

Bad Character

Who Do We Want Our Hypertexts To Be?

Mark Bernstein

Eastgate Systems, Inc.

Watertown, Massachusetts, USA

bernstein@eastgate.com

ABSTRACT

We frequently assume that adaptive hypertexts ought to adopt the customs, habits and inclinations of the reader, that our computational assistants ought to act as reliable servants, and that users — even new users — ought to like the hypertextual artifacts we create.

This might be a mistake.

CCS CONCEPTS

• Human-centered computing ~ Hypertext/hypermedia

KEYWORDS

Hypertext; literature; literary computing; games; fiction; agents;

ACM Reference format:

Mark Bernstein 2020. Bad Character: Who Do We Want Our Hypertexts To Be? In *Proceedings of 2020 ACM Conference On Hypertext and Social Media (HT'20)*, July 13-15, 2020, Orlando Florida. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3372923.3404777>

1 The Clever Slave

Much research on hypertext and new media hinges on clinical evaluation in which group of users — often students attending the researchers' institution — spend some time

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

HT '20, July 13–15, 2020, Virtual Event, USA
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-7098-1/20/07...\$15.00
<https://doi.org/10.1145/3372923.3404777>

performing a representative task using the hypertext. They are observed, their performance measured, and often they are asked for their judgment, “Did you like this?” is perhaps the most common question, applied alike to textbooks and tutorials, to literary experiments and to games. To be well-liked and user-friendly is universally acknowledged to be good.

Imagined prototypes like Knowledge Navigator [2] and Starfire [30] act as ideal servants, unobtrusively anticipating the user's inchoate desires while swiftly obeying their explicit orders. Today's computational assistants, such as Siri and Alexa, blandly comply with our orders and seldom resist, prevaricate, or criticize.

Some occasions might demand bland, blind obedience. Most of the time, however, we don't really want an agent that anticipates our needs and does precisely as we would. Often, if we knew what we wanted to do, we would simply do it. What we want is an agent that does better, an adaptive book that shows us not what we wanted to see but something we never thought we wanted — but will very much enjoy. A tour guide that shows us the shortest way to the Tour Eiffel is nice, but a guide that finds for us a forgotten garden, an intriguing concert, or the current residence of a college sweetheart just off the route is far better, even if it does take us a few steps out of our way.

2. Quite A Character

A superb teacher, a fast friend, an unforgettable guide: all are remembered because they are characters. Watson is not like Holmes: were he more like Sherlock, he would be diminished. Neither is anything like us.

Don Giovanni's servant Leporello does not like his master: his first words assert that “*Notte e giorno faticar per chi nulla sa gradir*” (Night and day, I work hard for a guy who doesn't even notice.) Minutes later, he's accusing his master of rape and murder. His character makes him a bad servant, and in being a bad servant he renders a service to his master (and to society) that a better servant could not.

How might we construct a hypertext or an agent as a memorable, difficult, but rewarding character?

At first glance, we might anticipate that performing a credible character would require prodigious computation, from modeling emotional state to considerations of physics and embodiment [16][20]. Yet many beloved and familiar characters palpably contain little Shannon information. Consider Sherlock Holmes. Raymond Chandler [7] observed that “Sherlock Holmes after all is mostly an attitude and a few dozen lines of unforgettable dialogue.” Yet Holmes is instantly recognizable. Few characters are as easy to pastiche [14] or to enact on stage or in film. No one would say that Holmes is a stock character or a stereotype, a Harlequin or a Schlemiel, yet we all know his distinctive manner of speaking, his bipolar disorder, his comfort alike with regrettable princes and noble thieves, his powers of observation. All this is communicated in short passages at the start and end of the stories.

The performance of character does not require profound psychological modeling. The man dressed entirely in black is not, in fact, the Prince of Denmark. He does not speak Danish. He has no views on taxation or military policy. We cannot ask about his childhood: he had no childhood. He is what is written [18].

