TOWARD FORMALIZING CULTURE: FIRST STEPS

Selmer Bringsjord
Department of Cognitive Science
Department of Computer Science
Lally School of Management & Technology
Rensselaer Polytechnic Institute (RPI)

Abstract
After a quick autobiographical look back at some of the main cultural forces in my own early life, desiderata that perhaps any acceptable, rigorous theory of culture must satisfy are proposed. Following that, reaction to March's theory of culture is provided, which thereby points the way toward a more cognitively robust and realistic theory of culture that would make crucial use of formal logic.

Keywords: culture, theory, March’s model, formal logic.

1. Pre-Analytic Thoughts on “My” Cultures

One of the great joys of my job, indeed my life, is the privilege of traveling to new places, to interact with folks — predominantly — on matters of the mind; specifically on whether (and if so, how) minds can at least in part be given to machines. Such travel of course exposes me to cultures very different than the one in which I was initially raised, and the subsequent one in which I have long lived.

I was raised, in my early years, in a hybrid culture: that of Norway on the one hand, and New York City and its suburbs on the other. Both my parents were Norwegian, spoke Norwegian at home, and, invariably fueled by Norway’s distinctive cuisine, socialized with Norwegians — especially with Norwegians in Brooklyn, a borough of New York City, the Norwegian community of which has since, alas, evaporated. As to the cuisine, my mother was a devout devotee, and maintained not only that Norwegian food is preeminent, but that it is specifically better than what you can get in rival Sweden. I remember an early lecture from her as to why Swedish meatballs, though thought by many the world over to be peerless, are in actuality dreadfully drab and ridiculously small compared to their Norwegian counterparts. As to my father, he seemed to prefer the liquid side of my ancestral diet, especially powerful liquids.

At about the age of five, a second culture began to establish firm roots in my life, and gradually grew to overshadow the first. This second culture was definitely not Scandinavian: I remember realizing that my English-speaking classmates in 2nd grade had only the vaguest idea where the Scandinavian countries were on the big, spinnable globe in our classroom, and indeed literally no idea what countries composed the relevant region. In fact, they believed that America had been discovered not by the man from whom I am descended (Leif Erikson), but rather five centuries later by a latecomer named ‘Columbus’ who serendipitously landed in the “New World” rather far south of New York.
My classmates also celebrated only one independence day: July 4; May 17 meant nothing to them. But my family annually made the pilgrimage to Brooklyn for “our” parade on that May day.

So my second culture required no knowledge of Scandinavian geography and history, let alone Nordic drama\(^2\) and war; rather, my second culture was firmly and at once New York Cityish, and corporate and technological. I say ‘corporate’ and ‘technological’ because the suburbs of the Big Apple, and Gotham itself, and indeed the entire Hudson Valley, have long been singularly boosted by the great powerhouse of the union of both concepts, in the service of fast-moving commerce. I remember vividly when IBM, the oldest still-standing American pillar of corporate information-processing prowess, moved its world headquarters to Armonk, the town outside of New York City that my parents (along with some other Norwegians from the City) had decided to move to. IBM’s move meant that the small ski area in Armonk disappeared (since it was on Big Blue’s new land), replaced with a tow-less hillside that we could now ski only if willing to climb for each descent. I remember feeling that the disappearance of even a small ski area was catastrophic, in no small part because, as you probably know, Norwegian culture includes a passion for skiing and snow, and the inculcation in my case had been been thoroughly effective. But the culture of IBM, and the larger American techno-culture of which it was, and still is, in symbiosis, gradually enveloped me, and it may be no accident that today I still cherish the famous exhortation of Thomas Watson Sr.: THINK!.\(^3\) This became for a very long time the ubiquitous one-word slogan of IBM, and in many ways, given my still-vibrant interest in the marriage of human thinking on the one hand, and computing on the other, my early affirmation of and affinity for the slogan has persisted to this day.

I began by saying that I get to drop in on other cultures in the course of my job. An example is a most memorable trip I took a few years ago to the marvelous country of Romania, with my son. It was during that trip that I first began to ponder the possibility of formalizing culture. The specific catalyst of this reflection was the simple realization that Romanian culture includes knowledge of various propositions about Ceaușescu, and of communism generally. Just as anyone ignorant of the fact that the sun never sets in Tromsø in the summer is probably outside Norwegian culture, any adult ignorant of Ceaușescu and his dark rule is in all likelihood outside Romanian culture. Not only that, but it was clear to me that Romanian culture today is in no small part a function of Ceaușescu’s yesterday. Can such change be understood ahead of time, and managed, on the strength of a formal theory of culture?

