Transportability and Data Fusion in Causal Inference

Junyi Liao

1 Transportability

How to transfer causal knowledge across heterogeneous domains?

Domain discrepancy. Let π and π^* be domains associated, respectively, with SCMs $\mathcal{M}$ and $\mathcal{M}^*$ compatible with a causal DAG $\mathcal{G}$. We denote $\Delta \subseteq \mathbf{V}$ as a set of variables such that, for each $V \in \Delta$, there might exist a discrepancy; either $f_V \neq f_V^*$, or $p(\mathbf{U}_V) \neq p^*(\mathbf{U}_V)$.

Selection diagram. Given the discrepancies Δ with respect to graph $\mathcal{G} = \langle \mathbf{V}, \mathbf{E} \rangle$, let $\mathbf{S} = \{S_V : V \in \Delta\}$ be the set of selection variables. Then, the selection diagram $\mathcal{D}$ is defined as

$$\mathcal{D} := \langle \mathbf{V} \cup \mathbf{S}, \mathbf{E} \cup \{S_V \to V\}_{V:S_V \in \mathbf{S}} \rangle.$$

I list the types of transportability below.

- **Transportability.** From one domain (experiment) to another (observation)
- **Z-Transportability.** Experiments in source domain are limited.
- **Meta-Transportability.** Across multiple domains.
- **Mz-Transportability.** Across multiple domains, with limited experiments.
- **G-Transportability.** General Case.
- **Soft g-transportability.** From atomic interventions to soft interventions.

1.1 Transportability

Definition (Transportability). Given two domains π and π^* characterized by shared causal DAG $\mathcal{G}$ and probability distributions p and p^*, respectively. A causal relation R is said to be transportable from π to π^* if $R(\pi)$ is estimable from the set I of interventions on π, and $R(\pi^*)$ is identified from p, p^*, I in any model that induces $\mathcal{D}$.

From this definition, given a causal query $p^*(\mathbf{y}|\mathrm{do}(\mathbf{x},)\mathbf{z})$, it can be expressed in terms of interventional distributions on π (p with do-operators) and observational distributions on both π and π^* (p and p^* without do-operators). An equivalent statement is presented below.

Theorem 1 (Pearl and Bareinboim, 2011). Let $\mathcal{D}$ be the selection diagram characterizing π and π^*, and $\mathbf{S}$ a set of selection variables in $\mathcal{D}$. The relation $R = p^*(\mathbf{y}|\mathrm{do}(\mathbf{x}), \mathbf{z})$ is transportable from π to π^* if and only if the expression $p(\mathbf{y}|\mathrm{do}(\mathbf{x}), \mathbf{z}, \mathbf{s})$ 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.

A more convenient kind of transportability is given below.

Definition (Direct transportability). A causal relation R is said to be directly transportable from a domain π to another π^*, if $R(\pi) = R(\pi^*)$.

The direct transportability can be recognized via S-admissible set. A variable set $\mathbf{Z}$ is said to be S-admissible with respect to the causal effect of $\mathbf{X}$ on $\mathbf{Y}$, if

$$(\mathbf{Y} \perp\!\!\!\perp \mathbf{S} \mid \mathbf{Z})_{\mathcal{G}_{\overline{\mathbf{X}}}}.$$

Theorem 2 (Pearl and Bareinboim, 2011). For two population π and π^*, suppose we have the selection diagram $\mathcal{D}$ with selection variable set $\mathbf{S}$. The stratum-specific causal effect $p^*(\mathbf{y}|\text{do}(\mathbf{x}), \mathbf{z})$ is transportable from π to π^* if $\mathbf{Z}$ is an S-admissible set with respect to the causal effect of $\mathbf{X}$ on $\mathbf{Y}$.

1.2 Z-Transportability

Definition (z-Transportability). Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be sets of disjoint variables and $\mathcal{D}$ a selection diagram relative to domians $\langle \pi, \pi^* \rangle$. Denote the available interventional distributions available in π as

$$I_z = \bigcup_{\mathbf{Z}' \subseteq \mathbf{Z}} p(\mathbf{v}|\text{do}(\mathbf{z}))$$

A causal relation R is said to be transportable from π to π^* if $R(\pi) = p^*(\mathbf{y}|\mathbf{do}(\mathbf{x}))$ is z-transportable from π to π^* if it is identified from p, p^*, I_z in any model that induces $\mathcal{D}$.

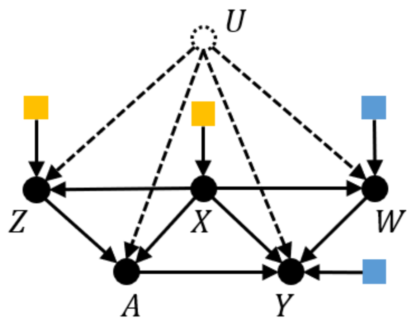
Compared with the transportability discussed in the previous subsection, z-transportability restricts the interventions on variable set $\mathbf{Z}$. Similarly, we have the following theorem.

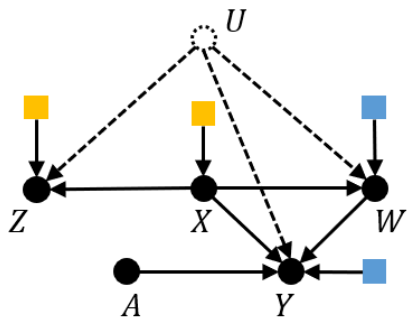
Theorem 3 (Bareinboim and Pearl, 2013a). For two population π and π^*, suppose we have the selection diagram $\mathcal{D}$ with selection variable set $\mathbf{S}$. The relation $R = p^*(\mathbf{y}|\text{do}(\mathbf{x}))$ is z-transportable from π to π^* if and only if the expression $p(\mathbf{y}|\text{do}(\mathbf{x}), \mathbf{s})$ is reducible, using the rules of do-calculus, to an expression in which all do-operators are applied to subsets of $\mathbf{Z}$, and the S-variables are separated from these do-operators.

1.3 Meta-Transportability

Transportability techniques are particularly valuable in situations that allow to combine empirical knowledge from multiple source domains. While transportability is not realizable in any single diagram, meta-transportability (or μ-transportability, for short) can be feasible.

Definition (μ-transportability). Let $\mathscr{D} = \{\mathcal{D}_1, \cdots, \mathcal{D}_n\}$ be a collection of selection diagrams relative to source domains $\Pi = \{\pi_1, \cdots, \pi_n\}$ 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 $R = p(\mathbf{y}|\text{do}(\mathbf{x}))$ is said to be μ-transportable from Π to π^* in $\mathscr{D}$ if $p^*(\mathbf{y}|\text{do}(\mathbf{x}))$ is uniquely computable from $\bigcup_{i \in [n]} \langle p^{(i)}, I^{(i)} \rangle \cup p^*$ in any model that induces $\mathscr{D}$.





