Introduction to the Fact Ledger

Charles Hoffman, CPA (Charles.Hoffman@me.com)

Andrew Noble, PNA, BBus (Andrew@nobleaccounting.com.au)

October 8, 2018 (DRAFT)

ABSTRACT: Accountants are familiar with the notion of the ledger and in particular the general ledger. This paper explains the utility of what I am calling the fact ledger. A fact ledger, or something that provides similar functionality, is necessary for any accounting process automation implementation. Intermediate information “containers” are necessary to connect processing steps and tasks. Rather than individual proprietary peer-to-peer implementations a standard fact ledger provides leverage. Think of how software applications stored data prior to the invention of the relational database; essentially each application had to write a custom data storage scheme. The notion of the fact ledger helps to solve a similar type of problem which exists in accounting process automation. Add the capabilities of a distributed digital ledger and fact ledgers become even more interesting. Essentially, an XBRL instance is a fact ledger. An XBRL Formula processor allows one to chain process steps together. Agreeing on the specific XBRL syntax alternatives and business logic (semantics) can create a standard fact ledger format.

Copyright (full and complete release of copyright)

All content of this document is placed in the public domain. I hereby waive all claim of copyright in this work. This work may be used, altered or unaltered, in any manner by anyone without attribution or notice to me. To be clear, I am granting full permission to use any content in this work in any way you like. I fully and completely release all my rights to any copyright on this content. If you feel like distributing a copy of this work, you may do so without attribution or payment of any kind. All that said, attribution is appreciated should one feel so compelled. The copyrights of other works referenced by this work are established by the referenced work.
A BRIEF HISTORY OF LEDGERS

Between 5,000 and 10,000 years ago farmers in Mesopotamia, where agriculture was born, used physical objects to count crops and animals\(^1\). The distinction between types of crops or animals was made by using different types and shapes of objects. Then, in about 3200 BC, around 5,000 years ago, the first spreadsheet was invented.

These farmers began documenting information using clay tablets in the earliest form of human writing ever discovered called Cuneiform. They partitioned their clay tablet into rows, columns, and cells. These farmers used single-entry accounting. The spreadsheet below documents an account of barley distribution\(^2\):

\[\begin{array}{|c|c|}
\hline
\text{Barley} & \text{Units} \\
\hline
1 & 2 \\
\hline
\end{array}\]

---

\(^1\) Denise Schmandt-Bessersat, *On the origins of writing*, YouTube.com, [https://www.youtube.com/watch?v=kidWPJFb0](https://www.youtube.com/watch?v=kidWPJFb0)

\(^2\) Metropolitan Museum, Proto-Cuneiform tablet with seal impressions: administrative account of barley distribution with cylinder seal impression of a male figure, hunting dogs, and boars, [https://www.metmuseum.org/art/collection/search/329081](https://www.metmuseum.org/art/collection/search/329081)
In 1211 AD a bank in Florence was the first documented use of double-entry accounting\(^3\). Between 1299 AD and 1300 AD double-entry accounting came of age. In 1494 AD during the Renaissance, Venetian mathematician and Franciscan friar Luca Pacioli\(^4\) published a book, *Summa de arithmetica, geometria, Proportioni et proportionalita* (Sum of Arithmetic, Geometry, Proportion and Proportionality)\(^5\). That book documented an approach to accounting now called double-entry bookkeeping and recommended that others use this approach. The approach allowed for better error detection and the ability to differentiate unintended errors from fraud. Accountants adopted that new approach.

Accountants have a special name for the spreadsheets, or tables, that these farmers invented. Accountants call these *ledgers*\(^6\). A ledger is simply a place where you record information such as transactions. There are a number of different types of ledgers that accountants use. General ledgers\(^7\) tend to be double-entry and capture numeric information. Special ledgers, or subsidiary ledgers, are also generally double-entry and are basically special purpose ledgers for recording specific classes of transactions. For example, a sales ledger is used to record sales transactions of an economic entity and a purchases ledger is used to record purchases of an economic entity.

