Ethics for a Combined Human-Machine Dialogue Agent

Ron Artstein
University of Southern California
Institute for Creative Technologies
12015 Waterfront Drive
Playa Vista CA 90094-2536, USA

Kenneth Silver
University of Southern California
Department of Philosophy
3709 Trousdale Parkway
Los Angeles CA 90089, USA

Abstract

We discuss philosophical and ethical issues that arise from a dialogue system intended to portray a real person, using recordings of the person together with a machine agent that selects recordings during a synchronous conversation with a user. System output may count as actions of the speaker if the speaker intends to communicate with users and the outputs represent what the speaker would have chosen to say in context; in such cases the system can justifiably be said to be holding a conversation that is offset in time. The autonomous agent may at times misrepresent the speaker's intentions, and such failures are analogous to good-faith misunderstandings. The user may or may not need to be informed that the speaker is not organically present, depending on the application.

Introduction

The amalgamation of humans and machines has long been the realm of science fiction and the popular press. The term *cyborg* was proposed by Clynes and Kline (1960) for a biological organism, adapted to a different environment with artificially enhanced or altered bodily functions. Such artificial regulation and enhancement are commonplace today, from pacemakers to artificial limbs to cochlear implants. Machines taking over people's cognitive functions, however, are still the realm of fiction, and ethical implications are explored in the literature, sometimes humorously – for example in a recent series from the *Dilbert* comic strip, where Dilbert is imprisoned for a crime committed by his body while under the control of an external brain stimulator (August 19–21, 2015: http://dilbert.com/strip/2015-08-21).

While it is not presently possible for machines to take control of human minds, we already have systems that combine human and machine cognitive functions in a representation outside the human body. Friedman, Salomon, and Hasler (2013) have developed virtual proxies for teachers – avatars that are projected on a screen in front of a class, delivering lectures and answering routine questions, while handing off more complex tasks to the teacher who is present remotely. Our own *New Dimensions in Testimony* prototype (Artstein et al. 2014; Traum et al. 2015b) allows users to conduct a conversation with a speaker who is not present in real time, through a computer algorithm that selects appropriate responses to user utterances from a set of pre-recorded statements. The prototype is intended to replicate a conversation with Holocaust survivor Pinchas Gutter. Holocaust education today relies to a great extent on survivors talking to audiences in museums and classrooms, relating their experiences directly and creating an intimate connection with their audiences (Bar-On 2003). However, the youngest survivors are in their seventies today, and in a few years there will be no more survivors left to tell the story in person. The prototype will afford future generations the opportunity to engage in such unmediated conversation, talking to Pinchas Gutter and asking him questions about his life before, during and after the Holocaust. We call this technology "time-offset interaction", and it can have a wide range of applications, such as preserving the memory of a person for the future (historical figures as well as ordinary people who wish to converse with their descendants); enabling conversation with family and friends who are temporarily unavailable (traveling, deployed overseas, or incarcerated); and allowing popular speakers (leaders, celebrities) to engage in conversation with multiple people at the same time.

The technology raises a philosophical and ethical question: under what conditions can an artificial agent be said to adequately represent a real person? More specific questions include determining who is acting when the system produces an utterance, identifying the speaker and addressee of such speech acts, consent and fair representation of the speaker, veracity of the statements and responsibility for errors, and truth in representation of the speaker to the user. The next section describes the dialogue agent in more detail, and the remainder of the paper explores the ensuing philosophical and ethical considerations.

The Dialogue Agent

In the *New Dimensions in Testimony* prototype, users talk to a persistent representation of a Holocaust survivor presented on a video screen, and a computer algorithm selects and plays individual video clips of the survivor in response to user utterances. The result is much like an ordinary conversation between the user and the survivor. The system has been described in detail in previous publications, covering the proof of concept (Artstein et al. 2014), the content elicitation process (Artstein et al. 2015), the language process-

