fig. 3, and the task of deciding whether there exists an unbiased estimand for the relation $R(\pi^*) = P^*(y|do(x))$. It is not difficult to show that $R(\pi^*)$ is not (separately) transportable from the domains π_a and π_b, however, it turns out that this relation is μ-transportable from the domains when treated in conjunction. A more involved analysis is required in this case though, which yields the following μ-transport formula for $R(\pi^*)$ [1,6] :

$$\sum_{w_1, w_2, w_3, z} P^*(y|z) P^{(a)}_{x,w_2,w_3}(w_1, z) P^*(w_2|w_1) P^{(b)}_{x,w_1,w_2}(w_3) \tag{1}$$

In this case we have a witness showing that $R(\pi^*)$ is μ-transportable from the combination of the two sources together with the target domain, but the question arises how to perform a systematic decomposition guided by a guarantee that when it fails, there is no alternative way to decompose the target relation $R(\pi^*)$ in order to μ-transport it.

[5]Indeed, the impossibility follows from the completeness of the do-calculus for ordinary transportability.

[6]We show explicitly in Section 6 how the proposed algorithm entails this μ-transport formula.



(a) (b)

Figure 3: Selection diagrams illustrating a more involved analysis that yields an estimand (Eq. (1)) for the target quantity which combines information from three domains, the two sources π_a and π_b together with the target π^*.

Consider now Fig. 1(b) (called the "s-bow arc") which is known to be the smallest possible graph where $R(\pi^*) = P^*(y|do(x))$ is not transportable (Theorem 2 in (Bareinboim and Pearl, 2012)). This structure can be trivially extended to the μ-transportability case assuming that there are two domains with identical selection diagrams. It is obvious that $R(\pi^*)$ cannot be obtained from the available data; note that there is no possible alternative decomposition for R, and R is neither trivially nor directly μ-transportable from any of the domains. The reduction of μ-transportability to a transportability test can be justified for any causal relation and collection of domains where (1) the selection diagrams coincide and (2) the target quantity is not pairwise-transportable, which implies that the target relation is also not μ-transportable.

This, however, does not exhaust the possible cases of non-μ-transportability. Consider Fig. 4 in which the source domains do not share selection diagrams and the target quantity is $R(\pi^*) = P^*(y|do(x))$. If an oracle claims that $R(\pi^*)$ is not μ-transportable, it is still not trivial to show that this claim is true. Formally, we need to display two models M_1, M_2 such that the following relations hold (by Lemma 1):

$$\begin{cases} P^{(i)}_{M_1}(X, Z, Y) = P^{(i)}_{M_2}(X, Z, Y), \\ P^{(i)}_{M_1}(X, Y|do(Z)) = P^{(i)}_{M_2}(X, Y|do(Z)), \\ P^{(i)}_{M_1}(X, Z|do(Y)) = P^{(i)}_{M_2}(X, Z|do(Y)), \\ P^{(i)}_{M_1}(Y|do(X), do(Z)) = P^{(i)}_{M_2}(Y|do(X), do(Z)), \\ P^*_{M_1}(X, Z, Y) = P^*_{M_2}(X, Z, Y), \end{cases}$$

for $i = \{a, b\}$ and all values of X, Y, Z, and also,

$$P^*_{M_1}(Y|do(X)) \neq P^*_{M_2}(Y|do(X)),$$

for some value of X and Y.

We formally show how to construct such a certificate, which will be instrumental for the general proof of completeness for any collection of selection diagrams. See the proofs in (Bareinboim and Pearl, 2013b).

Theorem 2. *$P^*_x(y)$ is not μ-transportable in the selection diagrams in Fig. 4.*

5 Characterizing μ-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 several 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 for μ-transportability.[7]

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

The existence of a sC-component does not prevent μ-transportability between domains by itself. For instance, the graphs in Fig. 2 are single sC-components but the effects of interest are μ-transportable there.

Consider now a more specific type of sC-component carrying three extra features that prevents the μ-transportability of the causal effects in more general cases: the structure is closed under ancestral (i.e., consist of ancestral sets), there exists an S-node pointing to some of its elements, and it is a forest.

