Cognitive goods, normal goods and the market for information

Leigh Caldwell
Irrational Agency/Inon
January 2019

Introduction

Lassie died one night. Millions of viewers, not all of them children, grieved…the mourners knew that Lassie didn't really exist…Did they enjoy the episode?

Thomas Schelling, "The Mind as a Consuming Organ"

In this essay of 1987, Schelling explores the reward and emotions generated by experiences that we know will have no actual, permanent impact on our lives. Examples are easy to find. People care about events in the news that can have no measurable effect on their own life outcomes. They care about income relative to their peer group, when (given a particular level of purchasing power for the currency) it is only absolute income that allows them to buy things. (Indeed, many of us would rather not know what our colleagues earn.) People avoid finding out medical information that could be important to optimising future health outcomes.

If I hadn't seen such riches, I could live with being poor.

James, "Sit Down"

In short, people experience emotions – and utility – from mental states that do not relate to material outcomes. But there is something that stops us simply choosing whatever mental states make us happy (Ainslie 1992).

It is no use trying to model these processes simply by assigning a utility function to Lassie's death, my relative income or my lack of awareness of health. There can be no well-behaved utility function because these are not well-behaved economic goods. A (dis)utility function over Lassie's death cannot apply if the agent has a consistent model of the world. Preferences over relative income cannot simultaneously obey convexity and transitivity[1]. Lack of awareness of medical status does not obey the irrelevance axiom. And even though we may put a positive value on these mental states, they cannot be priced, because they are neither scarce nor tradable.

If instead we define a new type of object, the *cognitive good*, we can create a model that allows these objects to be valued consistently.

This paper proposes a set of axioms that cognitive goods should obey, proposes a model that obeys the axioms and allows further predictions to be made, and outlines

[1] A demonstration will be provided in a separate appendix. The only way to reconcile this with a standard utility function is to assume that people have general *dis*utility for other people's income, i.e. they prefer others to earn less regardless of their own income. This is possible in theory, but seems implausible in the real world.

consequences for some key markets: the market for information, cultural goods and money. A few potential policy consequences are explored.

Examples of mental objects that meet the definition of cognitive goods include:
- Experiences of fictional events [figure 5 below]
- Self-image and identity markers [figure 6 below]
- Belief-based utility [figure 7 below]
- Desire for an income level relative to a peer group in preference to an absolute amount [figure 8 below]
- States where the absence of information is preferred to its possession [figure 9 below]
- The results of sporting events [figure 10 below]

An important step in understanding the nature of cognitive goods is exploring the constraints that stop them from being arbitrarily created. To do that, I propose a plausible mechanism by which cognitive goods might be generated and valued.

The existence of cognitive goods in turn influences how agents value, and consume, normal goods. Therefore, even aside from the direct welfare that may be generated by cognitive goods, they are important for consumer welfare theory in general.

Cognitive goods are an important modelling artefact for the emergence of *cognitive economics*, defined by Kimball (2015) as "the economics of what is in people's minds" and by Mulgan (2017) as "the view of thought as involving inputs and outputs, costs and tradeoffs". The concept of cognitive goods unifies these two definitions, which approach the subject from complementary angles: what is the subject matter, and what are the mechanisms that animate it.

Proposed axioms for cognitive goods

Start with the following **axioms**:
1. Mental states are internal to the state of an individual agent, and have no direct effect on any other agent
2. A weak preference relation exists: agents may prefer one mental state to another (though they can also be indifferent between particular states)
3. An agent's mental state at time t+1 is a function of its mental state at t, and the agent's inputs from the external world at time t
4. To make a choice over normal goods, an agent must change mental state at least once. This claim is based on the idea that the agent has to mentally consider the options to evaluate them, this takes a finite amount of time, and that their state of mind is different after the choice than before – if in no other respect that they now know what choice they have made.
5. An agent's actions are determined by its mental state.
6. A *cognitive good* is defined as any part or subset of a mental state; hence, beliefs about the world are a type of cognitive good

Some immediate conclusions or lemmas can be drawn from these axioms.

Lemma 1: Agents are constrained in their ability to generate cognitive goods.

By axiom 3, mental state at any time is a function of previous mental state plus sensory input. Therefore, an agent cannot arbitrarily choose any mental state at any time.

Lemma 2: Mental states generate a continuous stream of utility or reward.

(because we prefer one mental state to another, and states change over time, any preferred state provides some kind of ongoing payoff)

Lemma 3: Mental states can affect the decisions that agents make about normal goods

From axiom 3, mental state is a function of sensory input (including sensory inputs about the goods being chosen). Also from axiom 3, mental state is a function of previous mental state. Therefore, different starting states can lead to different finishing states under the same sensory inputs. From axiom 5, the agent's actions (which includes its choices between goods) are influenced by its mental states. Therefore, a different starting mental state can lead to a different choice of goods.

Lemma 4: Beliefs can affect decisions

From lemma 3, mental state affects decisions. From axiom 6, beliefs are a type of mental state. Then, beliefs can potentially affect decisions. As a corollary, preferences – defined as the rules upon which decisions are based – can be seen as a kind of belief: the belief that "A is better than B" is equivalent to a preference for A over B.

Lemma 5: Each mental state has a natural successor state

From axiom 3, each state is a function of the preceding state plus sensory input. If no external sensory input is provided, the state is determined solely by the preceding state (up to a random error term).

Lemma 6: The absence of a cognitive good can be preferred to its existence

From axiom 2, one mental state can be preferred to another. From axiom 6, a cognitive good is part of a mental state. In principle, a mental state that does not contain the good can be preferred to one that does.

I now explore a possible mechanism that obeys these axioms and can illuminate the behaviours we would expect agents, and markets, to exhibit in the presence of cognitive goods.

The psychological background to this model

Clark (2015) and Hohwy (2013) argue that the primary function of human cognition is to predict the future. The ability to predict gives agents such immense power to make better decisions with limited resources, that this capability provides an overwhelming evolutionary advantage.

Gilbert and Wilson (2007) propose that the ability to *prospect* (imagine the future), is moreover one that people enjoy, at least in the case of positive potential outcomes; our imagined future happiness gives us pleasure, and we are motivated to engage in daydreaming and planning those potential futures.