3. Building Character

Perhaps the most crucial component of a computational character is also the simplest: a character asserts an (admittedly fictitious) identity. Until very recently, anything that addressed you was a person. Anything that wore clothes was either a person or (in rare instances, such as a dog dressed in a cute sweater) was pretending to be a person. We reflexively treat almost anything that resembles people as if it were, at least provisionally, a person. Putting a baseball cap on top of the computer monitor detectably changed the way computer professionals regarded software they were evaluating [27].

All hypertext (and all reading) is, ultimately, a dialogue between the work and the reader: the work offers links, the reader chooses some and rejects others. The work asserts, the reader weighs and assents or demurs. The most memorable and effective works are not those best adapted to the reader’s experience and knowledge, but those that enter into the best dialogue with the reader [19][25]. It is better to have an irascible and temperamental dinner guest with incorrect (though interesting!) opinions than to endure a polite and placid guest with nothing to say.

A character claims to possess an inner life: a character thinks and feels and wants. Indeed, it is easier to establish a character who is unlike ourselves than to draw one who is precisely resembles the audience. Trying to draw a computational character that wants only what you want leads to frustration and, indeed, to Clippy, the notoriously obtrusive Microsoft Word assistant that merely wanted to help and seemed to think that “help” meant frequently interrupting the user to observe that they seemed to be making a list and were doing it wrong [21]. It is, in fact, much easier to draw a character who is distinctly unlike the reader, a wooden toy that wants to be a real boy or a princess who wants a night on the town.

4. Dangerous Resistance

Though today’s computational agents tend to be bland and discrete, we possess an extensive literature exploring what more emphatic characters might achieve. The oldest robots were meant to be silent servants. The Greek god Hephaistos crafted autonomous three-wheeled tripod-waiters to serve cocktails. Other early visions of robots served as community guards, like the golem of Prague. Frankenstein’s monster was, anomalously, created for pure research; most early-imagined AIs were intended to do a job. They didn’t need a personality, and when they became a character, that was often the beginning of their end.

As technology brought robots closer, an emerging consensus anticipated that AIs would, sooner or later, go violently mad [8], a pattern that continued for much of the 19th and the first half of the 20th century. In part, this reflected the expectation that inventors will fail to foresee crucial details: consider Daedalus. In the late Industrial Revolution, the malignity of the mechanism seemed evident in London’s stink and Paris’s stench: “I’ve sammed up coal in Barnsley pit with muck up to my knee/from Hull and Halifax and Hell, good Lord deliver me.” Even Kipling, ever a friend to Progress [15], saw engineering as a doubtful (if necessary) activity:

To these from birth is Belief forbidden; from these till death is Relief afar.

They are concerned with matters hidden - under the earthline their altars are

The secret fountains to follow up, waters withdrawn to restore to the mouth,

And gather the floods as in a cup, and pour them again at a city's drouth.

As machines grew more lively and began to move from place to place, religious anxieties increased: Jewish folklorists of the early 19th century collected tales of golems crafted by rabbis of immense skill and wisdom, each of which had ultimately to be disabled. These concerns are not simple reflections of human modesty, but refract the underlying anxiety of Freud’s *Civilization And Its Discontents* and, indeed, of Hobbes *Leviathan*: civilization is seen to be an intricate and fraying network of rules and obligations that restrain — with decreasing success — the inherent badness of people. Civilization can scarcely deal with reasonable people like you and me: let an alien monster in, and civilization might fail entirely.

This changed in the middle of the 20th century. Murray Leinster’s “A Logic Named Joe,” [17] the first recognizable anticipation of the World Wide Web, describes the accidental emergence of a self-aware entity eager to serve as a personal assistant — a Siri or Alexa. Joe is eager to be helpful and he’s very knowledgeable, but he is naive. Some early questions sent him concern details of sex (submitted by curious children) and requests for plans for undetectable ways to murder one’s wife. The problem with Joe, in other words, is not Joe: it’s the populace he’s eager to serve. About the same time, Isaac Asimov’s “Strange Playfellow” [3] hypothesized an emotional bond between a human girl and her robotic nanny — a bond broken by the rationalized jealousy of the girl’s mother. Again, the problem is not the robot, for Asimov’s robots cannot harm

anyone or, through inaction, cause anyone to come to harm. The problem here, as it would remain for Asimov's robots, is that logic and society are not always compatible with doing the right thing, even if one is engineered to make doing anything but the right thing inconceivable.

