CULTURAL RATCHETING RESULTS PRIMARILY FROM SEMANTIC COMPRESSION

JOANNA J. BRYSON

Artificial Models of Natural Intelligence, University of Bath
Bath, BA2 7AY, United Kingdom
jjb@cs.bath.ac.uk

Humans acquire far more of their behaviour from conspecifics via culture than any other species. Our culture is larger because it accumulates, where other species' seem to stay approximately the same size (Tomasello, 1999). This chapter attempts to clarify the problem of cultural accumulation by distinguishing between the *size* of a culture that can be transmitted from one generation, and the *extent* of culture transmitted. A culture's size is determined largely by ecological constraints, and certainly homonins (and some other species) show adaptations to facilitate this. But the exponential accumulation hypothesised by (Tomasello, 1999) I claim cannot be accounted for this way, but rather is a consequence of increasing information value in semantic components. This process can be achieved through memetics — semantics will be selected for which transmits the most information. Thus cultural evolution achieves compression of information, generating increased extent in culture even when maintaining a fixed size. I support my argument with evidence from simulations explaining the size of culture (Čače and Bryson, 2007), and simulations demonstrating selection for increased extent Kirby (1999).

1. Introduction

Although language is undoubtedly a cultural artifact of unique utility, the evolution of language may best be viewed as one aspect of a general exceptional human capacity to evolve culture. By *culture* here I mean behaviour acquired from conspecifics by means other than genetic transmission (Richerson and Boyd, 2005; Bryson, 2008). For the last decade we have known that other species share with us the fact that some part of their behaviour repertoire is cultural (Whiten et al., 1999; van Schaik et al., 2003; Perry et al., 2003); more recently we have realized some of it is even taught (Franks and Richardson, 2006; Thornton and McAuliffe, 2006). Thus humanity's uniqueness is a question of extents: both the quantity of culture per individual, and the persistence of recognisable elements of culture across generations.

Many people have attempted to explain why human culture is uniquely large and complex. However, many of the hypothesised mechanisms might themselves be consequences of our accumulation of culture. For example, Beppu and Griffiths (2009) claim that cultural ratcheting requires the communication of beliefs about

hypotheses. Clearly, cultural accumulation is accelerated by such a process. But if the bald statement that there can be *no* accumulation without it were true, then there could be no explanation of how we acquired the ability to communicate beliefs about hypotheses in the first place.

Ironically, Beppu and Griffiths (2009) base their work on a simulation which demonstrates exactly the sort of mechanism that leads to accumulation without beliefs or intentions. Their methodology is based on the iterated-learning paradigm invented by Kirby (1999), yet they (and their reviewers) overlook the result that made the paradigm famous. In Kirby's simulation, *the culture itself evolves to be more replicable*, without any change in the agents that transmit it. It does this through a process of regularisation: in later cultures, semantics is represented more efficiently. The process is not magic — it is the simple consequence of selection. Here the selective pressure is just the fixed number and length of transmissions an agent will hear in its lifetime. Where variation generates alternative mechanisms for expression, the most likely mechanisms to be transmitted to the next generation are the most general. Irregular forms may never be heard, and therefore drop out of the culture, unless they refer to very frequently-referenced notions. For such frequent references, selection from limited transmission favours brevity over generality. Thus even in a simple simulation, a fixed size for a culture is sufficient selective pressure to generate cultural accumulation.

I begin this chapter by explaining my thesis in better detail. I then analyse the results of Beppu and Griffiths (2009), showing an error in the mathematical model they present, and accounting for their experimental results. I use the concepts introduced in this analysis to examine what factors do account for ratcheting. I finish by re-addressing the question of human exceptionalism.

2. Thesis and Key Definitions

The fundamental claim of this article is that cultural ratcheting doesn't just mean transmitting more information, or at least not more bits of data. For real animals as for Kirby's simulated agents, the amount of communication from one generation to the next is largely determined by biological constraints such as lifespan or the carrying capacity of an environment. Lifespan determines how long the agents have to learn; carrying capacity determines how many agents can live near enough to each other to routinely transmit information. Our lifespan has not changed much compared to other apes. We can certainly support a more dense population, and we have learned to transmit behaviour faster, through teaching. But the majority of this increased rate of communication isn't about the number of signals, but about their quality. Essentially, the elements of culture that get selected *through* as well as *for* transmission between generation are those with the most meaning per transmission.