Neuroscientists including Stachenfeld et al (2017) have examined a potential mechanism by which the brain makes predictions and acts on them: the *successor representation*. This theoretical concept, proposed by Dayan (1993) is now supported by neural and behavioural evidence (Gershman 2018). It represents potential states of the world in terms of the states that are likely to come after them. A state representing me opening my office door might be succeeded by a state that represents me stepping out into the hallway, which in turn is followed by me reaching the kitchen.

Caldwell (2017), building on arguments in Ainslie (1992), proposed that this capability to imagine, and be rewarded for imagining, the future is essential to making intertemporal tradeoffs. Future benefit cannot directly influence an agent's present choices; only the agent's present beliefs about future benefits can do so. The proposed mechanism for this is that agents are rewarded in the present for visualising and imagining future reward, and this present reward is a substitute for the reward available from immediate consumption. The ability to visualise and act on future reward is proposed as the basis of a "System 3", the counterpart to Systems 1 and 2 proposed by Stanovich and West (2000).

The causal graph model

Causal models, as described by Sloman and Lagnado (2015), are a way for agents to mentally understand the world and the relationships between events or objects within it. Pairs of objects A and B are causally related if A causes B; B can in turn cause C, implying that A indirectly causes C. Agents can use these models to predict the outcomes of their actions: if the agent takes action A, B will follow, then C in turn. Based on the agent's evaluation of outcome C, they can decide whether to proceed with A.

Learning theories discuss two kinds of representations of the world: model-free and model-based. In model-free representations, only the initial choice (A) and the ultimate outcome (C) are considered: for example, I may decide to get up from my desk and go to the shops because it is, on average, a rewarding activity (perhaps I have learned from experience that I will usually find something enjoyable there). A model-based representation instead takes into account the structure of this task and the intermediate states that I may go through before achieving an outcome. For example, I might think through my likely actions in the shop, the objects I might purchase, and try to consider whether they are worth the effort.

The successor representation is a version of model-based reasoning. However, I argue that the individual steps in the successor representation are learned in a model-free way. The step from the desk to the door, or the door to the hallway, do not have any internal structure – they are the “atoms” of the learning process.

Each of these atoms can be thought of as a causal relationship: *if* I go through the door, *then* I will be in the hallway. *If* I eat this chocolate, *then* I will be happy.

In Sloman and Lagnado’s model, the whole causal network must be evaluated in order to decide whether to take the first step – it is a fully model-based representation. However, Daw and Dayan (2014) present some of the algorithmic challenges inherent in model-based reasoning, which suggests that model-based learning and reasoning must be constrained in certain ways to be computationally feasible.

I propose a constraint that can resolve this challenge: model-based reasoning and representations use *reward* to operate, and this reward in turn creates model-free caches within the overall network. An agent’s mental model of the world is composed of individual causal beliefs, each of which is a model-free “atom” in the larger causal network. Each belief has a level of reward associated with it.

Formally:

A causal graph G = (V, E, r) where:
V = a set of nodes, each representing a single *interpretation of stimuli*
E = a set of edges $\mathbf{e_{1,2}} = (\boldsymbol{v_1}, \boldsymbol{v_2})$: agent’s belief that v_1 causes v_2
r(v) = the reward associated with node v

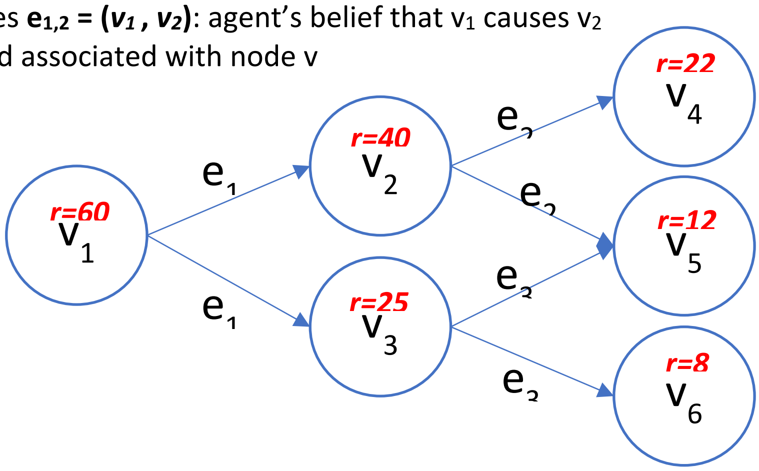

Figure 0: How a causal graph might be defined.

In order to make a decision between two options, an agent mentally simulates the causal consequences of each option. Following the causal network from the option to its immediate and successive consequences generates a certain amount of reward at each node in the network. The total reward the agent accumulates by following each option guides them in choosing the preferred option.

$$R(v) = r(v) + \sum_{e(v,b)\in E} R(b)\delta$$

Where 0 < δ < 1, the *causal discount rate*. In Figures 1-4, the causal discount rate is set to 1; subsequently 0.8.