This is certainly a gigantic question. One must no doubt start with smaller ones. For instance: Does culture consist merely in a collection of propositions, or is there more to it? My experience in Romania answered this question on the spot, for it was clear to me there that not only is assent to certain propositions essential for membership in a culture, but certain actions are crucial as well. My hosts in Romania didn’t just happen to address me by my title (‘Professor’) and family name; their behavior seemed to be the default in Romania. By contrast, in the States, whether I like or not, even youth who have never met me before, but know full well who I am and what I do for a living, routinely address me by my given name.
2. Desiderata Derived from Pre-Analytic Reflections

The previous section has been in significant part about cultures, considered briefly and impressionistically. But what is a culture, rigorously speaking? My goal is to answer this question from a logico-mathematical-computational perspective. That’s quite a hyphenated mouthful. Put in simpler terms, I want to represent culture in formal logic. And, in addition, I want to implement that representation in computation, so that culture can be at least simulated, so that perhaps in turn a computing machine can, in some limited sense, have cultural intelligence. Such intelligence would seem to be a prerequisite for a machine able to teach culture, which is a specific applied aim of mine. (Figure 1 is a snapshot of a robot in my lab designed to teach students the culturally rich Chinese tea ceremony.) A second aim is one I’ve already divulged: the engineering and use of machines with cultural intelligence that can predict and manage cultural change.

![Figure 1. RAIR Lab Robot Teaching the Chinese Tea Ceremony. (The engineering here is made possible by Dr. Naveen Sundar Govindarajulu, working with Michael Garber-Barron and Dr. Mei Si. Funding provided by RPI.)](image)

Put in different terms, my project falls under both logic-based artificial intelligence⁴ and computational cognitive modeling⁵. According to the methodology propounded in these and other such publications, individual cognizers, whether of the human or machine variety, are conceived as believing and knowing propositions about their environments, as reasoning over these propositions, and as agents that perform culturally appropriate actions in these environments. In the case of the tea ceremony, the difference between appropriate and inappropriate actions is so tangible and clear that a robot can process the difference, and educate humans accordingly. Of course, the robot in Figure 1 has no general human-level sense of what culture is, at all. But that doesn’t mean the robot can’t be very pedagogically useful.

Another important part of my methodology is an emphasis on proof. In general, “results” in cognitive science are almost invariably not in the form of theorems; that, in my view, is most unfortunate, since the absence of theorems means that very little is ever settled to the point of not only consensus, but verification⁶. I suspect that physics is in no small part successful because in large measure it is piloted by those who discover and communicate proofs.
But where to start in search of a general theory of culture that could be imparted to a robot, and that is ripe for progressing on the shoulders of proof? My first step is confessedly naïve: namely, make some observations about my own experiences. I started to do that in the previous section, which ended with the observation that part of culture is purely propositional, while part pertains to certain customary behaviors; and the distinction has been made in the present section as well. For example, with respect to IBM, its culture still includes a deep optimism about the efficacy of information-processing technology to make the world a better place (propositional), and it once included a firm adherence to the sartorial rule for men that required wearing a tie with a white shirt (behavior).

Let us try to bring things into clearer focus, by asking and answering the following question: In light of my informal remarks, which I believe would be generally echoed by anyone seriously reflecting upon their own culture, what desiderata, informally expressed, would apparently need to be satisfied by any formal account of culture? The following list seems to me to be quite conservative and accurate, and includes categories that I claim any theory of culture must accommodate.

**The Real:** This is simply a set of propositions that sums up all that holds. The set therefore includes that 2+2=4, that Earth is spherical, that human beings exist, and Goodstein’s Theorem, and so on *ad infinitum*.

**The Book:** Each culture can be said to have a veridical — as I shall call it — *Book*. The Book is composed of certain true propositions. It is a fact that Norway, in the summer, is — as it is said — “the land of the midnight sun.” (Even the southern tip, where most of my family resides, in summer, never really gets dark, as I remember learning firsthand as a little child lying wide awake in disbelief in the town of Lyngdal, in the middle of the night.) The Book contains certain historical propositions as well. In theory, through time, the Book can only get larger: nothing can be retracted from the Book. The Book is a proper subset of the Real.

