

Design suggestions for an Epistocracy

Prakash Chandrashekar, April 2014

History - theoretical ideals and practical considerations

Epistocracy or Noocracy is a theoretical ideal for government where the wise rule. Humanity has attempted to create structures where people fit for government end up ruling and almost always, has ended up with some group having discretionary powers. It is not a stretch of imagination to say that those with discretionary power, without accountability, are those who have *power*, period.

On a parallel track, where philosophers were obsessing about the ideal and theoretical, the practical concerns of actually running organizations, has resulted in bureaucracy becoming the de-facto ruling structure. Governments, militaries, private firms, religious organizations, non-profits, NGO's, every successful organization with a replicable model are some species of hierarchical bureaucracies.

The mindspace in current political theory and practice is ruled by democracy. Almost every top leader or forum in the world praise democracy as a way to organize humanity. No one disses democracy, but actual democratic functioning is limited to small cooperatives and communes, organizations that do not even approach the limit set by Dunbar's number. Even democracies have their day to day functions run by a civil service that is unelected. Very few government processes actually change with a change in government. Economists would call this revealed preference. It is not a stretch of imagination to say that bureaucracy has won in the arena that matters the most, the real world.

If nothing more were imaginable, one could wind up this essay right here and answer the question, how will humanity steer itself in the future? -"with a bureaucracy", and be 100% true. Not a glamorous answer, but the *right* one. Why did bureaucracy succeed? We'll return to that question. And, is that the human future, a vagon bureaucracy, forever?

Element 1 - Open Source

Fortunately for the idealists, that need not be so. As humanity progressed due to material development, a cognitive surplus was created that resulted in many new theoretical forms of organizing society. The 20th century was a testament to the many failures of such visions, especially communism and fascism, partly due to the failure to shake off bureaucracy and partly due to other wrong incentives. Later in time, with the revolution in computers and connectivity, more promising visions have come up, prominent among these has been the open source movement in software. People freely create code and share it, with consumers of the code picking and choosing whichever version suits their particular needs. Wikipedia is operated by volunteers and provides a very good resource to the world.

So, if one were to open source government, how would one go about doing the same?

Cryptocurrencies

The open source projects that promise to be the most self-sustaining are the cluster of p2p currencies, beginning with Bitcoin. Bitcoin involves a breakthrough in practical byzantine fault tolerance, known as the blockchain. As long as a majority of the participating nodes are honest, one can trust the data in the

blockchain. This created unforgeable tokens which can be used as currency. This still did not give the tokens value. That was achieved by a rather cheeky distribution of coins, with the initial amount of coin generation being very generous and getting slower with time, until the overall cap is reached. This gave all the initial adopters of the technology incentive to get more participants in the initiative, as investors and merchants. The dramatic rise of Bitcoin's value is a testament, depending on whom you ask, either humanity's stupidity or recognizing the genius of the idea of crypto-currency.

One may object that bitcoin is not the most successful open source project. Apache, Linux or Wikipedia would be much better contenders. However, one must note that most open source projects do not have that many volunteers and even where they do, it is very difficult to maintain the quality over the discussion since literally anyone can join in. The phenomena of uninformative contributions and plain trolling are very common on internet forums and it is very difficult to maintain quality of discussion over time. And all of this is happening when the stakes are much lower than when we are talking about government, which is all about power over people. When people see any open source project get towards having power over people, they would flood the forums to make their side win. Hence, the moderators of the project have to shun people from the forums and setup a *hierarchy* of the people they need to listen to, and the people they would listen to, until you get people trying to filter out the masses. Thus, in even the most abundant scenario, software that can be copied, we can see hierarchy re-emerging, due to the simple considerations of attention allocation. Michael Goldhaber's and Esther Dyson's thoughts on the attention economy, on attention becoming the truly scarce resource, may be notable.

So, that explains why hierarchical bureaucracy emerges as a phenomenon. And here is where unforgeable tokens of value, created by a voluntarily accepted, neutral mathematical algorithm have a use - attention allocation in a world of abundance.

Element2 - Prediction markets

"Prediction markets (also known as predictive markets, information markets, idea futures, event derivatives) are speculative markets created for the purpose of making predictions. The current market prices reflect the predictions of the probability of the event or the expected value of the parameter. For example, a prediction market security might reward a dollar if a particular candidate is elected, such that an individual who thinks the candidate had a 70% chance of being elected should be willing to pay up to 70 cents for such a security.

