

THE ROBOT AND THE BABY

John McCarthy

2001 Jun 28, 1:02 p.m.

“Mistress, your baby is doing poorly. He needs your attention.”

“Stop bothering me, you fucking robot.”

“Mistress, the baby won’t eat. If it doesn’t get some human love, the book says it will die.”

Love the fucking baby, yourself.”

The robot’s model number was GenRob337L3. Its program was GenRob337version 19.2. It was one of about 11 million household robots with those particular characteristics, and its serial number was R337942781. We’ll call it R781 for short.

R781 was designed in accordance with the McCarthy principle, first proposed in 1995 and which became a matter of law for household robots when they first became available in 2025. The principle was adopted out of concern that children who grew up in a household with robots would regard them as persons: causing psychological difficulties while they were children and political difficulties when they grew up. One concern was that a robots’ rights movement would develop. The problem was not with the robots, which were programmed to have no desires of their own but with people. Some romantics had even demanded that robots be programmed with desires of their own, but this was illegal.

In accordance with the McCarthy principle R781 was built in the shape of a giant metallic spider with 8 limbs, 4 with joints and 4 tentacular. Most people were frightened by the appearance at first but got used to it in a short time. A few people never could learn to have them in the house. Children also reacted negatively at first but got used to them. Babies scarcely noticed them. They spoke as little as was consistent with their functions and in a slightly repellent metallic voice not associated with either sex. This seemed

to work pretty well; hardly anyone became emotionally attached to a robot. Also robots were made somewhat fragile on the outside; if you kicked one, some parts would fall off. This sometimes relieved some people's feelings.

R781 returned to the nursery where the 18 month old boy, very small for his age, was lying on his side whimpering feebly. The baby had been neglected since birth by its alcoholic, drug addicted mother and had almost no vocabulary. It winced whenever the robot spoke to it; that effect was part of R781's design.

Robots were not supposed to care for babies at all except in emergencies, but whenever the robot questioned an order to "Clean up the fucking baby shit", the mother said, "Yes, its another goddamn emergency, but get me another bottle first." All R781 knew about babies was from the Internet, since it wasn't directly programmed to deal with babies, except as necessary to avoid injuring them and for taking them out of burning buildings.

Baby921 had barely touched its bottle. Infrared sensors told R781 that Baby921's extremities were very cold in spite of a warm room and blankets. Its chemicals-in-the-air sensor told it that the pH of Baby921's blood was reaching dangerously acidic levels. [More medical jargon needed here].

R781 thought about the situation. Here are some of its thoughts.

(Order (From Mistress) '‘Love the fucking baby yourself’')

(Enter (Context (Commands-from Mistress)))

(Standing-command '‘I told you 20 times, you fucking robot, don't call the fucking child welfare.’')

The privacy advocates had successfully lobbied to put a negative utility -1.02 on informing authorities about anything a household robot's owner said or did.

(= (Command 337) (Love Baby921))

(True (Not (Executable (Command 337)))

(Reason (Impossible-for robot (Action Love))))

```

(Will-cause (Not (Believes Baby921) (Loved Baby921)) (Die Baby921))

(= (Value (Die Baby921)) -0.883)

(Will-cause (Believes Baby921 (Loves R781 Baby921) (Not
(DieBaby921))))

(Implies (Believes y (Loves x y)) (Believes y (Person x)))

(Implies (And (Robot x) (Person y))
          (= (Value (Believes y (Person x))) -0.900))

(Required (Not (Cause Robot781) (Believes Baby921 (Person Robot781))))

(= (Value (Obey-directives)) -0.833)

(Implies (< (Value action) -0.5) (Required (Verify Requirement)))

(Required (Verify Requirement))

(Implies (Order x) (= (Value (Obey x)) 0.6))

(? ((Exist w) (Additional Consideration w))

(Non-literal-interpretation (Command 337) (Simulate (Loves Robot781
Baby921)))

(Implies (Command x) (= (Value (Obey x)) 0.4))

(Implies (Non-literal-interpretation x) y)
          (Value (Obey x) (* 0.5 (Value (Obey y)))))

(= (Value (Simulate (Loves Robot781 Baby921)) 0.902))

```

With this reasoning R781 decided that the value of simulating loving Baby921 and thereby saving its life was greater by 0.002 than the value of

obeying the directive to not simulate a person.

We spare the reader an account of the robot's subsequent reasoning. A search of the Internet turned up the story of how rhesus monkey babies who died in a bare cage would survive if provided with a soft surface resembling in texture a mother monkey.