A proxy variable example. Consider the diagram shown above. The left hand side shows the causal diagram $\mathcal{G}$ with selection variables $\mathbf{S}_1 = \{S_W, S_Y\}$ (blue squared nodes) relative to $\langle \pi_1, \pi^* \rangle$ and $\mathbf{S}_2 = \{S_Z, S_X\}$ (gold squared nodes) relative to $\langle \pi_2, \pi^* \rangle$. We want to estimate $p^*(y|\text{do}(a), x)$, which is not directly transportable from both π_1 and π_2. However, $p^*(y|\text{do}(a), x)$ is μ-transportable:

$$p^*(y|\text{do}(a), x) = \sum_{z \in \mathcal{Z}} p^*(y|\text{do}(a), z, x) p^*(z|\text{do}(a), x) \tag{1}$$

$$= \sum_{z \in \mathcal{Z}} p^*(y|\text{do}(a), z, x) p^*(z|x) \tag{2}$$

$$= \sum_{z \in \mathcal{Z}} p^{(2)}(y|\text{do}(a), z, x) p^{(1)}(z|x), \tag{3}$$

The equality (1) is the law of total probability, the equality (2) is follows from $(Z \perp\!\!\!\perp A \mid X)_{\mathcal{G}_{\overline{A(X)}}}$, and the equality (3) follows from $(Z \perp\!\!\!\perp S_W, S_Y \mid X)_{\mathcal{G}_{\overline{A}}}$, and $(Y \perp\!\!\!\perp S_Z, S_X \mid Z, X)_{\mathcal{G}_{\overline{A}}}$.

The principle of decomposition can be applied to solve the μ-transportability problems:

Theorem 4 (Bareinboim and Pearl, 2013b). Given a set of domains $\Pi = \{\pi_1, \cdots, \pi_n\}$ and target domain π^* characterized by selection diagrams $\mathscr{D} = \{\mathcal{D}_1, \cdots, \mathcal{D}_n\}$ relative to target domain π^*. A relation $R(\pi^*)$ is μ-transportable from Π to π^* if it can be decomposed into a set of subrelations of the form:

$$R_k = p^*(\mathbf{V}_k|\text{do}(\mathbf{W}_k), \mathbf{Z}_k) \quad k = 1, \cdots, K$$

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

This decomposition is used in the algorithm for deriving μ-transportability. Moreover, Bareinboim and Pearl (2013) proposed a graphical criterion (μs-hedge) to characterize non-μ-transportability in collection of selection diagrams.

1.4 Mz-transportability

This definition combines z-transportability and μ-transportability.

Definition (mz-Transportability). Let $\mathscr{D} = \{\mathcal{D}_1, \cdots, \mathcal{D}_n\}$ be a collection of selection diagrams relative to source domains $\Pi = \{\pi_1, \cdots, \pi_n\}$, and target domain π^*, respectively, and $\mathbf{Z}_i$ (and $\mathbf{Z}^*$) be the variables in which experiments can be conducted in domain π_i (and π^*). Let $\langle p^{(i)}, I_z^i \rangle$ be the pair of observational and interventional distributions of π_i, where $I_z^i = \bigcup_{\mathbf{Z}' \subseteq \mathbf{Z}_i} p^{(i)}(\mathbf{v}|\text{do}(\mathbf{z}'))$, and in an analogous manner, $\langle p^*, I_z^* \rangle$ be the observational and interventional distributions of π^*. The causal effect $R = p^*(\mathbf{y}|\text{do}(\mathbf{x}))$ is said to be *mz*-transportable from Π to π^* in $\mathscr{D}$ if $p^*(\mathbf{y}|\text{do}(\mathbf{x}))$ is uniquely computable from $\bigcup_{i=1}^{n}\langle p^{(i)}, I_z^i\rangle \cup \langle p^*, I_i^*\rangle$ in any model that induces $\mathscr{D}$.

Analogously, $p^*(\mathbf{y}|\text{do}(\mathbf{x}))$ is *mz*-transportable from Π to π^* if the expression $p(\mathbf{y}|\text{do}(\mathbf{x}), \mathbf{s}_1, \cdots, \mathbf{s}_n)$ is reducible, using the rules of do-calculus, to an expression in which (1) do-operators that apply to subsets of I_z^i have no $\mathbf{S}_i$-variables or (2) do-operators apply only to subsets of I_z^*.

Algorithm. Bareinboim and Pearl (2014) derived the algorithm for deriving *mz*-transportability.

- **Eliminate redundant do-operators.** If $\mathbf{V}\backslash An(\mathbf{Y})_{\mathcal{G}} \neq \emptyset$, then

$$p(\mathbf{y}|\text{do}(\mathbf{x})) = p(\mathbf{y}|\text{do}(\mathbf{x} \cap An(Y)_{\mathcal{G}}), \mathbf{x}\backslash An(Y)_{\mathcal{G}})$$

  by Rule 1 of do-calculus. Then $\mathbf{x} \leftarrow \mathbf{x} \cap An(Y)_{\mathcal{G}}$, $\mathcal{G} \leftarrow \mathcal{G}(An(\mathbf{Y}))$, and the conditional term can be averaged out using the observational distributions $\sum_{\mathbf{V}\backslash An(Y)_{\mathcal{G}}} \mathcal{P}$.

- **Introduce other experiments.** If $\mathbf{W} = (\mathbf{V}\backslash\mathbf{X})\backslash An(\mathbf{Y})_{\mathcal{G}_{\overline{\mathbf{X}}}} \neq \emptyset$, then

$$p(\mathbf{y}|\text{do}(\mathbf{x})) = p(\mathbf{y}|\text{do}(\mathbf{x}, \mathbf{w}))$$

  by Rule 3 of do-calculus. Then $\mathbf{x} \leftarrow (\mathbf{x}, \mathbf{w})$.

- **Decompose the query by C-components.** If $\mathcal{C}(\mathcal{G}\backslash\mathbf{X}) = \{C_0, C_1, \cdots, C_k\}$, then

$$p(\mathbf{y}|\text{do}(\mathbf{x})) = \sum_{\mathbf{V}\backslash\{\mathbf{Y},\mathbf{X}\}} \prod_{i=0}^{k} Q[C_i].$$

  For each $Q[C_j]$, $\mathbf{X} \leftarrow \mathbf{V}\backslash C_j$, $\mathbf{Y} \leftarrow C_j$.

- **Identification.** Transfer each $Q[C_j]$ to some $Q^{\pi_i}[C_j]$ if $(\mathbf{S}_i \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{X})_{\mathcal{G}^{(i)}_{\overline{\mathbf{X}}}}$ and $\mathbf{Z}_i \cup \mathbf{X} \neq \emptyset$. Then, identify each $Q^{\pi_i}[C_j]$ in domain π_i.

1.5 G-Transportability

This definition generalizes the *mz*-transportability. In the previous discussion, both $z-$ and $mz-$transportability borrows the concept of *z*-identifiability, which indicates that each domain is associated with experiments based on every subset of manipulable variables $\mathbf{Z} \subseteq \mathbf{V}$. In g-transportability, experiments can be conducted on an arbitrary collection of subsets of $\mathbf{V}$.

