Application of Quantum-Bayesian analysis of block chain and textual data in Reverse Oracle Transforms for revealing hidden identities

C. J. Plooy (corne@bitonic.nl)

april 01, 2019

Abstract

Using a quantummechanical generalization of logical consistency of the truth as a starting point, we introduce an abstract definition of a universal Truth space. Then, we show how, using Bayesian analysis, the truth can be approximated arbitrarily closely, given sufficient data. This is then used to reverse the relationship between a black-box oracle model and the truth, yielding a Reverse Oracle Transform function to determine the answer to any question, based on the supplied input data.

As a demonstration, several implementations of the Reverse Oracle Transform function were created to analyze the available data on Satoshi Nakamoto, the creator of Bitcoin. Using these implementations, Satoshi Nakamoto's identity is revealed. As a side-effect, the Bitcoin fork that most closely resembles Satoshi's vision is also identified.

Contents

1 Introduction

There have been several attempts at revealing the identity of Bitcoin's creator, Satoshi Nakamoto[1, 2, 3, 4, 5, 6, 7, 8, 9]. However, none so far has proven to be conclusive, and the results contradict each other. What is really needed is a generic method for processing all available evidence.

2 The Truth space

All knowledge about the truth can be represented using boolean statements. For instance, knowledge about the value of π can be represented using statements like $\pi > 3.1415$ and $\pi < 3.1416$. There are an infinite number of statements (both true and false), and the same information may be conveyed by several (possibly infinite) different statements (redundancy).

Initially, we are not taking redundancies, contradictions and other logical relationships into account. So, initially, all permutations of truth values of these boolean statements must be considered.

The Truth space is defined as the infinite-dimensional space that has one axis for each permutation of truth values. For instance, if there were only two statements A and B, this would be a four-dimensional space, with the axes being $\neg A, \neg B$, $\neg A, B$, $A, \neg B$ and A, B. For N boolean statements, there are 2^N dimensions.

If there is such a thing as "the truth", then it holds one position in this space. Traditionally, any boolean statement is considered either to be true, or to be false. So, the possibilities of it being true and false simultaneously, as well as it being neither true nor false are excluded. In quantum mechanics, the situation is different. One way of interpreting quantum mechanics is in terms of probabilities of observed values: the sum of probabilities of finding it to be true and finding it to be false add up to one. Still, under this constraint, in quantum mechanics the truth can be any linear combination of true and false. As is common in quantum mechanics, we distinguish between the amplitude along a certain axis (which is the truth position projected onto that axis) and the probability of observing that value; the probability being the square of the absolute value of the amplitude.

A geometrical visualisation of such infinite-dimensional space is hard to make, and even low-dimensional analogues are hard to find: even to illustrate the interaction between only two statements, we need at least four dimensions. For a single statement, two dimensions are sufficient ($\neg A$ and A). Classically, only unit-positions on the axes, such as $(1, 0)$, $(-1, 0)$, $(0, 1)$ and $(0, -1)$, would be permitted, with the first two resulting in a probability of one for "true" and zero for "false", and the other way around for the last two. Quantummechanically, all positions on the unit circle are permitted. For instance, $(1/\sqrt{2}, -1/\sqrt{2})$ results in a probability of 1/2 for "true" and 1/2 for "false". In higher-dimensional spaces, this relation between the truth and falsehood of the same statement gets extended across all permutations, so that the truth position must lie on a

hyper-sphere with radius one.

2.1 Notation

We can start with the assumption there is some deterministic canonical ordering of statements, for instance by writing them out in English or with mathematical symbols or whatever may be appropriate, and sorting them alphabetically. This also offers a deterministic canonical ordering of coordinates, for instance using standard binary counting.

The truth T may be seen as an infinite-dimensional vector. Alternatively, using the canonical coordinate-ordering, the truth may be seen as function $T(n)$, which gives the n-th component of the truth vector as a function of the coordinate number n. The probability that the corresponding permutation of statement values is found is then $p(n) = |T(n)|^2$. Of course, the sum of all probabilities is one: $\sum_n p(n) = \sum_n |T(n)|^2 = 1$.