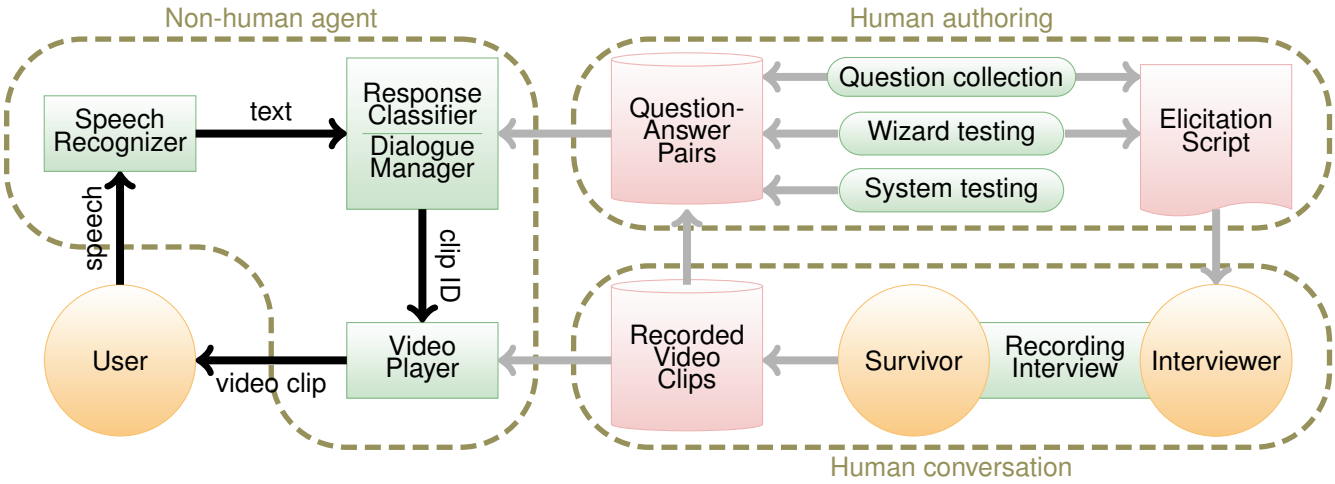

Figure 1: Information flow in the *New Dimensions in Testimony* prototype. Black arrows indicate information flow at runtime, while gray arrows show information flow during system preparation and training.

ing (Traum et al. 2015a), and the full prototype (Traum et al. 2015b). Here we briefly describe the overall architecture of the system, concentrating on the sources of information for the various components, in order to set the stage for the discussion of human and machine action, and its implications.

Information flows through the prototype at several distinct times. At runtime, a user talks to a computer system, which selects and plays prepared video clips in response to user utterances. The video clips were recorded at an earlier elicitation phase. Preceding and following the elicitation are several authoring phases, where human authors prepare materials for the interview and for the automated system. These phases are outlined in Figure 1 and detailed below.

At the heart of the runtime computer system is a response classifier and dialogue management component called NPCEditor (Leuski and Traum 2011), which selects a response to each user utterance. The input to NPCEditor is the text of the user's utterance (identified by a speech recognizer), and the output is the ID of a video clip that is played back to the user as a response. NPCEditor employs a statistical classifier which is trained on linked questions and responses; for each new user utterance, the classifier ranks all the available responses. The dialogue manager functionality chooses which response to play back to the user: typically it will choose the top ranked response, but it may choose a lower ranked response in order to avoid repetition, and if the score of the top ranked response is below a predefined threshold (determined during training), it will instead select an off-topic response that indicates non-understanding (such as "please repeat that" or "I don't understand").

The video clips themselves were recorded in an interview, designed to elicit useful material for responding to user questions: the survivor talked to an interviewer, who was following a script of elicitation questions and other prompts. The recordings therefore reflect the intentions of the survivor in answering the interviewer's questions. The video clips were only minimally processed, with the deliberate intention of keeping the content authentic and true to the source.

Human authoring is responsible for the interview scripts, the classifier training data, and the dialogue manager. The initial interview script included common questions asked of Holocaust survivors as identified by experts in the field, questions collected from audience members who had seen a film about the survivor and experienced a live question-answer session, and devised questions intended to elicit specific stories. Following the initial recording, a mock system was created where participants talked to the survivor, and responses were played back by human "wizards" in real time; questions from this wizard testing were incorporated into a second interview script, used in a second recording session. Subsequent to the second recording, questions from the interview scripts and wizard testing were manually linked to appropriate responses in order to train an initial classifier and assemble an interactive system; this system was tested with live participants, and questions collected during this testing were added to the training data, manually linked to responses, and used to retrain the classifier. The dialogue manager, responsible for avoiding repetition and handling non-understandings, is completely hand-authored.