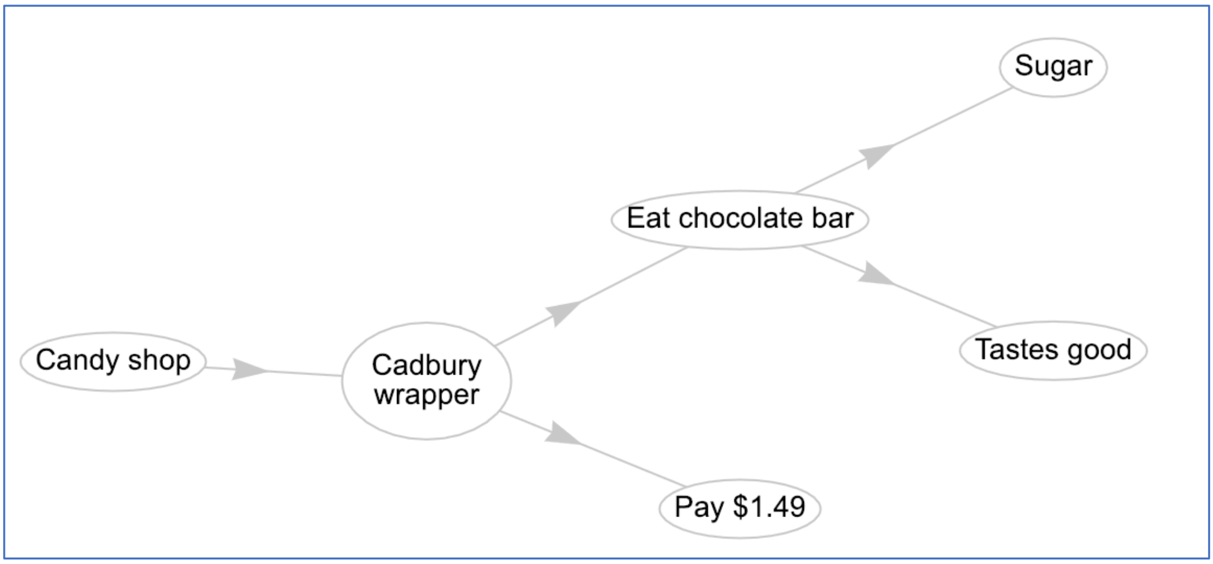

Figure 1: a graph of causal relations to do with eating chocolate

While mentally navigating this network, the agent has the opportunity to learn which nodes in the network are predictive of high reward. For example, I may visit a candy shop several times and end up with a chocolate bar, a bag of sweets, a box of Junior Mints – soon enough I will learn that the 'candy shop' node in my causal graph usually predicts a rewarding outcome. To save on future computation, I learn to directly associate 'candy shop' with reward – a model-free supplement to the model-based network. This enables me to make more efficient predictions: instead of mentally working through all the actions I could take when I go into to the shop, I only need to know that when I get there, something good will happen.

