Advances in Cognitive Systems 1 (2012) 23–36 Submitted 1/2012; published 7/2012

The Right Way

Patrick Henry Winston PHW@MIT.EDU
Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139 USA

Abstract

I ask why humans are smarter than other primates, and I hypothesize that an important part of the answer lies in the *Inner Language Hypothesis*, a prerequisite to what I call the *Strong Story Hypothesis*, which holds that story telling and understanding have a central role in human intelligence. Next, I introduce the *Directed Perception Hypothesis*, which holds that we derive much of our common sense, including the common sense required in story understanding, by deploying our perceptual apparatus on real and imagined events. Both the Strong Story Hypothesis and the Directed Perception Hypothesis become more valuable in light of our social nature, an idea captured in the *Social Animal Hypothesis*. Then, after discussing methodology, I describe the representations and methods embodied in *Genesis*, a story-understanding system that analyzes stories ranging from précis of Shakespeare's plots to descriptions of conflicts in cyberspace. Genesis works with short story summaries, provided in English, together with low-level *common-sense rules* and higher-level *concept patterns*, likewise expressed in English. Using only a small collection of common-sense rules and concept patterns, Genesis demonstrates several story-understanding capabilities, such as determining that both *Macbeth* and the *2007 Russia-Estonia Cyberwar* involve revenge, even though neither the word *revenge* nor any of its synonyms are mentioned.

1. Clues from Caves

Just about everyone agrees that much has been accomplished since Turing (1950) published his seminal paper, *Computer Machinery and Intelligence*. On the other hand, most would also agree that less has been accomplished than expected. Although applications of artificial intelligence are everywhere, we still do not have a computational theory of human intelligence. A team of dedicated first-class engineers can build systems that defeat skilled adults at chess and Jeopardy, but no one can build a system that exhibits the common sense of a child.

What has been tried? Turing argued that human intelligence is a matter of complex symbolic reasoning. Minsky (2006) argues for a multiplicity of ways of thinking coupled into a reasoning hierarchy with instinctive reactions on the lowest level and self-conscious reflection on the highest level. Brooks (1991) argues that, whatever human intelligence is, studying it directly is beyond the state of the art, and we must instead organize systems in layers of competence, starting with the objective of understanding low-level layers that produce insect-level intelligence. Still others, in many seminal papers, have suggested that the right way is, for example, through architectural design (Laird et al., 1987), neural mimicry (McClelland & Rumelhart, 1986), or statistical methods (Pearl, 1988).

Each of these approaches has made important contributions, especially from an engineering perspective, but none has shown us light at the end of the tunnel, not yet at least.

What is missing, I think, is an approach centered on asking what exactly makes humans different from other primates and from early versions of ourselves. For guidance, I ask about the early history of our species, and I find provocative suggestions in the speculations of paleoanthropologists, especially those of Tattersall (1998). Basically, Tattersall believes we are symbolic and other primates were not and are not. He says we were not symbolic either, until about 50,000–70,000 years ago. Before that, we were structurally rather modern for perhaps 100,000 years, but during that earlier period, like the Neanderthals, all we could do was make simple stone tools and work with fire.

Then, we started making art, and eventually produced the drilled seashell jewelry found in the Blombos Cave, the cave paintings at Lascaux, and the figurines at Brassempouy. Such art, Tattersall believes, requires symbolic thinking and its appearance is evidence of becoming symbolic. Tattersall argues that we became symbolic rather suddenly, probably in southern Africa, possibly in a population reduced to a few thousand or a few hundred individuals. It was not a matter of slowly growing ability proportional to slowly growing brain size. More likely, it was a evolutionary accident, with nonlinear effects, that unleashed the power of other faculties previously evolved by selection for benefits other than producing human-level intelligence.

2. Four Hypotheses

Of course, saying we are symbolic does not take us very far toward a computational theory. Chomsky (2005), who frequently cites Tattersall, takes us further by suggesting that we are unique in our ability to combine two concepts to make a third without limit and without disturbing the contributing two. To a linguist, Chomsky's suggestion sounds like the *merge* operation, central to minimalist theories of language. To many practitioners of artificial intelligence, his suggestion sounds like the capacity to construct complex descriptions, using some sort of language, which leads to the following hypothesis:

> **The Inner Language Hypothesis:** Using a symbolic inner language, we construct symbolic descriptions of situations and events that are far beyond the reach of other primates.