We thus have a system that is intended to replicate the experience of talking to an actual person, where each conversational output by the system is the result of both an intentional action by the speaker (when recording the response), and a decision by an autonomous software component (selecting the response in context). The decisions of the agent are made by a statistical classifier with some human authoring. This joint human and non-human action, which is intended to simulate an actual person, raises questions about the source and nature of action, obligations towards the speaker and the user, and responsibility for the system's operation.

Philosophical Considerations

The outputs of the system at runtime have a mixed source: each utterance is associated with the speaker's intention during the original recording, but in the runtime context the utterance may or may not convey the same intention. This

raises questions about the extent of the system's autonomy, the sources of action, and the nature of the conversation.

Autonomy

The *New Dimensions in Testimony* prototype has some degree of autonomy because it can react to novel inputs. This is in contrast with a system like *Ask the President*, deployed at the Nixon presidential library in the early 1990s, where users selected a question from a predefined set, and the system would play back a recording of President Nixon in response (Chabot 1990). *Ask the President* had fixed reactions to fixed inputs, so it could not be considered autonomous; whereas the present dialogue agent is designed to identify the best reaction to any input, which is a measure of autonomy. At the same time, there are limits to what the dialogue agent can do: for example, it cannot forge a new utterance, and the dialogue manager is primarily reactive. The agent is an example of the *selection approach* to dialogue management (Gandhe and Traum 2010); selection from a corpus allows a dialogue agent to issue complex and human-like utterances with little risk of producing strange or unusual (or even impossible) verbal or non-verbal behaviors. For a system intended to represent an actual person, the selection approach also helps with managing the ethical issues of fair representation and veracity of the speaker's statements, discussed below. But our analysis is not limited to this type of system. An agent with more autonomy, for example one that could synthesize utterances (verbal and non-verbal), or one with more ability to take initiative in the conversation, would be subject to the same ethical considerations, though meeting these considerations (and demonstrating that they have been met) would be more difficult.

A separate issue related to autonomy is the system's intended function, which is to capture as faithfully as possible the behavior of the speaker. If the system were capable of doing so perfectly, then in some sense we might say that it lacked autonomy altogether, as it would be a complete slave to the speaker's action. But the system is not trying to replicate any actual action of the speaker, but rather how the speaker *would have* acted in a novel situation. This raises a separate question, namely to what extent the system's operation represents actions of the speaker.

Action

The speaker clearly acted during the elicitation interview: he answered questions and responded to prompts. He could have declined to answer any of the questions, and he was not coerced to answer as he did. But is he still acting when the system responds to the user? Although most of our actions seem to occur just where we are, many actions seem to continue on in the world in some sense (Thomson 1971). For example, a person can set an alarm for the next day, and it can sound appropriate to say that when the alarm rings, the person woke herself up. Similarly, one person could guide another on a journey with clues that the first person has already laid out, even if they are not around for the journey itself. A common means we have of extending our actions is through the use of technology – we reach farther with a grasper; we shout farther with a bullhorn; we communicate farther with email. Given this, we might maintain that the system is a way for the speaker to extend his actions to holding conversations with future generations.

(A more radical view might argue that the speaker himself extends into the future through the use of the system. It has been recently argued that our bodies can in various ways extend beyond our natural limbs if something is able to play a functional role played by those limbs or else reliably allows us to act (Clark 2008). We will not explore this notion here.)

To maintain the view that the speaker is acting through the system, two things must be true: the speaker must intend or be trying to communicate with a user of the system, and he has to have some knowledge or control over the system's outputs. The speaker needn't have a particular user in mind, but if for example the system uses recordings made without the speaker's knowledge, then the outputs of the system will not be considered actions of the speaker. As for control over the system's functioning, we take ourselves to reliably control or guide the means with which we act (Frankfurt 1978). Obviously the speaker can't directly control the system outputs; once the recordings are done and the system is assembled, control over how the system will answer novel questions is ceded to the speech recognition, response classification, and dialogue management algorithms. Still, if the system pairs answers with questions as the speaker would, then the speaker can reliably predict the outputs. So, if the speaker intends to communicate with a user and has some measure of control over the outputs of the system, then a strong case can be made that the speaker will be able to perform an extended action through the system.