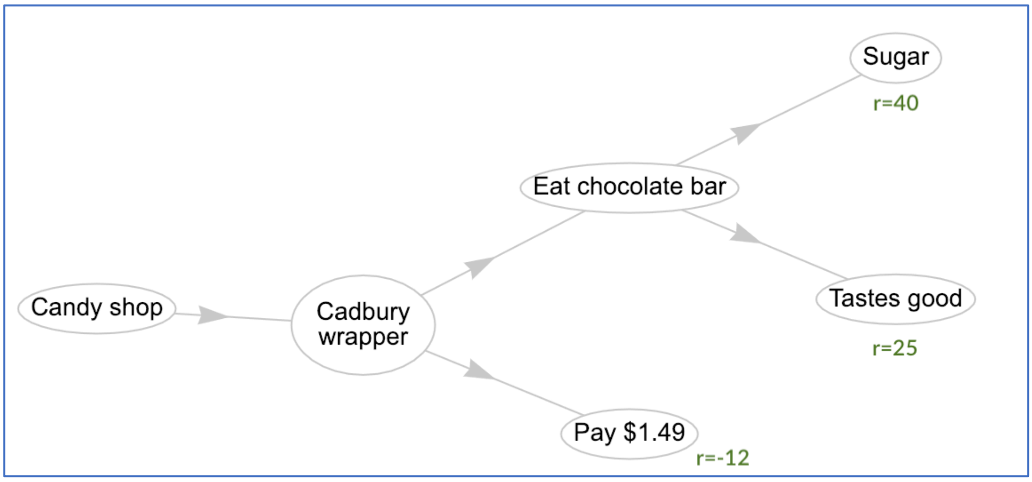

Figure 2: a graph of primary reward and the causal relations that lead to it

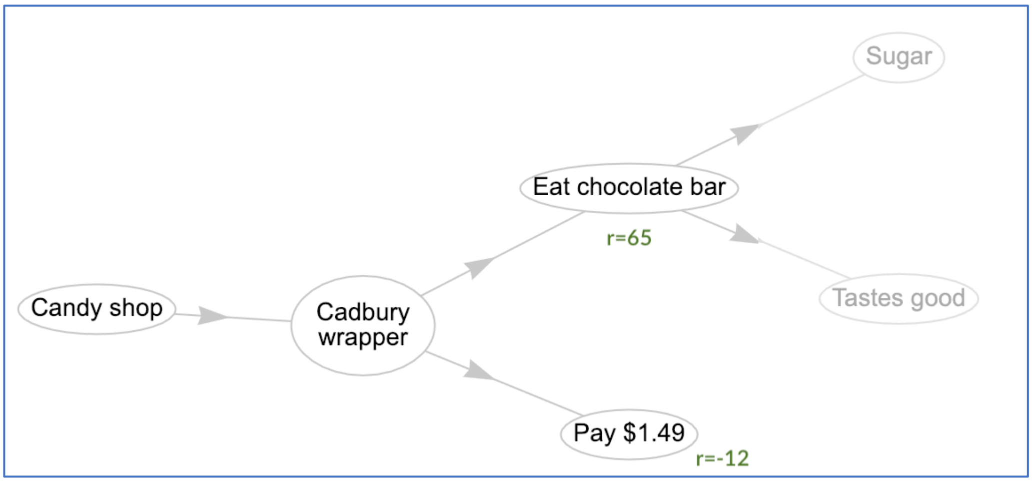

Figure 3: reward is cached one level up from the primary reward

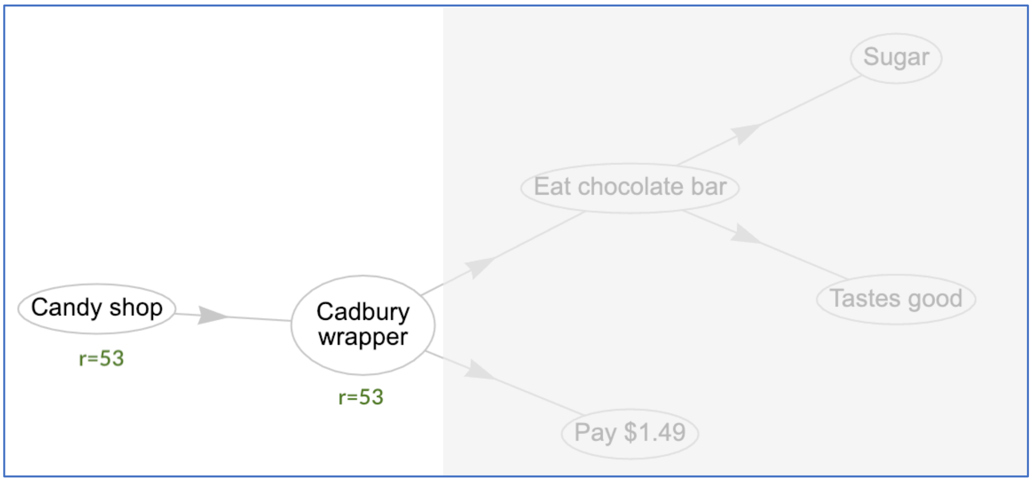

Figure 4: reward is cached two levels up and the agent learns that going to the shop is, in general, a good idea.

This process can be summarised as learning to associate rewards with intermediate states. The shop does not start out as intrinsically rewarding (only the things that happen *after* I get there are rewarding in themselves). But I learn to associate it with reward. Ultimately, I will feel happy just when I get to the shop, even before I buy, open and eat the chocolate.

Formally, if a node is not cached, the usual recursive process occurs:

$$R(v) = r(v) + \sum_{e(v,b)\in E} R(b)\delta | v \notin C$$

If the node has a cached reward value, that value is used:

$$C(v) | v \in C$$

And if the total discount is less than a cutoff value, zero reward is produced and no further evaluation takes place.

$$0 | \delta < \epsilon$$

With sufficient reinforcement, the mental state of *imagining* the candy shop can become rewarding in itself. Similarly, states like thinking about a loved one, a political belief, or watching Lassie on TV, can all become rewarding.

In this model, reward is generated whenever a node in the graph is mentally *active* – which we might see as roughly equivalent to having attention from the decision maker's executive function. An active node automatically activates its causal successors in the graph. The agent's external behaviour is likely to be driven by seeking out stimuli that will (directly or indirectly) activate the most rewarding nodes. Its internal processing will consist of choosing to activate nodes in the network that, insofar as it can calculate, will result in high levels of reward.

A number of key psychological questions remain about how this mechanism is likely to be implemented. These could be explored with further empirical work:

- Activation would be expected to decay as it moves through the graph; an active node will not always fully activate its successor nodes (especially if it has multiple successors, for example in cases of uncertainty)
- The learning algorithm needs to implement a cutoff in order to avoid infinite recursion through the mind. This cutoff could be based on the number of nodes navigated, the level of reward generated by the node, or some other mechanism.
- The parameters of the learning mechanism need to be measured empirically. Reinforcement learning models tend to use a variant of the Rescorla-Wagner model (see for example Miller, Barnet, Grahame 1995) which indicates that learning occurs when the reward generated from an action differs from the predicted level of reward. Implemented in this model, this would suggest that the cached reward amount at a node in the graph would gradually approach the true weighted amount of downstream reward generated from its successor nodes. However, the speed of approach may be fast or slow, and this has differing implications especially when the agent must learn in a fast-changing environments.
- Uncertainty is modelled in this graph by multiple effects for a single cause. It is unclear whether the brain's autonomous mechanisms can represent probabilities in a fine-grained way; will all possible outcomes of an event receive equal weight in the reward calculation, or will they be weighted by likelihood (perhaps with approximation)?
- Similarly, it is not clear whether the graph can meaningfully handle quantitative calculation. Is it more rewarding to imagine four chocolate bars than three? How about forty chocolate bars? If so, how much more? I suspect the answer is that numeric weights above 2 or 3 must be imposed by a second-order calculation, perhaps through algorithmic "System 2" rules that the agent applies outside of the reward graph. However, empirical work would be needed to explore this further.
- It is unclear whether pain and reward are neurally implemented in the same 'currency' or there are two different neural mechanisms at play. In the current model I assume that they can be traded off against each other – sufficient pleasure can outweigh a certain amount of pain – but this may need to be revisited.
- The agent needs a mechanism to recognise external stimuli and associate them with nodes in its internal model. This mechanism may involve activating multiple nodes

simultaneously, setting off parallel reward calculation streams. The exact type of mechanism that brains might use to do this is not well-understood psychologically, but neural network models provide one possible route.

The different examples of cognitive goods given above can all be implemented in this model. Examples are shown in figures 5-11.

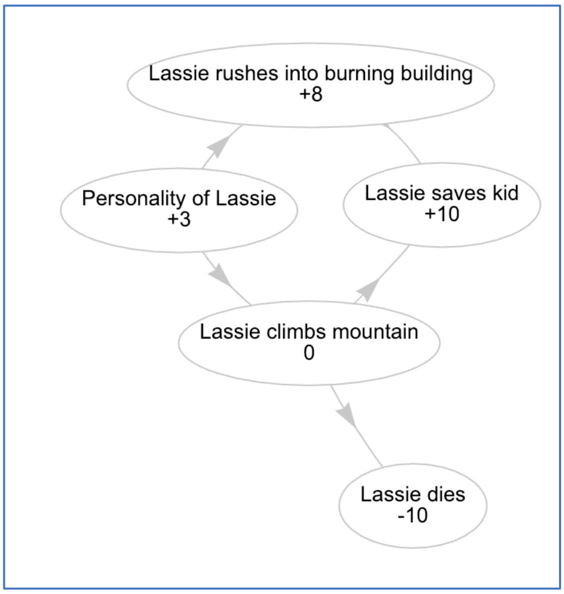

Figure 5: Reward from watching (and imagining) the adventures of Lassie. The 'Personality of Lassie' is the origin node of this graph, and reward ultimately accrues to the activity of imagining and understanding that personality.

