Errata to
Information Modeling and Relational Databases, Second Edition
Halpin, T. and Morgan T. 2008
(Morgan Kaufmann)

Last updated: 2020 April 7

If you spot any other errors in this second edition, please e-mail them to Terry Halpin at t.halpin@live.com.

The following corrections should be made to the *first* printing. To determine the print version of your copy of the book, look at the row of digits below the copyright notice on the verso of the main title page: the final digit denotes the print version (e.g. 1 = first printing, 2 = second printing).

The symbol "⇨" means "is replaced by".

p. xxiv, Online Resources, para. 1: Replace by: "To reduce the size and hence cost of the book, supplementary material has been made available for downloading by accessing the following link to the relevant Website: https://booksite.elsevier.com/9780123735683/?ISBN=9780123735683. There are five appendices. Appendix A provides an overview of the evolution of computer hardware and software. Appendix B discusses two kinds of subtype matrix that can be used to determine subtype graphs from significant populations. Appendices C, D and E discuss advanced aspects of SQL, focusing on set-comparison queries, ranking and extrema queries, and implementing derivation rules using triggers, generated columns, stored procedures, views, user defined functions, and common table expressions.".

p. 35, para. 3, line 5: "Counties" ⇨ "Countries".

p. 36, para. 3, lines 4 and 5: "city" ⇨ "country".

p. 41, Ex. 2.2, Q. 3(a): Delete "This level is concerned with providing".

p. 42, Ex. 2.2, Q. 4 (c) and (e): insert "**add**: " before the fact statements.

p. 54, Table 2.3: "Physic-Persistent" ⇨ "Physical-Persistent".

p. 57, para. 4, sentence 2: "The Zachman Institute …(see *www.zifa.com*)" ⇨ Zachman International promotes the use of the Zachman framework in industry (see *www.ZachmanInternational.com*).

p. 57, para. 4, last 2 sentences: "At the … writing." ⇨ "At the DAMA International Symposium in March, 2008, John Zachman announced a new version of his framework, called Zachman Framework[2] (see diagram at *http://zachmaninternational.com/images/stories/The%20Zachman%20Framework.pdf*).

p. 65, para. 2, line 2: "and plays" ⇨ "so plays".

p. 69, para. 4, line 2: "Adam and Jim" ⇨ "Terry and Tony";
                              "Eve and Mary" ⇨ "Norma and Gwen".

p. 70, Table 3.3: "1971" ⇨ "1970".

p. 70, Facts (25) and (27): "1971" ⇨ "1970".

p. 70, para. 5, 3rd sentence: "In order to split (25) into (26) and (27) we probably relied on our" ⇨ "A sufficient reason for splitting (25) into (26) and (27) would be an".

p. 70, para. 6, first sentence: append "(e.g., Wirth did most of Pascal's design in 1968-69, but did not successfully implement it or finalize its definition till 1970)".

p. 70, para. 6, last sentence: delete "just".

p. 70, last sentence: "elsewhere" ⇨ "in various ways".

p. 71, Fact (27a): "1971" ⇨ "1970".

p. 71, Fact (27a) with reference scheme omitted: "1971" ⇨ "1970".

p. 71, last para., line 1: delete "for us".

p. 76: "Jones" ⇨ "Bloggs" and "ES" ⇨ "FB" (2 occurrences).

pp. 76-77, Figures 3.5 and 3.6: surname and initials data should not be in italics.

p. 79, Fig. 3.7: Delete the arrow connection from Seattle to Chicago, and the bottom-most arrow connection from Los Angeles to Chicago.

p. 85: "XEU" ⇨ "EUR" (2 occurrences).

p. 92, Ex. 3.4, Q. 2: "Q. 10." ⇨ "Q. 10; (i) Q. 11.".

p. 103, Ex. 3.5, Q. 2: Change the Manager column data to: Smith J, Jones E, Brown C, Smith J, Collins A. Change the salary and birth year data for Project P4 to: 50000, 1946.

p. 103, Ex. 3.5, Q. 5: Replace the three male staff entries by the following two entries: (John Jones, Sales); (Bob Smith, Admin).

