Ishaan Jhaveri
ENGL 4903: Science Fiction and the Value of Utopia/Dystopia
Instructor: Ricardo Wilson
May 24, 2017

1 Introduction

James Baldwin once wrote, "The history of America is the history of the Negro in America" (Peck). In Colson Whitehead's *Zone One,* the protagonist Mark Spitz leads a group of "sweepers" through the streets of Zone One, (formerly Manhattan, New York) removing "stragglers" from the streets. Stragglers are harmless zombies that seek only to cling to places they held dear while alive. Sweepers attempt to remove all traces of them, effectively removing them from the history of the new Zone One region. Whitehead may be offering us a glimpse into the way the colonizers of America "blanked out" the history of the region that took place before they inhabited it. He might also be echoing Baldwin's sentiments in *I Am Not Your Negro,* that the perpetrators of the slave trade robbed African-Americans of their history. Baldwin's theoretical work and Whitehead's literary work make two distinct points about blanking out history. First, they argue that the colonizers of America committed actual physical atrocities on African slaves to wipe out their history. Baldwin talks about how the slave traders transplanted slaves, forbade them from speaking their ancestral language, and separated them as much as possible from other slaves from the same region in Africa. Whitehead talks about the casual air with which Mark Spitz and the other sweepers "pop and drop" stragglers, and remove them from places they hold dear, forcefully removing all association between the stragglers and these places (Whitehead 118). Second, they implicitly indicate that the very fact that these acts of blanking out themselves are not focused on in dominant historical discourses suggests that they are still actively conducted as more of America's history is written. Thus Baldwin claims that to see what is missing in popular depictions of American

history one must look at the treatment of the Negro in America throughout its history. Whitehead's Mark Spitz begins to sympathize more and more with the stragglers as he considers, contrary to his orders from Buffalo, their past lives and how real they were, to the same end. Thus Baldwin and Whitehead lament the blanking out of Black History in America.

Octavia Butler exhibits a keen awareness of this idea in *Kindred*. Her story depicts the physical realities of slavery in painful detail. She doesn't "blank" anything out. However, she also appears to offer a route by which African-Americans can reclaim this lost history. In *Power and Repetition: Philosophies of (Literary) History in Octavia E. Butler's* Kindred, Christine Levecq argues that Butler is engaging with the idea of recovering history. Butler's protagonist Dana, an African-American woman living in 1970s California, has a series of episodes where she is transported back to the antebellum south and made to be a slave. In her first episode in the past she is almost shot, her second almost raped and in her third, beaten to within an inch of her life.

Butler evokes these near-death encounters in palpable detail. In Dana's first episode, she "[finds herself] looking down the barrel of the longest rifle [she] had ever seen. [She hears] a metallic click, and [freezes]…" (Butler 14). In her next, she is watching a black men getting beaten by white men, "[she can] literally smell his sweat, hear every ragged breath, every cry, every cut of the whip" (Butler 36). She is almost raped by one of the white men shortly after witnessing this. "When [she] tries to scramble away, he pulls [her back]… He [leans] down close to [her], pinning [her] flat on [her] back. In her third episode when she describes the whip as a "hot iron across [her] back,

burning into [her] through [her] light shirt, searing [her] skin…" (Butler 107). The descriptions are rife with sensual minutiae. Deeply painful sights, sounds, smells and physical feeling characterize her experience of the past. As Levecq points out, this is in stark contrast to how Butler depicts Dana's relationship with the present. "[During] the only extended period of eight days she spends there she lives off aspirins and sleeping pills, feels 'disoriented' and 'confused' (Butler 115), and acts as if 'nothing was real' (Butler 116)" (Levecq 531). In the present, she has only seen violence on television. Dana even admits, "I felt as though I were losing my place here in my own time. Rufus's time was a sharper, stronger reality. The work was harder, the smells and tastes were stronger, the danger was greater, the pain was worse" (Butler 191). Finally, Dana is only able to leave the past for good after ensuring that Alice's children (her ancestors) are free and safe, to ensure that she will be born, that he present can even exist.

Thus Butler writes of a history that is infinitely more real than the present. She presents the present for Dana and Kevin, both writers surrounded by books, as a textual remnant of its past. She "retells" the slave narrative by writing about the realities of it, and declares that the present is only possible as a result of the past. Butler suggests that the way to reclaim lost Black History is to write about it. In evoking powerful details of the past, she reverses the blanking out of Black History in dominant history. That this past is more real than the present is because the present is a remnant of dominant history, which has blanked out Black History. Only by retelling history with an appreciation of Black History can the present be made real, validated. The book ends with a conversation between Dana and Kevin about needing to feel "sane" again. Kevin concludes that "now

that the boy is dead, [they] have some chance of staying [sane]" (Butler 264). Only now that they have had this real experience of the past, come to appreciate Black History and seen to it that Rufus Weylin, a white man that committed many atrocities on black people, is dead, can the present be validated. Only once dominant history is dead and Black History revived, can the present be validated.