We use an alternative definition of selection diagram. Given a collection of discrepancies $\boldsymbol{\Delta} = \{\Delta_1, \cdots, \Delta_n\}$ with respect to graph $\mathcal{G} = \langle \mathbf{V}, \mathbf{E}\rangle$, let $\mathbf{S} = \{S_V : V \in \bigcup_{i\in[n]} \Delta_i\}$ be the set of selection variables. Then, the selection diagram $\mathcal{G}^{\boldsymbol{\Delta}}$ is defined as

$$\mathcal{G}^{\boldsymbol{\Delta}} := \langle \mathbf{V} \cup \mathbf{S}, \mathbf{E} \cup \{S_V \to V\}_{V:S_V\in\mathbf{S}}\rangle.$$

We denote the domain specific selection variable set by $\mathbf{S}^{(i)} = \{S_V : V \in \Delta_i\}$, and the rest by $\mathbf{S}^{(-i)} = \mathbf{S}\backslash\mathbf{S}^{(i)}$. Selection variables work like switches selecting the domain of interest. The state space of $S_V \in \mathbf{S}$ is the index set $\{0\} \cup \{i : V \in \Delta_i\}$. Hence, a selection diagram can be viewed as the causal diagram for a unifying SCM representing heterogeneous SCMs where

$$p_{\mathbf{x}}(\mathbf{y}|\mathbf{w}, \mathbf{S}^{(i)} = i, \mathbf{S}^{(-i)} = 0) = p_{\mathbf{x}}^{(i)}(\mathbf{y}|\mathbf{w}).$$

Definition (g-transportability). Let $\mathcal{G}^{\mathbf{\Delta}}$ be a selection diagram relative to source domains $\Pi = \{\pi_1, \cdots, \pi_n\}$ and a target domain π^*. Let $\mathscr{Z} = \{\mathcal{Z}^{(i)}\}_{i=1}^n$ be a specification of available experiments, where $\mathcal{Z}^{(i)}$ is the collection of sets of variables for π_i in which experiments on each set of $\mathbf{Z} \in \mathcal{Z}^{(i)}$ can be conducted. Given disjoint sets of variables $\mathbf{X}, \mathbf{Y}$ and $\mathbf{W}$, the conditional causal effect $p_{\mathbf{x}}(\mathbf{y}|\mathbf{w})$ is said to be g-transportable with respect to $\langle \mathcal{G}^{\mathbf{\Delta}}, \mathscr{Z} \rangle$ if $p_{\mathbf{x}}(\mathbf{y}|\mathbf{w})$ is uniquely computable from $\mathcal{P}^{\Pi}_{\mathscr{Z}} = \{p^{(i)}_{\mathbf{z}} \mid \mathbf{Z} \in \mathcal{Z}^{(i)}, \mathcal{Z}^{(i)} \in \mathscr{Z}\}$.

Theorem 5 (Lee, Correa and Bareinboim, 2019). A causal effect $p^*_{\mathbf{x}}(\mathbf{y}|\mathbf{w})$ is g-transportable with respect to $\langle \mathcal{G}, \mathscr{Z} \rangle$, if the expression $p_{\mathbf{x}}(\mathbf{y}|\mathbf{w}, \mathbf{S})$ is reducible, using the rules of do-calculus, to an expression in which every term of the form $p_{\mathbf{z}}(\mathbf{b}|\mathbf{c}, \mathbf{S}')$ satisfies $\mathbf{Z} \in \mathcal{Z}^{(i)}$ for some domains $\pi_i \in \Pi$, and

$$(\mathbf{S} \backslash \mathbf{S}') \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C} \text{ in } \mathcal{G}^{\mathbf{\Delta}} \backslash \mathbf{Z},\ \mathbf{S}^{(i)} \cap \mathbf{S}' = \emptyset.$$

Lee et al. (2020) has proposed the conditions for g-transportability in both unconditional and conditional case. For the conditional case, they established the connection between a graphical structure called *s*-thicket and the non-g-transportability of a unconditional causal query, *e.g.* $p(\mathbf{y}|\mathrm{do}(\mathbf{x}))$. For the conditional case, they proved the equivalence between the g-transportability of $p(\mathbf{y}|\mathrm{do}(\mathbf{x}), \mathbf{w})$, and of $p(\mathbf{y}, \mathbf{w}|\mathrm{do}(\mathbf{x}))$. Hence, a conditional g-transportable causal query can be computed by

$$p(\mathbf{y}|\mathrm{do}(\mathbf{x}), \mathbf{w}) = \frac{p(\mathbf{y}, \mathbf{w}|\mathrm{do}(\mathbf{x}))}{\sum_{\mathbf{y}' \in \mathcal{Y}} p(\mathbf{y}', \mathbf{w}|\mathrm{do}(\mathbf{x}))}.$$

1.6 G-transportability of Soft Interventions

The discussions above focuses on atomic interventions represented by do-operators, while in real scenarios the interventions of interest can respond to a collection of variables in a stoachstic manner.

We consider four types of interventions summerized below. An intervention on variable X replace f_x with another function f'_x. In the conditional and stochastic cases, $\mathbf{Pa}'_x$ are not necessarily contained by $\mathbf{Pa}_x$ as long as they do not include any descendant of X. $\mathbf{U}'_x$ is different from $\mathbf{U}_x$ and $\mathbf{U}'_x \cap \mathbf{U}_x = \emptyset$.

Type	Strategy	Function
Idle	$\sigma_X = \emptyset$	$f'_x = f_x$
Atomic/do	$\sigma_X = x$	$f'_x = x$ for some $x \in \mathrm{Dom}(X)$
Conditional	$\sigma_X = g(\mathbf{Pa}'_x)$	$f'_x = g(\mathbf{Pa}'_x)$
Stochastic	$\sigma_X = p'(X|\mathbf{Pa}'_x)$	f'_x s.t. $p'(x|\mathbf{Pa}'_x) = \sum_{\mathbf{u}'_x} p(f'_x(\mathbf{Pa}_x, \mathbf{u}_x) = x) p(\mathbf{u}_x)$

Table 1: Summary of the types of interventions considered.

Given an intervention $\sigma_{\mathbf{X}}$, a new model can be defined as

$$\mathcal{M}_{\sigma_{\mathbf{X}}} = \langle \mathbf{U} \cup \mathbf{U}'_{\mathbf{x}}, \mathbf{V}, (\mathbf{F} \backslash \{f_x\}_{X \in \mathbf{X}}) \cup \{f'_x\}_{X \in \mathbf{X}}, p(\mathbf{U}, \mathbf{U}'_{\mathbf{x}}) \rangle,$$

which induces a probability distribution

$$p(\mathbf{v}; \sigma_{\mathbf{X}}) = \sum_{\mathbf{u}, \mathbf{u_x}'} \prod_{i: V_i \in \mathbf{X}} p(v_i|\mathbf{Pa}_{v_i}, \mathbf{u}_{v_i}, \mathbf{u}'_{v_i}; \sigma_{\mathbf{X}}) p(\mathbf{u_x}'; \sigma_{\mathbf{X}}) \prod_{i: V_i \in \mathbf{V} \backslash \mathbf{X}} p(v_i|\mathbf{Pa}_{v_i}, \mathbf{u}_{v_i}) p(\mathbf{u})$$

and a causal graph $\mathcal{G}_{\sigma_{\mathbf{X}}}$ which contains a node σ_X for each $X \in \mathcal{X}$ with an edge $(\sigma_X \to X)$.