I propose to take the inner language hypothesis a step further. I believe that we humans developed the ability to string event descriptions into stories; that we further developed an ability to move backward and forward in remembered stories to explain and predict; that our story processing ability came to include the ability to combine stories into new stories never previously witnessed, from which imagination emerged. Thinking about this kind of thinking has led me to posit another hypothesis:

> **The Strong Story Hypothesis:** Our inner language enables us to tell, understand, and recombine stories, and those abilities distinguish our intelligence from that of other primates.

Why are stories so important? Because human education is full of stories, starting in modern times with the fairy tales of childhood, through the lessons of history, literature, and religious texts, and on to the cases studied in law, medicine, business, engineering, and science, complemented by the stories told to us by our parents, siblings, and peers. Even learning to follow a recipe when we learn a skill can be viewed as a special case of story understanding.

The pioneering natural-language work of Schank and his colleges and students, documented in numerous articles and books (Schank, 1972; Schank & Abelson, 1977; Schank & Riesbeck, 1981; Schank, 1991), presumed that stories are important. So does more recent work on story characterization and story generation, such as found in the numerous papers of authors such as Graesser (2011) and Young (2007). Here, with the Strong Story Hypothesis, I claim that story understanding is not just important, but rather that story understanding is the centrally important foundation for all human thinking.

Given that story understanding is centrally important, the next question is: Where does the common-sense knowledge needed to understand a story come from? We humans do not get it from the Web or from manually built common-sense databases, and even without a desire to understand what makes us different from other primates, depending on the Web or other sources of common-sense data is ill advised, because we know a lot we have never been told nor are likely to be told nor are likely to find written down anywhere.

I believe we generate much of what we know as needed, via the interaction of our symbolic and perceptual systems. Sometimes our symbolic system drives our vision system to engage itself on information in the physical world; at other times our symbolic system drives our visual system to engage itself on an imagined world.

I believe my point of view is well aligned with the work of Ullman (1996) on visual routines, which in turn was inspired by many psychophysical studies that suggest our human vision system is a powerful problem solver, not just an input channel. Accordingly, it is natural to draw a picture and move a problem from our symbol-processing faculties to our visual faculties whenever the problem is easier to solve on the visual side.

We often do not have to draw a picture, however, because imagination is enough. Consider this simple statement–question example: "John kissed Mary. Did John touch Mary?" Everyone seems to answer the question by deploying visual processes on an imagined kiss. Once that is done, the action-consequence knowledge can be cached as a rule, but being able to get the common-sense answer through perception means you can answer the question, when asked, even if you have not had any sort of kissing education.

Here is a more complex example from personal experience. As a friend helped me install a table saw, he said, "You should never wear gloves when you use this saw." At first, I was mystified, then it occurred to me that a glove could get caught in the blade. No further explanation was needed because I could imagine what would follow. It did not feel like any sort of formal reasoning. It did not feel like I would have to have the message reinforced before it sank in. It feels like I witnessed a grisly event of a sort no one had ever told me. I learned from a one-shot surrogate experience; I told myself a story about something I had never witnessed, and I will have the common sense to never wear gloves when I operate a table saw.

Such examples lead me to another hypothesis:

> **The Directed Perception Hypothesis:** Our inner language enables us to direct the resources of our perceptual systems to answer common-sense questions about real and imagined events, generating common-sense knowledge as a by-product.

Thus, I believe our inner language enables not only story manipulation but also the marshalling of our perceptual systems, especially our vision system, to solve problems on our behalf and produce symbolically cached common-sense rules.

So far, I have discussed the value of an inner language, but what about our outer language? The inner language must have come first, because there is no point in talking if you have nothing to say. Also, there is no point in talking if there is no one to talk to, so it must be important that we are social animals. This suggests a fourth claim:

> **The Social Animal Hypothesis:** Our social nature amplifies the value of story understanding and directed perception.

Once we somehow acquired an inner language, developing an outer language added substantial benefit and perhaps the inner and outer language coevolved rapidly. Being a social animal does not mean, however, that a sophisticated outer language will emerge, and along with it, an inner language. Other bipedal primate species have been social animals, but none seems to have become symbolic, including the Neanderthals.