Definition 7 (sC-forest)**.** *Let D be a selection diagram, where $\mathbf{Y}$ is the maximal root set. Then D is a $\mathbf{Y}$-rooted sC-forest if D is a sC-component, all observable nodes have at most one child, and there is a selection node pointing to some vertex of D.*

The sC-forest are special structures and we introduce below a composite structure that witnesses non-μ-transportability characterized by a pair of sC-forests. μ-Transportability will be shown impossible whenever such structure exists as an edge subgraph of the inputted collection of selection diagrams.[8]

[7] Departing from results given in (Spirtes, Glymour, and Scheines, 1993, Galles and Pearl, 1995, Pearl and Robins, 1995, Halpern, 1998, Kuroki and Miyakawa, 1999), the advent of C-components complements the notion of *inducing path*, which was earlier introduced in (Verma and Pearl, 1990), and opened the path for several observations that culminated in the breakthrough result proving completeness of the *do-calculus* for non-parametric identification of causal effects by (Huang and Valtorta, 2006, Shpitser and Pearl, 2006).

[8] μs-hedges generalize s-hedges (Bareinboim and Pearl, 2012), which are based on hedges (Shpitser and Pearl, 2006).

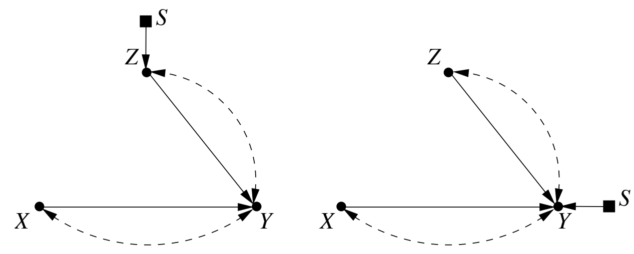


Figure 4: Collection of heterogeneous selection diagrams in which the target relation $P^*(y|do(x))$ is not μ-transportable from both domains (see Theorem 2).

Definition 8 (μs-hedge)**.** *Let $\mathbf{X}, \mathbf{Y}$ be two sets of variables in domain π^*, and let $\mathcal{D} = \{D_1, ..., D_n\}$ be a collection of selection diagrams relative to π^* and the domains $\Pi = \{\pi_1, ..., \pi_n\}$. Let F, F' be $\mathbf{R}$-rooted sC-forests such that $F \cap \mathbf{X} \neq 0$, $F' \cap \mathbf{X} = 0$, $F' \subseteq F$, $\mathbf{R} \subset An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$ in every $D_i \in \mathcal{D}$. Then F and F' form a μs-hedge for $P^*_{\mathbf{x}}(\mathbf{y})$ in $\mathcal{D}$.*

We can see a μs-hedge as a growing sC-forest F', which does not intersect $\mathbf{X}$, to a larger sC-forest F that do intersect $\mathbf{X}$ in all source domains at the same time. For instance, in Fig. 1(b), the sC-forests $F' = \{Y\}$, and $F = F' \cup \{X\}$ form a μs-hedge to $P_x(y)$; in Fig. 4, the sC-forests $F' = \{Z, Y\}$, and $F = F' \cup \{X\}$ form a μs-hedge to $P_x(y)$. Remarkably, there are similar sC-forests relative to the effects $P_x(y)$ in Fig. 3, but there is no a μs-hedge since the structures are not shared across domains.

Finally, we state below the formal connection between μs-hedges and non-μ-transportability.

Theorem 3. *Assume there exist a pair of sC-forests F, F' that form a μs-hedge for $P^*_{\mathbf{x}}(\mathbf{y})$ in $\mathcal{D}$. Then $P^*_{\mathbf{x}}(\mathbf{y})$ is not transportable from Π to π^*.*

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

6 Complete Algorithm for Deciding μ-Transportability of Joint Effects

We construct an algorithm for deciding μ-transportability called **μsID** (see Fig. 5), which extends previous analysis of transportability given in (Bareinboim and Pearl, 2012). We build on three particular observations articulated along the paper:

1. Decomposability (Theorem 1): The target relation can be broken into smaller subexpressions.

2. μ-Transportability (Def. 3-5): Causal relations can be partitioned into trivially and directly μ-transportable.

3. Non-μ-transportability (Theorem 3): The existence of a μs-hedge as an edge subgraph of the inputted diagrams can be used to prove non-μ-transportability.

The algorithm μsID simplifies and decomposes the inputted selection diagrams $\mathcal{D}$ (which is also a causal diagram over π^* when we disregard the S-nodes), partitioning the original problem into smaller blocks until either the entire expression is μ-transportable, or it runs into the problematic μs-hedge structure. For each block, $\boldsymbol{\mu}$**sID** tries to directly or trivially μ-transport it. Whenever $\boldsymbol{\mu}$**sID** exhausts the possibility of applying these operations, it exits with failure with a counterexample for μ-transportability – i.e., the graph local to the faulty call witnesses the non-μ-transportability of the query given that it will contain a μs-hedge as edge subgraph of the inputted diagrams.