Definition (Effect transportability). Let $\mathbf{Y}, \mathbf{X}, \mathbf{W} \subset \mathbf{V}$ be disjoint variable sets. The (conditional) effect of intervention $\sigma_{\mathbf{X}}$ on a set of outcome variables $\mathbf{Y}$, conditional on $\mathbf{W}$, $p^*(\mathbf{y}|\mathbf{w};\sigma_{\mathbf{X}})$, in a target domain π^*, is said to be transportable from $\langle \mathcal{G}^{\boldsymbol{\Delta}}, \mathscr{Z} \rangle$, if it is uniquely computable from the set of distribution $\mathscr{Z}$ for every assignment $(\mathbf{y}, \mathbf{w})$ and every set of models $\{\mathcal{M}^{(i)}\}_{\pi_i \in \Pi}$ inducing $\mathcal{G}^{\boldsymbol{\Delta}}$ and $\mathscr{Z}$.

Furthermore, Correa and Bareinboim (2022) proposed the conditions for g-transportability of soft interventions in unconditional and conditional cases.

Theorem 6. Let $\mathbf{Y}, \mathbf{X} \subseteq \mathbf{V}$ be any two set of variables, and let $\sigma^*_{\mathbf{X}}$ be an atomic, conditional or stochastic intervention. Then. the effect of $\sigma^*_{\mathbf{X}}$ on $\mathbf{Y}$ can be written as

$$p^*(\mathbf{y}|\sigma_{\mathbf{X}} = \sigma^*_{\mathbf{X}}) = \sum_{\mathbf{d}\backslash\mathbf{y}} p^*(\mathbf{d}\backslash\mathbf{x}; \sigma^*_{\mathbf{x}} = \mathbf{x}) \prod_{X \in \mathbf{X}\cap\mathbf{D}} p^*(x|\mathbf{Pa}_x; \sigma_{\mathbf{X}} = \sigma^*_{\mathbf{X}}).$$

Moreover, this effect is transportable from $\langle \mathcal{G}^{\boldsymbol{\Delta}}, \mathscr{Z} \rangle$ if and only if $p^*(\mathbf{d}\backslash\mathbf{x}; \sigma^*_{\mathbf{x}} = \mathbf{x})$ is transportable from $p^*(\mathbf{d}\backslash\mathbf{x}; \sigma^*_{\mathbf{x}} = \mathbf{x})$, where $\mathbf{D} = An(\mathbf{Y})_{\mathbf{G}_{\sigma_{\mathbf{X}}}}$.

This theorem reduces the problem of transporting $p^*(\mathbf{y}|\sigma_{\mathbf{X}} = \sigma^*_{\mathbf{X}})$ to that of transporting the effect of an atomic intervention.

Theorem 7. Let $\mathbf{Y}, \mathbf{X}, \mathbf{W} \subset \mathbf{V}$, $\mathbf{W} \cap \mathbf{Y} = \emptyset$, $\sigma_{\mathbf{X}}$ be any intervention, and $\mathcal{G}_{\sigma_{\mathbf{X}}}$ the corresponding interventional causal graph. Then, the effect of $\sigma_{\mathbf{X}}$ on $\mathbf{Y}$ conditioned on $\mathbf{W}$ is given by

$$p(\mathbf{y}|\mathbf{w};\sigma_{\mathbf{x}}) = p(\mathbf{y}|\mathbf{w}_{\mathbf{y}};\sigma_{\mathbf{x}}, \sigma_{\mathbf{w}_{\overline{\mathbf{y}}}} = \mathbf{w}_{\overline{\mathbf{y}}}) = \frac{\sum_{\mathbf{a}\backslash(\mathbf{y}\cup\mathbf{w}_{\mathbf{y}})} Q[\mathbf{A};\sigma_{\mathbf{X}}]}{\sum_{\mathbf{a}\backslash\mathbf{w}_{\mathbf{y}}} Q[\mathbf{A};\sigma_{\mathbf{X}}]},$$

where $\mathbf{W}_{\mathbf{y}} \subseteq \mathbf{W}$ is the set of variables connected to any $Y \in \mathbf{Y}$ by any path in $\mathcal{G}_{\sigma_{\mathbf{X}}[\mathbf{D}]\underline{\mathbf{W}}}$, with $\mathbf{D} = An(\mathbf{Y})_{\mathbf{G}_{\sigma_{\mathbf{X}}}}$, $\mathbf{W}_{\overline{y}} = \mathbf{W}\backslash\mathbf{W}_{\mathbf{y}}$, and $\mathbf{A} = An(\mathbf{Y} \cup \mathbf{W}_{\mathbf{y}})_{\mathcal{G}_{\sigma_{\mathbf{X}}\underline{\mathbf{W}}}}$. Furthermore, this effect is transportable from $\langle \mathcal{G}^{\boldsymbol{\Delta}}, \mathscr{Z} \rangle$ if and only if $Q[\mathbf{A};\sigma_{\mathbf{X}}]$ is transportable from $\langle \mathcal{G}^{\boldsymbol{\Delta}}, \mathscr{Z} \rangle$.

Then one can use the C-component decomposition and the rules of σ-calculus to derive the g-transportability for soft interventions.

2 Data Fusion

We discuss the methods to combine data collected from different sources, *e.g.* randomized experiments and observational studys.

Settings. We use the potential outcome framework. Suppose we have data collected from RCTs and observational studies. Each individual in the RCT or observational population is described by a random tuple $(X, Y(0), Y(1), A, S)$, where

- X is a p-dimensional vector of covariates,
- A is a dichotomous treatment assignment,
- $Y(a)$ is the potential outcome (also referred to as counterfactual outcome, which would be observed had the subject been given treatment assignment $A = a, a \in \{0, 1\}$), and
- S is an indicative dichotomous variable for trial eligibility and willingness to participate.

For RCT samples, $S = 1$ and for observational samples, S is unknown.

Suppose we have an RCT sample of size n identically distributed according to $(X, Y(0), Y(1), A, S) \mid S = 1$, and an observational sample of size m identically distributed according to $(X, Y(0), Y(1), A, S)$. Denote $\mathcal{R} = \{1, \cdots, n\}$ and $\mathcal{O} = \{n+1, \cdots, n+m\}$ as the index sets of samples from RCT and observational study, respectively.