p. 128, Fig. 4.31 derivation rule: Insert "**some**" before the second instance of "Car".

p. 139, Figs 4.45 and 4.46: "(sq_km)" ⇨ "(sq_km:)".

p. 145, para. 4, line 3: "4.47" ⇨ "4.50"; "4.50" ⇨ "4.51".

p. 168, para. 4: Move the text "(this default can be over-ridden by declaring the object type independent—see next chapter)" to the end of the paragraph.

p. 141, para. 4, line 5: change font for "Table 4.9" to match the paragraph font.

p. 169, Fig. 5.12: An inclusive-or constraint over two roles hosted by the same object type may also be displayed by connecting the constraint shape with a single dotted line to the junction of the two roles.

p. 185, para. 3: A shorter version of the SQL query is: **select count**(*) **from** Room **where** roomNr **like** '67-3%'.

p. 188, para. 1: Reword as follows: "The external uniqueness constraint over three roles indicates that each (genus, species, infraspecies) tuple refers to at most one plant kind. The other external uniqueness constraint declares that the (rank, inframame) pair identifies infraspecies. In the reference scheme for Plantkind, only the genus is mandatory. Moreover, outer join semantics applies in this case (see Section 10.1), so for each genus there is at most one plant kind with only a genus name, and for each genus-species combination there is at most one plant kind with no infraspecies. In a relational database, each plant kind maps to a sequence of four values, three of which may be null, and each quadruple is unique (treating null just like any other value)."

p. 190, Figure 5.30: change "(.code" to "(.code)".

p. 196, Table 5.13, last line: Widen final column to fit "1103.50" on one line.

p. 199, para. 1, last line: "Bank)" ⇨ "Bank".

Figure: Icon for Swimming should be the icon shown for Main Beach.

p. 206, Ex. 5.5, Q. 1, first table: "JM (Joan Coffee)" ⇨ "JC (Joan Coffee)".

p. 217, Fig. 6.6: Remove square brackets from most examples, and change caption to "Value constraints list the possible values of a value type or referencing role.".

p. 217, para. 2, replace the third sentence onwards by: "By default, ranges are assumed to be closed, so the end values are included. For example, {50..100} includes 50 and 100 as well as the numbers in between. Inserting *a left parenthesis "(" before a start value excludes that value*, and *appending a right parenthesis ")" after an end value excludes that value*. If desired, a square bracket may be used to explicitly indicate inclusion (the default). For example, "(0..100" and "(0..100]" each denote a range of positive (above 0) numbers up

to and including 100. One may also combine enumerations and/or ranges into a single constraint, as in Fig. 6.6(c).".

p. 217, final sentence: Append a period.

p. 218, before final paragraph, insert new paragraph: "As a fine point, a value constraint displayed next to an entity type with a reference mode is actually a role value constraint on the referential role played by the value type. In the NORMA tool, you can display it this way by toggling ExpandRefMode to True.".

p. 219, combine paragraph starting with "Although choosing" with previous paragraph.

p. 219, 2nd last para., replace by: "A cleaner approach is to list all existing countries in a *reference table* for Country. A country may appear in this table without playing any elementary fact roles in the schema. Such a country is said to be *independent*, since it exists in our model independently of doing anything. An **independent entity type** is a *primitive entity type whose elementary fact roles (if any) are collectively optional*. An *independent value type* may contain some values that play in no facts (elementary or existential). The term "independent" is not used for subtypes. If an independent object type plays some fact roles, the disjunction of these roles is not mandatory (implicitly or explicitly).".

p. 225, Fig. 6.13: "(name)" ⇨ "(.name)".

p. 229, Fig. 6.20: Graphically, applying an inclusive-or or exclusive-or constraint to the junction of two roles applies it to the role pair, so the xor constraint may also be displayed with just one dotted line connected to the junction of the two roles.

p. 235, Ex. 6.4, Q. 2 Figure: Role connectors from the exclusion constraint should be dashed lines.

p. 238, Ex. 6.4, Q. 6(a): "Transaction (code)" ⇨ "Transaction(.code)".