Before showing the more formal properties of $\boldsymbol{\mu}$**sID**, we demonstrate how it works through the μ-transportability of $Q = P^*(y|do(x))$ in Fig. 3.

Since $D = An(Y)$ and $\mathcal{C}(D \setminus \{X\}) = \{C_0, C_1, C_2, C_3\}$, where $C_0 = D(\{W_1, Z\})$, $C_1 = D(\{W_2\})$, $C_2 = D(\{W_3\})$, and $C_3 = D(\{Y\})$, we invoke line 4 and try to μ-transport individually $Q_0 = P^*_{x,w_2,w_3,y}(w_1, z)$, $Q_1 = P^*_{x,w_1,w_3,z,y}(w_2)$, $Q_2 = P^*_{x,w_1,w_3,z,y}(w_2)$, and $Q_3 = P^*_{x,z,w_1,w_2,w_3}(y)$. Thus the original problem reduces in trying to evaluate the equivalent expression $\sum_{z,w_1,w_2,w_3} P^*_{x,w_2,w_3,y}(w_1,z)P^*_{x,w_1,w_3,z,y}(w_2)$ $P^*_{x,w_1,w_2,z,y}(w_3)P^*_{x,z,w_1,w_2,w_3}(y)$.

First, $\boldsymbol{\mu}$**sID** evaluates the expression Q_0 and triggers line 2, noting that the nodes that are not ancestors of $\{W_1, Z\}$ can be ignored. This implies that $P^*_{x,w_2,w_3,y}(w_1, z) = P^*_{x,w_2,w_3}(w_1, z)$ with induced subgraph $G_0 = G_{x,z,w_1,w_2,w_3}$. $\boldsymbol{\mu}$**sID** goes to line 5, in which in the local call $\mathcal{C}(D \setminus \{X, W_2, W_3\}) = \{W_1 \leftarrow\!\!-\!-\!-\!\!\rightarrow Z\}$. In this call, ordinary identifiability would fail, but $\boldsymbol{\mu}$**sID** proceeds to line 6 testing whether $(\mathbf{S} \perp\!\!\!\perp W_1, Z|X, W_2, W_3)_{D^{(i)}_{\overline{X, W_2, W_3}}}$ in some domain π_i. The test comes false in π_b, but true for π_a, which makes $\boldsymbol{\mu}$**sID** to directly μ-transport Q_0 with data from the experimental domain π_a, i.e., $P^*_x(z) = P^{(a)}_{x,w_2,w_3}(w_1, z)$.

Second, $\boldsymbol{\mu}$**sID** evaluates the expression Q_1 triggering line 2, which implies that $P^*_{x,w_1,w_3,z,y}(w_2) = P^*_{w_1}(w_2)$ with induced subgraph $G_1 = \{W_1 \rightarrow W_2\}$. $\boldsymbol{\mu}$**sID** goes to line 5, in which in the local call $\mathcal{C}(D\setminus\{W_1\}) = \{W_2\}$.

function $\boldsymbol{\mu}$**sID**$(\mathbf{y}, \mathbf{x}, P^*, \mathcal{I}, \mathcal{D})$
INPUT: $\mathbf{x}, \mathbf{y}$: value assignments; P^*: observational distribution in the target π^*; $\mathcal{I}$: collection of all interventional distributions in the sources Π; $\mathcal{D}$: collection of selection diagram; D: causal diagram in π^* (despite S-nodes); $\mathbf{S}$: set of selection nodes (per domain).
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 μsID$(\mathbf{y}, \mathbf{x} \cap An(\mathbf{Y})_D, \sum_{\mathbf{V}\setminus An(\mathbf{Y})_D} P^*, An(\mathbf{Y})_D)$

3 Set $\mathbf{W} = (\mathbf{V} \setminus \mathbf{X}) \setminus An(\mathbf{Y})_{D_{\overline{\mathbf{X}}}}$.
if $\mathbf{W} \neq \emptyset$, **return** $\boldsymbol{\mu}$**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 $\sum_{\mathbf{V}\setminus\{\mathbf{Y},\mathbf{X}\}} \prod_i \boldsymbol{\mu}$**sID**$(c_i, \mathbf{V} \setminus c_i, P^*, D)$

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

6 **for** $i = 1, ..., |\mathcal{D}|$:

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

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

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

10 **if** $(\exists C')C_0 \subset C' \in \mathcal{C}(D)$, **return** $\boldsymbol{\mu}$**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 5: Modified version of transportability algorithm capable of recognizing μ-transportable relations.