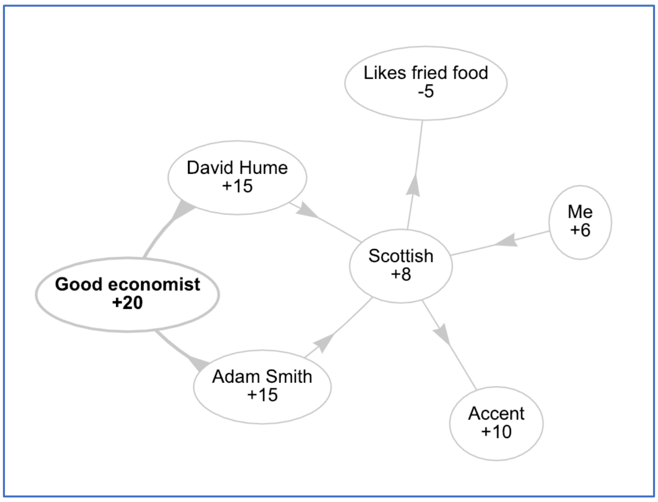

Figure 6: Reward from an identity marker – which could originate from fuzzy pattern recognition, associating the identity of being Scottish with being similar in other respects to Adam Smith or David Hume.

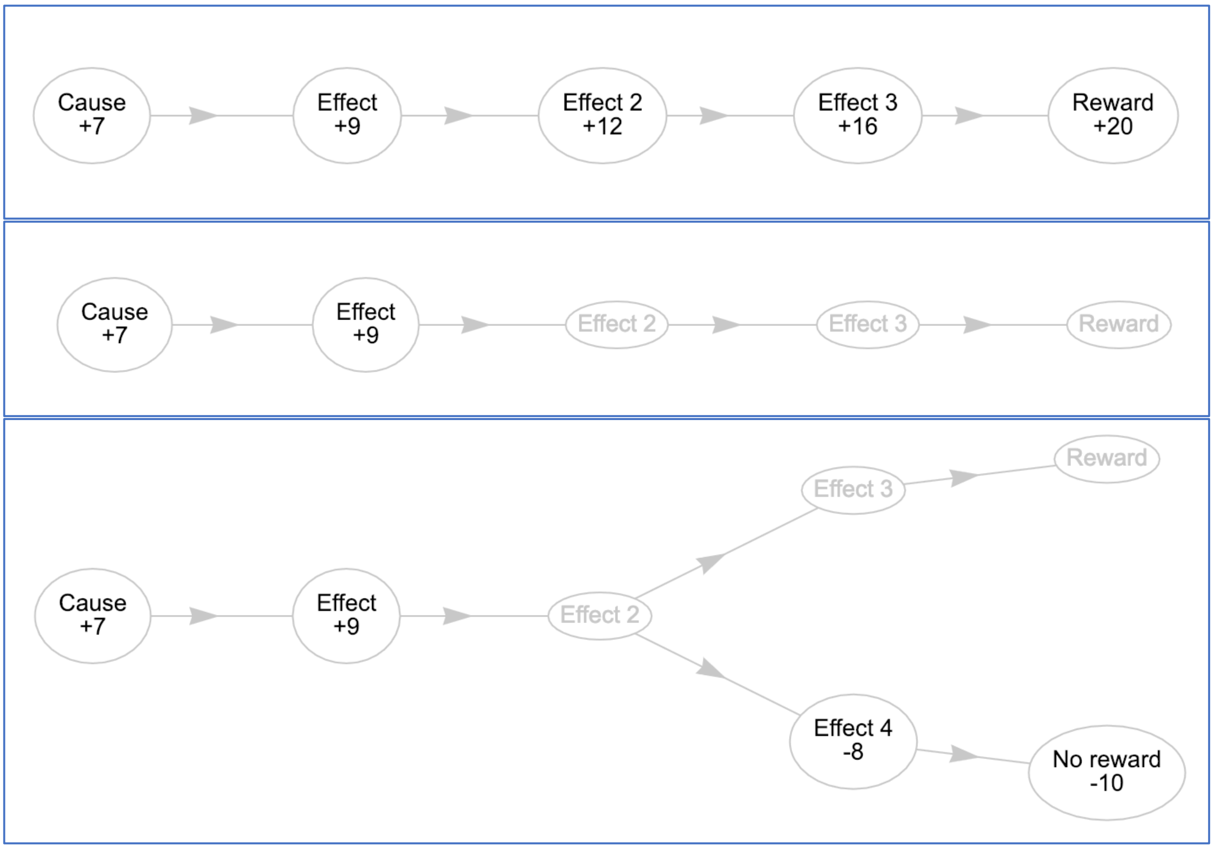

Figure 7: Belief-based utility. A belief, when first learned, is accurately associated with positive reward. The agent's mind caches reward at the original belief and truncates calculation. When new information is subsequently learned, the original belief is no longer updated.