Thus Butler presents *Kindred* as an attempt to reclaim lost history, and as a guide for how other writers can do the same. In this project I am first interested in testing the hypothesis that many works of Science Fiction (SF) can be read as responses to Butler's direction in *Kindred*, where writers, particularly black writers attempt to reclaim lost history by writing about it. After conducting initial research to establish this, I then want to ask a deeper question about how they have chosen to do this. If I can establish that many SF writers are attempting to reclaim lost history, an infinitely more interesting result of this discovery would be how they choose to do this: whether they retell events in history with a focus on the consequences of those events on black people, whether they read the aftermath of those events differently than dominant history does, whether they see different factors leading to these events than dominant history does, and so on.

The second question is much better tackled when data from answering the first question is available, so I held off on more structured hypotheses about the second question until I had data. In the rest of this paper I will describe my attempts at testing my first hypothesis computationally, where I fell short, and what I propose to do differently in a future attempt at answering these questions.

2 Approach

Wikipedia defines Historical Revisionism as:

> In historiography, the term historical revisionism identifies the re-interpretation of the historical record. It usually means challenging the orthodox views held by professional scholars about a historical event, or introducing new evidence, or of restating the motivations and decisions of the participant people. The revision of the historical record reflects the new discoveries of fact, evidence, and interpretation, which produce a revised history.

Butler's project of writing about Black History, rewriting historical events with a focus on uncovering the Black History component to them resembles this definition closely. Therefore I decided to structure my computational analysis along the lines of finding evidence of Historical Revisionism in SF text, particularly in the SF of black writers. I reasoned that if I could find evidence of Historical Revisionism in the SF of black writers, I could then evaluate the character of this Historical Revisionism to answer questions about how these writers have attempted to reclaim their history. I honed in on the part of the above definition that says Historical Revisionism challenges orthodox views about a historical event. I reasoned that if I could find evidence that a text written by a black SF author is talking about a historical event, I could evaluate the language and other aspects of the text to determine whether it suggests a different interpretation of the event than the dominant historical one. The burden before me was now to find a body of SF texts by black writers that determinedly talk about a specific event.

3 Data Collection

After meeting with reference librarians, on their recommendation I began by trying to find SF books organized by references to an event, with a mind to then pick those books that were written by black writers. I decided to try and find books that deal with the American Civil War, because of its significance to Black History and the fact that it happened a long enough while ago for there to be a good body of work referencing it. Finding such a list proved to be very difficult. I looked in the Routledge and Cambridge Companions to Science Fiction for such a list, through all of Cornell's online research guide to research in Science Fiction and Fantasy, on the popular online Science Fiction Encyclopedia as well as the fan fiction website Worlds Without End. None of these anthologies of SF research had such a list directly, or material that helped me create such a list indirectly.

Frustrated, I decided instead to manually search for books I knew to be about a certain historical event. I was forced to sacrifice the focus on Black History and focus just on Historical Revisionism. It is worth noting here that the lack of scholarship on this specific topic, SF writing about the Civil War or another event that is so ostensibly relevant to Black History is testament to the blanking out of Black History discussed previously. I decided therefore to find books about the dropping of the Atomic Bomb, a very popular topic in SF. To find a wealth of such books, I needed to look no further than Wikipedia. Wikipedia has a list of the 100 most popular SF books under the subgenre "nuclear holocaust fiction". Of these I managed to find digitized, computer-readable versions of about 10. I could have spent longer hunting down digitized versions of the remaining 90 and these 100 books would have been a sufficient corpus to perform at least some preliminary data analysis on, but with time running short I decided to just do what I

could with data I already had, 1,210 SF books in digitized, readable form from the Hathi Trust Digital Library, compiled by Professor Wilson and Laure Thompson a Cornell PhD student.

4 Data Analysis

Through my initial searches of the various resources I listed above for a list of SF books by references to an event, though I didn't find what I was looking for, I found many references to "Alternate History". The *Routledge Companion To Science Fiction* defines alternative history as "that branch of nonrealistic literature that concerns itself with history turning out differently than we know to be the case" (Bould 453). It talks about SF writers' use of alternate history to examine moral issues. Derek J. Thiess' *Relativism, Alternate History and the Forgetful Reader* connects "relativist historiography" with alternate history. In this book, Thiess makes the case that SF writing can often be read as Historical Revisionism, which is close to the hypothesis I am testing. It leaves out the important dimension of Black History, but since time was short I modified the purpose of this project to be a test of whether this kind of computational analysis could uncover anything about Historical Revisionism at all. I decided therefore to try and answer that question, and leave the Black History question as a next step, out of the scope of this immediate project.

