Veritas Blackpaper

21 Jan 2019

A free, decentralised platform for provable and self-sustainable discourse.

I. Introduction

Truth has always been a popular topic of debate. Nowadays, you can even hear people talk about the death of truth, post-truth, or something of the kind. The sentiment, overall, remains overwhelmingly pessimistic, and you can feel it, as it's signalled from virtually everywhere. There's a reason for that, although it has little to do with death. We get exposed to insanely large amounts of unfiltered, doctored and manipulated information, tuned for advertisement and political propaganda on the daily basis. No wonder people get upset. And then truth again, could it be that someone nicked our truth away? Globalisation, post-secularism, social media—only but *some* of the suspects. Whoever's to blame, one thing is certain—it doesn't feel like we still possess truth as such. Not only is it unclear whether something can or can not—be proved to be true—it's questionable if certain things can even truly be articulated. Then again, what exactly is truth?

In the past, multiple theories of truth have emerged. Correspondence theory, for example, proclaims: "said judgement is true if it reflects the objective reality." You can never go wrong with this, if you know objective reality very well and can probe it reliably. Yet, in some sense, this articulation of truth falls victim to the simple idea that things aren't necessarily what they seem. And sometimes things seem precisely what they seem, as opposed to what they are. There's only so much you can expect from any observer of any sort of reality. In the theory of coherence, on the other hand, coherence with a set of beliefs is in itself a test of truth and therefore does not require correspondence. Even though there are doubts to whether if objective reality can be confined within a coherent set of beliefs as the complexity of the perceived reality goes up, it still works remarkably well. Till structures don't grow too big, of course.

"Tell me," Wittgenstein's asked a friend, "why do people always say, it was natural for man to assume that the Sun went around the Earth rather than that the Earth was rotating?" His friend replied, “Well, obviously because it just looks as though Sun is going around the Earth.” Wittgenstein then said, “Well, what would it have looked like if it had looked as though the Earth was rotating?”... According to Nathan Salmon, this anecdote is completely true.[1] And it wouldn’t be daring to say it’s also true in a way that it so effortlessly cuts into the Achilles’ heel of correspondence theory. The fault is, undoubtedly, in the language. And it’s barely surprising. Language is hard. There’s only so much linguistic and logical capability you can expect from any observer. So we work around that, we add more observers, contrapose all available accounts, work out emerging contradictions, and all is well—coherence fills the gaps. But how long can we keep adding more and more observers? Unfortunately, as the number of distinguishable accounts grows, so does the number of inconsistencies between them. Complexity is of course bound to go over the top, straight into the Achilles’ heel of coherence. (There’s a good reason we have two feet.)

Now, truth—is absolutely the matter of discourse. Moreso, for it to stand, the truth-seeking process must be efficient, and to a certain degree, scalable. So how do we scale it?... Poorly, at least insofar the death of truth manifests itself. Internet took us by surprise, by far surpassing traditional means of communication. It concealed overwhelming, by human standards, complexity—required to satisfy the needs—of what we now regard as a bare necessity. But make no mistake to think a bare necessity would suffice. Even though computers are quite good at computing, they are still are completely oblivious to language. For example, this allowed for computing to be successfully used in science, but not in scientific *discourse*. In recent years, the developer community addressed most of the issues, standing in the way of computing. Suddenly, there’s new technology being engineered—and it’s solid, robust software—that would now allow us to create systems that were in not-so-distant past considered inconceivable. Unlike a whitepaper—highly nuanced specification intended to inform—instead, this document explores these new, emerging forms of discourse.

II. The Black Box

A content feed (as in TV programme, social media feed, etc.) is the prime example of a widely used abstraction to conceal unbelievable amount of complexity. *Feeds* have arguably done more to sales and advertisement—due to machine learning and nuanced data analysis—than any other abstraction ever before. It’s hardly transparent or “free” in any sense of the word, but already it shows just how much we rely on computing as a whole. Yet the idea of using electronic systems in, say, political, or administrative discourses, is met with reluctance at best, and often put off the table immediately.