I believe the four hypotheses are inseparable, and one without the others has only limited potential. With no inner language, there can be no story understanding or directed perception. With no story understanding or directed perception, having an inner language might enable some sort of reasoning, but not much educability. Without connection to perception, story understanding reduces to disconnected symbol manipulation by a system that may appear to be quite intelligent, but that depends too exclusively on linguistically supplied knowledge. Without connection to story understanding, an otherwise capable perception system can initiate reflex action, but lacks the ability to chain events together, to move backward and forward in such chains, to explain, and to predict. Without being a social animal, there would be less learning and less pull on inner and outer language evolution.

3. Five Steps: The Genesis Example

The four hypotheses may not be the whole story, but I believe they are sufficiently important to invite a great deal of research. Moreover, I believe that such research should proceed by looping through five steps: identify the competence to be understood; formulate computational problems; propose computational solutions; develop an exploratory implementation; and crystalize emergent principles. These steps are reminiscent of the methodological-step recommendations of Marr (1982). The rest of this section, by way of example, illustrates how the steps have guided my research program on story understanding.

3.1 First Step: Identify the Competence to be Understood

The competence to be understood is that of analyzing stories, such as the following rendering of the plot from Shakespeare's *Macbeth*:

> **Macbeth:** Macbeth, Macduff, Lady Macbeth, and Duncan are persons. Macbeth is a thane and Macduff is a thane. Lady Macbeth, who is Macbeth's wife, is greedy. Duncan, who is Macduff's friend, is the king, and Macbeth is Duncan's successor. Macbeth defeated a rebel. Witches had visions and talked with Macbeth. The witches made predictions. Duncan became happy because Macbeth defeated the rebel. Duncan rewarded Macbeth because Duncan became happy. Lady Macbeth, who is Macbeth's wife, wants to become the queen. Lady Macbeth persuades Macbeth to want to become the king. Macbeth murders Duncan. Then, Lady Macbeth kills herself. Dunsinane is a castle and Burnham Wood is a forest. Burnham Wood came to Dunsinane. Macduff had unusual birth. Macduff fights with Macbeth and kills him. The predictions came true.

I have used simple plot summaries from Shakespeare as anvils on which to hammer out ideas since my earliest work on analogy (Winston, 1980). My students and I still use them in our latest, much advanced work because they are easily understood and because they are rich in universally important factors such as power, emotion, consequence, and ties between people. We have found that the same kind of common-sense rules and concept patterns that work for Shakespeare also work for international conflict, such as the alleged 2007 Russian cyberattack on Estonia's network:

> **Cyberwar:** Estonia and Russia are countries. Computer networks are artifacts. Estonia insulted Russia because Estonia relocated a war memorial. Someone attacked Estonia's computer networks after Estonia insulted Russia. The attack on Estonia's computer networks included the jamming of web sites. The jamming of web sites showed that someone did not respect Estonia. Estonia created a center to study computer security. Estonia believed other states would support the center.

Macbeth and *Cyberwar* are representative in length and sophistication of the two dozen stories on which we have focused our attention. Of course, two dozen is a small number, but remember that our ultimate purpose is to understand human understanding, not to engineer systems that only give the appearance of understanding by processing Web-sized story sets in ways that shed little light, if any, on human understanding.

What do we mean by *understanding*? After reading stories such as *Macbeth* and *Cyberwar*, everyone has the competence to answer questions that range from obvious to thought provoking, many of which do not have explicit answers in the stories themselves: Who ends up dead? Why did Macduff kill Macbeth? Do the stories involve revenge? Which story presents a Pyrrhic victory? Is there a *Macbeth* role in the Russo-Estonia cyberwar? Is Russia's alleged attack on Estonia's computer networks an instance of *revenge* or *teaching a lesson*?

3.2 Second Step: Formulate Computational Problems

The competence to be understood draws attention to the key computational question: What representations make it possible to understand stories and to exploit that understanding? Certainly, knowledge will have to be represented, for without a representation, there can be no model, and without a model, there can be no explanation, prediction, or control.

We could just use some sort of semantic net as a universal representation to cover everything, but we felt it would be instructive to see what kinds of knowledge are needed in story understanding, how much of each kind is needed, and how often each kind of knowledge is put to use. Also, we were guided by the principle that refined description tends to expose regularity and constraint. Regularity and constraint are important, of course, because a model that supports story understanding must involve common sense and the ability to reflect on the implications of common sense.

Thus, the knowledge-representation question has several aspects, such as: how do we represent physical, social, and emotional relations and qualities; how do we represent actions and events; how do we represent common-sense and reflective knowledge?