Of the texts I had, 125 were written by authors known to write alternate histories. I deemed an author "known to write alternate histories" if the author was featured in *The Mammoth Book of Alternate Histories*, the largest collection of alternate histories I could find in Cornell's libraries or if the author was featured on the online Encyclopedia of Science Fiction's page on Alternate History. The authors I found were Kim Stanley

Robinson, James Morrow, Suzette Haden Elgin, Harry Harrison, Pat Cadigan, Fritz Leiber, Paul McAuley, Keith Roberts, Robert Silverberg, Gregory Benford, Pamela Sargent, Frederik Pohl, Stephen Baxter, Terry Bisson and Jack Williamson.

Had I had more time I would have grouped each of the 125 texts by the event that they presented an alternative history of, and used that to structure my analysis, but since I didn't, in my analysis I look for how these writers treated topics related to World War II and the dropping of the Atomic Bomb differently than how these topics are treated by all the SF writers in the larger 1,210 text strong corpus.

I used a program called MALLET that generates topic models for a corpus of texts. A topic is a collection of words (in this case 20) that co-occur together. In this analysis, co-occur means the words are on the same page in a given text. Briefly, the way MALLET works is by assigning a random topic number (0-249) to each unique word. It then performs several sampling iterations. For each word, it first erases the word's current topic and then guesses a new one. The probability that MALLET will choose a topic, *x*, for the word is proportional to the number of times topic *x* occurs in the current text and to the number of other occurrences of this word in this and other texts that have been assigned topic *x*. This process occurs 1000 times, and the topic assignments at the end of 1000 sampling iterations are reported.

4.1 Results

I was looking for different ways that words related to World War II and the atomic bomb are treated. I have included some results I found in the table below. The leftmost column has the word I searched for, the middle column contains the related words (other words in the same topic) in the topic model generated from the 125 alternate history texts,

and the rightmost column contains the same for the topic model generated from all 1,210 texts.

Word	**Related Words (Alternate History)**	**Related Words (All 1,210 SF texts)**
nuclear	george said the his brat suit sverre holly wengernook scopas war aquinas randstable bonenfant asked justice would paxton morning	power bomb energy atomic bombs radiation explosion fuel plant radioactive use fusion war hydrogen atom used enough any uranium
war	Same as above	war against men fight man battle those fighting fought peace death enemy than great power people killed own kill blood
american	the and people world this government for their	new york san city francisco california los north

	work new all power science its which project scientific political american these	washington angeles west american states where chicago state south town america east

Nuclear. To make a more insightful analysis of this data would take an understanding of what the texts in the alternate history corpus were about and what words like “aquinas”, “bonenfant” etc. entailed. But it is interesting to notice that in the general texts nuclear is associated more with the atomic bomb and “uranium”, whereas in the alternative histories it is associated with names and conversations between people, suggesting that perhaps these texts focus more on readings of the nuclear bomb event rather than the technicalities of it. That the word “war” is associated with nuclear more strongly in the alternate history corpus suggests that these authors might have been having conversations about nuclear war, perhaps as a result of World War II.

War. In the general text corpus “war” is associated with “man”, “men” and “fighting”, suggesting war as physical combat, but in the alternate history corpus war is associated more with “nuclear” and names, hinting again at possible conversations about nuclear war. “Justice” is associated with war (and nuclear) in the alternate history corpus but not in the general corpus, perhaps signaling conversations about the morality and outcome of the dropping of the atomic bomb by the alternate history writers, which would shed some light on the kind of revisionist project we are interested in.

American. In the general text corpus "american" is associated only with American city and state names. In the alternate history corpus it is associated with "government", "science", "power" and "political", indicative of conversations about the outcomes of and debates surrounding the dropping of the Atomic Bomb. This association is possible evidence that writers of alternative histories are carrying out a revisionist project evaluating the outcomes of the bomb on American power, politics and science, and perhaps offering novel interpretations of these outcomes.

5 Reflection

As of now my various approaches to data collection and data analysis seem quite haphazard and shallow. I had to adjust my initial research question quite a bit in order to fit it to the data I could find. Because I was strapped for time the modified data analysis task I did perform was shallow, and didn't shed much light on anything. It is a good starting point but incomplete without other computational data and better literary context. It is clear to me that undertaking a project like this, attempting to find what it takes for a computer to understand blackness and Black History enough to present meaningful data about it, requires more time than three weeks and more deliberate approach. This is something I wholeheartedly want to work on in the future. For now, I propose the following as a somewhat more deliberate approach to tackling a project like this, given more time.