p. 241, final para., line 3: "*subtype* of *B* and *C*" ⇨ "*subtype* of *C* and *D*".

p. 243, Fig. 6.35 and para. 1, line 3: "Female" ⇨ "FemalePerson".

p. 252, Fig. 6.48: "PersonKind(code)" ⇨ "PersonKind(.code)".

p. 253, para. 2, replace final sentence by: "Just based on age, we could specify any set of natural numbers within the possible age range that included 13, 17, and 19 but excluded 12 and 20 (e.g., {13, 17, 19, 1}, {13, 17, 19, 5, 63} and so on).".

p. 256, Fig. 6.53: "Period > 0" ⇨ "PeriodRate > 0" (2 occurrences).

p. 260, Ex. 6.5, Q. 6, Planets table, NrMoons of Saturn: "60" ⇨ "62".

p. 271, Sec 7.1 Title: "Step 6" ⇨ "Step 7".

p. 290, second para, first sentence: "four comparison operators <, ≤, >, ≥" ⇨ "six comparison operators <, ≤, >, ≥, =, ≠".

p. 294, Ex. 7.4, Q. 8: "Question 4" ⇨ "Question 5".

p. 299, Ex. 7.5, Q. 3: Add the predicate reading "… located in … ordered … of …".

p. 341, Table, Loan date column: "/00" ⇨ "/2000" (3 occurrences).

p. 341, Ex. 8.5, Q.2 Figure: "Teacher **and an** Academic" ⇨ "Teacher **and a** Professor".

p. 347, Table 9.1: Remove shading from right bottom cell (Sequence etc.).

p. 366, para. 3: Replace all but the first and last sentences by: "In this case, the cycle arose because of the pair-subset constraint combined with the mandatory role constraint on is-chaired-by. Given the pair-subset constraint, the cycle would also apply if Chair is independent or plays some other mandatory, functional role.".

p. 366, final paragraph: Replace first two sentences by: "UML 2 introduced a *subsets property modifier* to indicate that for each instance that is the context for a property (i.e. an attribute or association role) the population (extension) of that

property for that instance must be a subset of the population of another compatible property for that same instance. For example, adorning the citizen role in Figure 9.26(a) with {subsets resident} means that each citizen of a country is a resident of the same country.".

p. 367, Fig. 9.26: Move the subset constraint in (b) to connect the mid-points of the two predicates, and reword the Figure caption to read: "A pair-subset constraint in (a) UML and (b) ORM.".

p. 367, para. 2, line 5: "Test" ⇨ "TestNr".

p. 367, Fig. 9.27(a): "Test" ⇨ "TestNr".

p. 367, Fig. 9.27(b): Replace the entity type Test(.nr) by the value type TestNr.

p. 368, Figs. 9.28 and 9.29: Add directional arrow tip after "Define".

p. 371, Fig. 9.33: Add spanning uniqueness constraints to all the ORM predicates.

p. 373, Fig. 9.35(a): Replace lower left constraint by "{complete, overlapping}".

p. 374, para. 2, last line: "Party" ⇨ "Organization".

p. 379, para. 3, line 2: "null values" ⇨ "nulls".

p. 380, Fig. 9.44(a), derived association: "member" ⇨ "/member".

p. 380, para. 3: Replace by "Early versions of the UML specification declared that "both kinds of aggregation define a transitive, antisymmetric relationship (i.e., the instances form a directed, noncyclic graph)". But antisymmetry does not imply irreflexivity, much less acyclicity. Moreover, transitive closure is normally best achieved by derivation rather than assertion. Such property declarations for aggregation appear to be absent in UML 2.2.".

p. 380, para. 4: Replace by: "At any rate, some restriction on aggregations is needed to stop cases like that shown in Figure 9.45. If a person is part of a team, and a team is part of a club, it doesn't make sense to say that a club is part of a person. The example should be remodeled, as shown in Figure 9.44(a).".

p. 380, Fig. 9.45 caption: Replace by: "Illegal UML model. Aggregation association chains should not form a cycle.".

p. 381, final para: format should be fully justified.