People who buy low and sell high are rewarded for improving the market prediction, while those who buy high and sell low are punished for degrading the market prediction. Evidence so far suggests that prediction markets are at least as accurate as other institutions predicting the same events with a similar pool of participants." - mildly edited Wikipedia quote

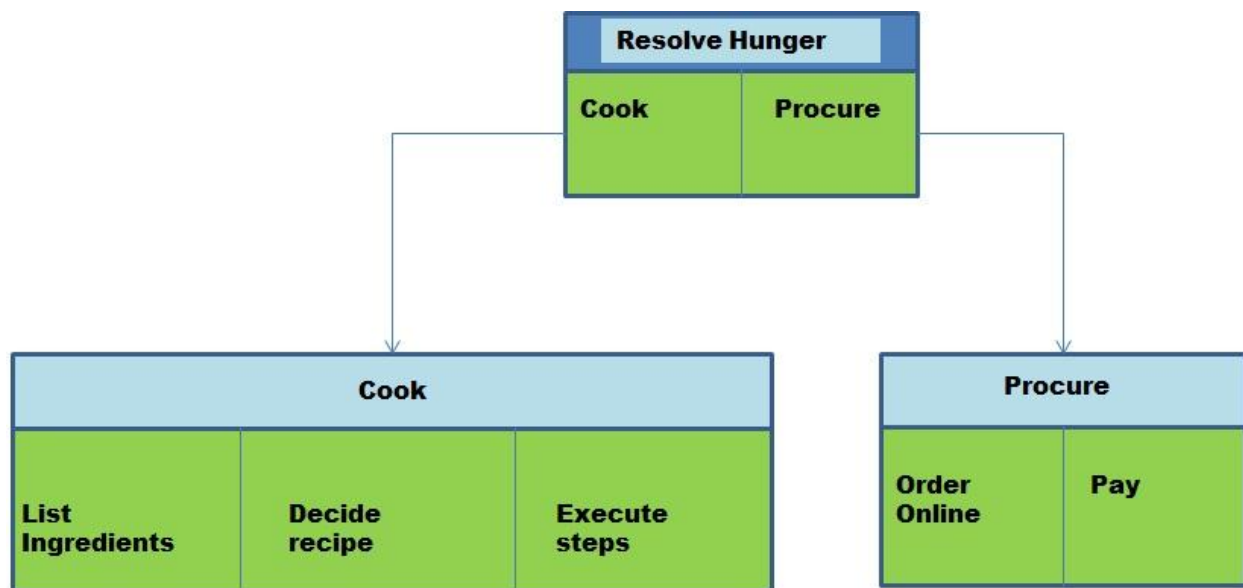
An interesting variation on the blockchain and prediction markets concept is truthcoin, a decentralized system of prediction markets, where the participants gain value if their predictions are more accurate and lose vice versa.

A detour into organizations

Getting back to bureaucracy for a minute, we need to take a brief detour into the world of business organization and business information systems. The body of knowledge known by the heading “Theory of Constraints” (TOC) is one of the most successful ways of thinking about business and organizations. The idea behind TOC is very simple. What is the goal behind a business - making more money now and in the future. What is preventing a business from making more money? That implies that there is a constraint somewhere. In other contexts, one can think of Liebig’s law of the minimum.

The prescription of TOC is simple - Identify this constraint, exploit the constraint, **subordinate all other decisions in the organization to the same**. It is simple, but not easy, since in traditional organizations, there are many **heuristics and rules that were made up in earlier times and other contexts that are no longer applicable to the current situation**. The main example listed in the TOC literature is that of product cost. Product cost as a criteria for what product to manufacture, was a heuristic that worked well when workers were paid by piece. It is no longer a good heuristic when the production processes are well set and the workers are paid a monthly salary. It makes more sense to check the time spent on the bottleneck process and see if displacing the current products are worth it.

The current state of art in TOC is the idea of strategy and tactics tree. Every strategy is a “why?” and every tactic is a “how” to achieve that “why?”. There is an entire tree of strategies and tactics to achieve those strategies that derive from one goal. Every process and procedure is a derivative of that.



Let's say there is an overall goal - to resolve personal hunger

That goal can have 2 basic strategies to achieve the same.

Cook Food, Procure food from Outside,

Under the strategy cook food, there would be the various tactics of collect list of ingredients available, decide recipe of dish, execute cooking steps.

Under the strategy procure food from outside, there would be various tactics like order online, pay, etc. Thus, almost every process in the organization can be thought of in the service of an overall goal. The strategy and tactics tree is just a simplistic way of designing systems so that everything derives from the common goal.