**The Hope:** Whereas the book corresponds to reality, what I call the *Hope* need not. But at least most members of the culture in question nonetheless believe that all the propositions in the Hope hold. Though I do firmly believe that Ibsen is a truly great dramatist, it may not be the case that he is the equal of Shakespeare, yet that he is is in my experience believed by more than a few Norwegians. And though I do think the part of Grieg’s ouvre that taps directly into Norwegian folk songs is breathtakingly beautiful, it may not be the case that Grieg has outdone Mozart. But one can “hope” that such propositions hold, if one is Norwegian. The Hope includes less literary propositions: One time when my mother complained about my purchase of a Volvo automobile, and I retorted that, well, Norway doesn’t make any cars, she instantly asserted the subjunctive that if Norway *were* to make cars, I could be sure that they would be much better than Volvos and Saabs from Sweden. Her assertion here is a mundane member of the relevant Hope, for her. The Hope is somewhat person-relative in a given culture, and can shrink and expand through time: propositions can be dropped, and new ones can be added. This is of course a gradual process.

**The Habits:** I’m an habitual skier. As mentioned, that is more than consistent with being a traditional Norwegian: at least in the younger years, it’s a prescription. There are obviously analogues in other cultures; indeed, in *every* culture. What I’m calling
“habits” includes tea ceremonies in China, consuming aquavit with fish in Norway, or the appreciation of Pálinka, generous glasses of which I was introduced to during my aforementioned trip to Romania. Habits include customs. Like the Hope, the Habits are subject to change through time, and some members of the class can in fact fade away.

**The Inculcation, or Not:** One doesn't instantly become a member of a particular culture: membership requires an education, and a gradual one at that. Someone had to tell me that Leif Erikson was the “real” discoverer of America. I had to learn how to ski. Someone had to provide traditional Norwegian sweaters to me, and explain the historical basis of the distinctive patterns woven into them. And so on. Of course, sometimes there is rejection, on the part of members of a culture, of propositions in the Book (which is a bad idea, and often irrational: the Book, remember, is veridical), and/or of propositions in the Hope. This is why I say ‘or Not.’ For instance, I was never convinced that the Swedes are bad automotive engineers, or that the many dishes in their diet are across the board inferior to the counterparts found in the cuisine of their neighbors to the West. Hence I rejected some members of the Hope.

**The Reasoning:** On the cognitive side of culture, if the pre-analytic data I’m seeking to mine is a sound guide, there is more than knowledge and belief: there is also, connected to this knowledge and belief, reasoning. This reasoning comes in various forms. Most hopes are defended with reasoning. For instance, I remember challenging my mother’s assertion that Norway won World War II for the Allies. She proceeded to present an argument to the effect that, were Hitler to have achieved an ability to build and use nuclear weapons, he would not have been defeated (first premise), and — here the second premise — it was a group of Norwegians who prevented him from reaching this capability. Therefore, she concluded (with at least some tongue in cheek), Norway did in fact win WW II. Of course, reasoning is applied in a seemingly infinite number of ways in association with the Book and the Hope.

I do not pretend that this is an exhaustive list; I assert only that some of the chief dimensions of culture are here pointed to, albeit pre-analytically. And I gladly admit that much more needs to be said about each category in the list, even at the informal level. For instance, religious views are a very important part of the Hope, in many cultures. Even the tea ceremony the robot in Figure 1 (partially!) understands is wrapped up with, indeed has its historical roots in, Buddhism. In fact, my goal of understanding culture, formally and computationally, and of enabling a machine to achieve a degree of this understanding as well, is motivated by a desire to model the clashing of cultures specifically in the area of religion (since so much conflict seems to arise out such clashes), and to use modeling and simulation to find futures in which such clashes can be resolved, and in some cases outright prevented. This of course means that mathematizing culture will require mathematizing religion, and ethics (since in most cultures that have religious aspects, morality is tied to those aspects); the road ahead is not easy, and for those not sharing my Leibnizian confidence that all of cognition can be formalized in logic, it will doubtless appear to be one that simply cannot be traveled.
March's Non-Logicist Model as a Starting Point

I pointed out above that culture involves not just countries, regions, and cities, but also corporations, such as IBM. My suspicion is that corporate culture is probably a more reasonable place to gain a formal foothold than the culture of an entire country or region. The latter scale is dauntingly complicated. Fortunately, it turns out that some rather clever thinkers have considered how corporate culture might be formalized, at least to a degree, and we can consider how the work of such thinkers measures up to the desiderata listed in the previous section. Chief among the thinkers in question is March, who offers a formal, and highly influential, if inexpressive, model that can be summarized as follows.