For the RCT data, we observe $\{(X_i, A_i, Y_i, S_i = 1)\}_{i=1}^n$, where $Y_i = A_i Y_i(1) + (1 - A_i) Y_i(0)$ according to the SUTVA and consistency assumption. Moreover, the randomization implies

$$\{Y(1), Y(0)\} \perp A \mid X, S = 1.$$

Notations. Define the avergae treatment effect (ATE) in and conditional (CATE) observational and RCT population:

$$\tau = \mathbb{E}[Y(1) - Y(0)],\ \tau(x) = \mathbb{E}[Y(1) - Y(0)|X = x],$$
$$\tau_1 = \mathbb{E}[Y(1) - Y(0)|S = 1],\ \tau_1(x) = \mathbb{E}[Y(1) - Y(0)|X = x, S = 1].$$

Propensity score and conditional mean outcome:for $x \in \mathbb{R}^p,\ a \in \{0, 1\}$,

$$e(x) = \mathbb{P}(A = 1|X = x),\ e_1(x) = \mathbb{P}(A = 1|X = x, S = 1),$$
$$\mu_a(x) = \mathbb{E}[Y(a)|X = x],\ \mu_{a,1}(x) = \mathbb{E}[Y(a)|X = x].$$

Denote by $\alpha(x)$ the conditional odds that an individual with covariates x is in the RCT or observational sample:

$$\alpha(x) = \frac{\mathbb{P}(i \in \mathcal{R} \mid \exists i \in \mathcal{R} \cup \mathcal{O}, X_i = x)}{\mathbb{P}(i \in \mathcal{O} \mid \exists i \in \mathcal{R} \cup \mathcal{O}, X_i = x)} = \frac{\pi_{\mathcal{R}}(x)}{\pi_{\mathcal{O}}(x)} = \frac{\pi_{\mathcal{R}}(x)}{1 - \pi_{\mathcal{R}}(x)}.$$

Finally, we denote the distribution of covariates in observational data as f. Correspondingly, the distribution of covariates in RCT can be obtaine by $f_1(\cdot) = f(\cdot|S = 1)$.

2.1 Observational data with no treatment and outcome

We first suggest two assumptions to generalize the findings in RCT to a target population.

Assumption 1 (Transportability of CATE). $\tau(x) = \tau_1(x)$ for all x.

Assumption 2 (Positivity). There exists some constant $c > 0$ such that $\mathbb{P}(S = 1|X) \geq c$ almost surely.

Suppose that in our observational data $\mathcal{O} = \{n+1, \cdots, n+m\}$ we only observe the covariates X_i. Some commonly used estimators, all identifiable and consistent, are listed below.

Inverse probability of sampling weighting (IPSW). Under Assumption 1 and 2, and exchangeability holds for RCT, the ATE can be identified:

$$\tau = \mathbb{E}\left[\frac{n}{m\alpha(X)}\tau_1(X)\middle| S = 1\right] = \mathbb{E}\left[\frac{n}{m\alpha(X)}\left(\frac{A}{e_1(X)} - \frac{1 - A}{1 - e_1(X)}\right)Y\middle| S = 1\right].$$

Let $\alpha_{n,m}$ be an estimate of the odds of the indicatrix of being in the RCT, the IPSW estimator is given by

$$\hat{\tau}_{\text{IPSW},n,m} = \frac{1}{n}\sum_{i=1}^{n}\frac{nY_i}{m\hat{\alpha}_{n,m}(X_i)}\left(\frac{A_i}{e_1(X_i)} - \frac{1 - A_i}{1 - e_1(X_i)}\right).$$

Plug-in g-formula. If exchangeability holds for RCT and Assumption 1,2 holds, then ATE can be identified with another approach:

$$\tau = \mathbb{E}[\mu_{1,1}(X) - \mu_{0,1}(X)] = \mathbb{E}[\tau_1(X)].$$

Then, we can fit an estimator $\hat{\mu}_{a,1}(\cdot)$ of $\mu_{a,1}(\cdot)$ for $a = 0, 1$, using the RCT data. Note that a correctly specified regression model consistently estimates the mean outcome:

$$\begin{aligned}
\mathbb{E}[Y(a)|X = x] &= \mathbb{E}[Y(a)|X = x, S = 1] \text{ (Mean exchangeability)} \\
&= \mathbb{E}[Y(a)|A = a, X = x, S = 1] \text{ (Exchangeability)} \\
&= \mathbb{E}[Y|A = a, X = x, S = 1]. \text{ (Consistency)}
\end{aligned}$$

Note that the assumption of mean exchangeability is stronger than Assumption 1. With an observational distribution of X, plug-in the estimated (conditional) mean outcome to g-formula, we obtain

$$\mathbb{E}[Y(a)] = \int \mathbb{E}[Y|A = a, X = x, S = 1] f(x) \mathrm{d}x.$$

Take the empirical version, we obtain the plug-in g-formula estimator:

$$\hat{\tau}_{\mathrm{G},m,n} = \frac{1}{m} \sum_{i=n+1}^{n+m} \left(\hat{\mu}_{1,1,n}(X_i) - \hat{\mu}_{0,1,n}(X_i) \right).$$

Doubly-robust estimator. The IPSW and outcome models can be combined to construct an Augmented IPSW (AIPSW) estimator:

$$\begin{aligned}
\hat{\tau}_{\mathrm{AIPSW},n,m} = &\sum_{i=1}^{n} \frac{1}{m \hat{\alpha}_{n,m}(X_i)} \left(\frac{A_i(Y_i - \hat{\mu}_{1,1,n}(X_i))}{e_1(X_i)} - \frac{(1 - A_i)(Y_i - \hat{\mu}_{0,1,n}(X_i))}{1 - e_1(X_i)} \right) \\
&+ \frac{1}{m} \sum_{i=n+1}^{n+m} \left(\hat{\mu}_{1,1,n}(X_i) - \hat{\mu}_{0,1,n}(X_i) \right).
\end{aligned}$$

This estimator is doubly robust, i.e., it is consistent when either one of the two models for $\hat{\alpha}_{n,m}(\cdot)$ and $\mu_a(\cdot)$ $(a = 0, 1)$ is correctly specified.

Calibration weighting. Let $\mathbf{g}(X)$ be a vector of functions of X to be calibrated. Then the weight ω_i of each individuals in the RCT sample can be solved via minimizing the negative entropy (so that the empirical distribution of calibration weights is not too far away from the uniform distribution) under the balancing constraint (so that the weighted mean of $\mathbf{g}$ in RCT sample is equal to the mean in the observational sample):

$$\begin{aligned}
&\min_{\omega_1, \cdots, \omega_n} \sum_{i=1}^{n} \omega_i \log \omega_i, \\
&\text{subjected to } \omega_i \geq 0, \; i = 1, \cdots, n, \\
&\qquad \sum_{i=1}^{n} \omega_i = 1, \; \sum_{i=1}^{n} \omega_i \mathbf{g}(X_i) = \frac{1}{m} \sum_{j=n+1}^{n+m} \mathbf{g}(X_j).
\end{aligned}$$

Based on the calibration weights, the CW estimator is given by

$$\hat{\tau}_{\mathrm{CW},n,m} = \sum_{i=1}^{n} \omega_i Y_i \left(\frac{A_i}{e_1(X_i)} - \frac{1 - A_i}{1 - e_1(X_i)} \right).$$

Moreover, an augmented CW estimator can be derived by combining it with the plug-in g-formula estimator.