That ends the detour we have had to take into the world of organizations.

I thank the reader for persisting with the essay till here, since I will be revealing the answers to the questions posed in the essay contest.

What is the best world I imagine for humans

I hope for a world where people love their lives, where any sadness can be mitigated, where all pain and death is self chosen, a world where no one feels that they have lost out in any way, a world full of fun and play, learning and loving.

What would guide us there

A peer to peer network maintaining a strategy, tactics, assumptions and risks database, with the overall goals for humanity being at the top of the tree would guide humanity. (There can be multiple top goals, or one goal “the good life” being right at top and other values below it.). This peer to peer network would be incentivized by provision of tokens, generated more in the earlier years of the network and smaller amounts later. The database would be protected from corruption by the proof of work that will be needed to obtain these tokens. The overall structure would incentivize voluntary sharing of knowledge and deepening of the network as successful predictors of various outcomes and successful designers of new processes would gain tokens that they can use to guide the overall network to their preferred values.

Design considerations

- ❖ The network will begin with the tokens getting generated. Discussions can begin on the main goals in a separate board or on the network itself depending upon the repressiveness of the outside regime.
- ❖ Suggestions made would float to the top if accompanied by some tokens. Readers can assign tokens to suggestions and increase their visibility. There would be a buffer area that would be cleared for any suggestions not accompanied by tokens.
- ❖ To make changes to the database one would need the tokens.
- ❖ To create new tactics, one would need new tokens, to create new goals - lots of tokens. The number of tokens needed has to be a floating value, but similar in percentage terms so that when the network becomes very deep and large, it takes extraordinary amounts of tokens to make top level changes.
- ❖ One can hang tokens onto a goal, sub-goal or strategy.
- ❖ To associate new tactics to a sub-goal, one has to lay out the assumptions behind it. Each of these linkages will be given weight by the tokens attached to it.

- ❖ Every goal would have acceptance criteria. Acceptance criteria will depend on the real world, hence some node has to become a trusted source of information for each acceptance criteria. Trusted sources could be decided by a vote of token holders.
- ❖ There needs to be a design mechanism in place that incentivizes narrow acceptance criteria and penalizes very wide acceptance criteria. This is similar to truthcoin's design.
- ❖ This will provide an automatic incentive to everyone in the network to resolve that goal in ways that don't contradict other goals. Resolving those goals in a manner acceptable to everyone concerned on a permanent basis, mostly by the design of a new process, will give a few of those tokens to the winner.
- ❖ One can list risks or unintended consequences against every possible strategy. Sometimes these may be acceptable, like medical side-effects, sometimes they may not be. Absolutely unacceptable consequences can be listed, which when triggered, will render that strategy or sub-goal as invalid.
- ❖ Proving that assumptions linking a currently active tactic to a strategy are invalid, will result in a reward. This can prevent the wait for the acceptance criteria to get triggered.



- ❖ On the explicit failure of a tactic to meet the acceptance criteria for fulfilling its strategy, there will be preference given to other tactics that have lost out in the voting, in the order that they were voted. This will ensure that the brain power that went into thinking about these solutions is not lost in time. But even other solutions will need to be evaluated in the light of the new facts and weights of values.
- ❖ In this manner, we eliminate one of the primary issues with any bureaucratic organization. Almost all bureaucratic organizations seek to maximize their budgets and perpetuate themselves. Hence, they are currently incentivized to never actually completely resolve the problem that they are setup to solve. If the problem is resolved, there ends their usefulness. Goldratt, the originator of TOC, believed that this reluctance to completely solve a problem by “cause-and-effect”, can be mitigated by a resolution from the corporation's side never to fire workers, always redeploying them to other tasks that will add value. Here, we are envisaging automated processes that are running on their own.
- ❖ The change of the weights will have to happen as per periodic processes, with the values attached to higher goals being refreshed at slower rates and the values corresponding to lower tactics changing sooner. This parallels the planning that happens in organizations or even elections in a democracy. Weights of values changing may render certain processes obsolete. Some may not.
- ❖ There needs to be a very high cost in terms of tokens to change acceptance criteria. This is to mitigate one of the biggest problems with current organizations, the moving of goalposts.
- ❖ Tokens will be freely tradeable and they will be created by submitting proof-of-work. People can freely give and trade tokens to others for valuable contributions made to various debates.

Theoretical advantages

In the structure described above, there is an attempt to create automatic incentives to permanently resolve issues.

