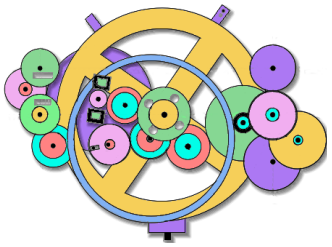
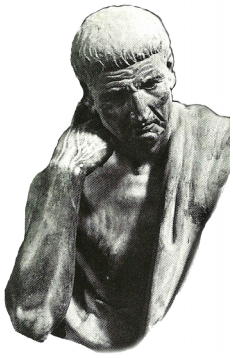
Antikythera Publications



Relational Database Design
http://www.AntikytheraPubs.com
asoler @ AntikytheraPubs.com

DATABASE DESIGN NOTE SERIES

Why Aristotle?



Prepared by: Frank Oberle and Antonio ("Tony") Soler

Information Technology professionals and so-called "Computer Scientists" are never lacking for something to occupy their time.

So why would anyone even suggest that time spent reading (much less studying) the writings of a grizzled old philosopher who lived over two thousand years before the computer was invented would aid them in their application and database designs?

You'd be surprised. Read on.

25 July 2014

See page 5 for information on other material from Antikythera Publications.

Copyright © 2014 by the Authors

Permission is granted to distribute unaltered copies of this document, so long as this is not done for commercial purposes.

THIS PAGE IS INTENTIONALLY BLANK

Database Design Note Series – Why Aristotle?

Preface

There is a wide variety of both printed and on-line material available that purports to provide guidance concerning the design of Relational Databases. As the book *Business Database Triage*[1] and others point out however, many of these sources offer questionable advice at best – some even propose approaches to database design that are almost guaranteed to cause trouble down the road as the business attempts to expand, or as its software developers begin their inevitable attempts to add new features or capabilities to their programs.

IT folks – enamored as they are by the technology that brought us from punched cards to SSDs[2] in less than one hundred fifty years – seem to believe that data management is an *Art* with an even shorter history, perhaps one that didn't come about until many businesses had already begun developing computer databases to support their operations. This idea of data management as an *Art*, much less a recently developed one, is a wildly mistaken and quite unfortunate belief.

History

Data organization techniques began well before the computer age! In fact, the Science[3] of data categorization and organization was mature enough by 350bce that Aristotle's treatise on this subject, the *Organon*, is still the most complete and exhaustive coverage of database design ever written.[4] Aristotle, unfortunately, wasn't able to obtain time on any early Athenian mainframes to give us an explanation of how we can (and should) apply these principles to modern relational database design, but chapter 5 of *Business Database Triage* will get you started.

When looking into Aristotle's legacy, it doesn't take long before you will encounter some commentary disparaging his scientific legacy to one degree or another. Aristotle never received a Nobel prize (forgetting that neither did Archimedes, Galileo or Newton). And it turned out that Aristotle was quite wrong in some of his conclusions and teachings but, given the astonishing breadth of his writings, that shouldn't be terribly surprising! The naysayers, displaying just a bit of envy or jealousy, seem to ignore that Aristotle's few errors – confined mostly, by the way, to the physical sciences – weren't revealed until the invention of such things as telescopes and microscopes, the availability of space travel, DNA[5] analysis, and other much later discoveries.

Not surprisingly, the *Organon* has been continuously available since the days of Alexander the Great, and can be easily located if you have any interest. The *Organon* is, to be sure, a tough slog, even with an English translation, but once Aristotle's approach (explained by examples in *Business Database Triage*) is grasped, logical solutions to many thorny database design issues become apparent – whether and under what conditions is it permissible to define a column as "NOT NULL" or the always thorny discussions about multiple inheritance, for example, all have logical answers (implied if not always explicit) in the Organon.

The Organon

The *Organon* is composed of six "Books" which, in the terminology of our own time, are equivalent to "chapters." In their common English translations, the titles of these are:

- **Categories**. Usually considered to be Book I.
- **On Interpretation**. Usually considered to be Book II.
- **Prior Analytics**. Usually considered to be Book III.

1 See page 5 for more information on this book.
2 "Solid State Disk" or "Solid State Drive" - two of more than one hundred fifty known meanings of this acronym.
3 Notice the shift from the word "Art" to the word "Science."
4 With all due respect to Codd, Date, Barker, Hay and the many other authors we recommend.
5 Deoxyribonucleic Acid – one of more than two hundred known meanings of this acronym.