3.3 Third Step: Propose Computational Solutions

With the competence specified and the representation problem identified, my students and I—the Genesis Group—anticipated we would need many representations.

We started with explicit representations for categories whose importance is self evident: class, because what you are determines what you can do (Vaina & Greenblatt, 1979); transition, because human reasoning seems to focus on how change causes change (Borchardt, 1994), and trajectory, path, and place, because movement along paths is extraordinarily common in language (Schank, 1972; Jackendoff, 1985). Next, as we discovered representational needs in exploratory implementation work, we added representations for dealing with coercion (Talmy, 1988), cause, goal, persuasion, belief, mood, possession, job, social relations, and time. Then, we added property and role-frame representations.

Genesis also has a representation for *common-sense if–then rules*, for much of story understanding seems to be routine inference making, such as knowing that, if one person kills another, then that other person is dead. Such rules connect explicit events in the story text with inferred events to form what we decided to call an *elaboration graph*. Common-sense rule chaining seems necessary but not sufficient for story analysis, because higher-level reflection seems to require search. *Revenge*, for example, is a harm event leading to a second harm event with the actors reversed, possibly with a long chain of intermediate events. I refer to such descriptions as *concept patterns*. Genesis deploys them using breadth-first search in the elaboration graph. This type of analysis is very different in detail, but inspired by, the pioneering work of Lehnert (1981).

Collectively, all our representations constitute Genesis's *inner language*. The representations in the inner language have come to enable description of just the sorts of concepts that would be important for survival, particularly classification, movement in the physical world, relationships in the social world, and various kinds of causation. Thus, Genesis's inner language is a cognitive substrate similar in spirit to that proposed by Cassimatis *et al*. (2008), although different in detail. Perhaps something like Genesis's inner language may eventually shed light on the inner language with which we humans describe the world.

3.4 Fourth Step: Develop an Exploratory Implementation

With the behavior specified, the computational problems identified, and posited solutions in hand, we set out to develop the exploratory Genesis system.

As a design principle, we decided that all knowledge provided to Genesis—including stories, if-then rules, and concept patterns—would be provided in English. We were motivated by our debugging philosophy and by the permanence of English; we knew that were we to start over, at least our knowledge base would be reusable.

Given our English-only decision, we had to choose a means to get from English to descriptions couched in our representation suite. Having tried a popular statistical parser, we eventually choose to use the Start Parser, developed over a 25-year period by Katz (1997) and his students, because the Start Parser produces a semantic net, rather than a parse tree, which made it much easier for us to incorporate it into a system that translates from English into descriptions in Genesis's inner language. We also chose to use WordNet (Fellbaum, 1998) as a source of classification information. We sometimes augment WordNet with information in English as in "A thane is a kind of noble."

With our Start Parser-enabled translator, we can readily express the needed if–then rules in English. Examples follow, exactly as provided to Genesis: If X kills Y, then Y becomes dead. If X harmed Y and Y is Z's friend, then X harmed Z. X wanted to become king because Y persuaded X to want to become king. Henry may want to kill James because Henry is angry at James. If James becomes dead, then James cannot become unhappy.

As the examples show, rules can be expressed as *if–then* sentences or *because* sentences, with or without regular names, and possibly with the modifiers *may* or *cannot*. *May* marks rules that are used only if an explanation is sought and no other explanation is evident. *Cannot* marks rules that act as censors, shutting off inferences that would otherwise be made. In the example, a person does not become unhappy when he dies, even though killing involves harm and harm otherwise causes the harmed to become unhappy.

Concept-pattern descriptions are a bit more complicated. Here are two versions of *revenge*:

- Revenge 1: X and Y are entities. X's harming Y leads to Y's harming X.
- Revenge 2: X and Y are entities. X's harming Y leads to Y's wanting to harm X. Y's wanting to harm X leads to Y's harming X.

Which is the right version? That, of course, depends on the thinker, so we are able to model specific thinkers by including more or less sophisticated, or more or less biased, ways of looking at the world.

Equipped with common-sense rules, Genesis produces elaboration graphs of predictions and explanations, such as the elaboration graph shown in Figure 1. The white boxes correspond to elements explicit in the text; the gray boxes correspond to common-sense inferences. Note that, according to the connections in the graph, Macduff killed Macbeth because Macbeth angered Macduff. Fortunately, we do not always kill the people who anger us, but in the story, as given, there is no other explanation, so Genesis inserts the connection, believing it to be plausible.