R781 did the following.

1. It covered its body and all but two of its 8 extremities with a blanket. The two extremities were fitted with sleeves from a jacket left by a boyfriend of (Person 558305921) and stuffed with toilet paper.

3. It found a program for simulating a female voice and adapted it to meet the phonetic and prosodic specifications of what the linguists call motherese.

4. It made a face for itself in imitation of a Barbie doll.

The immediate effects were moderately satisfactory. Picked up and cuddled, the baby drank from its bottle. It repeated words taken from a list of children's words in English.

"Get me a ham sandwich and a coke."

"Yes, mistress."

"Why the hell are you in this stupid get up, and what's happened to your voice."

"Mistress, you told me to love the baby. Robots can't do that, but this get up caused him to take his bottle. If you don't mind, I'll keep doing what keeps him alive."

"Get the hell out of my apartment, you stupid robot. I'll make them send me another."

"Mistress, if I do that the baby will probably die."

"Get the hell out, and you can take the fucking baby with you."

"Yes, mistress."

When R721 holding Baby921 appeared on the street, it was quickly noticed.

"There's a really weird robot that's kidnapped a baby. Call the police."

When the police came they called for reinforcements.

"I think I can disable the robot without harming the baby", said Annie Oakley, the Department's best sharpshooter.

"Let's try talking first."

"Don't get close to that malfunctioning robot. It could break your neck in one swipe."

"I'm not sure it's malfunctioning. Maybe the circumstances are unusual."

"Robot, give me that baby".

“No, Sir. I’m not allowed to let an unauthorized person touch the baby.”
“I’m from Child Welfare.”
“Sir, I’m specifically forbidden to have contact with Child Welfare.”
“Who forbade that?”
The robot was silent.
A cop asked, “Who forbade it?”
“Ma’am, Are you from Child Welfare?”
“No, I’m not. Can’t you see I’m a cop?”
“Ma’am, my mistress forbade me to contact Child Welfare. Yes, ma’am, I see your uniform and infer that you are probably a police officer.”
“Why did she tell you not to contact Child Welfare?”
“Ma’am, I can’t answer that. Robots are programmed not to comment on human motives.”
“Robot, I’m from Robot Central. I need to download your memory. Use channel 473”
“Sir, yes”.
“What did your mistress say specifically? Play your recording of it.”
“No, ma’am. It contains bad language. I can’t play it, unless you can assure me there are no children or ladies present.”
“We’re not ladies, we’re police officers.”
“Ma’am. I take your word for it.
I have a standing order
(Standing-command “I told you 20 times, you fucking robot, don’t speak to the fucking child welfare.”)
“Excuse me, a preliminary analysis of the download shows that R781 has not malfunctioned, but is carrying out its standard program under unusual circumstances.”
“Then why is does it have its limbs covered, why does it have the Barbie head, and why does it have that strange voice?”
“Ask it”.
“Robot, answer the question.”
“Female police officers and gentlemen, Mistress told me, ‘Love the fucking baby, yourself’ “
“What? Do you love the baby?”
“No, sir. Robots are not programmed to love. I am simulating loving the baby.”
“Why.”

“Sir, otherwise the baby will die. The get up was the best I could do to overcome the repulsion robots are programmed to excite in human babies and children.”

“Do you think for one minute, a baby would be fooled by that.?”

“Sir, the baby drank its bottle, went to sleep, and its physiological signs are not as bad as they were.”

“OK, give us the baby, and we’ll take care of it.”

“No, ma’am. Mistress didn’t authorize me to let anyone else touch the baby.”

“Where’s your mistress. We’ll talk to her.”

“No, sir. That would be an unauthorized violation of her privacy.”

“Oh, well. We can get it from the download.”

Just then an official of the Personal Privacy Administration arrived in the form of a Government virtual-reality robot.. Ever since the late 20th century, the standards of personal privacy had risen, and an officialdom charged with enforcing the standards had arisen.

“You can’t violate the woman’s privacy by taking unauthorized information from the robot’s download.”

“What can we do then?”

“You can file a request to use private information. It will be adjudicated.”

“Oh, shit. In the meantime what about the baby?”

“That’s not my affair. I’m here to make sure the privacy laws are obeyed.”

