CAUSALITY CORRECTIONS IMPLEMENTED IN 2nd PRINTING

Updated 9/26/00

page v  *insert* “TO RUTH” centered in middle of page

page xv  *insert* in second paragraph “David Galles” after “Dechter”

page 2  *replace* “2000” with “2004” in 2nd paragraph, line 8 of 1.1.2

page 3  *insert* “,” after “(−)” in footnote 1
  *replace* “and not” with “not, and implies,” in footnote 1

page 19  *append* (continue italics) to end of Theorem 1.2.7, “(We exclude $X_i$ when speaking of its “nondescendants”.)”

page 30  *replace* in line 7 from top “mutually” with “jointly”
  *insert* “parental” before “Markov” in first line of 2nd paragraph after Theorem 1.4.1
  *append* to end of footnote 16 “but I am not aware of any nonparametric version.”

page 52  *insert* “stable” after “IC*, that takes a” in 2nd paragraph after Theorem 2.6.2,
  *replace* “sampled” with “stable” in Input line of IC* Algorithm.
  *append* “(with respect to some latent structure)” to same line

page 68  *replace* in line 1 after Eq. (3.2), “mutually” with “jointly”

page 72  *replace* “(1990, 1999)” with “(1990, 2001)” on line 6

page 89  *replace* in paragraph starting “Indeed, if condition...”. Should be “conditions require” not “condition require”


page 130  *replace* the “P” with “E” in the formula (second line of section 4.5.4.)
  Should read “$E(Y[\hat{x}, \hat{pa}_Y \setminus x])$”
  *replace* “we can compute the difference” with “we should replace the controlled difference” in last line of page
page 131 *replace* from top of page through the end of section 4.5.4 with the following:

\[ P(\text{admission | male, dept}) - P(\text{admission | female, dept}) \]

with some average of this difference over all departments. This average should measure the increase in admission rate in a hypothetical experiment in which we instruct all female candidates to retain their department preferences but change their gender identification (on the application form) from female to male.

In general, the average direct effect is defined as the expected change in \( Y \) induced by changing \( X \) from \( x \) to \( x' \) while keeping the other parents of \( Y \) constant at whatever value they obtain under \( do(x) \). This hypothetical change is what law makers instruct us to consider in race or sex discrimination cases: “The central question in any employment-discrimination case is whether the employer would have taken the same action had the employee been of a different race (age, sex, religion, national origin etc.) and everything else had been the same.” (In Carson versus Bethlehem Steel Corp., 70 FEP Cases 921, 7th Cir. (1996)).

The formal expression for this hypothetical change involves probabilities of (nested) counterfactuals (see Section 7.1 for semantics and computation) that cannot be written in terms of the \( do(x) \) operator.\(^9\) Therefore, the average direct effect cannot in general be identified, even from data obtained under randomized control of all variables. However, if certain assumptions of “no confounding” are deemed valid,\(^10\) then the average direct effect can be reduced to

\[
\Delta_{x,x'}(Y) = \sum_{p|y_{\neg X}} [E(Y|\hat{x}', p\hat{a}_{Y|X}) - E(Y|\hat{x}, p\hat{a}_{Y|X})]P(p|y_{\neg X}|\hat{x}), \tag{4.11}
\]

and the techniques developed in Section 4.4 for identifying control-specific plans, \( P(y|\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) \), become applicable.

---

\(^9\) Using the counterfactual notation of Section 7.1, the general expression for the average direct effect is

\[ \Delta_{x,x'}(Y) = E(Y_{x'Z_x}) - E(Y_Z) \]

where \( Z = p\hat{a}_{Y|X} \). The subscript \( x'Z_x \) represents the operation of setting \( X \) to \( x' \) and, simultaneously, setting \( Z \) to whatever value it would have obtained under the setting \( X = x \). This general expression reduces to (4.11) if \( Z \perp Y_{x'Z_x} \) holds for all \( z \). Likewise, the average indirect effect is defined as \( E(Y_{Z|x'}) - E(Y_Z) \).

page 164  *replace* ... “do(x, y, w)” with “do(x, z, w)” in line 8 after Definition 5.4.3.

page 165  *replace* last two sentences of section 5.4.2 with:
The expressions corresponding to these policies are \( P(y|do(x), do(z)) \) and \( P(y|do(x)) \), and this pair of distributions fully represents the policy implications of indirect effects. Similar conclusions have been expressed by Robins and Greenland (1992). (But see Chapter 4, footnote 9, page 131.)

page 177  *delete* “tormented” in paragraph 3, line 2

page 184  *append* to end of Definition 6.2.1 (continue italics - except ‘unbiased’):
“If (6.10) holds, we say that \( P(y|x) \) is unbiased.”

page 236  *replace* “& (X \rightarrow Y)” with “& (X \not\rightarrow Y)” in first formula of Theorem 7.3.8

page 240  *replace* the last sentence in the last paragraph of section 7.4.1 with:
However, this effectiveness is partly acquired by limiting the counterfactual antecedent to conjunction of elementary propositions. Disjunctive hypotheticals, such as “if Bizet and Verdi were compatriots,” usually lead to multiple solutions and hence to nonunique probability assignments.

page 246  *insert* in footnote 26 after “(see Section 5.4.3).” “Epidemiologists refer to (7.46) as “no-confounding” (see (6.10)).”