Thus it proceeds to line 6 testing $(\mathbf{S} \perp\!\!\!\perp W_2|W_1)_{D_{\overline{W_1}}}$. The test comes false, which makes $\boldsymbol{\mu}$**sID** move to line 7. The test again returns false since $\mathcal{C}(G_1) \neq \{G_1\}$. Finally, line 8 evaluates true since $\mathcal{C}(\{W_2\}) \in \mathcal{C}(G_1)$; $\boldsymbol{\mu}$**sID** trivially μ-transport Q_1 using the data from the target domain π^*, i.e., $P^*_{w_1}(w_2) = P^*(w_2|w_1)$.

Third, evaluating the expression Q_2, $\boldsymbol{\mu}$**sID** goes to line 2, which implies that $P^*_{x,w_1,w_2,z,y}(w_3) = P^*_{x,w_1,w_2}(w_3)$ with induced subgraph $G_2 = G_{X,W_1,W_2,W_3}$. $\boldsymbol{\mu}$**sID** goes to line 5, and in this local call $\mathcal{C}(D \setminus \{X, W_1, W_2\}) = \{W_3\}$. Note that ordinary identifiability would fail here, but $\boldsymbol{\mu}$**sID** proceeds to line 6 testing whether $(\mathbf{S} \perp\!\!\!\perp W_3|X, W_1, W_2)_{D^{(i)}_{\overline{X, W_1, W_2}}}$ in some domain π_i. The test comes false for ordinary transportability in π_a, but positive in π_b, which allows the direct μ-transportability of Q_2 using data from the experimental domain π_b (i.e., $P^*_{x,w_1,w_2}(w_3) = P^{(b)}_{x,w_1,w_2}(w_3)$).

Forth, $\boldsymbol{\mu}$**sID** evaluates the expression Q_3 and triggers line 5, $\mathcal{C}(D \setminus \{X, W_1, W_2, W_3, Z\}) = \{Y\}$. In turn, both tests at lines 6 and 7 fail, but the test in line 8 succeed since $\mathcal{C}(\{Y\}) \in \mathcal{C}(D)$. Then, $\boldsymbol{\mu}$**sID** trivially μ-transport Q_3 using the data from the target domain π^*, i.e., $P^*_{x,w_1,w_2,w_3,z}(y) = P^*(y|z)$ (after simplification). The composition of the return of these calls generates the expression provided in Section 4.

Finally, we prove next soundness and completeness of the algorithm $\boldsymbol{\mu}$**sID**.

Theorem 4 (soundness)**.** *Whenever $\boldsymbol{\mu}$**sID** returns an expression for $P^*_{\mathbf{x}}(\mathbf{y})$, it is correct.*

Theorem 5. *Assume $\boldsymbol{\mu}$**sID** fails to μ-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 $\boldsymbol{\mu}$**sID** contain as edge subgraphs sC-forests F, F' that form a μs-hedge for $P^*_{\mathbf{x}'}(\mathbf{y}')$.*

Corollary 1 (completeness)**.** *$\boldsymbol{\mu}$**sID** is complete.*

Next, we show that the graphical criterion of μs-hedge indeed characterizes μ-transportability.

Corollary 2. *$P^*_{\mathbf{x}}(\mathbf{y})$ is μ-transportable from Π to π^* in $\mathcal{D}$ if and only if there is not μs-hedge for $P_{\mathbf{x}'}(\mathbf{y}')$ in $\mathcal{D}$ for any $\mathbf{X}' \subseteq \mathbf{X}$ and $\mathbf{Y}' \subseteq \mathbf{Y}$.*

Furthermore, we show below that the do-calculus is complete for establishing μ-transportability, which means that failure in the exhaustive application of its rules implies the inexistence of a mapping from the available data to the target relation (i.e., there is no μ-transport formula), independently of the method used to obtain such mapping.