There are many other types of ledgers used that are related to or unrelated to keeping the accounts of an economic entity. Some of these ledgers are public ledgers, others are private ledgers\(^8\). Most of these tend to be single-entry. Here are some examples of ledgers:

- Property titles
- Chain of custody
- Building permits
- Delivery records
- Digital rights management
- Insurance
- Proof of authenticity
- Anti-money laundering

Most recently there are discussions about immutable digital distributed ledgers using cryptographic technologies\(^9\). The notion of triple-entry accounting\(^10\) was first articulated in about 1998\(^11\). The

\(^8\) InterChainZ, Sharing Ledgers for Sharing Economies, YouTube.com, https://www.youtube.com/watch?v=Hwhigpr4720
\(^11\) Yuji Ijiri, Interview about Triple Entry Bookkeeping, YouTube.com, https://www.youtube.com/watch?v=7YE8lWI3tAA
blockchain community and others\textsuperscript{12} seem to have borrowed that term and changed what it means\textsuperscript{13}. Others can debate who owns the term “triple-entry”. Some say that triple-entry is the most important invention in 500 years\textsuperscript{14}. People are already experimenting with triple-entry accounting systems\textsuperscript{15} (if you watch the video, note the reference to the link to the XBRL document about 4 minutes into the video). The way the blockchain people tend to use the term “triple-entry” is taking two double-entry systems and putting a publically available link between the two double-entry systems documenting that the transactions go together. I don’t know if the idea has been proven, but intuitively it seems like a very useful tool.

The bottom line is that ledgers are useful tools particularly in the area of accounting.

**DEscribing the Fact Ledger**

In solving issues related to accounting process automation I stumbled upon what I am referring to as a fact ledger. A fact ledger is a ledger in that it is used to record information. The information that a fact ledger is used to record is facts. In my particular case I am interested in recording facts related to a financial report. But a fact ledger could be used to record a single fact, some set or sets of facts, or an entire database of information such as the complete set of all public company financial reports that have been submitted to a regulator such as the Securitas and Exchange Commission. Many different fact ledgers might be used in the process of creating a financial report.

Consider the graphic below which shows the steps, processes, and tasks that are involved in creating a financial report. The process accounts for transactions, actions, events, situations, occurrences, and other phenomenon/circumstances that impact an economic entity. Then, the economic entity communicates information in the form of facts about the financial position and financial condition of the entity to interested parties. Quantitative and qualitative information is reported in the form of facts. During the report creation process quality control checks are performed to detect and correct errors that may have occurred. Information comes from various sources and has intermediate stops as a financial report is constructed. The margin for error is low because there are risks related to noncompliance with statutory and regulatory reporting rules. Scrutiny is particularly high for public companies whose facts are there for all to see.

\textsuperscript{12} Ian Grigg, Triple Entry Accounting, http://iang.org/papers/triple_entry.html
\textsuperscript{14} Daniel Jeffries, Why Everyone Missed the Most Important Invention in the Last 500 Years, https://hackernoon.com/why-everyone-missed-the-most-important-invention-in-the-last-500-years-c90b0151c169
\textsuperscript{15} YouTube.com, DEVCON1: Balanc3 - Triple Entry Accounting, https://www.youtube.com/watch?v=NEYTpMoQy0
What is a fact? A fact is simply a single, reportable piece of information contained within a financial report, distinguishable from all other facts, contextualized for unambiguous interpretation or analysis by one or more distinguishing characteristics. Facts can be numbers, text, or prose. A fact can be financial or non-financial. A fact might have run through the general ledger or not. Generally facts are not used in isolation but rather in sets that make up the fragments of a financial report.

A fact ledger is simply a standardized approach to managing the facts that make up financial report. An XBRL instance is an instantiation of a fact ledger using global standard syntax. But a collection of XBRL instances can also be a fact ledger.