There is also a motive to explicitly lay out the values so that people can understand the reasons for various processes much better. For e.g., One can see that workers did much better in capitalist economies than they actually did in communist economies. A theoretically honest communist country would have had to resolve this dilemma. North Korea explicitly lies. They are beyond the pale. However, a theoretically honest communist society would reconsider its policies and undertake reform or explicitly mark a “unacceptability of bourgeois ways of trade” as one of its important values, precluding capitalist reform. Similarly, one can imagine countries slowing down growth so that all the population can catch up, if harmony or outcome equality is heavily weighted making it a primary value.

Irrespective of your personal politics, I hope it is possible to see the value in a system that makes values explicit.

Rituals that are followed without giving thought to their origin have a certain power, but beyond a time, the human effort that goes into performing them, is effort wasted. These were great solutions in a time when humanity was apart and messages took a long time to reach. However, most of humanity is currently on a network where we can communicate with each other instantly. We can re-examine the processes that we follow in our day to day life and see their current relevance.

Objections and FAQ

No one can possibly map all human desires	Yes, that is why it is an incentivized crowd sourced project. Everyone can contribute insight or computing power and gain tokens.
Why the choice of a P2P network. Wouldn't a centralized network be more efficient?	A centralized network would be more efficient, but it would be vulnerable to attacks. The opponents of any new system would have advantages if they had only one place to aim for. In contrast, A P2P system is far more resilient. Napster was shut down, Bittorent is still alive and streaming.
I can spam the structure with nonsensical additions	The proof of work is explicitly in place to place a cost on doing such a thing. If the growth of such a blockchain mirrors the growth of bitcoin, it will become stronger with time, more difficult to spam
The tokens are worthless, why would anyone spend computing power on their creation	I'm thinking of the same logic as bitcoin. If you believe that in the future this network will have value, then these tokens will have value. You could play a part in sculpting the future.
Bitcoin's current structure, where every node maintains all data is ridiculous	There has to be a certain amount of replication to maintain integrity of the data by cross referencing other nodes. But this need not be complete. There are tree-structures envisaged which store this data more efficiently. However, there is always a trade-off between resilience and efficiency.
You need to get into the actual meaning of words	This structure should also incentivize making explicit tactics and the meaning behind the word. “Rule of Law” is a generic term. “Explicitly published laws with free access to courts with judges who are paid salaries that cannot be held back by the government

	<p>in charge” is a better described tactic. This is perhaps why I think that this entire project may actually need a previously existing ontology, but we can use the structure to refine the ontology that we need, using the tokens as incentive.</p>
<p>You’re remaking central planning in a nice format. Haven’t you read Hayek’s fatal conceit?</p>	<p>Actually, I have read a lot of free market literature and admire the ability of the price mechanism to resolve issues. This is why at the base of it, the tokens being tradable is a price mechanism. The values that people seek to attach to various moral goals are their personal choice. If I want to increase the weight of a certain value, I need to gain more tokens.</p> <p>But on a deeper level, I think the correct answer is - the classic liberal prescription is also finally a heuristic, it is a tactic to resolve issues under certain conditions. It is important for a person to feel that his property in his car remains unviolated in today’s world, but if he is granted free access tomorrow to a variety of rental options, then does it remain that important anymore? Is it in the person’s attention span? Is there a point in writing down in stone, the inviolability of property for all eternity? Maybe, just maybe we could do better.</p>
<p>The decisions of the programmer are set in stone and his mistakes will echo through eternity.</p>	<p>There are multiple replies to this objection. Any person of knowledge or power who, ever wrote a book or created a structure or organization has faced this issue.</p> <p>The creators of this software are expected to be humble and have contingencies in place. As a backstop, with a 51% control of the network, one can slowly undo the latest value changes that were the cause of concern. One could also imagine that beyond certain percentages, one could alter the meta values of the structure. (values that trigger various changes, various percentages, time values, etc.)</p>
<p>How will those not intelligent enough contribute?</p>	<p>That is unfortunately true of the growth of the robotic economy as well. Even in the current scenario, no one has solutions that will not make the labourers ultimately dependent on the mercy of the thinkers.</p> <p>If the values are structured in a manner that dignity is an important value, then intelligence amplification will arise as a possible long term solution.</p>

Notes on the possible evolution of the network

Talking about a theoretical organization structure is all fine. Why would anyone else subject themselves to the rule of this network? I envisage that the spread of the network will occur similar to the rise of parliaments in previous centuries. They started out as advisory boards and later ended up with power. Somewhere in time, people thought that it was better to be subject to the consensus rule of a lot of elected people than the whims of a monarchical line. Similarly, sometime people will think that the predictability created by the mathematical structure will be a better bet than relying on the whims of

people's rule. The network starts with a core of smart altruists with free time. As the network grows, proposing interesting and innovative solutions to public issues, more and more political parties around the world will start copying solutions from it. After a while, due to the osmosis of ideas, some parties would rise that would simply declare that they would openly hand power over to the epistocratic network. Even if they do not, (Humans are sometimes rightfully called homo hypocritus) the real conflicts will shift to the assignments of weights in the network. Legitimacy would soon belong to the network.