While the system involves an autonomous agent, it does not qualify as acting under its own free will. Even if it were able to perfectly replicate the speaker's behavior, it is debatable whether that would count as meaningful communication (see Searle 1980), but the present system certainly falls short of meaningful understanding. Kane (1989; 2003) argues that free will requires that at least some actions come about via some non-deterministic process, and while the response classifier has a statistical component, it is ultimately deterministic (after being trained). The response classification and dialogue management algorithms prescribe what response the system will give, and the system is not able to give a different response or refrain from giving a response. Pereboom (2001:113–114) argues that if an entity's interactions with the environment are completely determined by programming done by others, then it is necessarily not free and not morally responsible for its actions. Therefore, it would be hard to maintain that the system is itself acting freely.

The relationship is between freedom and responsibility is also controversial. Many philosophers take free will to be necessary for moral responsibility, although there has been significant pushback in the wake of Frankfurt (1969), which purports to give a counterexample to the thought that responsibility requires the ability to do otherwise. Nevertheless, we think it is also unlikely that the system would fulfill other necessary conditions that have been given for responsibility apart from the requirement of having been able to do otherwise. Therefore, the system as such cannot be held

morally responsible. When examining moral and ethical implications of the system's operation, we will be careful not to ascribe moral responsibility to the system itself.

Conversation

Since the actions performed by the user and the speaker are attempts at communication, it is worth clarifying who is communicating with whom. The user is talking to the system, but it is less clear who the user is addressing – whether it is the speaker or the system itself. What we think will depend in part on how the user conceives of the system: the user may address the speaker if she is unaware that the speaker is not organically present to hear her. If she knows, however, then it is not clear whether she can address the speaker directly, other than in the sense of addressing a person who is absent or deceased, or an inanimate object.

As for the speaker, during the elicitation interview he is directly addressing the interviewer. But the speaker also intends his message to go further, namely to future users of the system, who therefore constitute indirect (or lateral) addressees (Clark and Carlson 1982). At runtime, however, the user, who was an indirect addressee when the utterance was recorded, becomes a direct addressee of the utterance played by the system. But who plays the speaker role of the speech act? If we can consider the speaker as being able to perform an extended action in this case, then perhaps we may consider him to be performing a speech act. We would then be justified as literally considering the interaction as a conversation that is offset in time: the speaker has prepared his part in advance, and is performing his prepared act at the time of the interaction with the user. A problem with this view is that the system sometimes fails in its response selection, delivering an utterance that cannot be reasonably construed as something the speaker would have chosen to say in that context. Under the criteria described above, such a response would not be considered to be under the speaker's control, and therefore not an extended action of his.

Could we consider the system to play the speaker role of the speech act addressed to the user? Talking about dialogue systems as performing speech acts and even modeling this explicitly (e.g. Traum et al. 2003) is appropriate for systems that represent themselves, or a fictional character identified with the system. However, our system represents a real person, and each utterance carries an original intention by the real person. It appears, then, that when an utterance cannot be considered an action of the speaker, it is some effect that arises jointly from the speaker's past action and the system's operation.

Civility of discourse may serve as a test for whether users perceive interaction with the system as equivalent to interacting with a person. When users interact with dialogue systems that depict fictional animated characters, we observe some behaviors that are not common when talking to a live human, such as testing the system (e.g. "Do you know about the Dodgers?" or "Can you turn around?"), but also hazing and flaming, that is testing the system not in good faith or being outright rude (Robinson et al. 2008). It has been argued that joint action (like conversation) can be a source of obligations for the agents involved (Bratman 1992; 2014). Discourse obligations, that is an obligation to say something (Traum and Allen 1994), naturally arise in conversations between users and dialogue systems, but behaviors like hazing and flaming suggest that moral obligations such as treating one's interlocutor with respect may not arise. If users exhibit hazing and flaming towards the system, it may suggest that they view the interaction as similar to interacting with a fictional character; conversely, if users consistently refrain from hazing and flaming, they may view the interaction as talking to a person. We have not observed instances of hazing or flaming with the *New Dimensions in Testimony* prototype, though this may be because such behaviors typically arise in interactions without a moderator or tester present, and so far we haven't had many users interact with the system under those conditions.