Corollary 3. *The rules of do-calculus, together with standard probability manipulations are complete for establishing μ-transportability of all causal effects of the form $P^*_{\mathbf{x}}(\mathbf{y})$.*

7 Conclusions

Informal discussions concerning the difficulties of generalizing results across populations have been going on for almost half a century (Cox, 1958, Campbell and Stanley, 1963, Heckman, 1992, Hotz, Imbens, and Mortimer, 2005, Manski, 2007) and appear to accompany every textbook in experimental design. By and large, these discussions have led to the obvious conclusions that researchers should be extremely cautious about unwarranted generalization, that many threats may await the unwary, and that extrapolation across studies requires "some understanding of the reasons for the differences" (Cox, 1958, p. 11).

This paper embeds these discussions in a precise mathematical language, and then provides researchers with theoretical guarantees that, if certain conditions can be ascertained, generalization across populations can be accomplished, protected from the threats and dangers that the informal literature has accumulated.

The paper goes further and provides a complete algorithm for computing the correct transport formula, namely, the proper way of modifying the experimental results so as to account for differences in the populations. These transport formulae enable the investigator to select the essential measurements in both the experimental and observational studies, and combine them into a bias-free estimand of the target quantity.

Of course, our analysis is based on the assumption that the analyst is in possession of sufficient background knowledge to determine, at least qualitatively, *where* the populations may differ from one another. In practice, such knowledge may only be partially available and, as is the case in every mathematical exercise, the benefit of the analysis lies primarily in understanding what knowledge is needed for the task to succeed and how sensitive conclusions are to knowledge that we do not possess.

This paper complements a recent work on a task called *z-transportability* (Bareinboim and Pearl, 2013a), which deals with transferring causal information when just limited experiments are available for use.

8 Acknowledgment

The authors would like to thank the reviewers for their comments that help improve the manuscript. This research was supported in parts by grants from NSF #IIS-1249822, and ONR #N00014-13-1-0153 and #N00014-10-1-0933.

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, 293 –301.

Bareinboim, E. and J. Pearl (2012): "Transportability of causal effects: Completeness results," in J. Hoffmann. and B. Selman, eds., *Proceedings of The Twenty-Sixth Conference on Artificial Intelligence (AAAI 2012)*, 698–704.

Bareinboim, E. and J. Pearl (2013a): "Causal transportability with limited experiments," in M. desJardins and M. L. Littman, eds., *Proceedings of The Twenty-Seventh Conference on Artificial Intelligence (AAAI 2013)*, in press.

Bareinboim, E. and J. Pearl (2013b): "Meta-transportability of causal effects: A formal approach," Technical Report R-407, Cognitive Systems Laboratory, Department of Computer Science, UCLA.

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

Cox, D. (1958): *The Planning of Experiments*, NY: John Wiley and Sons.

Galles, D. and J. Pearl (1995): "Testing identifiability of causal effects," in P. Besnard and S. Hanks, eds., *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI 1995)*, San Francisco: Morgan Kaufmann, 185–195.

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

Halpern, J. (1998): "Axiomatizing causal reasoning," in G. Cooper and S. Moral, eds., *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI 1998)*, San Francisco, CA: Morgan Kaufmann, 202–210, also, *Journal of Artificial Intelligence Research* 12:3, 17–37, 2000.

Heckman, J. (1992): "Randomization and social policy evaluation," in C. Manski and I. Garfinkle, eds., *Evaluations: Welfare and Training Programs*, Cambridge, MA: Harvard University Press, 201–230.

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

Höfler, M., A. Gloster, and J. Hoyer (2010): "Causal effects in psychotherapy: Counterfactuals counteract overgeneralization," *Psychotherapy Research*, DOI: 10.1080/10503307.2010.501041.

Hotz, V. J., G. Imbens, and J. H. Mortimer (2005): "Predicting the efficacy of future training programs using past experiences at other locations," *Journal of Econometrics*, 125, 241–270.

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

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

Kuroki, M. and M. Miyakawa (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.

Morgan, S. and C. Winship (2007): *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research)*, Cambridge University Press, New York, NY.

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

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

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

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

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

Pearl, J. and J. Robins (1995): "Probabilistic evaluation of sequential plans from causal models with hidden variables," in P. Besnard and S. Hanks, eds., *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI 1995)*, San Francisco: Morgan Kaufmann, 444–453.

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

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

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

Spirtes, P., C. Glymour, and R. Scheines (1993): *Causation, Prediction, and Search*, New York: Springer-Verlag.

Spirtes, P., C. Glymour, and R. Scheines (2000): *Causation, Prediction, and Search*, Cambridge, MA: MIT Press, 2nd edition.

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

Verma, T. and J. Pearl (1990): "Equivalence and synthesis of causal models," in *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence (UAI 1990)*, Cambridge, MA, 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.