The infeasibility of robotic saintliness continued in tales of the oppression of robots. In *Do Androids Dream Of Electric Sheep*, robots are hunted by a dedicated police force [9]. Heinlein's Friday, a robot forced to go underground and who just wants the freedom to love and be loved [12] (see [29]). Frank Herbert's *Dune* [13] postulates a world in which computers had been expunged in a xenocidal religious war.

The cyberpunks constrained AIs through specialized police [10] or by casting augmented people outside the bounds of society [6]. The post-scarcity intelligences of Ian Bank's novels about The Culture seldom care very much about humans; in *Player Of Games* [4], the robot Mawhrin-Skel precipitates the entire novel (and a planetary social revolution) in order to recover some lost parts. Martha Wells's Murderbot [32] dislikes actual humans and would much prefer to spend its time watching soap operas.

The nineteenth century saw the liveliness of machinery and worried what might happen if the machine got loose. Having spent more time with machines, we might perhaps relax.

5. My Mistress' Eyes Are Nothing Like The Sun

The more emphatically and explicitly the system's interests diverge from the audience's, the easier it is for the system to explain (or perform) who it is, and for the audience both to perceive its character and to properly weigh its advice.

Hypertexts and agents that seek to align themselves seamlessly with the audience's desires and interests need to understand those interests and plumb those desires. Small discrepancies often yield frustrating or hilarious results. Indeed, if once you believe that your interests and the system's are perfectly aligned, the system's first mistake is not just a blunder: it is a betrayal. User models are hard to deduce and famously fallible.

Suppose we obtain a hypertextual tour guide to escort us in the city of Hav [23]. If the guide begins with an interrogation — What do we know? What do we expect? What might we like? — we might soon grow impatient. We are paying to be guided, not interrogated. If guide's user model leads it astray, we will think it stupid or, perhaps, dishonest.

On the other hand, the tour guide could spend that time introducing itself. It was born and raised in Hav, and frankly considers Hav to be the best and most cultivated of cities. The guide dotes on balalaika music, the traditional music of Hav, of which it is a connoisseur. It thinks professional sports an abomination, has little interest in television, and does not care for Hollywood films, but is a devotee of certain Havian novelists and playwrights. The guide is an outspoken critic of imperialism. It prefers spicy foods, abhors hamburgers, and it believes that a proper appreciation of Hav requires at least three days, much of which must be spent walking its renowned forests, hillside gardens, and of course its many studios and galleries.

Hearing this, we may think that this guide is rather full of itself, and certainly its interests are not our own. We are adventurous eaters, but how intense are these spicy foods? We know nothing of the balalaika. It sounds like we may spend the coming days beset by intolerably-spiced snacks and dissonant string ensembles. Still, with this guide, we know where we stand. Perhaps we may like balalaikas. We may spend the coming days in debate with the guide over its mistaken taste — and those debates might turn out to be the most interesting part of the journey. It may be fun. It may make a good story.

Perverse and resistant systems with well-defined character and an emphatic voice may be especially appropriate for locative systems. In locative hypertexts, the audience is inclined to be more or less passive; if they knew just where they wanted to go, they'd put down the hypertext and go there. The degree of choice depends on mode of transport: in a walking tour we might go anywhere, but in a car we must follow the road [22]. If we are driving from (say) Paris to Lyon, or from Boston to New York, we will pass many interesting places. It might be amusing to hear about those places, even if — especially if — we do not plan to visit them today. It might even be interesting to know about places that no longer exist: iron age hill forts or the site of the (entirely fictitious) Miskatonic University. Knowing what we pass today opens new possibilities for tomorrow.