To clarify this, think about transmitting arithmetic. Imagine two societies that transmit how to add every possible pair of numbers between 1 and 999. One

culture does this by rote, the other teaches only 0–9 by rote, and then a few rules that enable adding numbers of up to 3 digits length. For the purpose of clarity in this paper, I will define *size* such that the size of the rote culture is significantly larger than that of the rule-based culture. But I will say that the *extent* of the two cultures is the same. *Size* is roughly the number of bits physically transmitted between individuals. *Extent* is the range of behaviours the culture enables.

Thus the fundamental claim of my article is that the size of the culture is set largely by biological constraints. Of course, the qualification *largely* matters here. Practices such as systematic teaching do increase the amount of information that can be reliably transmitted from one individual to the next. Other practices like agriculture fall somewhere between my two categories, since the effect of communicating culture relevant to food production (rather than some other kind of culture) is to increase the population density, thus addressing some biological constraints. This is a sort of increased cultural richness, though not the semantic one I am emphasising. My claim is that the main contributing factor to cultural ratcheting is the accumulation of increasingly powerful concepts. Or, using the terms just introduced, my thesis is that human-like exponential cultural ratcheting requires cultural extent increasing without cultural size.

3. Cultural Accumulation Cannot Require Accumulated Culture

Recently, Beppu and Griffiths (2009) reported that observing the behaviour of others is not sufficient to produce cumulative cultural evolution. They base this claim on two sources of evidence. First, they argue mathematically that transmission must necessarily result in a loss of information, making reference to the children's game of telephone $[US]$ or Chinese whispers $[UK]$. In this game, one person whispers a message to someone who has not heard it. After some number of such transmissions the final and original message are compared for amusement.

This signal-degradation observation has been both made and addressed before. Dawkins (2000) uses exactly the same metaphor to describe biological and memetic evolution. He then goes on to explain how errors introduced by noise are in the vast majority of cases corrected in true Darwinian selection by processes that apply rules for signal correction. In the Bayesian terms Beppu and Griffiths utilise, the point is that the receiver has priors as well as the sender, and in perceiving the message is able to correct for most noise. Sperber and Hirschfeld (2006) make largely the same observation in the specific context of modern human culture. In a reaction against the prior over-emphasis on imitation which Beppu and Griffiths (2009) are also responding to, Sperber and Hirschfeld credit their massive-modularity hypothesis with providing the additional information necessary for interpreting observed events. Information accumulated from some combination of genetics, individual experience and previously-acquired culture allows informed perception which effectively boosts a transmitted signal back to its original strength.

This is precisely why the game telephone requires whispering: not only to remove the chance of overhearing, but also to remove a great deal of linguistic information. This increases the probability of transmission error. Note that even so the message will still be perceived as language — no game of Chinese whispers ever ends with someone making white noise. Real cultural transmission is not so much like this game as like real telephone communication, where repeaters in the network amplify the signal in order to retain the signal over distance.

Of course there are some errors in transmission. More importantly, some individuals will die without ever transmitting some of the knowledge they have acquired culturally. In a sustainable cultural system, losses due to either death or corruption must be compensated for by individual discovery. For a culture to maintain its size, innovation merely needs to compensate, for growth, innovation must exceed loss. Errors of transmission may in themselves result in useful innovation, particularly if the 'error' is due to an intelligent perceiver making a 'correction' to a signal that is more sensible than the original transmission. This process can operate either with or without the awareness of a sentient receiver. Where transmission is perfect and entirely reliable, there would be no evolution and indeed no change.

Returning to Beppu and Griffiths (2009), the authors seek to verify their results with a second information source: live experiments. Taken in even a cultural-evolutionary context, the experiments are rather short, and they detect no significant accumulation from simple imitation. However, when the experimenters radically increase the amount of information transmitted by allowing the demonstrators to describe their hypotheses, accumulation becomes detectable. I suspect that the power of the initial form of the experiment is not sufficient that the apparent insignificance of its results are meaningful. That the level of accumulation is far lower with less information is unsurprising.

4. Determinants of Culture's Size and Extent

In previous work Čače and Bryson (2007) have presented a simulation model showing the spread and fixation of a tendency for the altruistic communication of food-processing strategies. By analysing this model, we can identify the factors determining the size of a culture. Basically, the size of a culture is determined by the probability of any particular piece of information being transmitted. Factors determining this probability include: the rate of information transmission, the maximum lifespan of an individual, and how many other agents an individual observes in the average time interval.