p. 387, para. 3, line 3: "role that specializes a role played by" ⇨ "that specializes an association involving".

p. 387, Fig. 9.52: change multiplicity on assignedCar to "0..1"; remove Rank fact type; "uses" ⇨ "is assigned"; change UC on company car fact type to span just the lefthand role; add the asserted subtype "Executive"; modify the textual constraint to read "**Each** Employee **who is an** Executive is assigned **some** CompanyCar.".

p. 388, last line: "binaries" ⇨ "unaries".

p. 389, Table 9.3, Step 4: "*m:n* and *n*-ary" ⇨ "remaining".
                  Step 7: "semi-derived" ⇨ "semiderived".

p. 390, Fig. 9.56: "familyName" ⇨ "birthdate".

p. 390, Fig. 9.57: "Title" ⇨ "PersonTitle" (2 occurrences).

p. 390, Fig, 9.58: "Booelan" ⇨ "Boolean".

p. 392, Table 9.4, last row: "Object cardinality" ⇨ "Value-comparison"; "Class multiplicity" ⇨ "Textual constraint".

p. 393: "semi-derived" ⇨ "semiderived".

p. 397, last line: "*www.puml.org*" ⇨ "*www.cs.york.ac.uk/puml/*".

p. 402, Role relation, column roleId: "R4" ⇨ "r4".

p. 426, Fig. 10.30: Add intransitive constraint to ring fact type.

p. 436, Fig. 10.39: "values" should be in bold type.

p. 436, Fig. 10.40: "values" should be in bold type (2 occurrences).

p. 436, Fig. 10.40(b): "… was written by … in …" ⇨ "… includes … in …".

p. 438, Ex. 10.4, Q. 1: "(sequence of railway cars)" ⇨ "(ordered set of railway cars)".

p. 440, para. 4: "sentence" ⇨ "declarative sentence".

p. 442, penultimate paragraph: Change to "Default predicate readings for the linking predicates are assigned, such as "involves", appended by numbers if needed to distinguish linking fact types that link back to the same object type (which plays more than one role in the fact type being objectified). The modeler may edit these default readings to provide more meaningful readings for the link fact types, as shown in Figure 10.45".

p. 450, Ex. 10.5, Q. 1 (a), (b): Reduce indentation to match that of (c).

p. 454, Figures 10.59 and 10.60: Move the multiplicities to after the attribute type declaration, e.g. code:string [1] {P}.

p. 454, Fig. 10.60(c): Change the line style for the unaries to a dashed line.

p. 465, line 4: "'F' for" ⇨ "'F' (for".

p. 476, para. 6, sentences 2 and 3: Replace by: "A superkey of a table is a set of one or more columns whose entries must be unique, thus providing a way to identify rows in the table. A *candidate key* is a *minimal* superkey.".

p. 482, Fig. 11.6: Add uniqueness constraint to left role of derived fact type.

p. 483, Ex. 11.2, Q. 1: "null value" ⇨ "null".

p. 496, Fig. 11.24: "{'start', end'} ⇨ "{'start', 'end'}".

p. 504, Fig. 11.32: The foreign key relationship arrow between Survey.channelNr and Channel.channleNr should point downwards, not upwards.

p. 505, Ex. 11.3, Q. 1: "points" ⇨ ".points".

p. 505, Ex. 11.3, Q. 2: Add spanning uniqueness constraint to Supply association.

p. 508, Ex. 11.3, Q. 18: "< was directed by" ⇨ "directed / was directed by".

p. 518, Replace 2nd paragraph by: "Recall that a direct supertype of a subtype is connected directly to the subtype (i.e., with no intermediate subtypes on the connecting path). Each subtype has one or more direct supertypes. A connection between a subtype and its direct supertype(s) is shown as a *solid arrow if it provides a path to the preferred identifier of the subtype*. If a subtyping connection appears as a *dashed arrow*, the subtype either introduces its own preferred reference scheme (this is called *context-dependent reference*), or it obtains its identification scheme via a subtyping connection to another supertype.".

p. 519, para. 4, line 6: "record" ⇨ "records".