2.2 Including logical relationships

A logical relationship between two statements simply means that the amplitude of certain permutations is zero. For instance, if two statements A and B are notationally different, but convey exactly the same information, there truth statements must be equal. In the simplified four-dimensional case$\neg A, \neg B$ ($n = 0$), $\neg A, B$ ($n = 1$), $A, \neg B$ ($n = 2$) and A, B ($n = 3$), this results in $T(1) = 0$ and $T(2) = 0$. Effectively, this reduces the space in which the truth is to be found to a lesser-dimensional space: in the example, this would be the space with the axes $\neg A, \neg B$ and A, B.

2.3 Extracting statement values

We define $S(A)$ as the set of coordinate numbers, for which the statement A is true in the permutation corresponding to that coordinate number. Naturally, $S(A)$ and $S(\neg A)$ are each other's complements: if U is the set of all coordinate numbers, then $S(A) \cup S(\neg A) = U$.

The probability of finding statement A to be true is then $p(A) = \sum_{n \in S(A)} p(n)$.

3 Bayesian analysis

Bayesian analysis can be used to place constraints on the probability values of certain permutations. Bayes' theorem states that

$$p(H|E) = \frac{p(E|H)}{p(E)} p(H)$$

Given a piece of evidence E_i,We can use this to update an a-priori truth estimate $Q_i(n)$ to a new estimate $Q_{i+1}(n)$. For this, each individual axis is treated as the hypothesis H_n. This gives

$$Q_{i+1}(n) = p(H_n|E_i) = \frac{p(E_i|H_n)}{p(E_i)}p(H_n) = \frac{p(E_i|H_n)}{\sum_n p(E_i|H_n)}Q_i(n)$$

The only thing left to do is to find $p(E_i|H_n)$. Luckily, since all statements are included in the truth space, E_i is too. This yields the trivial result

$$p(E_i|H_n) = \begin{cases} 1 & n \in S(E_i) \\ 0 & n \notin S(E_i) \end{cases}$$

Given sufficient evidence E_0,E_1,E_2,..., the sequence Q_o,Q_1,Q_2,... will converge to the truth T.

4 Reverse Oracle Transform functions

If we had some oracle function $O(A)$ that returns whether or not statement A is true, we could learn the truth of any statement very easily, simply by asking the oracle. Unfortunately, we have no a-priori access to such a function, at least not in its most generic form.

The key insight here is that there exists a transformation from oracle function O to truth T. Even though both are unknown a-priori, the transform itself is straightforward. So straightforward, in fact, that it can be reversed: given the truth T, we can reconstruct the oracle function O. This is the Reverse Oracle Transform.

We can apply this Reverse Oracle Transform to incremental truth estimates Q_o,Q_1,Q_2,..., to yield oracle functions $O_0(A)$,$O_1(A)$,$O_2(A)$,....

The precise details about generating these oracle functions depend on the evidence data provided (E_0,E_1,E_2,...). For the application of revealing the identity of Satoshi Nakamoto, two such transforms were developed: the Cryptographic Asset Release Reverse Oracle Transform (CARROT) and the Textual Analysis Reverse Oracle Transform (TAROT).

4.1 CARROT

The Cryptographic Asset Release Reverse Oracle Transform looks at data related to the mining of cryptocurrency, in this case Bitcoin. For each Bitcoin block, available data is collected, like

- The block height
- The timestamp
- The nonce value
- Any source of miner identity information, such as self-identification inside the block or on the internet (found by searching for the block ID), and the spending taint of the coinbase transaction

The implementation defines a number of logical consistency and Bayesian analysis elements. For instance, a block cannot be mined by one party and also be mined by another party. Two blocks with the same parent block are not likely to be mined by the same party, as that would amount to a miner orphaning itself. Also, a consistent time-rate of nonce increments between blocks makes it likely they are generated by the same miner.

4.2 TAROT

The Textual Analysis Reverse Oracle Transform gathers evidence from textual sources. This involves not only writing style, but also timestamps (people are likely to have a more-or-less fixed sleeping cycle, which reduces the likelihood of certain combinations of identities), and also the semantics of the writing, using the work of von Ludwig[10]. Sub-modules were created for natural language (English) and source code (C++). In the case of source code, things like the choice of programming language, libraries and architecture designs are automatically included due to the nature of the transform.

