Introduction

My aim in this book is to rethink the intricate, and increasingly intimate, configurations of the human and the machine. Human–machine configurations matter not only for their central place in contemporary imaginaries but also because cultural conceptions have material effects.¹ As our relations with machines elaborate and intensify, questions of the humanlike capacities of machines, and machinelike attributes of humans, arise again and again. I share with Casper (1994), moreover, the concern that the wider recognition of "nonhuman agency" within science and technology studies begs the question of "how entities are configured as human and nonhuman prior to our analyses" (ibid.: 4). Casper proposes that discussions of nonhuman agency need to be reframed from categorical debates to empirical investigations of the concrete practices through which categories of human and nonhuman are mobilized and become salient within particular fields of action. And in thinking through relations of sameness and difference more broadly,

¹ The word *imaginary* in this context is a term of art in recent cultural studies (see Braidotti 2002: 143; Marcus 1995: 4; Verran 1998). It shares with the more colloquial term *imagination* an evocation of both vision and fantasy. In addition, however, it references the ways in which how we see and what we imagine the world to be is shaped not only by our individual experiences but also by the specific cultural and historical resources that the world makes available to us, based on our particular location within it. And perhaps most importantly for my purposes here, cultural imaginaries are realized in material ways. My inspiration for this approach is Haraway's commitment to what she names "materialized refiguration (1997: 23), a trope that I return to in Chapter 13. The particular imaginaries at stake in this text are those that circulate through and in relation to the information and communication networks of what we might call the hyperdeveloped countries of Europe and North America.

2

Human–Machine Reconfigurations

Ahmed (1998) proposes a shift from a concern with these questions as something to be settled once and for all to the occasioned inquiry of "which differences matter, here?" (ibid.: 4). In that spirit, the question for this book shifts from one of whether humans and machines are the same or different to how and when the categories of human or machine become relevant, how relations of sameness or difference between them are enacted on particular occasions, and with what discursive and material consequences.

In taking up these questions through this second expanded edition of Plans and Situated Actions, I rejoin a discussion in which I first participated some twenty years ago, on the question of how capacities for action are figured at the human-machine interface and how they might be imaginatively and materially reconfigured. Almost two decades after the publication of the original text, and across a plethora of subsequent projects in artificial intelligence (AI) and human-computer interaction (HCI), the questions that animated my argument are as compelling, and I believe as relevant, as ever. My starting point in this volume is a critical reflection on my previous position in the debate, in light of what has happened since. More specifically, my renewed interest in questions of machine agency is inspired by contemporary developments both in relevant areas of computing and in the discussion of human-nonhuman relations within social studies of science and technology.² What I offer here is another attempt at working these fields together in what I hope will be a new and useful way. The newness comprises less a radical shift in where we draw the boundaries between persons and machines than a reexamination of how - on what bases - those boundaries are drawn. My interest is not to argue the question of machine agency from first principles, in other words, but rather to take as my focus the study of how the effect of machines-as-agents is generated and the latter's implications for theorizing the human. This includes the translations that render former objects as emergent subjects, shifting associated interests and concerns across the human-artifact boundary. We can then move on to questions of what is at stake in these particular translations-inprogress and why we might want to resist or refigure them.

² At the outset I take the term agency, most simply, to reference the capacity for action, where just what that entails delineates the question to be explored. This focus is not meant to continue the long-standing discussion within sociology on structure and agency, which I take to reiterate an unfortunate dichotomy rather than to clarify questions of the political and the personal, how it is that things become durable and compelling, and the like.

Introduction

Chapter 1 of this edition provides some background on the original text and reflects on its reception, taking the opportunity so rarely available to authors to respond to readings both anticipated and unexpected.³ Chapters 2 through 10 comprise the original text as published in 1987. In each of these chapters, new footnotes provide updated references, commentaries, and clarifications, primarily on particular choices of wording that have subsequently proven problematic in ways that I did not foresee. I have made only very minor editorial changes to the text itself, on the grounds that it is important that the argument as stated remain unaltered. This is true, I believe, for two reasons. First, the original publication of the book marked an intervention at a particular historical moment into the fields of artificial intelligence and human-computer interaction, and I think that the significance of the argument is tied in important ways to that context. The second reason for my decision to maintain the original text, and perhaps the more significant one, is that I believe that the argument made at the time of publication holds equally well today, across the many developments that have occurred since. The turn to so-called situated computing notwithstanding, the basic problems identified previously - briefly, the ways in which prescriptive representations presuppose contingent forms of action that they cannot fully specify, and the implications of that for the design of intelligent, interactive interfaces - continue to haunt contemporary projects in the design of the "smart" machine.