The questions I seek to answer are 1) Do black SF writers write about history in order to reclaim lost Black History, and 2) If so, how. The first burden before me in a computational analysis of how SF writers treat history is to conceptualize "history" in a way that is readable by a computer. A good place to start on this is to consider ways in

which other Digital Humanities (DH) researchers have conceptualized history in a form that a computer can understand, or analyzed the text of books in a way that belies how these books treat history.

5.1.1 Models to Emulate

5.1.2 Jurafsky, "Discovering Laws of Semantic Change"

Stanford DH Professor Dan Jurafsy's "Discovering Laws of Semantic Change" project provides an open-source mapping between word meanings and time. That is, you can input a word and a time period into his program and it will return the dominant meaning of that word at that time. This could be useful for establishing time period, because current DH techniques are pretty good at estimating word meanings in a text from context. So one way to determine the time period a text is set in, is to determine the meanings of certain key words in the text, and then use Jurafsky's program to see the time period in which these words carried the meaning they carry in this text. For example the word "guy" used to mean grotesque man until the 1800s, when it started acquiring the meaning it has today, "any man". A downside to this approach is that these changing meanings only apply to some words, which may not be commonly used in the text. Further, this kind of mapping might produce an approximate time period of a text, like "before the 1800s", but it might not be able to determine the time period with greater precision.

5.1.3 Lee, comparing words using a "TF-IDF representation"

Professor Lillian Lee at Cornell has considerable experience in DH and suggested that one way to understand the time period of a text is to compare words in that text to

words in general English fiction from different eras using a "TF-IDF representation". This representation isolates words that are representative of the entire text. So she suggested that I could create a TF-IDF for say the 1850s, 1870s, 1900s, 1940s etc. from general English fiction books from these eras and then create TF-IDFs for each text I am interested in, and see which era's TF-IDF is most similar to the TF-IDF of a given text. This is a way of isolating the time period a given text belongs to. This method could also be used to determine whether different sections of a text belong to different time periods. For example in Butler's *Kindred* the flashbacks to the 1800s might have different language than the language of the present day (1970s), and this TF-IDF method might show that.

5.1.4 Bamman, Smith and Underwood, "A Bayesian Mixed Effects Model of Literary Character"

This project infers "character types in a collection of 15,099 English novels published between 1700 and 1899". They create "character types" and model which character types are popular in different time periods. So using their methods I could infer the types of a few of the characters in a text I was interested in, and estimate its time period by seeing which time period was known for similar character types. A downside to this method is that there might not be significant difference in popular character types across some time periods.

These methods help to determine the time period of a text, which would help with the first question, of whether some SF authors indeed are attempting to reclaim lost

history by writing about it. If I were to run a text with a known publication date through one of more of these methods, I would expect the methods to correctly predict the text's approximate publication date. I would expect this to be the case for general English fiction. If I were to now run SF texts and general texts published around the same time through one or more of these methods, it would be very interesting if the SF texts (or some of them) were predicted to have a different publication date than their actual publication date. This would suggest that there is something in these texts that indicates they are talking about a different time period than the one they are writing in. If such texts were found, the answer to my first question, whether it can be said that some SF writers follow Butler's lead of reclaiming lost history by writing about it, would be in the affirmative. Isolating which texts showed signs of talking about history and then evaluating those texts would be the starting point to answering the second question, of how they go about rewriting history.

This approach would work with the 1,210 Hathi Trust text data I currently have. As for the general English fiction corpus, I could get that from Google's open source n-gram program that has a corpus of one million English fiction books from 1500 to 2008.

6 Works Cited

Butler, Octavia E. *Kindred*. Boston: Beacon press, 1988. Print.

Bould, Mark. *The Routledge companion to science fiction*. London: Routledge, 2009. Print.

Hamilton, William L, Leskovec, Jure and Jurafsky, Dan. *HistWords: Word Embeddings for Historical Text*. N.p., n.d. Web. 24 May 2017.

I Am Not Your Negro. Dir. Raoul Peck. Magnolia Pictures, 2017. Film.

Levecq, Christine. "Power and Repetition: Philosophies of (Literary) History in Octavia E. Butlers "Kindred"." *Contemporary Literature* 41.3 (2000): 525. Web.

Thiess, Derek J. *Relativism, Alternate History, and the Forgetful Reader: Reading Science Fiction and Historiography*. Lanham, MD: Lexington , 2015. Print.

Watson, Ian, and Ian Whates. *The mammoth book of alternate histories*. London: Robinson, 2010. Print.

Whitehead, Colson. *Zone one: a novel*. Vintage: London, 2012. Print.