In this, it is taken into account that published meta-data, such as timestamps and claimed identities (such as forum accounts) are not necessarily reliable: they could be forged, for instance, by forum operators. We assume The Bitcoin whitepaper, the first version of the Bitcoin source code and anything signed with Satoshi Nakamoto's PGP key to be written by Satoshi Nakamoto, and include other evidence using logical constraints and Bayesian analysis.

The body of relevant evidence available for this Reverse Oracle Transform is quite large, since Satoshi Nakamoto might have published other writings in various other places, before or after the introduction of Bitcoin, on subjects not necessarily related to Bitcoin in any way. For this purpose, the Internet Archive was used.

5 Results

An initial run of the algorithm returned an obviously false result. This was fixed by adding the logical constraint that this paper itself cannot be written by Satoshi Nakamoto.

The next run of the algorithm returned a clear answer for the identity of Satoshi Nakamoto, together with a list of publications published by, and cryptographic assets owned by him[1].

The list of cryptographic assets holds few surprises - it roughly corresponds to what is commonly believed to belong to him. Interestingly, the final oracle considers it roughly 20% likely that Nakamoto lost access to the keys holding

[1] The oracle reveals Nakamoto being a single, male person. This may be an artifact from the initial assumptions though; it is very well possible he actually received significant support from other people, so even though the publications and cryptographic assets can be attributed to this person, potentially large parts of Bitcoin's actual invention may actually have been done by others.

his early mining profits. This result was obtained even though the oracle only returns boolean values, by asking the oracle for the truth of a statement of the form $p(X) > y$.

The list of publications was more interesting. Of course, this includes the initial set of works assumed to belong to him, such as the Bitcoin whitepaper. It also included the Bitcoin Talk forum posts that were already attributed to him. Later posts, under other pseudonyms now also identified as Satoshi Nakamoto, reveal the Bitcoin fork that truly has his support: that turns out to be Bitcoin Classic.

Earlier publications by Nakamoto prove that he has been involved in attempts to create a cryptography-based economy for a long time. A talk attributed to him[11] by the algorithm was published as early as 1987, under a different pseudonym; it does not yet include ideas like the creation of a cryptographic currency or the use of proof-of-work, possibly because certain cryptographic primitives were not yet widely know at that time.

Unfortunately, Nakamoto's true name cannot be determined using available information. It turns out he is a time traveler, traveling from the year 2140 back to november 5th, 1984, and using pseudonyms ever since. His sudden disappearance can be explained by him traveling back to the time where he came from.

References

[1] Andy Greenberg, Gwern Branwen, 2015, https://www.wired.com/2015/12/bitcoins-creator-satoshi-nakamoto-is-probably-this-unknown-australian-genius/

[2] Joshua Davis, 2011, https://www.newyorker.com/magazine/2011/10/10/the-crypto-currency

[3] Adam L. Pennenberg, 2011, https://www.fastcompany.com/1785445/bitcoin-crypto-currency-mystery-reopened

[4] Ted Nelson, 2013, https://www.youtube.com/watch?v=emDJTGTrEm0

[5] Alec Liu, 2013, https://motherboard.vice.com/blog/who-is-satoshi-nakamoto-the-creator-of-bitcoin

[6] John Markoff, 2013, https://bits.blogs.nytimes.com/2013/11/23/study-suggests-link-between-dread-pirate-roberts-and-satoshi-nakamoto

[7] Andy Greenberg, 2014, https://www.forbes.com/sites/andygreenberg/2014/03/25/satoshi-nakamotos-neighbor-the-bitcoin-ghostwriter-who-wasnt/#7d96da914a37

[8] John Biggs, 2014, https://techcrunch.com/2013/12/05/who-is-the-real-satoshi-nakamoto-one-researcher-may-have-found-the-answer/

[9] Ben Wiseman, 2014, https://www.newsweek.com/2014/03/14/face-behind-bitcoin-247957.html

[10] Stephen C. von Ludwig, 2019, http://www.elsewhere.org/pomo/1244924019

[11] Satoshi Nakamoto, 1987, https://www.youtube.com/watch?v=dQw4w9WgXcQ