Ethical Considerations

The system as it has been used here – to provide future generations with the opportunity for direct interaction with a Holocaust survivor – appears quite ethical. A Holocaust survivor is able to relate his experiences to audiences that solicit that information, and the system is engineered to provide answers that are most germane to the questions given. But using an artificial agent to represent a real human carries risks of misrepresentation of the speaker or his intentions, or otherwise wronging the speaker or users of the system. Here we discuss some of these risks.

Consent

One fundamental right of the speaker is that they must consent to the use of their likeness for the system; if they are not able to consent then it should at least be clear that the speaker would have consented if they had been able to. It is obvious that our speaker, Holocaust survivor Pinchas Gutter, consented to his likeness being used in an interactive system, but other cases can become less clear. Take the exhibit *Ask the President* from the Nixon presidential library described above: the exhibit was in place at the time of the library's dedication, attended by Mr. Nixon, so we can assume that he consented to that exhibit. But the exhibit allowed only a fixed set of questions tied directly to video responses, so there was no chance of giving the wrong response to a question. What if the exhibit were later changed to allow users to ask anything, with the best response determined by a software agent? Can we assume Mr. Nixon would have consented to that use? Moreover, Nixon's video clips were of diverse origins, providing visual clues that the clips were not originally intended as responses to the user question. But Nixon also provided many hours of interview specifically meant to be viewed by future generations (Nixon-Gannon Interviews 1983). Would Mr. Nixon be happy with the presentation of snippets from these interviews as directly addressing a user in response to questions? Would his consent be required? Does is matter that Mr. Nixon is a deceased public figure?

Fair representation

The speaker also has the right to fair representation in the system: it should react to users as the speaker would. Fair

representation does not require directly addressing each user utterance, but rather giving an output that can be construed as a reasonable response by the speaker; this includes off-topic responses (such as "please repeat that" or "I don't understand"). Our most recent evaluation figures show that the system can give appropriate direct responses to about 64% of the user questions, with 20% off-topic responses and the remaining 16% being errors (errors can be reduced by increasing the share of off-topic responses, but this also reduces the rate of appropriate direct responses). In Traum et al. (2015a) we concluded that these numbers are sufficient to enable a reasonable conversation flow. But do these numbers also constitute a fair representation of the speaker?

The rate of errors alone may not be a good measure of fair representation, since not all errors are alike: in many cases where a system output cannot be construed as a reasonable response by the speaker to the question, the speaker's intention still comes through. Such an utterance might be a poor representation of the speaker's attention or conversation skills, but it is still a fair representation of what he has to say. The problematic misrepresentations are the ones where the system output, in context, conveys a different meaning than that intended by the speaker; while we have not been able to measure the rate of misrepresentation, it is necessarily lower than the measured error rate.

Because the system is based on selection, each utterance represents an actual intention of the speaker; misrepresentation only arises due to context. A future system may achieve broader scope and greater variety using generation, synthesizing verbal and nonverbal utterances, or a more elaborate dialogue manager with deep reasoning and speaker initiative. This could lead to misrepresentation in the form itself (unintelligible verbal utterances and strange or impossible nonverbal behavior). Explicit modeling of user intentions in the dialogue manager would require demonstrating that the system intentions reflect intentions that the speaker would have had, in addition to a demonstration that these intentions are conveyed by the utterances in context.

A separate question related to fair representation is assigning blame when things go wrong. A relevant analogy to making an utterance in an inappropriate context may be journalism ethics: is a system failure akin to quoting a person out of context, or maliciously editing their speech? We have established that a misrepresentation does not constitute an action of the speaker so he is not responsible, nor is the system a morally responsible entity. How about the developers? The most common causes of failure are deficiencies in the training data. As we noted above, development of the *New Dimensions in Testimony* prototype involved substantial effort to ensure, to the extent possible, that the training data are adequate, so the developers cannot be accused of lack of due diligence. If nobody is to blame, then it appears that misrepresentations of the speaker by the system should be thought of not as akin to malicious editing, but rather more like cases of not understanding the question.