First, reality $R$ is represented as a vector $(d_1, d_2, ..., d_m)$. Each $d_i$ represents a "dimension" of reality, and has a value of 1 or -1. We assume that every organization includes a set of $n$ agents $i_1, i_2, ..., i_n$. $R$ constitutes a kind of "ground truth," and the values of its dimensions are independent of the belief of agents. There is also in March's scheme the concept of the beliefs of agents with respect to the dimensions of reality at a given time $t_k$, which I write as $b(i_i, d_j, t_k)$, and which has a value of 1, 0, or -1 through time. In addition, every corporate culture is assumed to have an organizational code of received truth, which includes likewise a value of 1, 0, or -1 for each of $c(i_i, d_j, t_k)$.

March makes this model dynamic by first legislating that the code can affect the beliefs of agents. We can formulate simple update rules to capture March's ideas: first:

If $c(i_i, d_j, t_k) = 0$, then $b(i_i, d_j, t_{k+1}) = b(i_i, d_j, t_k)$.
If $c(i_i, d_j, t_k) = 1/\text{-1}$, then $b(i_i, d_j, t_{k+1}) = 1/\text{-1}$.

But these conditionals only regiment change in the direction of the code to agents. What about the other direction? How can the beliefs of agents impact the code? The second part of March's updating machinery includes that only "superior" agents can cause a change in the code. Superior agents are those who have beliefs that match reality on more dimensions than the code does. We can invoke a simple counting function $C'$ that computes, for the relevant inputs (viz., the values of $b$ and $c$) whether a given agent does exceed the veridicality of the code at any timepoint $t$. We invoke a second counting function $C^2$ that yields 1 for a given $c(i_i, d_j, t_k)$ if the majority of superior agents differ in their value for $b(i_i, d_j, t_k)$. Given this machinery, the update rule for the code requires that the values of a given $c(i_i, d_j, t_k)$ be changed to match those of the differential beliefs of the superior agents. It is thus seen that March can build some simple simulations, by picking a starting configuration.

Preliminary Steps Toward Expansion and "Logification" of March's Scheme

Unfortunately, as even casual study of my opening informal remarks about culture make plain, March's scheme is inadequate. The inadequacy becomes even starker when one places March's scheme alongside the desiderata listed above. I now sketch some of the steps that need to be immediately taken in order to address these inadequacies, by sketching some simple logico-mathematical-computational machinery.

To begin, reality, $R$, on March's model, has no declarative content. This is unacceptable, since, necessarily, to say that there is such a thing as external reality is to say that such-and-such holds. We thus make a "simple" change: instead of talking of a vector of
dimensions, we simply talk of a vast collection of formulae \(<p>\), each of which represents some proposition \(p\) in \(\mathbb{R}'\), which represents all that is in the Real. Ultimately, the formal language underlying \(\mathbb{R}'\) will be dizzyingly rich and expressive, and certainly no such thing has yet to be discovered and specified.\(^{13}\)

As to the beliefs of agents on March’s scheme, we will need to make another change in order to head in the direction of doing justice to the Book and the Hope: at a minimum, we shall need to say that beliefs have targets: namely, propositions. We need to do this because March’s framework is here again bereft of declarative content. Instead, then, of a belief \(b(i_j,d_k)\) simply returning a value of 1, 0, or -1, a belief will need to be an operator \(B\) ranging over an agent \(i_j\) who is a member of the culture in question, a formula \(<p>\) believed by that agent, and other things that are beyond the scope of the present paper (e.g., a timepoint at which the belief is held). We will thus write such things as \(B(i_j,<p>)\), which is certainly in line, as many readers will note, with formulas in epistemic logics. To model the Book, the objects of belief are in fact elements of the Real, and in addition the Book includes knowledge, which will be captured by formulae having the general form \(K(i_j,<p>)\).

But what about dynamism? How is that to be handled? This is certainly a challenge, but one that needs to be met. For our pre-analytic data unmistakably implies that as time flows on, changes ensue: Hopes and Habits, after all, come and go. In addition, while March’s framework includes no provision for communication between agents, such communication is obviously at the heart of culture. I would not be a member of a culture without the human communication between myself and other humans in that culture.

Given this observation, and given the methodology to which I am committed, a number of prominent pre-existing formalisms present themselves for consideration. For example, dynamic epistemic logic\(^{14}\) warrants consideration. In this family of logics, what agents believe and know can change through time in principled fashion on the basis of what is communicated. Unfortunately, the engine for dynamism that one finds in pre-analytic data about culture involves not just flat, stark communication of information (announcements, as they are often called in dynamic epistemic logic), but argumentation and discussion (as the example given above relating to Swedish versus (hypothetical) Norwegian cars reveals). Something much richer than dynamic epistemic logic is clearly needed.