There will be forks because this is open source software. There will be sub-domains where every goal will have different weights. But getting everyone to move with the fork or change will be a rough task. For the people who are not into the network, getting them to believe that stratnet2 is better than stratnet1 would be quite an uphill task, since they would have to understand every little difference. The energies of the secessionists would probably be better spent in improving the weights of the ongoing network.

Artificial intelligence will be a reality someday. And when it develops, every aspect for which the network is sought could be delivered faster by the AI. How the network reacts to an emergence of AI will depend on the values that the creators embed into the AI. The AI will have the entire structure available to it to peruse. It can choose to cooperate by offering suggestions for better processes, which will earn it tokens which it can use for gaining real world resources, which it can use to augment its own values in the network. It may choose to completely disregard the network and proceed to world domination in its own manner. If the network has gained credibility that it represents a mix of the values that most humans want and a good set of strategies to achieve the same, then there will be an increased probability of people programming their AIs with the goal matrix of the network. If it hasn't, then the AI would most probably win. Understanding this, a surveillance network to suppress potentially dangerous technologies may get created by the actions of the network and it may guide the future creation of artificial intelligences.

The Buddhists have a saying.

Buddham Sharanam gachhaami, Dhammam Sharanam Gachhaami, Sangham Sharanam Gachhaami

I take refuge in the enlightened one, I take refuge in Dharma/Dhamma, the law, the way of the universe, I take refuge in the community

Humanity evolved from taking refuge in people, in shamans, chiefs and kings to taking refuge in large bureaucracies and parliaments (Sangha - community). I hope that they will eventually move to taking refuge in "the law as software"(Dharma - the way of the universe). A community of enlightened people will still be there, but they will not be "ruling". They will be suggesting the way and the weights, winning or losing with their predictions and designs. The network described here is possibly one of the very few ways in which all of humanity can come together. It would guide our limited attention to the issues that matter most. It would be the beginning of our age of collective intelligence.

References

Noocracy - <http://en.wikipedia.org/wiki/Noocracy> [Last accessed 8Apr 2014]

Dunbar's Number - http://en.wikipedia.org/wiki/Dunbar's_number [Last accessed 8Apr 2014]

Bitcoin - https://en.bitcoin.it/wiki/Main_Page [Last accessed 8Apr 2014]

Practical Byzantine fault tolerance -

http://en.wikipedia.org/wiki/Byzantine_fault_tolerance#Practical_Byzantine_fault_tolerance [Last accessed 8Apr 2014]

Prediction Markets

Prediction Markets - http://en.wikipedia.org/wiki/Prediction_market [Last accessed 8Apr 2014]

Trading Uncertainty for Collective Wisdom <http://www.lumenogic.com/www/static/pdf/ci-servan-schreiber.pdf> [Last accessed 8Apr 2014]

Truthcoin - Decentralized prediction markets <https://github.com/psztorc/Truthcoin> [Last accessed 8Apr 2014]

Attention Economy

The Attention Economy and the Net - Michael Goldhaber (<http://firstmonday.org/article/view/519/440>) [Last accessed 8Apr 2014]

Esther Dyson (http://www.econtalk.org/archives/2013/01/esther_dyson_on.html) [Last accessed 8Apr 2014]

Liebig's Law of Minimum - (http://en.wikipedia.org/wiki/Liebig's_law_of_the_minimum) [Last accessed 8Apr 2014]

Theory of constraints - www.goldratt.com [Last accessed 8Apr 2014], www.tocico.org [Last accessed 8Apr 2014]

Theory of constraints, Strategy and tactics tree - www.tocico.org/?page=strat_tact_portal [Last accessed 8Apr 2014]

The Fatal Conceit - Friedrich Hayek - en.wikipedia.org/wiki/The_Fatal_Conceit [Last accessed 8Apr 2014]

Creating Friendly AI - <http://intelligence.org/files/CFAI.pdf> [Last accessed 8Apr 2014]

Buddhist quotes - various sources