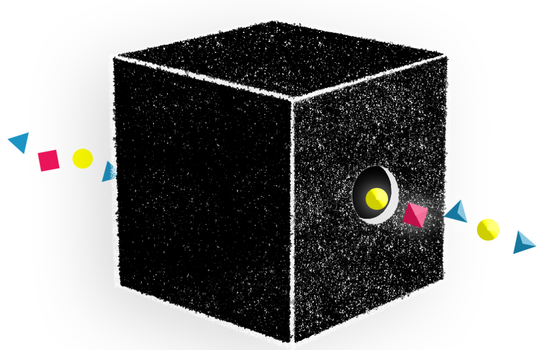


Fig. 1
Black boxes are able to conceal a lot of complexity.

Electronic voting, for example, is met with a great deal of suspicion, quite rightfully so. Decision-making being the cornerstone topic of all conversation. We don't trust the black boxes even remotely enough to allow the idea of computer systems to have extensive control over the means by which decision-making is carried away in regard of the corresponding discourses. We can't, for example, make government implement certain legislation, based on the fact that some excellent computer—maybe even capable of excellent analysis—would pronounce this legislation legitimate. Mainly, because the general public acknowledges that such computer would first have to be developed and then maintained, by someone. And it has to be allowed by the ones in actual power of course. Moreover, it's not like computers can do language, so these all are incredibly valid objections to-raise against the use of computers as far as decision-making is concerned.

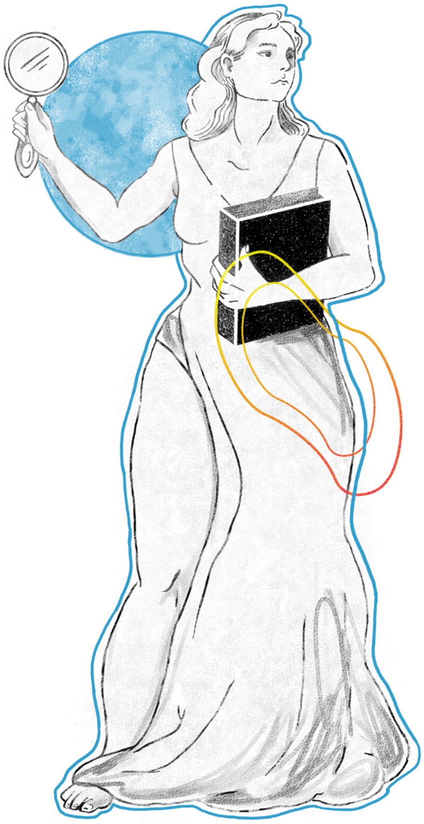
Fig. 2
The structure of decentralised boxes is easily inspectable.

Bitcoin proved in practice that its underlying technology, blockchain—is sufficiently reliable as means of computing even for something as sensitive as finance. In essence, blockchain is a data structure that is used to implement a secure, immutable, and decentralised database. *Secure* means you don't have to trust anyone for it to work. *Immutable* means it's write-only, once data is written in, it can't be tampered with later. *Decentralised*, finally, means there doesn't have to be a leader, and data can be consistently replicated, shared, and synchronised across multiple parties. This is called consensus. What consensus means, is that everyone can agree on the exact state of the affairs. It may be hard to believe, but we really couldn't quite do this before. And a lot of research and development was going to happen in the following years. Ethereum showed the possibility of even more sophisticated consensus networks. In fact, they went as far as to create a fully-fledged decentralised virtual machine, capable of running programs written in a Turing-complete programming language on top of it. Now, this allowed for computation to transcend the level of any single machine, or cluster, in a way that unlimited number of parties can now all unanimously agree on computation itself.

Whereas individual machines fall short at transparency, consensus provides much more solid, transparent, and safe ground for the black box foundation. From now on, in this paper—when referring to a black box—a data structure, algorithm, or even a complete application, powered by means of consensus, is assumed. Hence satisfying the aforementioned properties. Needless to say, when talking about consensus, one ought not to think of vanilla blockchain alone. There are alternative consensus algorithms offered by the likes of Hyperledger, Iota, Hedera Hashgraph, MimbleWimble—just to name a few. Countless black boxes tailored to specific problems are already there. What's unclear, on the other hand, is what is *the* box for the job and precisely what the system shall function like. No surprise nothing else but sustainability remains the biggest concern in regard to decentralised computation. In order to function freely, the box must be able to sustain itself in a reasonable, transparent, fashion. To achieve this, numerous "crypto" projects have gone with the way of token economics, in which the network is capable of producing value in the form of tokens of ownership, which can later be exchanged for other assets, or even fiat. And that is exactly what happens to Bitcoin, Ethereum, and other coins. The flat-out fiat value drives attention, but it also drives manipulation. As soon as fiat comes in play, any economic system itself encapsulates itself within a much bigger picture. The lack of actual intrinsic economic value that these projects offer—is another obvious obstacle on the way of free, decentralised computing to sustainability.