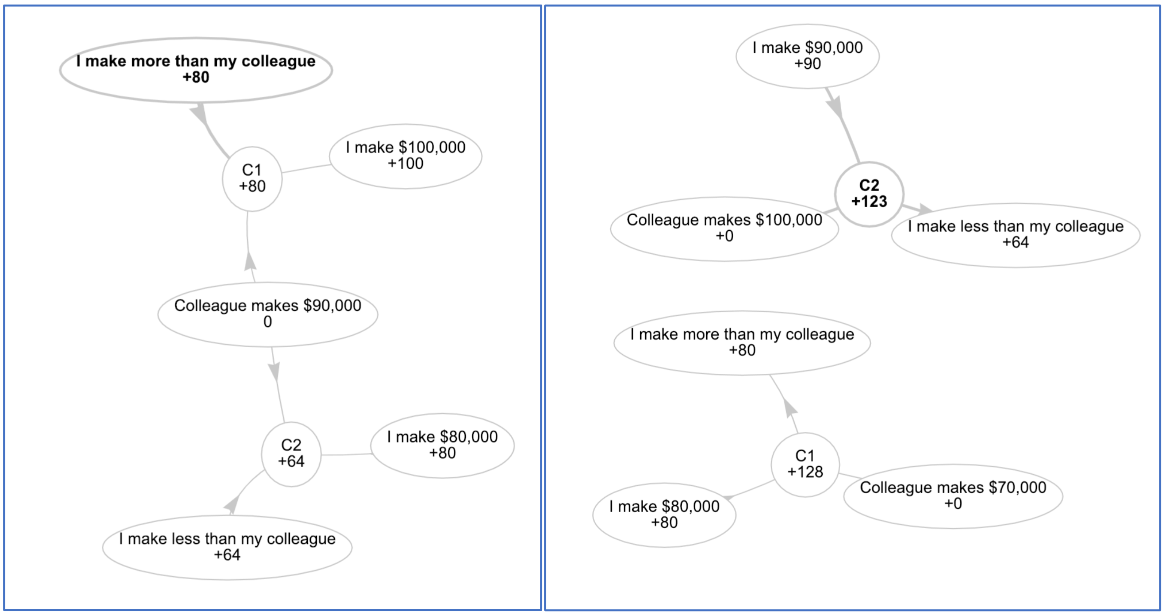

Figure 8: Relative income. If the agent first learns the value of relative income in an environment where their own income is the only variable, they may learn to attach reward to a higher relative income. In a new environment where that is not the case, this could lead to apparently irrational preferences (C1 being preferred to C2).

The definition of cognitive goods

If this process, or one like it, is indeed the origin of valuation of mental states, it has implications for the first lemma of cognitive goods: the constraints on their generation.

A *fundamental cognitive good*, in this model, is a single node of the graph – such that activation of this good produces direct reward according to solely the numeric reward value attached to that node. A *simple cognitive good* is an active connected subgraph of an agent's causal graph. A *composite cognitive good* is the union of multiple simple cognitive goods which need not be connected. A *social cognitive good* is a subgraph, a representation of which is shared between multiple agents and activates corresponding cognitive goods in each.

This terminology is partly consistent with Benjamin, Heffetz, Kimball, Szembrot 2014 (BHKS), who define fundamental aspects of welfare as those which cannot be broken down into smaller parts, composite aspects as consisting of several fundamental aspects (this includes both the 'simple' and 'composite' cognitive goods in the present definition) and public aspects, those related to a whole society's wellbeing rather than one individual's. BHKS's public aspects do not correspond directly with the social cognitive goods defined here, which could include privately owned brands or product concepts, but in practice there is likely to be an overlap between these two groups.

Normal economic goods are constrained or scarce due to their production function. The goods are made from inputs, which may be raw materials or other (intermediate) goods. The inputs are scarce, and as a consequence, production of the goods is limited.

Cognitive goods – in the model proposed here – are produced instead by activation of different parts of the causal graph that the agent holds in their mind. Although this activation consumes a small amount of energy, that is not the primary constraint. Instead, activating one part of the graph automatically activates the nodes "downstream" of this subgraph (axiom 3: mental state is a function of the previous mental state). These activated downstream nodes may not be positively rewarding. The desire to activate cognitive goods that produce only positive reward is a key constraint on the agent's consumption of mental states.

Conversely, in order to activate a subgraph of the agent's mental graph, the nodes that lie 'upstream' of it should be activated. Agents may not have sufficient knowledge of their own causal graph, or the mental self-control, to explicitly choose which nodes to activate. If so, they need to use sensory stimuli or other secondary objects to cause activation. As a result, they may buy goods or experiences that provide this stimulation. In other words, they acquire normal goods in order to indirectly manipulate the cognitive goods in their own mental state.

Cognitive goods may therefore have correspondence or similarities to normal goods, but need not always. A common class of cognitive good that does not necessarily correspond to any normal good, are goods transmitted through language instead of sensory experience. Language can activate the agent's causal graph directly, without the need for the referent concepts in the graph to be experienced. Language – and hence information – therefore can become directly rewarding, independently of the objects it refers to.

The market for information

Information is typically seen (Stigler 1961, Allen 1990) as a very different kind of object to a normal good. Information is satiable, non-rival, non-excludable and – without artificial institutional structures such as information property rights – cannot be priced above zero.

Examining the agent's utility in terms of cognitive goods shows that information is more similar to normal goods than is usually thought.

When an agent gains new information, they update their beliefs – or equivalently, their causal graph of the world. In the traditional view, this might change the agent's choices over goods because they now know more about the goods they are choosing between, and can make more accurate decisions.

In the cognitive goods model, since agents gain value from the structure of this causal graph, new information actually changes the real value that an agent generates from the world. The information does not just help the agent better calibrate to the real world, it affects the meaningful state of that world itself. The agent may not immediately update their beliefs on exposure to new information; a learning process might be needed, especially if the information is complex or contradictory to an existing belief.

Two results of this are:

- agents may seek out information not because it is accurate but because it is rewarding

For example, curiosity (Loewenstein 1994) motivates people to find out information that they believe will reward them. For this to hold, the agent must expect that new information, on average, will be more positive than negative. A prediction from this, which could be tested empirically, is that happy people will be more curious than unhappy people.

- Agents may avoid information that would change the content of the causal graph in such a way as to reduce total reward

People are known to actively avoid information (Ho, Hagmann, Loewenstein 2018, Golman & Loewenstein 2015) and this model here provides a potential mechanism by which this would operate. It should be possible to empirically measure the underlying causal model and compare the predictions of the model with actual behaviour.

Köszegi and Rabin (2009) find a consistent result with regard to preferences over good and bad news, and propose theoretical consequences that follow if agents receive anticipatory utility from planning their future consumption. The current model suggests a mechanism by which agents might generate that utility (although without the imposition of further assumptions, it does not necessarily predict the same loss averse informational preferences that they rely on.)

The key prediction of the model is that people cannot consciously avoid activating specific nodes in the causal graph. If certain causes point to unrewarding effects, it may be impossible to avoid navigating to those unrewarding (painful) locations in the graph. However, if the graph does not contain any edges that point to those locations, the agent cannot accidentally navigate to them. Agents could therefore be motivated to avoid learning this information.