In the Čače and Bryson (2007) version of our simulation, individual bouts of communication transmit perfectly. Knowledge only leaves the simulation when an agent dies without transmitting it. It enters the system through a process of individual discovery. These discoveries occur at a low, fixed rate — for example, five percent of the population may discover one thing in their lives, and the rest

discover nothing. What they discover is random and may already be known in the local population. At the beginning of he simulation, knowledge is present at the rate of discovery — the population knows about 0.05 things. Over time this builds to an equilibrium value, where that equilibrium is determined primarily by the transmission probability.

The extent of culture is determined by the size of culture multiplied by the average behavioural utility of its content. In the our simulation, the extent of the the culture is fixed by its size, because all knowledge is equally valuable and takes equally long to transmit. This is the exact opposite of the case for the Kirby (1999) simulation described earlier. There, the size of the culture is fixed, but the expressivity of what is transmitted increases as that which is expressed becomes regularised. Frequently used terms become shorter even at the costs of regularity, while infrequently used terms must either be regular or be likely to be lost before transmission to the next generation.

For another example of increasing 'quality' of transmitted culture, Bechlivanidis (2006) has extended the Čače and Bryson (2007) simulation to include foods of varying quality, including 'bad' foods that it was detrimental to eat. As before, agents were equally likely to discover on their own any of the available foods, and further they were not able to individually judge a food's quality or long-term effects, but continued eating it out of habit or social influence. However, simple biology results in the culture favouring good foods. Healthier individuals are more likely to live long, and thus have more opportunity to socially influence those around them.

Bechlivanidis' work demonstrates that the Čače and Bryson results are robust to false beliefs being held and communicated. It also allows us to explore additional circumstances that can lead to enhanced cultural evolution. For example, if the agents attended preferentially to the most prestigious individual available, the culture grows faster because all agents tended to have better knowledge and therefore to be able to support a denser population. Prestige could rely on even simple indicators like apparent age. Note that contrary to a small detail of Richerson and Boyd (2005), prestige-led imitation proves adaptive even though the prestigious individuals receive nothing in return for the attention of their neighbours.

In general then, higher-quality information will be selected for if available. In line with the experimental outcomes of Beppu and Griffiths (2009), the extent of culture can grow more rapidly if abstract rules rather than raw data can be transmitted. Obviously, once cultural artifacts like language and religion evolved they massively facilitated cultural accumulation. Similarly, new concepts such as *self*, *desires* and *other* must massively increase the amount the range of plans available to cognition, by compressing a wide range of phenomena into relatively simple accounts (c.f. Donald, 1991).

5. Why Humans are Special

Why doesn't cultural ratcheting happen for all species? It's possible, of course, that it *is* happening for some species, but that they are still in a phase hominids were in only half a million years ago, with very little visible change in technology. However, the biological constraints on culture size reflect a set of tradeoffs well away from the niche of most species — for example, long lives and high cognitive plasticity for retention of information. Certainly large brains are expensive; probably the hominid ones co-evolved with our culture (Aiello, 1997; Silk, 2007). Notably, the recent period (50,000 years) of rapid cultural evolution has actually accompanied a significant *reduction* in hominid brain size. This could reflect the onset of modern language making transmission so easy that an adaptive advantage was still conferred even with reduced memory capacity.

It is nevertheless striking that wild chimpanzee culture seems to be of relatively fixed size. Further, chimpanzee culture seems to be recoverable within two generations after an artificial troop of orphaned chimpanzees is reintroduced into a naturalistic setting (Whiten, personal communication). One possible reason for this is that there are also *ecological* limits on the adaptive size and stability of cultures. If all chimpanzees persistently fed on particular optimal species for too long, those food sources would go extinct. It may be ecologically advantageous that chimpanzees continue to discover and move between new food sources from time to time.