For any proper black box to function freely, it must to be intrinsically useful. In other words, not only the token in itself must represent value clearly, just so its direct or indirect long-time ownership can be justified, one ought to have a clear picture of how one would use it over time. Tokens must exhibit utility. Although this limitation adds a whole other level of design complexity, it also guarantees a certain degree of certainty to which the system will reflect the state of the affairs. Just how language limits what we can say, discourse limits what we can agree upon. (For that reason, when *what is the case* gets complicated, we proceed to simplify it.) To leverage decentralised computing for anything, to which decision-making is a concern, the premise has to be *agreeable* in the similar fashion the computing itself is agreed upon.

III. Veritas



In Roman mythology, Veritas, meaning truth, is the goddess of truth. Hence the name of a *free*, decentralised platform for provable and self-sustainable discourse, the subject of this paper. In Veritas, the whole of the critique is what determines the truth-value of text. With the help of the aforementioned black box concept, Veritas seeks to approach a diverse number of problems, such as fact-checking, proofreading and practical decision-making, at scale.

It soon becomes apparent that for this platform, due to its inherent epistemic complexity, to stand, or rather—survive—in the wild, some solid form of responsibility must be enforced upon its participants, or *actors*. However, hardly any form of responsibility can be expected to work at scale. Veritas interprets this necessity via personal, financial responsibility. Although it's still questionable whether large-scale discourse can be managed consistently, Veritas attempts to address this by making the very act of lying—as economically unviable, as possible: actors agree on the state of the network by engaging financially. In the context of newly-emerging non traditional economic systems, such as the ones discussed in the previous chapter, this objective no longer seems quite as impossible.

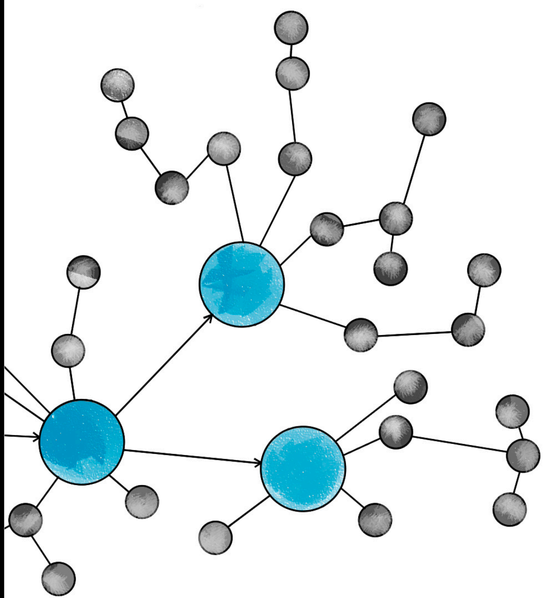


Fig. 3
Blue nodes represent knowledge, black-the emerging critique.

The foundation of Veritas lies in the *knowledge graph*, which is not barely meant to represent the relations between pieces of knowledge (in the form of text), but also to include the whole of the critique associated with it. This whole of the critique, in part, is what determines *confidence*—algebraic truth-value—of each and every text in the network. Graph structure allows for confidence to be calculated in the bottom-up fashion: starting with the most specific, atomic critique, all the way up the graph—towards somewhat well-established knowledge. As the value of the truth-function converges over time—that is required for discourse to take place—confidence is gradually mapped onto the actors' virtual balance, in the form of *tas* tokens, which can later be used in the network, or simply exchanged for fiat.