Veracity

An issue related to fair representation is the veracity of the speaker's utterances, as they are interpreted in the context of talking to the system. The purpose of the *New Dimensions in Testimony* prototype is to support Holocaust education, so its goal is not only to respond to questions as the organic agent would, but also for the users to develop an understanding that is historically accurate. The speaker's statements are generally accurate, but the context in which the system outputs a response can cause subtle implications, as in the following exchange from a test of the automated system.

USER: *How did you find food in the Warsaw ghetto?*

SURVIVOR: *The food in the beginning, uh, was scarce and spare. There wasn't very much, the bread was black, it wasn't anything specific, so there wasn't very much going. But as time came along, you know, within a few months, uh, the shops filled up and there was quite a, uh, there was reasonable food, provided you had money. If you didn't have money, uh, you know, it was difficult, but if you had money you could actually buy things. And things were rationed, but you, you could get a lot on the black market right from the beginning and, you know, if you went as a Christian boy to a shop and you asked for whatever, you could get it, a coffee or, but, but soon things started changing.*

The response is conversationally appropriate and appears to answer the question, but a historian observing the test noted that it talks about the procurement of food in Warsaw before the ghetto was formed (after the ghetto walls were erected, the speaker was no longer able to go to shops pretending to be a Christian boy). As a representation of the speaker it is not bad, and can easily be attributed to a slight mishearing or misunderstanding of the question, or even forgetting the context midway. But if the system is held to a higher standard of historical accuracy, then this could be considered a failure. The selection approach limits problems with veracity to implications drawn from context (assuming the speaker's original statements are true); a generation approach would need to also ascertain the veracity of generated statements.

Informedness

A separate ethical issue concerns the rights of the user: does the user have a right to know that they are interacting with a system of this kind, and that the organic person that is digitally represented is not present to see the user and hear their specific question? This is not a theoretical question: such misconceptions happened during our initial "wizard" testing, when users were seated in front of a screen and asked questions while video responses were manually selected by hidden human operators. Despite being informed of the nature of the conversation, some users thought that they were engaged in a live videoconference with the speaker.

Our impression is that whether a user is owed this right would depend on the application. A system that answers the door and receives deliveries probably does not owe the serviceman the knowledge that the speaker who signs for the package is only an artificial representation. On the other hand, if the speaker uses the system to represent herself to her boyfriend for the purposes of listening to his problems without having to be organically present, then it seems like it would be appropriate for the boyfriend to both be upset and the blame the speaker for this deception.

Conclusion

The technology of time-offset interaction has many advantages: through the use of a non-human software agent, people may be able to extend their being and actions beyond their body, location, or lifespan (metaphorically or metaphysically). And users of these systems can in some sense communicate with individuals with whom they would otherwise never have been able to communicate. These are benefits that would ostensibly make the world a better place and increase our overall happiness.

The main risk in this technology is that of misrepresentation – both misrepresentation of the speaker or the information he intends to convey, as well as misrepresentation to the user or the speaker about the nature of the interaction. The first type of misrepresentations are caused by local failures of the autonomous agent, and are therefore akin to honest misunderstandings, not attributable to any individual human agent (assuming no malice or negligence on behalf of the developers). We have not set specific standards for the amount of acceptable misrepresentation, as these will likely vary by application.

Representation of humans by autonomous machine agents is the future, in one way or another. Rather than viewing it as a diminishing of interpersonal contact, we feel that we should embrace this technology as an expanding and enriching of the means we have of communicating with one another across the generations.

Acknowledgments

The New Dimensions in Testimony prototype was made possible by generous donations from private foundations and individuals. We are extremely grateful to The Pears Foundation, Louis F. Smith, and two anonymous donors for their support. The USC Shoah Foundation provided financial and administrative oversight and support, and Conscience Display collaborated on content development. The Los Angeles Museum of the Holocaust, the Museum of Tolerance, New Roads School in Santa Monica, and the Illinois Holocaust Museum and Education Center offered their facilities for data collection and testing. We owe special thanks to Pinchas Gutter for sharing his story, and for his tireless efforts to educate the world about the Holocaust.