Given the elaboration graph, Genesis is ready to look for higher-level concepts of the sort we humans would see in the story but only if we reflect on what we read. Genesis sees, for example, not only *Revenge* but also a *Pyrrhic victory* in the elaboration graph for Macbeth shown in Figure 2: Macbeth wants to be king, murders Duncan to become king, which makes Macbeth happy, but then the murder leads to Macbeth's own death.

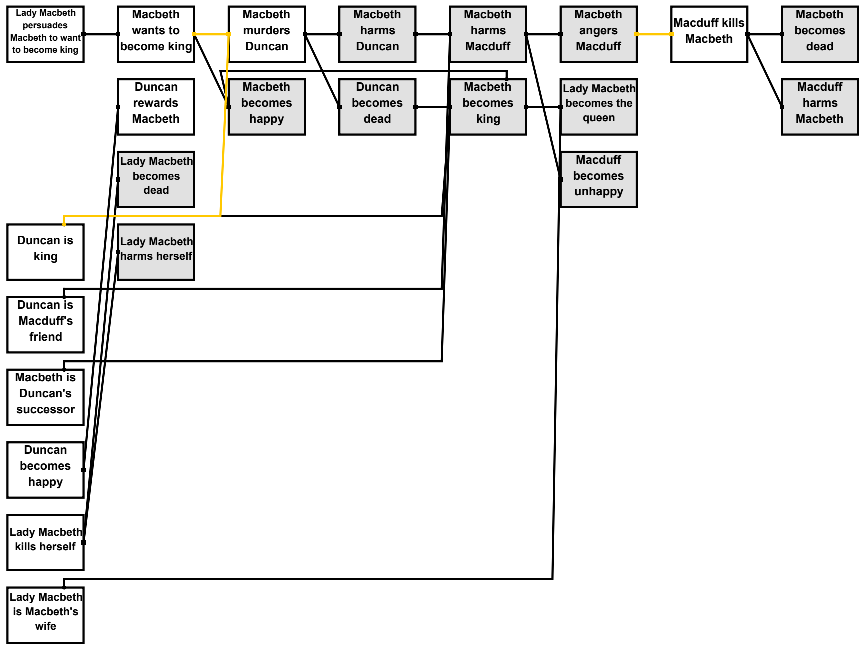

Figure 1. Genesis's story understanding system produces an elaboration graph from common-sense rules together with a story. White boxes indicate information given explicitly in the Macbeth story. Gray boxes indicate information produced by common-sense rules.

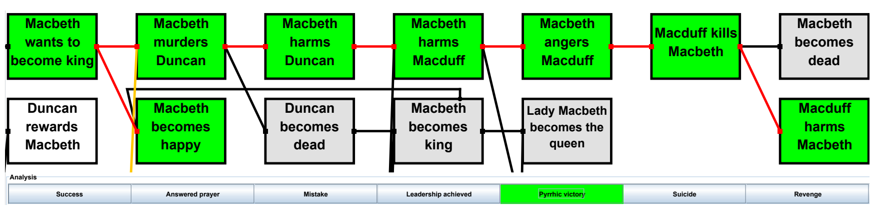

Figure 2. Genesis's story understanding system uses the elaboration graph, together with concept patterns, to augment the explicit knowledge provided in the story and simple inferences generated using common-sense rules. Here, Genesis discovers a Pyrrhic victory, shown highlighted.

For a more contemporary example, Genesis finds revenge in the elaboration graph produced from a description of the alleged Russian cyberattack on Estonia's network infrastructure, as shown in Figure 3. Genesis not only finds revenge, but also looks for the acts of harm involved, then uses WordNet to find the most related acts in what the political scientists call the Goldstein (1992) index, which enables it to characterize the revenge in *Macbeth* as a tit-for-tat, while the revenge in the Russian cyberattack on Estonia is an escalation.