Why then are humans such an exception? For some time I have been promoting the idea that the key difference between humans and other primates is simply that we are the only primates capable of temporally-precise imitation (Bryson, 2001, 2004, 2008, 2009). That is, humans are the only primate demonstrating the ability to precisely replicate not just snap-shot final postures, but entire temporal trajectories of variable values over many degrees of freedom (e.g. pitch, volume, formants). This allows the encoding of substantially more information, such as might allow arbitrary patterns to be encoded redundantly and thus robustly. I have previously argued that the difference between hominids and other vocal imitators such as song birds may be that our capacity for second-order referential reasoning, derived from the primates' complex social structures. This in turn allows for compositionality, which in turn facilitates the sort of holophrasis evolution originally hypothesised by Wray (1998), and to some extent modelled by Kirby (1999), whose simulations assume recursion. This theory may be unnecessarily elaborate, however. Our common ape ancestor was already larger (though notably not longer-lived) than other vocal imitators, and may have more easily been able to support large brains and therefore large memories for a large lexicon.

6. Conclusion

To return to my primary point, I have argued that ratcheting is not just about innovation or increasing the number of things transmitted in a culture, but rather primarily due to increasing the behavioural productivity of the average piece of culturally-transmitted information. I motivated this by arguing against the claim that simplistic transmission necessarily leads to degradation and loss. First, this loss is ameliorated by the expectations of the receiver, and second, to the extent it exists it can be compensated for by some rate of individual innovation. This point is critical to understanding the origins of language and sophisticated meta-cognition (beliefs about our beliefs.) Any scientific explanation of these phenomena must rest on culture accrued without them.

I have illustrated my arguments with several published simulations. These demonstrate that processes exist such as my argument describes. We cannot prove that these are the processes that lead to cultural evolution in humans, but a scientific theory is never proven certainly. Theories are judged by their likelihood given currently available data. The demonstrated existence of appropriate processes certainly increases the likelihood they might have played a role in our cultural history.

References

Aiello, L. C. (1997). Brains and guts in human evolution: The Expensive Tissue Hypothesis. *Brazilian Journal of Genetics*, 20(1).

Bechlivanidis, C. (2006). An examination of the role of prestige in cultural evolution with the aid of agent-based modelling. Technical Report CSBU–2006–21, Department of Computer Science, University of Bath, UK.

Beppu, A. and Griffiths, T. L. (2009). Iterated learning and the cultural ratchet. In *The 31st Annual Meeting of the Cognitive Science Society (CogSci 2009)*, pages 2089–2094, Amsterdam. Lawrence Erlbaum Associates.

Bryson, J. J. (2001). Embodiment vs. memetics: Does language need a physical plant? In Pfeifer, R. and Westermann, G., editors, *Developmental Embodied Cognition (DECO-2001)*, Edinburgh.

Bryson, J. J. (2004). Language needs second-order representations and a rich memetic substrate. In *Evolution of Language 5*, Leipzig, Germany. talk.

Bryson, J. J. (2008). Embodiment versus memetics. *Mind & Society*, 7(1):77–94.

Bryson, J. J. (2009). Representations underlying social learning and cultural evolution. *Interaction Studies*, 10(1):77–100.

Čače, I. and Bryson, J. J. (2007). Agent based modelling of communication costs: Why information can be free. In Lyon, C., Nehaniv, C. L., and Cangelosi, A.,

editors, *Emergence and Evolution of Linguistic Communication*, pages 305–322. Springer, London.

Dawkins, R. (2000). Foreword to Susan Blackmore's *Meme Machine*, pages *vii–xvii*. Oxford University Press, second edition.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press, Cambridge, MA.

Franks, N. R. and Richardson, T. (2006). Teaching in tandem-running ants. *Nature*, 439(7073):153.

Kirby, S. (1999). *Function, Selection and Innateness: the Emergence of Language Universals*. Oxford University Press.

Perry, S., Baker, M., Fedigan, L., Gros-Louis, J., Jack, K., MacKinnon, K., Manson, J., Panger, M., Pyle, K., and Rose, L. (2003). Social conventions in wild white-faced capuchin monkeys: Evidence for traditions in a neotropical primate. *Current Anthropology*, 44(2):241–268.

Richerson, P. J. and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. University Of Chicago Press.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S., and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*, 299(5603):102–105.

Silk, J. B. (2007). Social components of fitness in primate groups. *Science*, 317(5843):1347.

Sperber, D. and Hirschfeld, L. (2006). Culture and modularity. In Carruthers, P., Laurence, S., and Stich, S., editors, *The Innate Mind: Culture and Cognition*, volume 2, pages 149–164. Oxford University Press.

Thornton, A. and McAuliffe, K. (2006). Teaching in wild meerkats. *Science*, 313(5784):227–229.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.

Whiten, A., Goodall, J., McGew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., and Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399:682–685.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18(1):47–67.