p. 519, Replace last paragraph by: "In this example, we chose employeeNr as the preferred identifier for student employees in the context of roles specific to StudentEmployee. This choice is shown by the solid subtyping arrow from StudentEmployee to Employee (indicating a path to the subtype's preferred identifier). The subtyping connection from StudentEmployee to Student is dashed, since this does not provide a path to the subtype's identifier. The subtype connections from Student and Employee are also dashed, because these subtypes introduce new reference schemes in preference to that of their Person supertype.".

p. 519, Fig. 11.44: Subtype connection lines to Person should be dashed.

p. 521, Fig. 11.47: Replace second instance of "Height" by "Width" (2 cases).

p. 529, Fig. 12.1: "(firstname)" ⇨ "(.firstname)"; "(name)" ⇨ "(.name)".

p. 533, Fig. 12.9: In the selection result, remove the female gender rows, and add a height column, leaving just these two tuples: (David, M, 172); (Terry, M, 178).

p. 543, Fig. 12.22: "(usd:)" ⇨ "(USD:)".

p. 545, para. 1: Remove indent in first line below subsection header.

p. 545, para. 2, first sentence: Replace the text after "result to " by: "those *X* entries in *A* whose associated *Y* entries in *A* include *all* the row entries of *B*.".

p. 546, para. 1: Remove indent in first line below subsection header.

p. 553, Ex. 12.1, Q. 5(a): "MoneyAmt(usd:)" ⇨ "MoneyAmount(USD:)".

p. 561, para. 2, line 4: "to." ⇨ "to".

p. 566, Fig. 12.36: "where", "or", "and" ⇨ : "**where**", "**or**", "**and**".

p. 567, Fig. 12.37: "where" ⇨ : "**where**" (2 occurrences).

p. 569, Fig. 12.40, last line of query: "startsign" ⇨ "starsign".

p. 569, para. 2, line 5: "younger" ⇨ "older".

p. 569, SQL syntax: "select * | [**all** | **distinct**] …" ⇨ "select [**all** | **distinct**] * | …".

p. 580, SQL syntax: "[**outer**] {**left** | **right** | **full**}" ⇨ "{**left** | **right** | **full**} [**outer**]".

p. 606, Table 12.13: "distinct" ⇨ "[distinct]" (six occurrences).

p. 606, para. 3: Replace by "The count function may be used in two ways. `Count` (*) returns the number of rows in the specified table, whereas `count` (colname) returns the number of values in the named column. If `distinct` is specified with a bag function, any duplicates are excluded. If duplicates are wanted, the keyword `all` may be used; however, since this is the default, it is often omitted. For obvious reasons, adding `distinct` to `max` or `min` has no effect."

p. 607, end of third set of code examples: "Pupi   l" ⇨ "Pupil".

p. 609, Ex. 12.9: "The Log table …" ⇨ "2. The Log table …".

p. 610, Ex. 12.9 Q. 2(a): "retail price " cost price" ⇨ "retail price - cost price".

p. 611, final para.: Indent first line.

p. 617, SQL syntax: "select * | [**all** | **distinct**] …" ⇨ "select [**all** | **distinct**] * | …".

p. 617, SQL syntax: "[**outer**] {**left** | **right** | **full**}" ⇨ "{**left** | **right** | **full**} [**outer**]".

p. 634, SQL syntax: "select * | [**all** | **distinct**] …" ⇨ "select [**all** | **distinct**] * | …".

p. 634, SQL syntax: "[**outer**] {**left** | **right** | **full**}" ⇨ "{**left** | **right** | **full**} [**outer**]".

p. 644, Schema diagram: "*Car* (empNr, carRegNr)" ⇨ "*Drives* (empNr, carRegNr)".

p. 648, Fig. 13.2; Add arrows from Ann to David and from Bill to Chris.

p. 646, Sec. 13.3, para. 2, line 6: "implemetation" ⇨ "implementation".

p. 687: "Introduction" ⇨ "Schema Equivalence and Optimization".

p. 698, Fig. 14.6: "Each S_1 corresponds to … $B = b_1$" ⇨ "Each S_i corresponds to … $B = b_i$".