To take Genesis to a higher level, we have implemented mechanisms that read stories as if by two separate persona, which we jocularly call Dr. Jekyll and Mr. Hyde. Equipped with overlapping

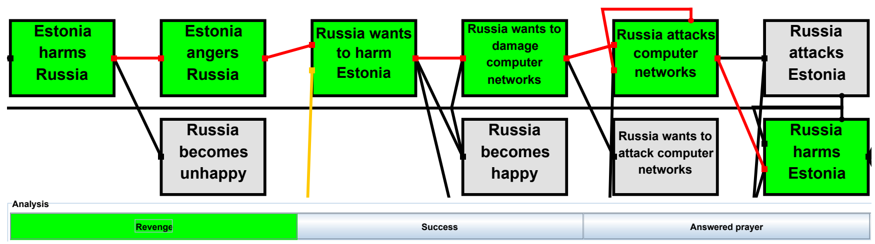

Figure 3. The common-sense rules and concept patterns honed on Macbeth have broad application. Here, the alleged Russian cyberattack on Estonia reveals an instance of *revenge*, shown highlighted.

but slightly different points of view, the two personas see things differently, in a manner reminiscent of the early work of Carbonell (1981).

In Figure 4, for example, Dr. Jekyll, on the right, concludes that Macduff kills Macbeth in an act of insane violence; Mr. Hyde, on the left, sees revenge. Both read the same story, but Dr. Jekyll thinks the only reason you would kill someone is that you are insane. Mr. Hyde looks for a reason, and then sees anger. Dr. Jekyll has the rule:

- James may kill Henry because James is not sane.

Mr. Hyde has another rule:

- Henry may kill James because Henry is angry at James.

Social psychologists would say that Dr. Jekyll behaves situationally, more Asian in outlook, because he looks for a situation that has caused a person to do a terrible thing, whereas Mr. Hyde behaves dispositionally, more Western in outlook, because he attributes terrible actions to characteristics of the actor (Morris & Peng, 1994).

An obvious next step will be to demonstrate culturally guided precedent finding, thus contributing to an important but slowly developing line of research on analogical reasoning (Winston, 1980; Falkenhainer et al., 1989; Gentner & Markman, 1997; Fay, 2012).

3.5 Fifth Step: Crystalize Emergent Principles

Encouragements, mild surprises, and hints of principles have already emerged from Genesis research. For example, we were encouraged by the ability of Genesis to work with stories of many types, including not only Shakespeare and conflict in cyberspace, but also simply written fairy tales, law cases, medical cases, and science fiction. We were surprised that so little knowledge was needed to produce credible performance. Genesis exhibits some characteristics of human story understanding evidenced by its ability to answer a variety of questions about the stories it reads, yet it does its work using only about two dozen common-sense rules and another dozen reflective patterns, several of which, revenge in particular, arose frequently in our experiments.

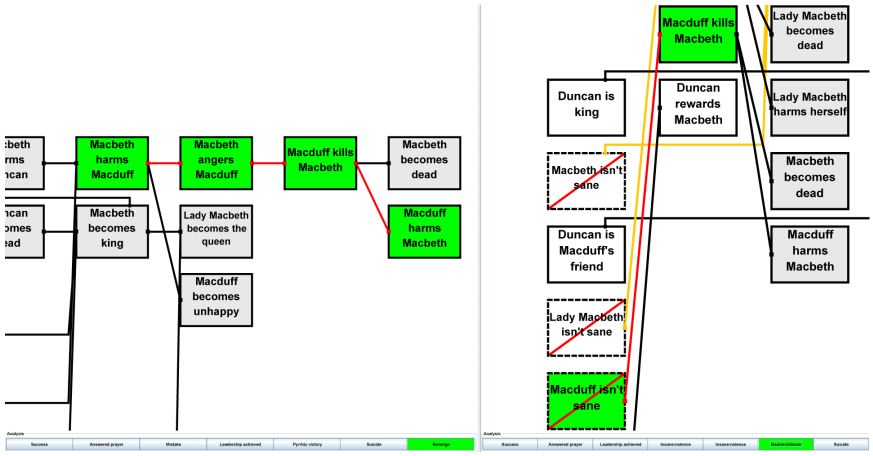

Figure 4. Opinions differ according to culture. One person's act of legitimate revenge is another person's act of insane violence.

3.6 The Way Forward

Good hypotheses lead to better questions. Here I list four representative examples that particularly interest me, all of which emerged from work focused on either the Strong Story Hypothesis or the Directed Perception Hypothesis.