2.2 Observational data with treatment and outcome

In observational data, the actual data generating process is unknown and there exists unmeasured confounding in general, i.e.

$$\{Y(1), Y(0)\} \not\perp A \mid X.$$

Some methods have been proposed leverage trial and observational studies subject to hidden confounding.

2.2.1 Linear Combination

Denote our estimand as θ, then we can obtain its estimator $\hat{\theta}^{\text{exp}}$ and $\hat{\theta}^{\text{obs}}$ from trial data and observational data, respectively. Usually, in a well-designed randomized experiment, $\hat{\theta}^{\text{exp}}$ is unbiased, however its variance can be large due to the restriction of sample size. $\hat{\theta}^{\text{obs}}$ is biased, but it often has smaller variance than $\hat{\theta}^{\text{exp}}$ since observational data is easily accessible.

A straight-forward linear combination strategy can be proposed to leverage the two estimators. Denote the combined estimator as

$$\hat{\theta}_\lambda = (1-\lambda)\hat{\theta}^{\text{exp}} + \lambda\hat{\theta}^{\text{obs}}.$$

The optimal λ is chosen by minimizing the theoretical MSE:

$$\lambda^* = \frac{\text{Var}(\hat{\theta}^{\text{exp}}) - \text{Cov}(\hat{\theta}^{\text{exp}}, \hat{\theta}^{\text{obs}})}{\delta^2 + \text{Var}(\hat{\theta}^{\text{exp}} - \hat{\theta}^{\text{obs}})}$$

where δ is the bias of observational estimator. Practically, it can be estimated as $(\hat{\theta}^{\text{exp}} - \hat{\theta}^{\text{obs}})^2$.

2.2.2 Bias model

This idea comes from Kallus et al. (2018). Define a bias function $\eta(x) \neq 0$ to model the discrepancy between the true CATE and the estimated CATE:

$$\eta(x) := \tau(x) - \tau^{\mathcal{O}}(x).$$

Here $\tau^{\mathcal{O}}(x) = \mathbb{E}[Y|X=x, A=1] - \mathbb{E}[Y|X=x, A=0]$. Our estimator $\hat{\tau}_m^{\mathcal{O}}(x)$ is obtained by plugging-in $\tau^{\mathcal{O}}(x)$ since it is identifiable. Suppose $\eta(x)$ can be well approximated by a function with low complexity. We use a linear model to simulate the bias:

$$\tau(x) = \tau^{\mathcal{O}}(x) + \theta^\top x,\ \theta \in \mathbb{R}^p.$$

Now, we use a reweighting approach to obtain the expression of τ:

$$\tau_i^* = \left(\frac{A_i}{e(X_i)} - \frac{1-A_i}{1-e(X_i)}\right) Y_i.$$

Then we can learn θ through a least-squares approach on the RCT sample:

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^p}{\text{argmin}} \sum_{i=1}^{n} \left(\tau_i^* - \hat{\tau}_m^{\mathcal{O}}(X_i) - \theta^\top X_i\right)^2$$

note that $\hat{\tau}_m^{\mathcal{O}}(\cdot)$ is learned on the observational data. Using the estimated $\hat{\theta}$, we can recover the causal effect:

$$\hat{\tau}_{n,m}(x) = \hat{\tau}_m^{\mathcal{O}}(x) + \hat{\theta}_{n,m}^{\top} x.$$

Under some regularity conditions, the $\hat{\theta}_{n,m}$ estimated through least squares is consistent, and $\hat{\tau}(\cdot)$ is consistent on its target population.

2.2.3 Confounding function

This idea comes from Yang et al. (2020). We use $S = 1$ to denote the RCT sample and $S = 0$ the observational data. Define the CATE $\tau(\cdot)$ and confounding function, both can be identified:

$$\begin{aligned}
\tau(x) &= \mathbb{E}[Y(1) - Y(0)|X = x] \\
&= \mathbb{E}[Y|A = 1, X = x, S = 1] - \mathbb{E}[Y|A = 0, X = x, S = 1], \\
\lambda(x) &= \mathbb{E}[Y(0)|A = 1, X = x, S = 0] - \mathbb{E}[Y(0)|A = 0, X = x, S = 0] \\
&= \mathbb{E}[Y|A = 1, X = x, S = 0] - \mathbb{E}[Y|A = 0, X = x, S = 0] - \tau(x).
\end{aligned}$$

Then we parameterize $\tau(\cdot)$ and $\lambda(\cdot)$ as follows:

$$\tau(x) = \tau_{\varphi_0}(x),\ \lambda(x) = \lambda_{\phi_0}(x),\ \psi_0 = (\varphi_0^{\top}, \phi_0^{\top})^{\top} \in \mathbb{R}^{p_1} \times \mathbb{R}^{p_2}.$$

A crude estimator of ψ_0 can be obtained by least squares, since $\tau_{\varphi_0}(x)$ and $\lambda_{\phi_0}(x)$ are idetified.

Assumption 3 (Transportability and strong ignorability of trial treatment assignment).
$\mathbb{E}[Y(1) - Y(0)|X, S = 1] = \mathbb{E}[Y(1) - Y(0)|X, S = 0] = \tau(X)$, and
$Y(a) \perp A \mid (X, S = 1)$ for $a \in \{0, 1\}$, and $e(X, S) > 0$ almost surely.

Then, we introduce the following variable to mimic $Y(0)$:

$$H_{\psi_0} = Y - \tau_{\varphi_0}(X)A - (1 - S)\lambda_{\phi_0}(X)(A - e(X)).$$

Under Assumptions 3, we have $\mathbb{E}[H_{\psi_0}|A, X, S] = \mathbb{E}[Y(0)|X, S]$.

Furthermore, the semiparametric efficiency score of ψ_0 can be derived. An estimating equation using this score is applied to solve a semiparametric efficient estimator of ψ_0.

3 Multi-Arm Bandits with Unobserved Confounders (MABUC)

In standard multi-armed bandit problems, an agent is faced with $K \geq 2$ arms, each associated with an unknown independent distribution of rewards. In each round the agent pulls an arm and receives a reward from the corresponding distribution. The goal of the agent is to maximize the cumulative rewards over a series of rounds.

In the MABUC, agents are faced with the same task, except that unobserved confounders modify the agent's arm-choice predilections and payout rates at each round, and the dimensionality and functional form of the unobserved confounders are unknown.

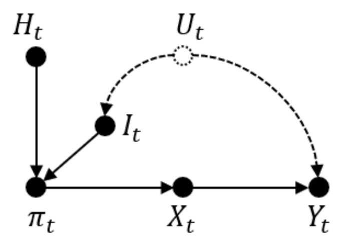
Definition 1 (Intent). Consider a structural causal model $\mathcal{M}$ and an endogeneous variable $X \in V$ which is determined by the structural equation $X = f_x(Pa_X, U_X)$ and can be intervene on. After realization $Pa_X = pa_x, U_X = u_x$, the output of the structural function given the current configuration of all unobserved confounders is said to be the agent's intent, i.e. $I = f_x(pa_x, u_x)$.