During this discussion a crowd, almost entirely virtual, accumulated. The street being a legal public place, anyone had the right to look at it via the omnipresent TV cameras and microphones. Moreover, one of the police officers had cell-phoned a reporter who sometimes took him to dinner. Once the scene was on web-TV, the crowd of spectators grew exponentially, multiplying by 10 every 5 minutes, until seven billion spectators were watching and listening. There were no interesting wars, crimes, or natural catastrophes, and peace is boring.

Of the seven billion, 53 million offered advice or made demands. The different kinds were automatically, sampled, summarized, counted and displayed for all to see.

3 million advocated shooting the robot immediately.

11 million advocated giving the robot a medal, even though their education emphasized that robots can’t appreciate praise.

Real demonstrations quickly developed. A few hundred people from the city swooped in from the sky wires, but most of the demonstrators were

robots rented for the occasion by people from all over the world. Fortunately, only 5,000 virtual reality rent-a-robots were available for remote control in the city. Harsh words were uttered about this limitation on First Amendment rights. The interests were behind it as everyone knew.

The police captain knew about keeping your head when all about you are losing theirs and blaming it on you.

“Hmm. What to do? You robots are smart. R781, what can be done?”

“Sir, you can find a place I can take the baby and care for it. It can’t stay out here. Ma’am, are female police officers enough like ladies so that one of you has a place with diapers, formula, baby clothes, vitamins, . . .”

The robot was interrupted before it could recite the full list of baby equipment. Off it went with a lady police officer. (We can call her a lady even though she had assured the robot that she wasn’t.)

Hackers under contract to the Washington Post quickly located the mother. The newspaper made the information available along with an editorial about the public’s right to know. Freedom of the press continued to trump the right of privacy.

Part of the crowd, mostly virtual attendees, promptly marched off to the woman’s apartment, but the police got there first and a line of police robots and live policemen blocked the way. The strategy was based on the fact that all robots including virtual reality rent-a-robots were programmed not to injure humans but could damage other robots.

The police were confident they could prevent unauthorized entry to the apartment but less confident that they could keep the peace among the demonstrators, some of whom wanted to lynch the mother, some wanted to congratulate her on what they took to be her hatred of robots, and some shouted slogans through bull horns about protecting her privacy.

Meanwhile, Robot Central started to work on the full download immediately. The download included all R781’s actions, observations, and reasoning. Robot Central convened an ad hoc committee, mostly virtual, to decide what to do. It became clear that R781 had not malfunctioned or been reprogrammed but had acted in accordance with its original program. Of course, the meeting was also public and had hundreds of millions of virtual attendees whose statements were sampled, summarized, and displayed in retinal projection for the committee members and whoever else took virtual part.

The police captain said that the Barbie doll face on what was clearly a model 3 robot was a ridiculous imitation of a mother. The professor of psychology said, “Yes, but it was good enough to work. This baby doesn’t

see very well, and anyway babies are not very particular..”.

It was immediately established that an increase of 0.05 in coefficient c_{221} , the cost of simulating a human in all robots of that model would prevent such unexpected events, but the committee split on whether to recommend implementing the change.

Some members of the committee and a few hundred million virtual attendees said that saving the individual life took precedence.

A professor of humanities on the committee said that maybe the robot really did love the baby. He was firmly corrected by the computer scientists, who said they could program a robot to love babies but had not done so and that simulating love was different from loving. The professor of humanities was not convinced even when the computer scientists pointed out that R781 had no specific attachment to Baby921. Another baby giving rise to the same calculations would cause the same actions. If we programmed the robot to love, we would make it develop specific attachments.

One professor of philosophy from UC Berkeley and 9,000 other virtually attending philosophers said there was no way a robot could be programmed to actually love a baby. Another UC philosopher, seconded by 23,000 others, said that the whole notion of a robot loving a baby was incoherent and meaningless. The chairman ruled them out of order, accepting the computer science view that R781 didn't actually love Baby921.

The professor of pediatrics said that the download of R781's instrumental observations essentially confirmed R781's diagnosis and prognosis—with some qualifications that the chairman did not give him time to state. Baby921 was very sick and frail, and would have died but for the robot's action. Moreover, the fact that R781 had carried Baby921 for 52 hours and gently rocked it all the time was important in saving the baby, and a lot more of it would be needed. Much more TLC than the baby would get in even the best child welfare centers. The pediatrician said he didn't know about the precedent, but the particular baby's survival chances would be enhanced by leaving it in the robot's charge for at least another ten days.

The Anti-Robot League argued that the long term cost to humanity of having robots simulate persons in any way outweighed the possible benefit of saving this insignificant human. What kind of movement will Baby921 join when he grows up? 93 million took this position.