To model interaction among agents in a culture, I am inclined to favor using the dynamic model of argument and counter-argument pioneered by John Pollock,\(^{15}\) but even sketching this approach is beyond the range of our current prolegomenon. In addition, my approach to formalizing culture requires that communication between human beings be ultimately cashed out as communication of information expressed in logic. But about these tricky topics I will say no more, and instead now end with a brief conclusion.

5. Concluding Remarks

This short paper has explored just the first propaedeutic steps toward a full computational formalization of culture, in order to ultimately not only slightly advance the science of culture, but to eventually enable relevant engineering (e.g., of technology for teaching culture, for predicting the effects of actions that impact a culture, etc.). What are the next steps?

First, an argument-centric set of mechanisms for agents to affirm or deny propositions in the Real and the Book must be designed. This will allow members of a given culture to
have changing epistemic profiles through time, as a function of communication in natural language.

Second, the application of these mechanisms must be driven by, and integrated with, the goals and plans of agents, including goals and plans in play in multi-cultural “markets.” Such mechanisms are completely absent from March’s scheme, but they are clearly crucial, as the pre-analytic review herein reveals. Whether it’s a company like IBM, or a country like Norway or the U.S. or Romania, agents not only perform actions given to them in rigid scripts, but they take actions in order to serve their more general goals. Hence, a planning formalism must be established.

With these two steps taken, it will be time to implement a simulator able to test hypotheses about how a culture is impacted by certain “inputs.” In the (very) long run, it might therefore be possible to predict what will happen to cultures as a result of major changes, such as Romania’s rapid change from dictatorial communism to a more free-market economy. Such predictive power would be the handmaiden of attempts to secure certain futures over others, by certain means over others.

Notes

1. I am indebted to seed funding from RPI for the purpose of exploring the possibility of AI systems able to “understand” and teach culture. This funding has helped launch the work of concretizing some of what is adumbrated in the present paper. Scare quotes are used because I claim to have shown, e.g. in S. Bringsjord, What Robots Can’t and Can’t Be, that — in light of my improved versions of Searle’s Chinese Room Argument (CRA) — no computational system can genuinely understand anything. For a further improved CRA, see S. Bringsjord and R. Noel, “Real Robots and the Missing Thought Experiment in the Chinese Room Dialectic.”

2. I remember from my earliest days a deep homage to Ibsen at home, and perhaps an even deeper veneration for Grieg, probably because he incorporated and elevated some of the very same folk music my mother played (piano and accordion).

3. Thomas Watson Sr. was CEO of IBM from 1914 to 1956, during which time computing was wed to punched card tabulation. For basic info, including some regarding the distinctive culture Watson created, see http://en.wikipedia.org/wiki/Thomas_J._Watson.


6. For more on this defect in cognitive science, see S. Bringsjord “Declarative/Logic-Based Computational Cognitive Modeling.”

7. One can find some information here: http://en.wikipedia.org/wiki/Norwegian_heavy_water_sabotage. There are a number of renditions of these heroics in popular media; e.g., see the film The Heroes of Telemark. For greater accuracy, and for information on survival skills that in fact still a not-insignificant part of Norwegian culture, see Mears, The Real Heroes of Telemark: The True Story of the Secret Mission to Stop Hitler's Atomic Bomb (London, UK: Hodder & Stoughton, 2003).

8. For a step in the direction of such daunting formalization, see S. Bringsjord and J. Taylor, “The Divine-Command Approach to Robot Ethics.” It is probably important to note that Simon, one of the founder of modern AI, and a nobelist in economics (and as it happens a collaborator with March), deserves much credit for at least suggesting that businesses can be computationally simulated. Ultimately it is this suggestion of Simon’s, combined with machine-reasoning in formal logic (another trajectory that Simon launched via his famous logic theorist program of 1956), that was rattling around in the back of my mind as I spent time in Romania. For Simon on computational simulation of organizations, see J. March and H. Simon, Organizations. Information regarding logic theorist can be obtained at http://en.wikipedia.org/wiki/Logic_Theorist.

10. I provide more notation than that given in J. March “Exploration and Exploitation in Organizational Learning,” in order to ease exposition. My syntax in no way extends the semantics of J. March, ibid..

11. To ease exposition, I leave aside the probability parameters in J. March, “Exploration and Exploitation in Organizational Learning.”

12. For a glimpse of what kind of formal language will be required, see S. Bringsjord and N. S. Govindarajulu, “Given the Web, What is Intelligence, Really?” *Metaphilosophy* 43.4 (2012): 361–532.


**Bibliography**