Intent can be viewed as an agent's chosen action before its execution, which is a proxy for any influencing unobserved confounders.

Definition 2 (K-armed bandits with unobserved confounders, MABUC) A K-armed bandit problem ($K \in \mathbb{N}, K \geq 2$) with unobserved confounders is defined as a SCM $\mathcal{M}$ with a reward distribution over $p(u)$ where, for each round $0 < t < T,\ t \in \mathbb{N}$:

1. Unobserved confounders: U_t represents the unobserved variables encoding the payout rate and unobserved influences to the propensity to choose arm x_t at round t.
2. Intent: $I_t \in \{x_1, \cdots, x_k\}$ represents the agent's intended arm choice at round t (prior to its final choice, X_t) such that $I_t = f_i(pa_{x_t}, u_t)$.
3. Policy: $\pi_t \in \{x_1, \cdots, x_k\}$ denotes the agent's decision algorithm as a function of its history and current intent, $f_\pi(h_t, i_t)$.
4. Choice: $X_t \in \{x_1, \cdots, x_k\}$ denotes the agent's final arm choice that is "pulled" at round t, $x_t = f_x(\pi_t)$.
5. Reward: $Y_t \in \{0, 1\}$ represents a Bernoulli reward (0 for losing, 1 for winning) from choosing arm x_t under unobserved confounder state u_t as decided by $y_t = f_y(x_t, u_t)$.

A graphical illustration of MABUC is given below.



3.1 Regret decision criterion (RDC)

In a MABUC instance with arm choice X, intent $I = i$, and reward Y, agents should choose the action a that maximizes their intent-specific reward, or formally:

$$a^* = \underset{a}{\text{argmax}}[Y_a | X = i],$$

where Y_a is the counterfactual had X been a. The counterfactual term $\mathbb{E}[Y_x | X = x_i]$ is referred to as the effect of treatment on the treated (ETT).

Theorem. The counterfactual ETT is empirically estimable for arbitrary action-choice dimension when agents condition on their intent $I = i$ and estimate the response Y to their final action choice $X = a$, i.e.

$$\mathbb{E}[Y_a | X = i] = \mathbb{E}[Y | \text{do}(X = a), I = i].$$

3.2 Data Fusion

Suppose that the agent in MABUC possesses:

- observation of arm choices and payouts from other players, in forms of $\mathbb{E}[Y|X]$;
- the randomized experimental results from an expert, in forms of $\mathbb{E}[Y|\text{do}(X)]$;
- the knowledge to employ intent in its decision-making for choosing arms by maximizing the counterfactual RDC $\mathbb{E}[Y_a | X = i]$ where i encodes information about unobserved confounders since $i = f_i(pa_x, u)$.

Since the ETT minimized by RDC is estimable, we can construct a counterfactual dataset in the form of $\{\mathbb{E}[Y_a|X=i]\}_{(i,a)\in\mathcal{X}\times\mathcal{X}}$, which can be obtained by the agent from its historical reward data. Here Y is a Bernoulli reward, an estimator $\hat{\mathbb{E}}[Y_a|X=i]$ can be computed from the frequencies of win or loss under intent i and action a. However, it does not utilize the information from observation and experimental dataset.

Some data fusion strategy is presented below.

Strategy 1. Cross-Intent Learning Consider the expansion of counterfactual quantity $\mathbb{E}[Y_x]$, a single cell in this system can be solve as:

$$\mathbb{E}_{\text{XInt}}[Y_{x_r}|x_w] = \frac{\mathbb{E}[Y_{x_r}] - \sum_{i=1,i\neq w}^{K}\mathbb{E}[Y_{x_r}|x_i]\mathbb{P}(x_i)}{\mathbb{P}(x_w)}.$$

This form provides a systematic way of learning about arm payouts across intent conditions, which is desirable because an arm pulled under one intent condition provides knowledge about the payouts of that arm under other intent conditions.

Strategy 2. Cross-Arm Learning Consider a single-cell quary $\mathbb{E}[Y_{x_r}|x_w]$. For any three arms x_r, x_s, x_w such that $r \notin \{s, w\}$, it holds

$$\mathbb{E}[Y_{x_r}] = \sum_{i=1}^{K}\mathbb{E}[Y_{x_r}|x_i]\mathbb{P}(x_i), \quad \mathbb{E}[Y_{x_s}] = \sum_{i=1}^{K}\mathbb{E}[Y_{x_s}|x_i]\mathbb{P}(x_i)$$

Then the query intent $\mathbb{P}(x_w)$ can be expressed as

$$\mathbb{P}(x_w) = \frac{\mathbb{E}[Y_{x_r}] - \sum_{i=1,i\neq w}^{K}\mathbb{E}[Y_{x_r}|x_i]\mathbb{P}(x_i)}{\mathbb{E}[Y_{x_r}|x_w]} = \frac{\mathbb{E}[Y_{x_s}] - \sum_{i=1,i\neq w}^{K}\mathbb{E}[Y_{x_s}|x_i]\mathbb{P}(x_i)}{\mathbb{E}[Y_{x_s}|x_w]},$$

and we can solve our query in terms of the paired arm x_s:

$$\mathbb{E}_{\text{XArm}}[Y_{x_r}|x_w] = \frac{\left\{\mathbb{E}[Y_{x_r}] - \sum_{i=1,i\neq w}^{K}\mathbb{E}[Y_{x_r}|x_i]\mathbb{P}(x_i)\right\}\mathbb{E}[Y_{x_s}|x_w]}{\mathbb{E}[Y_{x_s}] - \sum_{i=1,i\neq w}^{K}\mathbb{E}[Y_{x_s}|x_i]\mathbb{P}(x_i)}.$$

This formula allows our agent to estimate $\mathbb{E}[Y_{x_r}|x_w]$ from samples in which any arm $x_s \neq x_r$ was pulled under the same intent x_w. It can be viewed as information about $Y_{x_r}|x_w$ flowing from arm $x_s \neq x_r$ to x_r (under intent x_w).