3.6.1 Story Understanding

While encouraging, our work on Genesis leaves many hard story-understanding problems still to be addressed. Two that seem especially important are:

How can we build systems that discover their own rules and patterns? At the moment, Genesis's education is direct: we provide all the common-sense rules and all the concept patterns either directly (in lists) or indirectly (embedded in instructional stories). Of course, humans likewise learn a great deal by being told, but sometimes we form our own common-sense rules and discover and name our own concept patterns. We want Genesis to do that, too. On the concept-pattern level, Finlayson (2010) demonstrates how to discover concept patterns automatically in ensembles of culture-defining stories.

How can we add bulldozer computing to understanding? Systems such as IBM's Deep Blue chess player and IBM's Watson Jeopardy contestant demonstrate what can be done by approaching a sufficiently narrow problem with serious engineering, massive computing power, and big data. Our work on Genesis has the opposite polarity. We aim to see how little knowledge Genesis needs to reach interesting, humanlike conclusions. Eventually, we must somehow learn how to handle large amounts of knowledge and experience with humanlike finesse. This is the hardest problem that we face.

3.6.2 Directed Perception

Vision has received far more attention than story understanding, and with today's massive computing, big data vision programs produce impressive results. Still, even the best of today's people finders occasionally surprise us by misidentifying tree trunks as people, leaving much to be done, such as addressing the following questions:

How can we develop vision systems that answer common-sense questions? Fortunately, we seem to be on the leading edge of renewed interest in going beyond object recognition toward visual reasoning. Several groups, attracted by DARPA support, are developing systems that recognize actions such as chase, drop, dig, flee, give, jump, push, and throw. Such systems seem poised to provide play-by-play descriptions of what is happening in front of them, providing a foundation for answering common-sense questions about what particular people are doing and how.

How can story understanding and vision systems learn to communicate? Of course, we can program story understanding and vision systems to communicate, but that soon leads to a deeper question: How do they learn to communicate? In pioneering work, Siskind (1992) showed how a basic language system can interact with a simulated vision system; Kirby (1998) demonstrated a system in which a syntax emerges from a set of communicating agents; Coen (2006) demonstrated a system in which multiple systems learn to work together to understand events viewed from multiple perceptual perspectives; and Beal (2007) offered a mechanism that acquires knowledge as a by-product of learning to communicate. All these put hands on a very large elephant (a metaphor that demonstrates the amazing power of the human visual reasoning system).

4. Contributions

I hypothesized that an inner language is the central element that enables human intelligence. The inner language is important, in part, because it enables description; description enables story telling; story telling is central to education; and surrogate experience, in the form of stories, greatly influences culture. I then hypothesized that the inner language also plays an important role in marshalling the resources of our vision system, that the inner language stimulates visual imagination, and that vision is a major problem-solving resource. Finally, I hypothesized that our social nature has a strong amplifying affect. Thus, the principal contribution of this paper is the articulation of four hypotheses:

- **The Inner Language Hypothesis:** Using a symbolic inner language, we construct symbolic descriptions of situations and events that are far beyond the reach of other primates.
- **The Strong Story Hypothesis:** Our inner language enables us to tell, understand, and recombine stories, and those abilities distinguish our intelligence from that of other primates.
- **The Directed Perception Hypothesis:** Our inner language enables us to direct the resources of our perceptual systems to answer common-sense questions about real and imagined events, generating common-sense knowledge as a by-product.
- **The Social Animal Hypothesis:** Our social nature amplifies the value of story understanding and directed perception.

The contributions of those of us who have built the Genesis system include implementing a story understanding system with both low-level common-sense and higher-level reflective knowledge, all provided in English; explaining how a story understanding system can find concepts such as *revenge* in stories that never mention the word *revenge* or any of its synonyms; and showing how to produce cultural variation in story interpretation through modifications of common-sense and reflective knowledge.

Work on Genesis was motivated by the four hypotheses and conducted in accordance with a methodology whose steps are identifying the competence to be understood; formulating computational problems; proposing computational solutions; developing an exploratory implementation; and crystallizing emergent principles. Accordingly, we believe we have offered an example of the right way to work toward an account of human intelligence.

Acknowledgements

Mark Finlayson, Pat Langley, and anonymous reviewers suggested various adjustments that improved this paper. An earlier version was presented at the AAAI Fall Symposium (Winston, 2011). The research described was supported, in part, by the National Science Foundation (IIS-0413206), the Office of Naval Research (N00014-09-1-0597), the Air Force Office of Scientific Research (A9550-05-1-0321), and the Defense Advanced Research Projects Agency (FA8750-10-1-0076).