The book that follows comprises a kind of object lesson as well in disciplinary affiliations and boundaries. The original text perhaps shows some peculiarities understandable only in light of my location at the time of its writing. In particular, I was engaged in doctoral research for a Ph.D. in anthropology, albeit with a supervisory committee carefully chosen for their expansive and nonprogrammatic relations to disciplinary boundaries.⁴ Although the field of American anthropology in the 1980s was well into the period of "studying up," or investigation of institutions at "home" in the United States,⁵ my dissertation project

³ Part of the discussion in Chapter 1 is drawn from opportunities provided earlier, in two discussion forums in the journals *Cognitive Science* 17(1), 1993, and the *Journal of the Learning Sciences* 12(2), 2003.

⁴ My committee included Gerald Berreman and John Gumperz, from the Department of Anthropology, and Hubert Dreyfus, from the Department of Philosophy, all at the University of California at Berkeley.

⁵ For a founding volume see Hymes (1974). di Leonardo (1998) offers a discussion of the enduringly controversial status of "exotics at home" within the discipline.

4

Cambridge University Press 978-0-521-85891-5 - Human-Machine Reconfigurations: Plans and Situated Actions, 2nd Edition Lucy Suchman Excerpt More information

Human–Machine Reconfigurations

(with the photocopier as its object, however enhanced by the projects of computing and cognitive science) stretched the bounds of disciplinary orthodoxy. Nonetheless, I was deeply committed to my identification as an anthropologist, as well as to satisfying the requirements of a dissertation in the field. At the same time, I had become increasingly engaged, through my interests in practices of social ordering and face-to-face human interaction, with the lively and contentious research communities of ethnomethodology and conversation analysis. It was these approaches, more than any, perhaps, that informed and shaped my own at the time. Finally, but no less crucially, my position as a Research Intern at Xerox Palo Alto Research Center (PARC) meant that my text had to speak to the fields of AI and HCI themselves.

My task consequently became one of writing across these multiple audiences, attempting to convey something of the central premises and problems of each to the other. More specifically, Chapter 4 of this volume, titled "Interactive Artifacts," and Chapter 5, titled "Plans," are meant as introductions to those projects for readers outside of computing disciplines. Chapter 6, "Situated Actions," and Chapter 7, "Communicative Resources," correspondingly, are written as introductions to some starting premises regarding action and interaction for readers outside of the social sciences. One result of this is that each audience may find the chapters that cover familiar ground to be a bit basic. My hope, however, is that together they lay the groundwork for the critique that is the book's central concern. These chapters are followed by an exhaustive (some might even say exhausting!) explication of a collection of very specific, but, I suggest, also generic, complications in the encounter of "users" with an intendedly intelligent, interactive "expert help system." I attempt to explicate those encounters drawing on the resources afforded by studies in face-to-face human interaction, to shed light on the problem faced by those committed to designing conversational machines. As a kind of uncontrolled laboratory inquiry, the analysis is perhaps best understood as a close study of exercises in instructed action, rather than of the practicalities of machine operation as it occurs in ordinary work environments and in the midst of ongoing activities. With that said, my sense is that the analysis of human-machine communication presented in Chapters 8 and 9 applies equally to the most recent efforts to design conversational interfaces and identifies the defining design problem for HCI more broadly. To summarize the analysis briefly, I observe that human-machine communications take place at a very limited site of

Introduction

interchange; that is, through actions of the user that actually change the machine's state. The radical asymmetries in relative access of user and machine to contingencies of the unfolding situation profoundly limit possibilities for interactivity, at least in anything like the sense that it proceeds between persons in interaction.⁶ Chapter 10, the conclusion to the original text, provides a gesture toward alternative directions in interface design and reaffirms the generative potential of the human-computer interface as a site for further research.