Economically speaking, newly created knowledge emits money, and critique redistributes it as the computed truth-value of the corresponding texts fluctuates. The biggest concern with regards to confidence, of course, is that it must accurately approximate the truths of the world. To guarantee this, Veritas introduces a competence hierarchy. The actors' competence determines the extent to which their critique is able to influence the truth-value function. As trustworthy actors are overall more likely to spill the truth accurately, the network shall benefit them more for the same amount of work (in the form of critique.) On the other hand, as you look at the extremes of the competence distribution across all actors, you'll find both extremely competent and extremely *in*competent actors. To help the hierarchy stay intact, it can be normalised. Figure on the right shows how hyperbolic tangent—a simple nonlinearity—can be applied to generally favour actors occupying positions closer to the centre of the

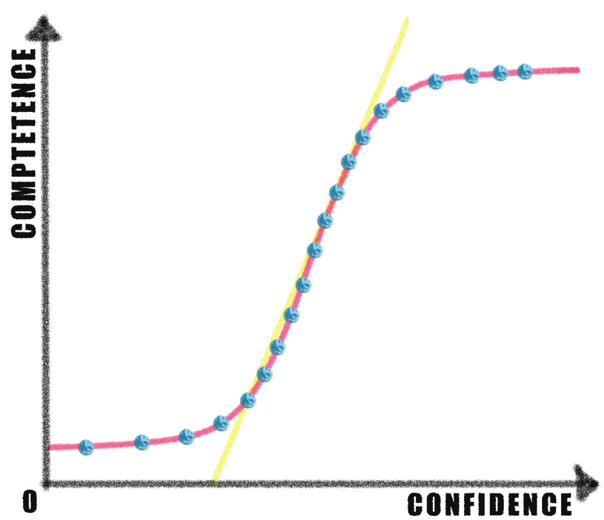
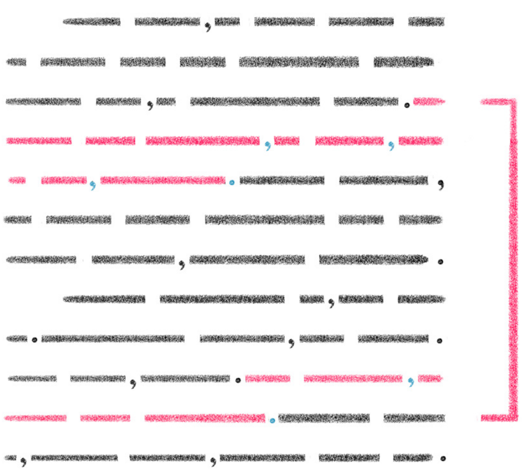


Fig. 4
Competence hierarchy can be normalised.

distribution. Starting at the very bottom, actors would face the decreasing degree of distrust, all the way up to the point where high-end responsibility starts to kick in, and the opposite takes place.

Every piece of critique is a text and text is subject to critique, therefore any critique is as well subject to critique. Criticism is all about correcting errors of the source text; it is also true that these errors can vary in their nature a lot. One of the reasons existing discourse platforms fall short when it comes to intertextuality of critique, is due to the fact that they fail to differentiate between different types of errors that can occur in a text. Veritas distinguishes five, somewhat overlapping categories of critique, such as the critique of



1. **CORRESPONDENCE.** This category includes a very specific set of errors associated with factual correctness. Correspondence critique explores to what extent the factual picture represents the objective reality.

2. **COMPLETENESS.** This includes both positive and negative critique of the source text, concerning its completeness, whether it can or can-not be proved.

3. **LOGIC.** Fallacies, contradictions, and incoherences that can be shown to exist within the logical picture of the text.

4. **LANGUAGE.** Any substantial critique of written language, including grammatical errors and mistakes, possible ambiguity of interpretation, and such.

5. **SINCERITY.** This type is specifically determined to explore all that goes *beyond* the structures of critique itself: matters of intent, malice, and manipulation.

To distinguish these errors is necessary, although by no means sufficient. You have to see that the knowledge graph, with the entirety of its critique is where “language meets algebra.” It isn’t easy to confine written language and computation—within a black box structure—unfortunately, we had only but to completely disregard the algebraic component in this paper, as it’s too specific to the implementation. That said, for any piece of critique—positive, negative, or anywhere in between—we can always think of its overall *sentiment*, ranging [-1; 1] from “agree completely” to “disagree completely.” Then, even though some form of manual sentimental input is unavoidable, the interface must help actors articulate it intuitively, as well as convey the critique itself. If we’re going to build a platform for discourse, it might as well “understand” text.

— badt@veritas.icu

[1] Yale Philosophy Review, 2008, Issue 4, p. 81.