Consider the following. Last year, a set of 5,374 public companies I was measuring for high-level information quality reported their year-end financial information as part of their 10-K to the Securities and Exchange Commission using the XBRL format. Within those 5,374 financial reports public companies provided 7,948,552 individual facts. Some facts were numbers, some were text (such as a policy), and some were prose (text block such as an entire policy, disclosure, or note). But all information reported accounted for the almost 8 million facts which were reported in these public 16

17 For my own benefit I want to remind myself that this information is contained in the extract general Microsoft Access database stored here: D:\Master-2018\EDGAR-Probe.
company financial reports which are provided as part of their 10-K filing. All those facts are individually identifiable. Those facts were all reported using XBRL instances which is a global standard syntax for representing business information in machine readable form.

Observing those reports I noted common patterns of the information represented in the reports. This is important because these 5,374 XBRL-based reports can be used to verify the logical model of information contained in the reports.

Essentially, each XBRL-based report is a fact ledger and all 5,374 reports can be combined into what amounts to one fact ledger which contains all information from every XBRL-based financial report submitted to the SEC. Or, a fact ledger could contain a single fact.

Fact ledgers are described by metadata. Documentation and references describe the meaning of each fact, labels provide additional details to humans perhaps in multiple languages, and business rules describe the numeric, mechanical, structural, logical, and other relations between facts. Metadata that describes this information is stored within an XBRL taxonomy. The facts and the metadata that describes the facts forms a knowledge base.

XBRL Formula processors have the capability of chaining XBRL instances together. I see this functionality as being similar to how SQL queries can be chained together.

**WORKING PROTOTYPE OF A FACT LEDGER**

While I am a certified public accountant and not a programmer, I can program to a degree. I used the skills that I do have to implement a working proof of concept of a fact ledger. While I implemented this instance of a fact ledger in a SQL database, Microsoft Access, I believe that a better implementation can be created within a graph database. My implementation included one primary table for the fact ledger and three lookup tables: one for noncore dimensions, a second for context information for each fact, and a third for unit information for each fact. Note that part of the reasoning behind this specific implementation is my limited ability to program.

XBRL instances can be generated from the fact ledger within the SQL database and transferred to some other fact ledger implementation. The XBRL is the export and import mechanism. Each fact ledger implementation will have the same meaning. The meaning in the exporting system, the XBRL instance transfer mechanism, and the importing system has the same meaning. That meaning is documented in the XBRL taxonomy.

---

18 An additional look up table is necessary to represent the one-to-many relation between a fact and it’s XBRL footnotes; I just did not implement that.
The following features have been implemented and tested within this working proof of concept implementation:

- Both double-entry transactions and single-entry information can be entered into the fact ledger. I have not addressed triple-entry at this point.
- XBRL Global Ledger formatted output and information formatted similar to XBRL instances which are submitted to the SEC by public companies can be output.
- While not implemented, creating the ability to import from either output format would be trivial to implement.
- XBRL dimensions, basically the ability to add noncore dimensions, is supported.
- Fact IDs are supported.
- XBRL footnotes are supported; however currently there is a limitation of only one XBRL footnote per fact. (To fully support XBRL, there should be a one-to-many relation between facts and footnotes rather than the current one-to-one relation.
- Numeric, text, and prose facts are supported.
- A basic interface for entering facts was created.

The XBRL taxonomy schema, XBRL linkbases, and XBRL formulas that make up the metadata that describes the facts, relations, and other rules related to the facts within the fact ledger can be created using standard off-the-shelf XBRL tools. However, specialized tools would be better and easier to use. The XBRL instance is simply generated from the information in the fact ledger database tables.

The following is a screen shot of the fact entry form:
The following are a number of examples of fact ledgers. Simple examples are shown first and the examples get more complex.