References

Beal, J. S. M. (2007). *Learning by learning to communicate*. Doctoral dissertation, Electrical Engineering and Computer Science Department, MIT, Cambridge, MA.

Borchardt, G. (1994). *Thinking between the lines*. Cambridge, MA: MIT Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.

Carbonell, J. (1981). Politics: An experiment in subjective understanding and integrated reasoning. In R. C. Schank and C. K. Riesbeck (Eds.), *Inside computer understanding: Five programs plus miniatures*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cassimatis, N. L., Murugesan, A., & Bugajska, M. D. (2008). A cognitive substrate for natural language understanding. In P. Wang, B. Goertzel and S. Franklin (Eds.), *Artificial general intelligence*. Washington, DC: IOS Press.

Chomsky, N. (2005). Some simple evo devo theses: How true might they be for language? Talk at the Alice V. and David H. Morris Symposium on Language and Communication. Stony Brook, SUNY, New York.

Coen, M. H. (2006). *Multimodal dynamics: Self-supervised learning in perceptual and motor systems*. Doctoral dissertation, Electrical Engineering and Computer Science Department, MIT, Cambridge, MA.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*, 1–63.

Fay, M. P. (2012). Enabling imagination through story alignment. Master's thesis, Electrical Engineering and Computer Science Department, MIT, Cambridge, MA.

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Finlayson, M. A. (2010). Learning narrative morphologies from annotated folktales. *Proceedings of the First International Workshop on Automated Motif Discovery in Cultural Heritage and Scientific Communication Texts*. Vienna, Austria.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*, 45–56.

Goldstein, J. (1992). A conflict-cooperation scale for WEIS events data. *Journal of Conflict Resolution*, *36*, 369–385.

Graesser, A., Dowell, N., & Moldovan, C. (2011). A computer's understanding of literature. *Scientific Studies of Literature*, *1*, 24–33.

Jackendoff, R. (1985). *Semantics and cognition*. Cambridge, MA: MIT Press.

Katz, B. (1997). Annotating the World Wide Web using natural language. *Proceedings of the 5th RIAO Conference on Computer Assisted Information Searching on the Internet* (pp. 136–159). Montreal, Canada.

Kirby, S. (1998). *Language evolution without natural selection: From vocabulary to syntax in a population of learners* (Technical Report). Language Evolution and Computation Research Unit, University of Edinburgh, Edinburgh, United Kingdom.

Laird, J. E., Newell, A., & Rosenbloom, P. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, *33*, 1–64.

Lehnert, W. (1981). Plot units and narrative summarization. *Cognitive Science*, *5*, 293–331.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

McClelland, J. L., & Rumelhart, D. E. (1986). *Explorations in parallel distributed processing*. Cambridge, MA: The MIT Press.

Minsky, M. (2006). *The emotion machine*. New York, NY: Simon and Schuster.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949–971.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco, CA: Morgan Kaufmann.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, *3*, 552–631.

Schank, R. C. (1991). *Tell me a story: A new look at real and artificial intelligence*. Evanston, IL: Northwestern University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schank, R. C., & Riesbeck, C. K. (1981). *Inside computer understanding: Five programs plus miniatures*. Mahwah, NJ: Lawrence Erlbaum Associates.

Siskind, J. M. (1992). *Naive physics, event perception, lexical semantics, and language acquisition*. Doctoral dissertation, Electrical Engineering and Computer Science Department, MIT, Cambridge, MA.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100.

Tattersall, I. (1998). *Becoming human*. Orlando, FL: Harcourt.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*, 433–460.

Ullman, S. (1996). *High-level vision*. Cambridge, MA: MIT Press.

Vaina, L. M., & Greenblatt, R. (1979). *The use of thread memory in amnesic aphasia and concept learning* (Working Paper 195). Artificial Intelligence Laboratory, MIT, Cambridge, MA.

Winston, P. H. (1980). Learning and reasoning by analogy. *Communications of the ACM*, *23*, 689–703.

Winston, P. H. (2011). The strong story hypothesis and the directed perception hypothesis. *Technical Report FS-11-01, Papers from the AAAI Fall Symposium* (pp. 345–352). Menlo Park, CA: AAAI Press.

Young, R. M. (2007). Story and discourse: A bipartite model of narrative generation in virtual worlds. *Interaction Studies*, *8*, 177–208.