page 255  *replace* in the 2nd line “pregnant” with “nonpregnant”

page 259  *insert* close parentheses after “(Sections 3.2 and 7.1), line 2 of Preface

page 284  *replace* “Michie in press” with “Michie 1999” in the last line of paragraph 4

page 329  *replace* “(1999)” with “(2000)” in last line of page

page 332  *replace* in paragraph starting “Even an erratic and ...”. Change “role” to “roll”

page 354  line 2 from bottom, *replace* “mediated by tar deposits” with “unmediated by tar deposits”


*replace* “Actual causality,” with “Causes and explanations.,” and *append* “www.cs.ucla.edu/~judea/”

page 364  *update* Hoover 1999 citation. *replace* “(1999)” with “(2001)”

page 366 *update* Michie citation. *replace* “(in press)” with “(1999)” and *insert* “pp. 60-86” before “vol. 15”

page 368 *update* Pearl 1999 citation. *remove* “To appear in” and replace “121” with “121:93–149.”

page 369 *insert* in Robins 1997 citation, “M. Berkane (Ed.),” before “Latent Variable Modeling...”


page 381 *insert* “27–8”, after “(examples) price and demand” and before “215-17” *replace* “245” with “245–7”, at end of “(exogeneity) controversies regarding...245” *combine* “explanation” and “explanations” to read “explanation, 25, 58, 221–3, 285, 308–9”

page 382 *insert* “131” after “indirect effects,” and before “165”
ADDENDUM TO CORRECTIONS  
IMPLEMENTED IN 2nd PRINTING  
Updated 12/14/00  

page 28 *replace* “income (Z)” with “income (I)” in the caption of Figure 1.5

page 48 *replace* in line before Definition 2.4.1, “when one of the coins becomes slightly biased.” with “when the coins become slightly biased.”

page 51 *append* to line 7, Rule R₄ to read:
Orient a – b into a → b whenever there are two chains a–c → d and c → d → b such that c and b are nonadjacent and a and d are adjacent.

page 231 *replace* Definition 7.3.4 and 2 lines following to read:
Definition 7.3.4 (Recursiveness)
Let X and Y be singleton variables in a model, and let X → Y stand for the inequality Yₓₓ(w)(u) ≠ Yₓ(w)(u) for some values of x, w, and u. A model M is recursive if, for any sequence X₁, X₂, . . . , Xₖ, we have

\[ X₁ → X₂, X₂ → X₃, . . . , Xₖ₋₁ → Xₖ ⇒ Xₖ ↗ X₁ \]  \hspace{1cm} (7.24)

Clearly, any model M for which the causal diagram G(M) is acyclic must be recursive.

page 382 *change* “Markov (assumptions underlying, 30)” to “Markov (assumption, 30, 69)”

page 382 *append* “69” after “causal, 30” in “Markov condition (causal, 30)”

page 384 *add* as subentry after “structural model, 27, 44, 202” “Markovian, 30, 69”.
PRINTING CORRECTIONS
TO BE IMPLEMENTED BY CAMBRIDGE

Updated 3/13/07

page 52 *replace* in line 17-18 "protection" with "projection"

page 73 *replace* equation between (3.11) and (3.12)

\[
\frac{P(p_{oi} \mid do(x'_i))}{P(p_{oi})} = \frac{P(s_{oi}, p_{oi} \mid do(x'_i))}{P(s_{oi}, p_{oi})},
\]

should be:

\[
\frac{P(p_{oi} \mid do(x'_i))}{P(p_{oi})} = \frac{P(s_{oi}, p_{oi}, x'_i \mid do(x'_i))}{P(s_{oi}, p_{oi}, x'_i)}.
\]

page 82 *replace* at end of paragraph 2: “since there is no back-door path from X to Z, we simply have” with “since there is no unblocked back-door path from X to Z in Figure 3.5, we simply have”

page 82 *replace* in Definition 3.3.3 (Front-Door): “(ii) there is no back-door path from X to Z; and” to “there is no unblocked back-door path from X to Z; and”

page 103 *replace* last paragraph on page 103 (including footnote 15) with:

To place this result in the context of our analysis in this chapter, we note that one class of semi-Markovian models satisfying assumption (3.62) corresponds to graphs in which all arrowheads pointing to \(X_k\) originate from observed variables. Indeed, in such models, the parents \(PA_k = L_k, X_{k-1}\) of variable \(X_k\) satisfy the back-door condition of Definition 3.3.1,

\[(X_k \perp \!
\perp Y \mid PA_k)_{GA_{X_k}},\]

which implies (3.62).15 This class of models falls under Theorem 3.2.5, which states that all causal effects in this class are identifiable and are given by the truncated factorization formula of (3.14); the formula coincides with (3.63) after marginalizing over the uncontrolled covariates.

---

15Condition (3.62) is too restrictive and lacks intuitive basis: a graphical, more general condition leading to (3.63) is formulated in (4.5), Theorem 4.4.1, read: \(P(y \mid x, x')\) is identifiable and is given by (3.63) if every action-avoiding back-door path from \(X_k\) to \(Y\) is blocked by some subset \(I_k\), of non-descendants of \(X_k\). (by “action-avoiding” we mean a path containing no arrow entering an \(X\) variable later than \(X_k\)) see [http://bayes.cs.ucla.edu/BOOK-2K/yudkowsky.html].
page 174  *replace* in 4th line from end of page: “is not a statement about $C$ being a positive causal factor for $E$, properly written” with “is not a statement about $C$ having a positive influence on $E$, properly written”

page 195  *replace* "Figure 6.1” with "Figure 6.3” in 5th line before end of 3rd paragraph