This work was supported in part by the U.S. Army; statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

Artstein, R.; Traum, D.; Alexander, O.; Leuski, A.; Jones, A.; Georgila, K.; Debevec, P.; Swartout, W.; Maio, H.; and Smith, S. 2014. Time-offset interaction with a Holocaust survivor. In *Proceedings of IUI*, 163–168.

Artstein, R.; Leuski, A.; Maio, H.; Mor-Barak, T.; Gordon, C.; and Traum, D. 2015. How many utterances are needed to support time-offset interaction? In *Proc. FLAIRS-28*, 144–149.

Bar-On, D. 2003. Importance of testimonies in Holocaust education. *Dimensions Online: A Journal of Holocaust Studies* 17(1).

Bratman, M. E. 1992. Shared cooperative activity. *The Philosophical Review* 101(2):327–341.

Bratman, M. E. 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press.

Chabot, L. 1990. Nixon library technology lets visitors 'interview' him. *Los Angeles Times* (July 21).

Clark, H. H., and Carlson, T. B. 1982. Hearers and speech acts. *Language* 58(2):332–373.

Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.

Clynes, M. E., and Kline, N. S. 1960. Cyborgs and space. *Astronautics* (September):26–27,74–76.

Frankfurt, H. G. 1969. Alternate possibilities and moral responsibility. *Journal of Philosophy* 66(23):829–839.

Frankfurt, H. G. 1978. The problem of action. *American Philosophical Quarterly* 15(2):157–162.

Friedman, D.; Salomon, O.; and Hasler, B. S. 2013. Virtual substitute teacher: Introducing the concept of a classroom proxy. In *Proc. 3rd European Immersive Education Summit*, 186–197.

Gandhe, S., and Traum, D. 2010. I've said it before, and I'll say it again: An empirical investigation of the upper bound of the selection approach to dialogue. In *Proc. SIGDIAL*, 245–248.

Kane, R. 1989. Two kinds of incompatibilism. *Philosophy and Phenomenological Research* 50(2):219–254.

Kane, R. 2003. Responsibility, luck, and chance: Reflections on free will and indeterminism. In Watson, G., ed., *Free Will*. Oxford University Press, 2nd edition. 299–321.

Leuski, A., and Traum, D. 2011. NPCEditor: Creating virtual human dialogue using information retrieval techniques. *AI Magazine* 32(2):42–56.

Nixon-Gannon Interviews. 1983. Walter J. Brown Media Archives & Peabody Awards Collection, University of Georgia Libraries, Athens, Georgia.

Pereboom, D. 2001. *Living Without Free Will*. Cambridge University Press.

Robinson, S.; Traum, D.; Ittycheriah, M.; and Henderer, J. 2008. What would you ask a conversational agent? Observations of human-agent dialogues in a museum setting. In *Proceedings of LREC*, 1125–1131.

Searle, J. R. 1980. Minds, brains and programs. *Behavioral and Brain Sciences* 3(3):417–457.

Thomson, J. J. 1971. The time of a killing. *Journal of Philosophy* 68(5):115–132.

Traum, D. R., and Allen, J. F. 1994. Discourse obligations in dialogue processing. In *Proceedings of ACL*, 1–8.

Traum, D.; Rickel, J.; Gratch, J.; and Marsella, S. 2003. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proceedings of AAMAS*, 441–448.

Traum, D.; Georgila, K.; Artstein, R.; and Leuski, A. 2015a. Evaluating spoken dialogue processing for time-offset interaction. In *Proceedings of SIGDIAL*, 199–208.

Traum, D.; Jones, A.; Hays, K.; Maio, H.; Alexander, O.; Artstein, R.; Debevec, P.; Gainer, A.; Georgila, K.; Haase, K.; Jungblut, K.; Leuski, A.; Smith, S.; and Swartout, W. 2015b. New Dimensions in Testimony: Digitally preserving a Holocaust survivor's interactive storytelling. In *Interactive Storytelling*, volume 9445 of *LNCS*. Springer. 269–281.