**XBRL Global Ledger Style, General Journal Entries** (different syntax, uses XBRL tuples as opposed to XBRL Dimensions which is used by all other examples)

1. Load this XBRL instance:
   

**XBRL Dimensions Style General Ledger Trial Balance**

1. Load this XBRL instance:
   
   [http://xbrbsite.azurewebsites.net/2018/Prototypes/Basic/TrialBalance/basic-SampleInstance-GeneralLedgerTrialBalance.xml](http://xbrbsite.azurewebsites.net/2018/Prototypes/Basic/TrialBalance/basic-SampleInstance-GeneralLedgerTrialBalance.xml)
**XBRL Dimensions Style Basic Financial Report**

1. Load this XBRL instance: [http://xbrlsite.azurewebsites.net/2018/Prototypes/Basic/Basic-XASB-ConsistentRF/Basic-SampleInstance.xml](http://xbrlsite.azurewebsites.net/2018/Prototypes/Basic/Basic-XASB-ConsistentRF/Basic-SampleInstance.xml)

2. Use this schema for reporting checklist and disclosure mechanics validation:

3. Report fragments and business rules results:

4. Reporting checklist and disclosure mechanics results:

**XBRL Dimensions Style More Complex Financial Report (XASB comprehensive example):**


2. Use this schema for reporting checklist and disclosure mechanics validation:

3. Report fragments and business rules results:

4. Reporting Checklist and Disclosure Mechanics results:

**XBRL Dimensions Style Actual Report Submitted to SEC by Microsoft**

1. Load this XBRL instance:

---

2. Use this schema for reporting checklist and disclosure mechanics validation:

3. Report fragments and business rules results for Microsoft:

4. Reporting checklist and disclosure mechanics results for Microsoft:

While each of these XBRL-based reports is 100% consistent with the XBRL technical specifications\(^\text{20}\), the XBRL technical specifications allow for different syntax alternatives. When, say, a regulator implements XBRL, the regulator specifies what amounts to a profile which indicates how XBRL-based reports are to be created when submitted to that specific regulator\(^\text{21}\).

But what if someone simply wanted to use XBRL and not submit information to some specific regulator. How would they create their XBRL-based reports? What exactly would guide the creation of the report given the alternatives allowed by the XBRL technical specifications?

The answer is no one; the creator of the report can pretty much do whatever they want as long as they stay within the guidelines of the XBRL technical specifications. But then ask yourself, why do the regulators that implement XBRL create profiles?

To say this in another way; what if someone created a general profile that anyone could use? What if that general profile was world-class in terms of design? What if the general profile took the best ideas of other profiles and used those ideas; but avoided all the areas where other XBRL implementations were experiencing problems? What if every aspect of the general profile was proven to work per deliberate, rigorous testing? Well, that was the purpose behind what I called the General Business Reporting Profile\(^\text{22}\).

The fact ledger standardizes the XBRL technical syntax and allows for a mechanism for defining the necessary business semantics for a fact ledger\(^\text{23}\).

---


\(^{21}\) XBRL-based Digital Financial Reporting Profiles and General Business Reporting Profile,

\(^{22}\) General Business Reporting Profile,

\(^{23}\) I am trying to get an implementation of transactions in a blockchain; something like this https://etherscan.io/tx/0x92bf246ea182a5275ac2f3d213f456b7ab149de616acb9516ee081620a509e and this http://certificates.b9lab.com/certificate.html?uuid=1d7c2be5-f9f3-45ed-aeb1-b397f2651136
BENEFITS OF A FACT LEDGER

One primary advantage of a fact ledger is a standard format for storing information that will be used in an accounting process automation project. Ultimately information has to be stored somewhere. What is stored in the fact ledger is information, not data. Databases feeding something such as the information which would end up in an XBRL instance based financial report comes from a wide variety of databases which store data. Some databases are internal to an organization, other information is external to an organization. What the fact ledger does is provide one common intermediate format that stores all information and all of the characteristics necessary to make use of the information within an accounting process automation project.

While I have not implemented all of these features necessary in this working proof of concept, it is contemplated that a fact ledger would have the following features.