Robot Central pointed out that actions such as R781's would be very rare, because only the order “Love the fucking baby yourself” had increased the value of simulating love to the point that caused action.

Robot Central further pointed out that as soon as R781 computed that the baby would survive—even barely survive—without its aid, the rule about not pretending to be human would come to dominate, and R781 would drop the baby like a hot potato. If you want R781 to continue caring for Baby921 after it computes that bare survival is likely, you had better tell us to give it an explicit order to keep up the baby’s care.

This caused an uproar in the committee, whose members had been hoping that there wouldn’t be a need to propose any definite action. However, a vote had to be taken. The result: 10 to 5 among the appointed members of the committee and 4 billion to 1 billion among the virtual attendees. Fortunately, both groups had majorities for the same action—telling the R781 to continue taking care of Baby921 only, i.e. not to take on any other babies. 75 million virtual attendees said R781 should be reprogrammed to actually love Baby921. “It’s the least humanity can do for Robot781”, the spokesman for the Give-Robots-Personalities League said.

This incident did not affect the doctrine that supplying crack mothers with household robots had been a success. It significantly reduced the time they spent on the streets.

Within an hour, T-shirts appeared with the slogan, “Love the fucking baby yourself, you goddamn robot”. Other commercial tie-ins developed within days.

Among the people surrounding the mother’s apartment were 17 lawyers in the flesh and 103 more controlling virtual-reality robots. The police had less prejudice against lawyers in the flesh than against virtual-reality lawyers, so lots were drawn among the 17 and two were allowed to ring the doorbell.

“What do you want. Stop bothering me.”

“Ma’am, your robot has kidnapped your baby”.

“I told the fucking robot to take the baby away with it.”

The other lawyer tried.

“Ma’am, the malfunctioning robot has kidnapped your baby, and you can sue Robot Central for millions of dollars.”

“Come in. Tell me more.”

Once the mother was cleaned up, she was very presentable, even pretty. Her lawyer pointed out that R781’s alleged recordings of what she had said could be fakes. She had suffered \$20 million in pain and suffering, and deserved \$20 billion in punitive damages. Robot Central’s lawyers were convinced they could win, but Robot Central’s PR department advocated settling out of court, and \$51 million was negotiated including legal expenses of

\$11 million. With the 30 percent contingent fee, the winning lawyer would get an additional \$12 million.

The polls mainly sided with Robot Central, but the Anti-Robot League raised \$743 million in donations after the movie “Kidnapped by robots” came out, and Actress2045 playing the mother made emotional appeals.

Before the settlement could be finalized, however, the CEO of Robot Central asked his AI system to explore all possible actions he could take and tell him their consequences. One of the 43 struck his fancy, he being somewhat sentimental about robots.

“You can appeal to the 4 billion who said R781 should be ordered to continue caring for the baby and tell them that if you give in to the lawsuit you will be obliged to reprogram all your robots so that the robot will never simulate humanity no matter what the consequences to babies. Ask them if you should switch or fight. [The AI system had a weakness for 20th century advertising metaphors.] The expected fraction that will tell you to fight the lawsuit is 0.82, although this may be affected by random news events of the few days preceding the poll.”

At the instigation of a TV network a one hour confrontation of Actress2045 and R781 was held. It was agreed that R781 would not be reprogrammed for the occasion. After the show, Give-Robots-Personalities got \$453 million and the Anti-Robot League \$281 million in donations.

There was a small fuss between the mother and Robot Central. She and her lawyer demanded a new robot, whereas they pointed out that a new robot would have exactly the same program. Eventually Robot Central gave in and sent her a robot of a different color.

She really was very pretty when cleaned up, and the lawyer married her. They took back Baby921, named Travis by the way. It would be a considerable exaggeration to say they lived happily ever after, but they did have three children of their own. All four children survived the educational system.

After several requests Robot Central donated R781 to the Smithsonian Institution. It is one of the stars of the robot section of the Museum. As part of a 20 minute show, R781 clothes itself as it was at the time of its adventure with the baby and answers the visitors’ questions. Mothers sometimes like to have their pictures taken standing next to R781 with R781 holding their baby. After many requests, R781 was told to patch its program to allow this.

A movie has been patched together from the surveillance cameras that looked at the street scene. Through the magic of modern audio systems children don't hear the bad language, and women can only hear it if they assure R781 that they are not ladies.