

# Errata and Updates For Database System Concepts, 6<sup>th</sup> Edition Silberschatz, Korth, and Sudarshan

Last updated: October 30, 2016

We list below errors, clarifications, and recent updates. NOTE: Some of the errata described here have been corrected in the 2nd and 3rd printing of the US edition of the book; these are tagged as “(1<sup>st</sup> pr.)”. If you own an international edition, note that these editions follow a different correction schedule, so your copy may still have errata that have been fixed in the US edition. Check your copy for the errata noted here, and ignore those that have been fixed in your copy.

## Errata for Parts 5 to 8, Chapters 17 to 26

### CHAPTER 20

1. (1<sup>st</sup> pr.) Page 897, definitions of Gini and Entropy: “ $\sum_{i=1}^k$ ”  $\rightarrow$  “ $\sum_{i=1}^k$ ”
2. (1<sup>st</sup> pr.) Page 898, definition of Information\_content: “ $\sum_{i=1}^r$ ”  $\rightarrow$  “ $\sum_{i=1}^r$ ”  
(Reported by: Daniel Sadogh Menasche)
3. Page 898, 2nd para: “All of this leads to a definition: The **best split** for an attribute is the one that gives the maximum **information gain ratio**, defined as: ”  
 $\rightarrow$   
“The **information gain ratio** of a split is then defined as:”  
Also add the following after the definition of Information\_gain:  
  
“Classifiers usually choose the split that gives the maximum information gain ratio. However, for binary splits, many classifiers choose the split that gives the maximum information gain, instead of the maximum information gain ratio. This is because the number of partitions is fixed, and moreover if a split results in most values being in one partition, its information content is very low, which can result in a high information gain ratio even with a small information gain.”
4. Page 898, para before the last para: “We then compute the information gain obtained by splitting at each value.”  
 $\rightarrow$   
“We then compute the desired metric (information gain or information gain ratio) obtained by splitting at each value.”
5. Page 898, 8 lines from page bottom: “ The best binary split for the attribute is the split that gives the maximum information gain.”  
 $\rightarrow$   
“The best split is the one that maximizes the desired metric.”

---

<sup>5</sup>**Errors reported by:** Deepak Aggrawal, G. Aishwarya, Jameel Al-Aziz, Scot Anderson, Yahui Chang, David Chiu, Jonghoon Chun, Matt Cremeens, Dona Dungan, Pham Nguyen Duc Duong, Helena Galhardas, Eric Gossett, Ravindra Guravannavar, Leon Ho, Pranav Jain, Jevitha K. P., Cheqing Jin, Minhua Kang, Celine Kuttler, Daniel Sadoc Menasche, Thimas Nielsen, Linda Null, Judi Paige, Donnie Pinkston, Subhasish Saha, Vemireddy Satish, Shan Shimin, Stan Thomas, Cam Hong Tran, Duc Tran, Daniel Vieira, and a few others. Their help, and in particular that of Daniel Sadoc Menasche, is deeply appreciated. Also thanks to Juha Haaga for suggestions on improvements for future editions.

6. Page 899, para 1, line 3: “that results in the maximum information-gain ratio”  
→  
“that maximizes the information gain, or information gain ratio.”  
(Reported by: *All above points related to best splits reported by Scot Anderson*)

## CHAPTER 21

1. (1<sup>st</sup> pr.) Page 918, line 1: “... where  $n(d)$  denotes the number of terms in the document ...” → “... where  $n(d)$  denotes the number of term occurrences in the document ...”

## CHAPTER 23

1. (1<sup>st</sup> pr.) Page 987, Figure 23.5, 3rd line from the bottom: “`coursr_id`” → “`course_id`”
2. (1<sup>st</sup> pr.) Page 992, Figure 23.10: Delete the “>” symbol at the end of “IID ID #REQUIRED >”
3. (1<sup>st</sup> pr.) Page 995, Figure 23.12: `<xs:complexType>` and `<xs:sequence>` elements should be added to enclose the subelements of the course element, in the same way as the department and instructor elements.
4. Page 1007: “`local:inst_dept_courses($i/IID)`” → “`local:dept_courses($i/IID)`”  
(Reported by: *Stan Thomas*)
5. (1<sup>st</sup> pr.) Page 1011, Section 23.6.2, first paragraph: “A tuple inserted in the *nodes* relation ...” → “A tuple is inserted in the *nodes* relation ...”.
6. (1<sup>st</sup> pr.) Page 1014, Figure 23.15, line preceding last line: “`<course>`” → “`</course>`”.

## CHAPTER 24

1. (1<sup>st</sup> pr.) Page 1058, Exercise 24.2(c): “... assuming your database supports index-only plans ...” → “... assuming your database supports covering indices (that is, indices that store extra attributes, in addition to the search key, at the leaf nodes) and index-only plans ...”
2. Page 1063, Figure 25.1: In the 2nd and 5th row, replace the value in the *to* column by “\*”.  
(Reported by: *Thomas Nielsen, who pointed out that none of the rows has a “\*” value.*)