- **Posterior Analytics**. Usually considered to be Book IV.
- **Topics**. Usually considered to be Book V.
- **On Sophistical Refutations**. Usually considered to be Book VI.

It isn't at all certain that the order shown above is that which Aristotle intended. Understanding Book I (Categories) is certainly easier for someone who is already familiar with Book V (Topics), but then the whole point of Topics isn't really evident without being familiar with Categories. Such is the process of cumulative learning.

We recommend beginning with **Categories** (Book I), particularly if this can be done in conjunction with Chapter 5, titled "Grammar, Sets, and (Predicate) Logic," a page from which is shown to the right.

Aristotle tells us several things about **Secondary Substances** in 2.2 of his Categories:

...that we predicate (that is, infer the existence of) a general Class such as "Person" (a Common Noun) from the existence of specific examples of a Person (such as "Lewis Carroll" listed above).

… that any Class (Secondary Substance: a Set described by a Common Noun), and in fact anything that is not a Primary Substance, is "predicable of" (that is "can be inferred from the existence of") at least one (even theoretical) instance of a Primary Substance.

… that it must be true that any Primary Substance (one identified by a Proper Noun) can be categorically classified as either "being in" or "not being in" the Secondary Substance (identified by a Common Noun), and if this isn't possible, the Set doesn't qualify as a Secondary Substance or a Class.

▼ ***Common Nouns*** identify a general class of persons, places, or things (for example "employee," "city," "building," "plane," "team"). In Logic, or in Database Design specifically, this type of Noun can refer to either a SET or a CLASS (we'll defer the important distinction between the two until we get to Collective Nouns below). In addition to the commentary on the left, Aristotle also tells us that, in order to qualify as a Secondary Substance (a "Thing" that we identify with a Common Noun), the "Thing" **must have no possible contrary** other than the theoretical set "not a Thing." For example, there is no contrary to "Person" other than the artificial "not-a-Person" set – which includes every other possible existing "Thing"; if such a contrary exists, the Noun is not a Common Noun, the Set being described is not a Class, and it doesn't represent a Secondary Substance. While there may be exceptions based on specific business reasons, Classes identified by Common Nouns should always be implemented as tables within a database unless there is a very good reason for not doing so.

Because of its critical importance to database design, assigning Names to Common Nouns will be discussed separately in "Naming our Things" below.

Aristotle Categories 2.3

Any Set is a "collection," and never a "composition." This of course applies to Classes and Common Nouns, since they are Sets. A Person may have two arms, two legs, and so forth, but such a list of body parts defines a composition, not a collection. A Person missing a leg is still (logically and legally) a Person.

▼ ***Collective nouns*** give a name to identifiable groups, but groups that are ultimately arbitrary or transient. A Collective Noun may refer to a group of persons (but not all persons), places (but not all places), or things (but not all things). In Logic, a Collective Noun always describes a Set, but since it doesn't necessarily include all members of the Set, it doesn't represent a

the place to discuss such arcana.

[Page 70] Chapter 5 - Grammar, Sets, and (Predicate) Logic – Part 1

Obtaining the Organon

The various books of the *Organon* are widely available on-line for reading and download, although some sites offer translations that vary in quality, and not all sites have the complete set. Our favorite site for obtaining good copies of these is the ePub collection of Australia's University of Adelaide, at:

https://ebooks.adelaide.edu.au/a/aristotle/

This collection of Aristotle's works, although incomplete, gives a good indication of the range and breadth of his interests. The various books of the Organon appear near the top of the list, and so are easy to locate.

Aristotle's Organon is a rather dense work that is probably less than suitable for casual reading at a personal computer. If you don't have a hand-held e-Pub Reader, you may want to consider any one of a variety of printed versions, although some only contain a few sections of the *Organon* – typically in a section titled "Logic" or something similar. For those with more than a casual interest in the application of the Science and Logic behind data organization to modern database design, our recommendation is the following inexpensive printed copy[6] of the full text in English translation:

https://www.createspace.com/3923345

This version (ISBN-13: 978-1478305620) is typically priced below $15 USD, and may also be ordered from most other U.S. and European on-line and physical booksellers.

But then, if you want just a simple introduction to Normalization, and how to form and use Predicates and Propositions[7], the mathematician Lewis Carroll (yes, *that* Lewis Carroll![8]) wrote a book titled *Symbolic Logic* that explains these particular issues in a more straightforward fashion than Aristotle did. It too has never been out of print, and is not difficult to locate.