A more robust estimate of the query can be obtained via inverse-variance-weighted average. Consider a function h_{XArm} such that $\mathbb{E}_{\text{XArm}}[Y_{x_r}|x_w] = h_{\text{XArm}}(x_r, x_w, x_s)$ and h_{XArm} performs the empirical evaluation of the RHS of the equation above. Moreover, let $\sigma_{x,i}^2 = \widehat{\text{Var}}(Y_x|i)$ indicate the empirical payout variance for each arm-intent condition. Then our estimator is

$$\mathbb{E}_{\text{XArm}}[Y_{x_r}|x_w] = \frac{\sum_{i=1,i\neq r}^{K} h_{\text{XArm}}(x_r, x_w, x_i)/\sigma_{x,i}^2}{\sum_{i=1,i\neq r}^{K} 1/\sigma_{x,i}^2}.$$

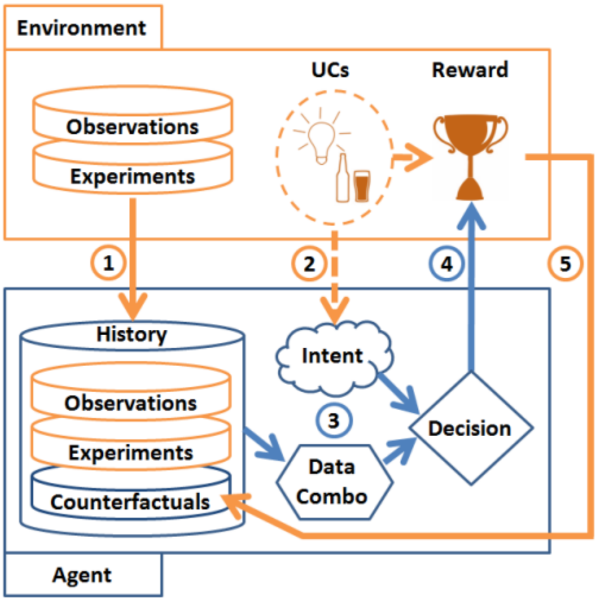
Strategy 3. The Combined Approach Until now, three methods are proposed to estimate an intent-specific counterfactual reward $\mathbb{E}[Y_{x_r}|x_s]$: + $\hat{\mathbb{E}}[Y_{x_r}|x_s]$ from conditionally randomized experiment. + $\mathbb{E}_{\text{XInt}}[Y_{x_r}|x_w]$ from cross-intent learning. + $\mathbb{E}_{\text{XArm}}[Y_{x_r}|x_w]$ from cross-arm learning.

While the variance of $\hat{\mathbb{E}}[Y_{x_r}|x_s]$ can be estimated directly as the conditional sample variance $\sigma_{x_r,x_s}^2 = \widehat{\text{Var}}(Y_{x_r}|x_s)$, the cross-intent and cross-arm estimates, as combinations of sample payout estimates, have roughly estimated variances:

$$\sigma^2_{\text{XInt}} = \frac{1}{2K-1}\left[\left(\sum_{i=1,i\neq w}^{K} \sigma^2_{x_r,x_i}\right) + \left(\sum_{i=1,i\neq w}^{K} \sigma^2_{x_s,x_i}\right) + \sigma^2_{x_s,x_w}\right],$$
$$\sigma^2_{\text{XArm}} = \frac{1}{K-1}\sum_{i=1,i\neq w}^{K} \sigma^2_{x_r,x_i}.$$

Then, an inverse-variance weighting scheme can be employed to leverage the three estimators:

$$\alpha = \hat{\mathbb{E}}[Y_{x_r}|x_s]/\sigma^2_{x_r,x_s} + \mathbb{E}_{\text{XInt}}[Y_{x_r}|x_w]/\sigma^2_{\text{XInt}} + \mathbb{E}_{\text{XArm}}[Y_{x_r}|x_w]/\sigma^2_{\text{XArm}},$$
$$\beta = 1/\sigma^2_{x_r,x_s} + 1/\sigma^2_{\text{XInt}} + 1/\sigma^2_{\text{XArm}},$$
$$\mathbb{E}_{\text{combo}}[Y_{x_r}|x_s] = \alpha/\beta.$$



Overview.

1. Assume that the agent has collected large samples of experimental and observational data from its environment.

2. Unobserved confounders (UCs) in the environment are realized by the agent, though their labels and values are unknown.

3. From these UCs and any other observed features in the environment, the agent's heuristics suggest an action to take, i.e., its intent. With its intent known, the agent combines the data in its history (in this work, by the prescription of Strategy 3 above) to better inform its decision-making.

4. Based on its intent and combined history, the agent commits to a final action choice.

5. The action's response in the environment (i.e., its reward) is observed, and the collected data point is added to the agent's counterfactual dataset.

3.3 Thompson Sampling with RDC

To address the MABUC problem using data fusion strategy, Forney et al. proposed an implementation of RDC using Thompson Sampling as its basis. In each round t, the agent performs as follows:

- Observe the intent i_t from the realization of unobserved confounders u_t in the current round;

- For each arm x_r, $r = 1, \cdots, K$, sample $\hat{\mathbb{E}}[Y_{x_r}|i_t]$ from distribution Beta$(s_{x_r,i_t}, f_{x_r,i_t})$ in which s_{x_r,i_t} is the number of successes ($Y = 1$) and f_{x_r,i_t} the number of failures ($Y = 0$).

- Compute the i_t-specific score for each arm using the combined datasets via Strategy 3.

- According to RDC, choose the arm x_a with the highest i_t-specific score.

- Observe the result and update $\hat{\mathbb{E}}[Y_{x_a}|i_t]$.

References

[1] Pearl, J. and E. Bareinboim (2011). Transportability of causal and statistical relations: A formal approach. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI 2011).*

[2] Bareinboim, E. and J. Pearl (2013a). Causal transportability with limited experiments. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI 2013).*

[3] Bareinboim, E. and J. Pearl (2013b). Meta-transportability of causal effects: A formal approach. In *Proceedings of the 16th International Conference on Artificial Intelligence Statistics (AISTATS 2013).*

[4] Bareinboim, E. and J. Pearl (2014). Transportability from Multiple Environments with Limited Experiments: Completeness Results. In *Advances in Neural Information Processing Systems 27 (NeurlIPS 2014).*

[5] Lee S., J. D. Correa, and E. Bareinboim (2019). General Identifiability with Arbitrary Surrogate Experiments. In *Proceedings of the Thirty-Fifth Conference Annual Conference on Uncertainty in Artificial Intelligence.*

[6] Correa J. D. and E. Bareinboim (2020). General Transportability of Soft Interventions: Completeness Results. In *Advances in Neural Information Processing Systems 33 (NeurlIPS 2020)*

[7] Colnet, B., I. Mayer, G. Chen, A. Dieng, R. Li, G. Varoquaux, I. P. Vert, J. Josse and S. Yang (2020). Causal inference methods for combining randomized trials and observational studies: a review. *arXiv preprint arXiv:2011.08047.*

[8] Lee D., S. Yang, L. Dong, X. Wang, D. Zeng, J. Cai (2023). Improving trial generalizability using observational studies. *Biometrics.* 2023 Jun;79(2):1213-1225.

[9] Kallus, N., A. M. Puli, and U. Shalit (2018). Removing hidden confounding by experimental grounding. In *Advances in Neural Information Processing Systems.*

[10] Yang, S., D. Zeng, and X. Wang (2020). Improved inference for heterogeneous treatment effects using real-world data subject to hidden confounding. *arXiv preprint arXiv:2007.12922.*

[11] Forney, A., J. Pearl and E. Bareinboim (2017). Counterfactual Data-Fusion for Online Reinforcement Learners. In *Proceedings of the 34th International Conference on Machine Learning*, in *Proceedings of Machine Learning Research 70:1156-1164.*