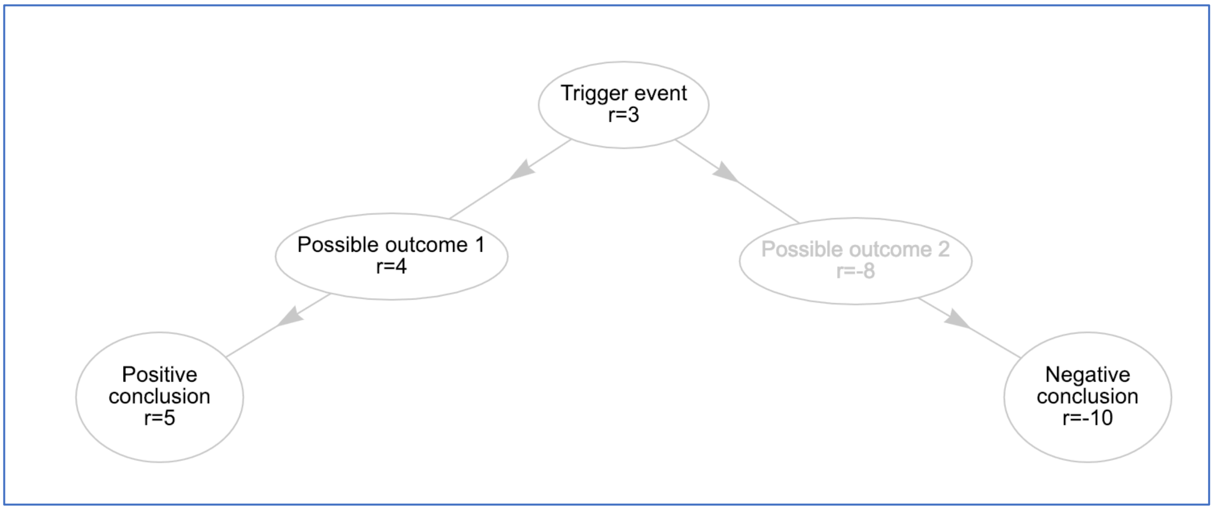

Figure 9a: before learning about possible outcome 2, the agent's overall valuation is positive

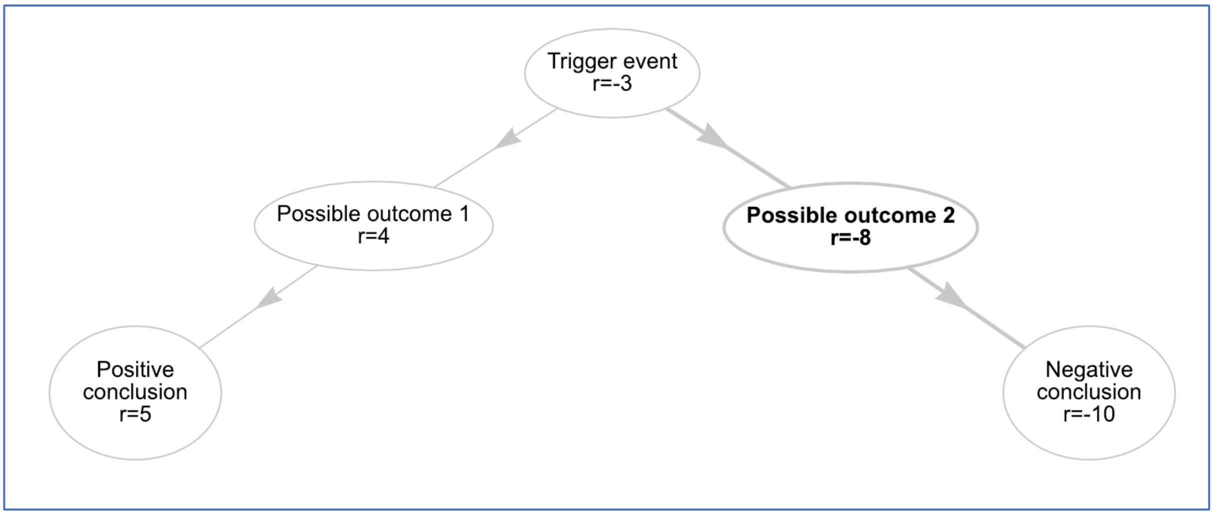

Figure 9b: after learning about possible outcome 2, the agent's overall valuation is negative

The way information is absorbed and communicated also creates supply constraints in the market for knowledge. Information is typically seen as a non-rival good, with no cost to copying or providing it once it has been created. However, if the meaningful component of information is not the external expression of it but the impact it makes on the agent's mental state, the cost of interpreting that expression must be considered part of the information's production function.

Conclusions from this argument include:

- Information markets are likely to be more partitioned and fragmented than we typically expect: with individuals gaining very different amounts (and even signs) of utility from the same external expression of a belief
- Information is non-satiable, with repetition of the same information still providing value
- Information can be priced, because the processes required to acquire and absorb it are still scarce
- Information is non-evaluable

The market for culture

Culture is usually consumed because the agent enjoys doing so – it produces rewarding cognitive goods. Agents may, in effect, choose to care about arbitrary nodes in their graph because doing so generates reward. A cultural object requires its viewer to willingly absorb its claims into their causal graph – knowing consciously that it may be fictional, but using *cognitive decoupling* to build the graph, hold it at arm's length and mentally navigate it for the duration of the consumption.

Other than the consumption of fiction (e.g. Lassie) a prime example of this is outcomes in sport. Although sport also has a social function, it seems likely that people attach significance to outcomes beyond just the utility of the social bonds involved. One mechanism by which this could occur is through an optimistic navigation of the causal graph to seek out potential positive outcomes. If agents have a bias (within this domain) towards imagining positive rather than negative outcomes, *or* if positive outcomes are simpler and less complex to imagine than negative ones, the model predicts that they will engage in thinking and speculation about sporting events.

