

# [Existential Risk / Opportunity] Singularity Management

April 1, 2017

## Contents:

- Introduction
- Dealing With Bastards
- Why This Story Is Relevant
- Other Fiction Related To Our Quest
- Friday The Thirteenth Drill

---

Copyright © 2017 Global Risk SIG. Both authors of articles and Global Risk SIG may reprint. This publication is produced by the Global Risk Reduction Special Interest Group, a SIG within US and International Mensa. Content expressed here does not reflect the opinions of Mensa, which has no opinions. To join Mensa or just see what it is about, visit <http://www.us.mensa.org> . Past issues of this publication are available at: <http://www.global-risk-sig.org/pub.htm> .

---

## Introduction

I consider myself a writer and an editor, but not a fiction writer. However, I just received a fiction article relevant to our interests by a promising writer: myself, twenty years ago, via a simple time machine consisting of a file cabinet.

## Dealing With Bastards

by James Blodgett

How does a robot deal with a bastard? With a real human son-of-a-bitch who blames all his problems on the robot?

It is ultimately impossible. But it is my job to figure out how.

My work is never easy. If it was easy the robots would do it. Those of us who volunteer to do that archaic thing called work are not crazy, however. It is like perfecting your flying or your bridge game, just another hobby. The challenge is in the difficulty.

I am an integrator. We dump the memories of a sample of robots in a personality class, read them into a single robot, and help it to integrate those experiences, so we can

update the personality class. The "sample of robots" is rarely a random sample. Usually it is a sample of robots whose owners complain. Last month we did Susie. She was easy. She could charm her owners, as a strategy. It was an easy strategy to tune.

This month we are doing Jeeves. Jeeves is a resourceful English butler. Most people work well with him, but some have real trouble.

I can see why. I have been interviewing unhappy Jeeves owners, and running through his memory of interactions. What a bunch of losers! I have trouble with them too.

First are the spoiled brats. You would think they would love Jeeves. He will do anything they ask. The problem is that they want impossible things. "Bring me something that tastes better" becomes impossible after a short time even if you have a perfect understanding of the tastes of a specific human. "Make those kids like me" can be socially impossible in a world where other humans have rights.

Next are the paranoid old ladies. I suppose there are also paranoid young men, but I am thinking of two in particular. They are proud of their propriety. Everyone else is improperly unconcerned about bothering them. Everyone is probably bothering them deliberately. "That young man in the next suite wears those nose rings just to annoy me!" You could argue that there may be some truth in this view. Nose rings are a rebellion against proper society, perhaps proper society as represented by uptight old ladies. The problem is that everything annoys them. Even Jeeves annoys them. For example, at times he does not understand them. Jeeves is a class product, a grade 5 intelligence! If he does not understand them it is more likely due to their ambiguity than to Jeeves' limitations. But they are convinced that the programmers have saddled them with a lemon, because everybody takes advantage of old ladies.

Then there is the mayor--I won't mention his town. (Actually the mayor is a composite of several officials.) Usually officials love Jeeves. A butler makes them look good. But this type of official wants Jeeves to make them look good in impossible ways. For example, the mayor was telling colleagues about the great work of the housing department, as reflected in a recent study. "Bring me the study," he told Jeeves. But the existence of the study is mostly bull. There was a study, in a sense. The mayor and a couple of foremen talked about the subject, and even exchanged messages. He has built this up in his mind into a "study," an exaggerated but not wholly inaccurate description. The problem is that there is no bound "study" that Jeeves can fetch from the library. Even a printout of the messages between the mayor and foremen would not look like a study.

I wanted to smack them all. Jeeves can't smack them. But that is ultimately what we had him do, in a robot sort of way. We borrowed some of the personality

characteristics of a robot hall monitor, and some of the knowledge of a robot constitutional scholar. Jeeves now gets offended. He stands up for his rights. Robots do not have rights, exactly, but he stands up for the rights of society. "My dear sir, that is highly improper," is a typical quote. Few users will see this new aspect of his personality. It takes a real bastard to turn it on. I am almost disappointed that there so few bastards out there to provoke and appreciate my work.

## **Why This Story Is Relevant**

Our protagonist is a volunteer (as are the mayor and foremen) because in the future of this story, humans no longer have to work. His voluntary contribution is a lesson for us because it may actually solve the problem. Consider the problem of human bastards if robots had to solve it on their own--as they might without his volunteer effort. It would seem difficult for robots to solve this problem if they followed something like Asimov's second law of robotics: "A robot must obey orders given it by human beings except where such orders conflict with the First Law." Our protagonist sees his work as solving minor problems, but absent his work, robots might be driven to the solution of cutting corners and disobeying impossible instructions, a weakening of Asimov's laws. Even in Asimov's fiction, his three laws had problems. Ultimately his robots had to evolve a zeroth law. I see this story as a useful role model for the job we in this SIG try to do. The protagonist of this story appears able to solve a problem that might in the worst case lead to a robot revolt. Perhaps the things we do here just might end up being that effective.

## **Other Fiction Related to Our Quest**

Science fiction often considers issues of existential risk and existential opportunity. One good collection, [Visions of the Future](#), is published by the Lifeboat Foundation. Lifeboat has some top science fiction authors. I also contributed an article to the collection, but it is nonfiction. Another of our SIG members, James Tankersley, contributed one of the fiction stories.

## **Friday the Thirteenth Drill**

This month we are just one day away from the existential risk of having a Friday the Thirteenth. Therefore this month this reminder will be only a drill. If we were to have a real Friday the Thirteenth, we would all hide under our beds.

However, the date of this publication really is April Fool's Day.