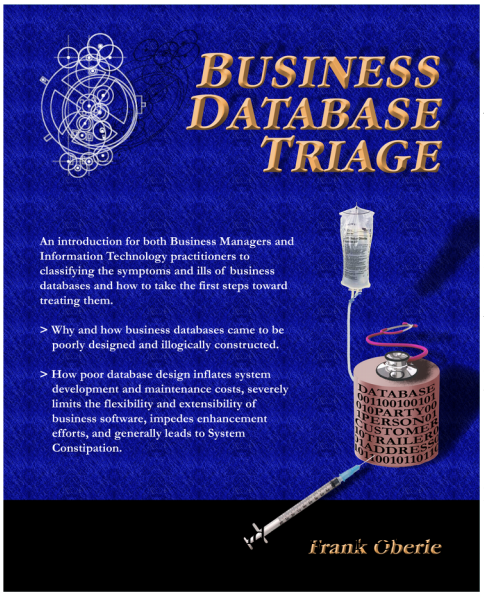
6 The producers of ISBN-13: 978-1478305620 have no relation to Antikythera Publications whatever, although we use the same printer.

7 Many IT practitioners seem quite surprised to learn that the words "Normalization," "Predicate Logic," and similar terms actually predate Ted Codd and the Relational Model by more than a century.

8 Carroll actually met Charles Babbage in 1867 but, by that time Babbage's chief programmer Ada Augusta Byron had already passed away. Had Lewis and Ada had the chance to work together, we'd likely be well into quantum computing by now, and the most advanced programming languages we have today would seem beyond quaint.

Other Publications

Antikythera Publications



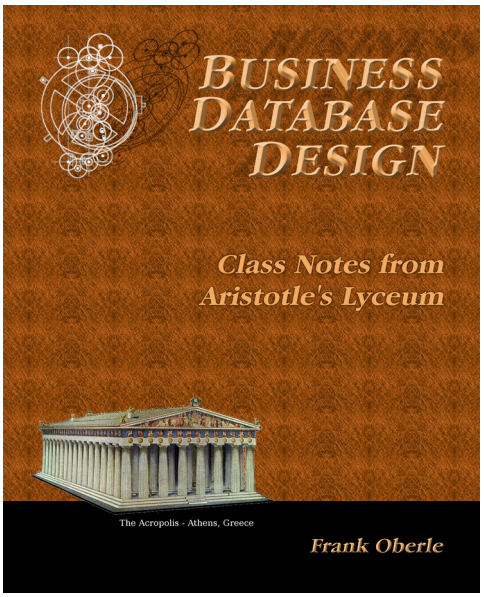
In addition to an ongoing series of Database Design Notes, Antikythera Publications recently released the book "*Business Database Triage*" (ISBN-10: 0615916937) that demonstrates how commonly encountered business database designs often cause significant, although largely unrecognized, difficulties with the development and maintenance of application software. Examples in the book illustrate how some typical database designs impede the ability of software developers to respond to new business opportunities – a key requirement of most businesses.

A number of examples of solutions to curing business system constipation are presented. Urban legends, such as the so-called object-relational impedance mismatch, are debunked – shown to be based mostly on illogical database (and sometimes object) designs.

"*Business Database Triage*" is available through major book retailers in most countries, or from the following on-line vendors, each of which has a full description of the book on their site:

CreateSpace: https://www.createspace.com/4513537

Amazon:
www.amazon.com/Business-Database-Triage-Frank-Oberle/dp/0615916937



A follow-up book, "*Business Database Design – Class Notes from Aristotle's Lyceum*" is due to be available in the early part of 2015.

"*Business Database Design*" leads the reader through the logical design and analysis techniques of data organization in more detail than the earlier work – which concentrated more on understanding and identifying problems caused by illogical database design rather than their solutions.

These logical approaches to data organization, espoused by Aristotle and an "A-List" of his successors, have formed the basis for scientific discovery over more than 2,400 years, and directly led to the technology we deal with today, notably including both relational and object theory.

"*Business Database Triage*" explained the reasons why these principles were virtually impossible to apply during the early years of our transition to the use of computers in business, but since the technology is now sufficiently mature that such compromises can no longer be justified, the time has come to relearn logical data organization techniques and apply them to our businesses.

To download the ERD_A TrueType Font used in this document, along with a tutorial and keyboard map, visit www.AntikytheraPubs.com, where a variety of other Database Design Notes are also available for download.