- Information as to the origin or provenance of all facts in the fact ledger would be available.
- A full audit trail of all information in the fact ledger would be available.
- Text, numbers, and prose data types are provided for.
- Single-entry, double-entry, or triple-entry functionality would be available.
- Information could be output into raw XBRL, Inline XBRL, and likely RDF format; with additional converters information might be directly output to HTML, PDF, Word, Excel, and other such human readable formats.
- Because information being worked with is at the appropriate level, accounting process automation is easier and an organized, consistent format can be worked with rather than ad hoc information formats.
- Because information being worked with is at the appropriate level, individual work flow steps are easily created.
- Because information being worked with is at the appropriate level, applying process control and monitoring techniques such as Lean Six Sigma are easier.
- Human readable reports are dynamic pivot tables rather than static two dimensional reports.

STANDARDIZED TRANSACTION SEMANTICS

A chart of accounts represents a model of the line items that are contained within a financial report. Not within the chart of accounts, but within other information which is used with that chart of accounts, the hierarchy into which those line items fit to make up the financial report is somehow provided. This financial report “tree” and “branches” and chart of accounts “leaves” come together to form a financial report model. Different economic entities have different financial report models but those

---

24 A financial report is actually a “graph” per network theory, but most accountants don’t understand the important difference between a tree and a graph.
25 Comparing Reporting Styles, Notes on Comparability,
financial report models follow common patterns. I call these report model patterns reporting styles\(^2\)\(^6\). These patterns are observable and explainable.

What many accountants tend to do is create a chart of accounts not considering the formal reporting style model that financial reports follow and then during the financial report creation process; because the established chart of accounts does not match the formal financial report model that they must follow; they have to make modifications to the roll up of information because of this miss-match. Other sorts of issues pop up also that are generally caused by being unconscious of how computers work with their chart of accounts or deficiencies in an accounting systems ability to create an adequate chart of accounts or issues related to the mechanisms used to create report hierarchy.

When considering an accounting system in isolation (which most accountants do), using a formal model might not make much sense. But, when you consider the bigger picture and that someone ultimately uses the information within that accounting system; the benefits of a formal model becomes more evident.

Think of the world of digital distributed shared ledgers. Imagine trying to actually share information on those ledgers. Can accounting system “A” built by one programmer “A” read and understand data from accounting system “B” built by some different programmer “B” and that other accounting system make sense of the information from accounting system “A”? The answer is clearly NO if the two accounting systems don’t share the same formal model of if all relevant context is not provided.

A formalized model provides leverage. For example, logical deduction can be used to find missing information if a formal model is used. Validation and transformation rules can be adopted across disparate data sets that meet requirements established by the formal model.

So you don’t think this is a problem or issue? Consider this. Several years ago I visited United Technologies and had a day long discussion with their external financial reporting team. United Technologies, at the time, consolidated the general ledgers of, at the time, about 900 subsidiaries into one financial report that is submitted to the U.S. Securities and Exchange commission. To each of the subsidiaries that have their ad hoc general ledger chart of accounts, there is no problem here. But consider the perspective of the external financial reporting department of United Technologies. For them this is a big problem and a lot of manual effort to fit the 900 charts of accounts together appropriately.

And here is another situation to consider. The U.S. Securities and Exchange Commission has about 10,000 public companies reporting information to them. I am not interested in “funds” or “trusts” or companies that report using “IFRS” (at the moment). I filtered this list to about 6,000 public companies that I have been measuring and analyzing. Imagine if you are the SEC and you had to analyze all of those companies and you did not have a formalized model which was used to create the individual financial

reports? How exactly would you achieve the ability to compare across about 6,000 public company reports?

To be very clear, I am not saying that financial reports need to be one common form. There are many different financial report models. However, I can prove that 68% of all public company financial reports fit into 13 reporting styles for US GAAP reports. If you can extract information from the reports using that model that is rather compelling evidence of the existence of the model. An additional 40 financial report models gets you to about 92% of all public companies. It is unclear exactly how many reporting styles the remaining 8% fit into but it is likely between perhaps 100 and up to 400 patterns if each company reported uniquely which is unlikely. The analysis is incomplete at this point which would clear this up.