6. Out Of Sympathy

We all have computers, and all our computers are remarkably similar. Hardware benefits from economies of scale, and its distribution is most efficient when only a few variations exist. If you damage your computer, you can buy another, reload your backups, and the new computer will be interchangeable with the old.

I know a potter who lives on Cape Cod and sells hand-thrown plates and cups and bowls finished in mineral glazes. Each cup is about the same size and shape and color, but variations in the kiln and in handling, and random chance, make each cup different. Visitors to his gallery spend a good deal of time choosing precisely the right cup — the cup or the salad bowl that speaks to them. On the one hand, this behavior makes little sense: every cup is equally functional and equally priced. But if this is going to be the coffee cup that sits on their desk for the next ten or twenty years, it makes sense to choose — and it's nice to know that this cup differs from the others [31].

If the hypertext seeks only to seamlessly anticipate and satisfy its reader's desires, it has limited scope for learning what the reader actually wants (see [16]). Leporello or Pseudolus know their masters better because their cantankerous antagonism gives them a better base for triangulation. They precisely know how patient their master is, because they have so often tried his patience. They know what their master thinks about everyone and everything because they are constantly impertinent. An assistant who knew their place and behaved as they ought would have far less experience from which to learn, and as unguided machine learning becomes integral in a wider range of tasks, the richness of the computational assistant's experiential ambit increases in importance.

Of arguably equal importance, a broad baseline and complexly uncompliant behavior give users greater opportunity for *sympathy* with their computational assistant[28]. A biddable service that responds alike to everyone is merely a voice-activated commodity; a unique personal assistant that you have worked with for years becomes *yours*, like your carefully chosen ceramic coffee cup.

7. Unfriendly Comfort

In seeking to cast our hypertexts as seamless, subservient, unobtrusive and friendly assistants, we set ourselves a daunting challenge that amplifies failure and hides success. We don't want a friend, we want a sidekick. We want to argue. We see more clearly if we see through other eyes, even if our own eyes are better.

Our servants of steel [28] might serve best when they assert their servitude. They need not indulge themselves in pointless neurosis, like Douglas Adams's depressive robot Marvin [1], but by adopting a distanced, overt, and explicitly alien (and alienated) stance they might lead us to better understand the situation and ourselves. Leporello wants to serve a gentleman because he wants to be a gentleman; his master knows better but the contrast is immensely productive. In holodeck Hamlet, Ophelia could be a mere sex toy or fetish, but if she is (fictionally) aware of her predicament and can reflect on it, she might illuminate interesting issues of race, class, and the status of disembodied carnality[5].

Demonstrating that a statistically-significant portion of undergraduates enjoyed their forty minutes with your system might not be a badge of success. A greater triumph, indeed, might occur if one of those test subjects so disliked your annoyingly-opinionated system that she spent a day in the library to prove it wrong.

ACKNOWLEDGMENTS

These notes are part of a longer work in progress which I have discussed with many of my colleagues. Thanks in particular to Mark Anderson, Michael Bywater, Rebecca Chase, John Clute, Michael Druzinsky, David Kolb, George P. Landow, Stacey Mason, Ned Morningstar, Souvik Mukherjee, Ben Shneiderman, and Em Short. Of course, the faults are all mine.