p. 698, Fig. 14.16, Corollary 4: Change to: "An equality constraint over *C*'s roles in the RHS *implies* a frequency constraint of $\geq n$ on *C*'s role in the LHS. If the equality constraint applies *and* each of *C*'s roles in the RHS has a UC then this is *equivalent* to a frequency constraint of *n* on *C*'s role in the LHS combined with an external UC over the roles of *B* and *C* in the LHS (thus extending Corollary 2).".

p. 728, Fig. 14.49: "*Old*.feeyear" ⇨ "*Old*.feeYear".

p. 731, Fig. 14.52: Add a mandatory role constraint to Judge's role.

p. 731, para. 2: Insert before second last sentence: "The inclusive-or component of the exclusive-or constraints on the Player and Judge roles is implied in the global schema, so is ignored in the relational mapping".

p. 732, Fig. 14.54: Strengthen the exclusion constraints on the Judge and Player roles to exclusive-or constraints.

p. 742, para. 1: Replace by: Codd's original work on 2NF and 3NF classified attributes as "prime" or "non-prime" depending on whether or not they belong to some candidate key. We now use the terms "key attribute" and "nonkey

attribute" for these concepts. So a key attribute belongs to *some* key (primary or alternate), and a nonkey attribute belongs to *no* key.

p. 745, para. 1: Replace the text by the following, and move this new version to become the first paragraph on p. 742: The principle underlying the refinement to second and third normal forms has been nicely summarized by Bill Kent (1983) as follows: "A nonkey field must provide a fact about the key, the whole key, and nothing but the key". This should be refined by replacing the final occurrence of "the key" by "a key".

p. 790, Fig. 15.19: "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)".

p. 791, Fig. 15.20: "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)".

p. 792, Fig. 15.21: "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)"

p. 793, Fig. 15.22: reverse the direction of the subset constraint; "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)", "(code)" ⇨ "(.code)".

p. 794, Fig. 15.23: "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)".

p. 795, Fig. 15.24: "(name)" ⇨ (.name)"; "(id)" ⇨ "(.id)", "(code)" ⇨ "(.code)".

p. 820, line 4: "BPML" ⇨ "BPMN".

p. 829, Ex. 15.7, Q. 2 Figure: "Name" ⇨ "ItemName"; "(code)" ⇨ "(.code)", "(nr)" ⇨ "(.nr)"; "Report" ⇨ "Report(.nr)"; remove ReferenceNumber fact type.

p. 840, para. 3, line 1: "Star" ⇨ "Store".

p. 864, Sec. 16.6, line 4: "change" ⇨ "change?".

p. 865, para. 1, line 8: "identifers" ⇨ "identifiers".

p. 866, Fig. 16.20: "Terry Halpin" ⇨ "Terry Halpin".

p. 868, final para., line 1: "decarations" ⇨ "declarations".

p. 868, final line: "<owl:Restriction>" ⇨ "</owl:Restriction>".

p. 887, Ex. 16.8, Q. 4 figure: "birthplace of" ⇨ "the birthplace of"; add to Country, Buss Pass and Parking Bay "# * country code", "# * pass nr", "# * bay nr".

p. 900, Subtyping, first para.: "for *A*" ⇨ "for *B*".

p. 911, first entry: "www.mkp.com/imrd/" ⇨ "www.mkp.com/imrd2".

p. 911, Fact-Oriented Modeling Tools, NORMA entry: delete final ")".

p. 911, 3rd last line: "zifa" ⇨ "ZachmanInternational".

p. 914, Carver and Halpin reference: "'Semantic Normalization', Neumont University Technical Report." ⇨ "'Atomicity and Normalization', *Proc. EMMSAD2008*, eds, T. Halpin, J. Krogstie, and H. Proper, CEUR, Montpellier, June 2008."

pp. 914-924: Page headers should read "Bibliography".

p. 918: "Halpin, T.2000c" ⇨ "Halpin, T. 2000c".

p. 936, "popular reference mode, 84. 105" ⇨ "popular reference mode, 84, 105".