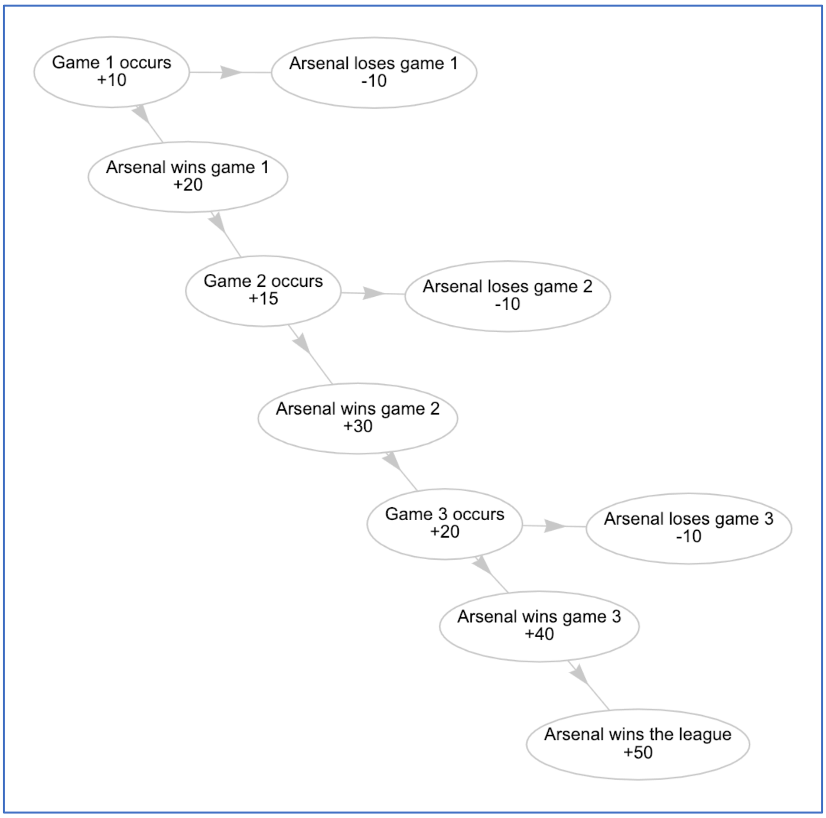

Figure 10: In this example, Arsenal needs to win three games in a row to win the league. If the agent truncates navigation of the graph at any point where this becomes impossible, they will focus only on the ultimately positive outcome and the overall experience of imagining the outcomes becomes rewarding – even though the probability of a positive outcome is low.

As well as providing subjective enjoyment to the consumer, the reward generated by these cognitive goods can also change their causal graph.

However, decoupling is never perfect. Our minds inevitably mistake some fictional claims for truth – not least because there is usually some truth in there. We willingly connect our reward centres to the decoupled version of the world – where is the fun in watching Lassie if we do not allow ourselves to care?

While my overt beliefs about Lassie's world may be clearly marked as fictional, my mind is wandering autonomously over those beliefs and my established model of the world, making hidden maintenance updates. Those updates can transfer parts of the Lassie model into my "real", permanent model.

If this does occur, cultural products can influence beliefs about the real world.

Valuation of normal goods

In a cognitive world, normal goods are no longer normal.

An agent's preferences over bundles of normal goods are no longer independent of the agent's mental state. All preferences are constructed based on the agent's causal graph. A bundle of goods activates nodes in the causal graph; the degree of reward generated by this activation guides the agent's choices.

The structure of the causal graph, and its state of activation prior to encountering the goods, both affect the degree of reward that the encounter generates. Therefore, preferences and choices over goods can in principle change as the agent's mental state varies.

Even more obviously, the agent's choices over normal goods are not only a function of the actual rewards those goods generate, but of the rewards the agent *predicts* they will generate prior to choice and consumption. Such predictions are firmly embedded in the agent's subjective model of the world.

The existence of a cognitive good in the agent's mind therefore affects their choices over normal goods. Cognitive goods act as complements or substitutes for normal goods.

For example: Awareness of dangers to the environment, a cognitive good, is a complement to recycled paper, a normal good. An increase in the total supply of that cognitive good will change the market price of recycled paper.

Marketing and advertising both directly create cognitive goods; this is the channel through which they change the market value of normal goods.

Money, as a normal good, is valued through the causal graph of agents throughout the economy. Agents' beliefs about both prices and availability of other goods affect their valuation of money. The cacheing of value through model-free learning then limits agents' ability to update this valuation.

Policy consequences

Policymakers may wish to affect the value that consumers put on money, as a monetary policy lever. The theory of cognitive goods suggests that they should:

- Encourage consumers to prospect over actual market prices, increasing the chance that they will update cached valuations
- Create new beliefs that directly influence the value of money: for example, increasing awareness of counterfeiting (to reduce valuation) or highlighting the highest potential returns available (to increase valuation). Being closer in the causal graph to the 'money' node, these beliefs will have an outsize influence on the valuation of that node.
- Encourage greater visibility of prices within transactions – for example discouraging the use of contactless payment or automatic direct debits from bank accounts.
- Focus these activities on agents whose valuation of money is most susceptible to change, or differs to the greatest extent to the policy-optimal valuation.
- Communicate directly about the value of money – possibly using pop-cultural tools.

Other potential policy implications include:

- Attempting to incorporate valuation of cognitive goods into overall consumer welfare functions. This model could provide support to the measurement of large-scale subjective welfare (BHKS) or the 'Gross National Happiness' concept (e.g. Diener 2000).
- Examining cognitive goods as a contributor to productivity. It is reasonable to expected that firms or individuals with a more accurate model of their customers' cognitive goods will be more efficient at producing wellbeing, whether or not this is reflected in priced market transactions.

Future research

Many of the themes mentioned above are developed only in outline form, and there is plenty of room for further theoretical exploration and empirical testing of the ideas proposed. If the process model gains further support from empirical neuroscientific or behavioural research, the implications for consumer microeconomic theory will be an interesting and rich area for working out.

Much work remains on the psychological foundations of the model. Although current directions in neuroscience research seem broadly supportive of the approach, it is too early to say with confidence that the mind definitely works on the basis proposed in this paper.

A number of the concepts discussed have previously been explored by both psychologists and economists qualitatively, but often without formal models. There has been a small body of cognitive economics literature (McCain 1992, Walliser 2008, Egidi and Rizzello 2003, Kimball 2015) some of which does propose more formal models, but it has rarely been connected explicitly to some of the empirically interesting problems that authors such as Loewenstein and Schelling have studied. Perhaps as a result, this field has never quite taken hold as a mainstream subfield in the economics literature. I hope that the model proposed in this paper will provide one possible source of momentum for future work.