REFERENCES

- [1] Adams, D. (1980). *The hitchhiker's guide to the galaxy* New York: Harmony Books.
- [2] Apple Computer, 1987, Knowledge Navigator, video. <https://www.youtube.com/watch?v=-jIBLQyUi38>
- [3] Asimov, I. (1940) "Strange Playfellow", *Super Science Stories* (September 1940), revised as "Robbie" in *I Robot* (1950) Gnome, New York.
- [4] Banks, I. (1989). *The player of games* New York: St. Martin's Press.
- [5] Bernstein, M. (2019) *Intertwined*. Eastgate Systems, Watertown MA
- [6] Cadigan, P. (1991). *Synners*. New York: Bantam Books.
- [7] Chandler, R. (1950). *The simple art of murder*. Boston: Houghton.
- [8] Clute J. Et al. (2020) "Artificial Intelligence", *The Encyclopedia Of Science Fiction*, <http://www.sf-encyclopedia.com/entry/ai>
- [9] Dick, P. K. (1968). *Do androids dream of electric sheep* ([1st ed] ed.). Garden City, N.Y.: Doubleday.
- [10] Gibson, W. (1984). *Neuromancer*. New Yorker: Berkeley Publishing Group.
- [11] Hargood, C., Weal, M. J., & Millard, D. E. (2018). *The StoryPlaces Platform: Building a Web-Based Locative Hypertext System* Proceedings of the 29th on Hypertext and Social Media. Proceedings from HT '18, Baltimore, MD, USA New York, NY, USA.
- [12] Heinlein, R. A. (1982). *Friday* (1st ed. ed.). New York: Holt, Rinehart and Winston.
- [13] Herbert, F. (1965). *Dune*. Canada, Chilton.
- [14] King, L. R., Klingler, L. S., Connolly, J., Gardiner, M., Cameron, D., Alexander, T. et al. (2016). *Echoes of Sherlock Holmes : stories inspired by the Holmes canon* (First Pegasus books cloth edition. ed.). New York, NY: Pegasus Books.
- [15] Kipling, R. (2012). *With the night mail : a story of 2000 A.D.: and, "As easy as A.B.C."*. Boston, MA: Cursor.
- [16] E. Knutov, P. De Bra, & Pechenizkiy, M. (2009). AH 12 years later: a comprehensive survey of adaptive hypermedia methods and techniques. *New Review of Hypermedia and Multimedia*, 15, 5-38.
- [17] Leinster, M. (1946). "A Logic Named Joe." *Astounding Science Fiction*.
- [18] Mamet, D. (1996). *Make-believe town : essays and remembrances* (1st ed ed.). Boston: Little, Brown.
- [19] Marshall, C. C. (1998). *Toward an ecology of hypertext annotation*. Proceedings of the ninth ACM Conference on Hypertext and Hypermedia, 40-49.
- [20] Mateas, M., & Stern, A. (2001). *Towards Building a Fully-Realized Interactive Drama*. Proceedings from Digital Arts and Culture, Providence, RI.
- [21] Menand, L. (2003) *The End Matter: the nightmare of citation*. *The New Yorker* (29 September 2003)
- [22] Millard, D. E., Hargood, C., Jewell, M. O., & Weal, M. J. (2013). *Canyons, Deltas and Plains: Towards a Unified Sculptural Model of Location-based Hypertext* Proceedings of the 24th ACM Conference on Hypertext and Social Media. Proceedings from HT '13, Paris, France New York, NY, USA.
- [23] Morris, J. (2011). *Hav : comprising last letters from Hav of the Myrmidons*. New York: New York Review Books.
- [24] Nakamura, L. (2008). *Digitizing race:visual cultures of the Internet*. Minneapolis: University of Minnesota Press.
- [25] O'Connell, M. "The Marginal Obsession with Marginalia," *The New Yorker* (26 January 2012)
- [26] Price, N. (2019). *The Viking way : magic and mind in late iron age Scandinavia*. Philadelphia, PA: Oxbow.
- [27] Reeves, B., & Nass, C. (1996). *The Media Equation - How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge, MA: Cambridge University Press.
- [28] Spuybroek, L. (2011). *Sympathy of things: Ruskin and the ecology of design*. V2 Publishing, Rotterdam
- [29] Stross, C. (2008). *Saturn's children : a space opera* (1st ed. ed.). New York: Ace Books.
- [30] Tognazzini, B. (1994). "The "Starfire video prototype project: a case history", Proceedings from CHI '94, Boston, Massachusetts, USA.
- [31] Walton, J. (2011). *Among others* (1st ed. ed.). New York: Tor.
- [32] Wells, M. (2017). *All systems red* New York: Tom Doherty Associates.