As I said financial reports are not random. Formal reporting models do exist. I have been measuring financial report consistency with those models for three years for US GAAP. These models exist for both US GAAP and IFRS.

Entropy is defined as a "lack of order or predictability; gradual decline into disorder". The opposite of entropy is occurring with these XBRL-based US GAAP financial reports of public companies. Why? This is because of US GAAP, the rules of financial reporting. Imperfections in US GAAP are being revealed. Imperfections of the understanding of professional accountants that create reports of US GAAP are being revealed. Imperfections in the representation of information using XBRL are being revealed. The issues are being resolved, XBRL-based reports of these 6,000 companies are becoming even more consistent, reports are becoming more consistent, and the patterns are becoming even clearer making the financial report models even easier to identify. Entropy is being reduced.

A fact ledger provides us with a framework for decreasing the level of entropy that typically erodes our financial knowledge when data is simply stored in traditional accounting systems that are not built on a formalized model. Formalized standardized fact ledgers or fact sets can be used around the globe.

The Object Management Group (OMG) and the financial services industry are leaders in this area. Ultimately the information collected within an accounting system will be used by someone and generally that someone will be outside that specific accounting system.

A fact ledger full of standardized information that can be explained by a formalized model is better than thousands, or millions, of ad hoc representations of information simply because accountants are

---

unconscious that others actually do want to make use of all that accounting information. Add to that, making the fact ledger available within a distributed ledger provides additional functionality.

CONCLUSION

Ledgers are tools. There is nothing natural about the ledger. A ledger was an invention of man. The columns in a ledger are abstractions. The numbers and other information that go into the columns are symbols. The ledger is a useful idea. Professional accountants have used ledgers for many, many years. A fact ledger is a new type of ledger that offers leverage.

A single-entry ledger is not much more than a glorified list. When you use a double-entry ledger what the transaction represents has to be explained by reasoning. A triple-entry ledger links two double-entry ledgers and is publically available for all to see the transaction still needs to be explained by reasoning but additionally the transaction is there for all to see which makes it very tough to lie since others are watching.

While it is easy to talk about the possibilities offered by accounting process automation using techniques such as structured information and artificial intelligence\(^\text{32}\), these ideas need to be implemented in some manner to actually make them work. Each implementation creating proprietary individual peer-to-peer solutions can be effective but might not be the most efficient approach, particularly when others use make use of your information. In such cases, global standard solutions such as the fact ledger should be considered.

ACKNOWLEDGEMENTS

We have done our best to figure out the fact ledger; we speculate that others will come up with good ideas to make this even better. I would like to thank the following individual that had a direct influence in the ideas behind the fact ledger. Rajib Doogar (doogar@uw.edu) is an accounting professor at the University of Washington Bothell Campus. His ideas about looking at the general ledger as “trees”, “branches”, and “leaves” and explaining the matrix algebra behind these relations was very helpful. Raynier van Egmond (raynier@xbrlc.com) is both a software engineer and I would say a business analyst. I met Rene at the first XBRL International meeting 20 years ago and we have been collaborating ever since. Hamed Mousavi (hamedmousavi@yahoo.com) is a software engineer with a very good background in accounting. Hamed and I have been collaborating to create the Pesseract application for five years. Andrew Noble (Andrew@nobleaccounting.com.au) is an accountant in Australia who created the LODGEiT system for submitting XBRL information to Australian regulators and the AccFx accounting software system which is built using Wolfram’s Mathematica. David Eichner, CPA (david@icansoftware.com) sells and installs accounting software and has been a long-time collaborator and colleague and has worked to implement what amounts to a fact ledger in Microsoft Dynamics and Acumatica. Ghislain Fourny (ghislain.fourny@gmail.com) is a data scientist who I met when doing a database project relating to XBRL for a company he used to work for, 28msec. His ideas related to cell stores and NOLAP where helpful to understanding fact ledgers.