Readers familiar with the original text of *P&SA* may choose to pass over Chapters 2 through 10 or to focus more on the footnotes that offer further reflections, references, and clarifications. The chapters that follow the original text expand and update the arguments. Chapter 11, "Plans, Scripts, and Other Ordering Devices," makes clear, I hope, that although the focus of the preceding chapters is on plans (as understood within dominant AI projects of the time), the research object is a much larger class of artifacts. In this chapter I review developments both in theorizing these artifacts in their various manifestations and in empirical investigations of their workings within culturally and historically specific locales. Chapter 12, "Agencies at the Interface," takes up the question of what specific forms agency takes at the contemporary human-computer interface. I begin with a review of the rise of computer graphics and animation, and the attendant figure of the "software agent." Reading across the cases of software agents, wearable, and socalled pervasive or ubiquitous computing, I explore the proposition that these new initiatives can be understood as recent manifestations of the very old dream of a perfect, invisible infrastructure; a dream that I locate now within the particular historical frame of the "service economy." Chapter 13, "Figuring the Human in AI and Robotics," explores more deeply the question of what conceptions of the human inform current projects in AI and robotics, drawing on critiques, cases, and theoretical resources not available to me at the time of my earlier writing. In both chapters I consider developments in relevant areas of research - software agents, wearable computers and "smart" environments, situated robotics, affective computing, and sociable machines - since the 1980s and reflect on their implications. Rather than a comprehensive survey,

⁶ I should make clear at the outset that I in no way believe that human–computer interactions broadly defined, as the kinds of assemblages or configurations that I discuss in Chapters 14 and 15, are confined to this narrow point. Rather, I am attempting to be specific here about just how events register themselves from the machine's "point of view."

6

Human–Machine Reconfigurations

my aim is to identify recurring practices and familiar imaginaries across these diverse initiatives.

Finally, Chapter 14, "Demysitifications and Reenchantments of the Humanlike Machine," and Chapter 15, "Reconfigurations," turn to the question of how it might be otherwise, both in the staging of humanmachine encounters and through the reconfiguration of relations, practices, and projects of technology design and use. As will become clear, I see the most significant developments over the last twenty years, at least with respect to the argument of this book, as having occurred less in AI than in the area of digital media more broadly on the one hand (including graphical interfaces, animation, and sensor technologies) and science and technology studies (STS) on the other. The first set of developments has opened up new possibilities not only in the design of socalled animated interface agents but also - more radically I will argue - in mundane forms of computing and the new media arts. The further areas of relevant change are both in the field of STS, which has exploded with new conceptualizations of the sociotechnical, and also in my own intellectual and professional position. The latter has involved encounters since the 1980s with feminist science studies, recent writings on science and technology within cultural anthropology, and other forms of theorizing that have provided me with resources lacking in my earlier consideration of human-machine relations. During that same period, I have had the opportunity with colleagues at PARC to explore radical alternatives to prevailing practices of system design, informed by an international community of research colleagues. Engaging in a series of iterative attempts to enact a practice of small-scale, case-based codesign, aimed at creating new configurations of information technologies, has left me with a more concrete and embodied sense of both problems and possibilities in reconfiguring relations and practices of professional system design. I have tried in these chapters to indicate my indebtedness to these various communities and the insights that I believe they afford for innovative thinking across the interface of human and machine. Inevitably, both my discussion of new insights from science and technology studies and of new developments in computing is partial at best, drawing selectively from those projects and perspectives with which I am most familiar and that I have found most generative or compelling. Drawing on these resources, I argue for the value of research aimed at articulating the differences within particular human-machine configurations, expanding our unit of analysis to include extended networks of social and material production, and recognizing the agencies, and attendant

Introduction

responsibilities, involved in the inevitable cuts through which bounded sociomaterial entities are made.

The expansion of the text in terms of both technologies and theoretical resources is accompanied by a commitment to writing for new audiences. In particular, the new chapters of this book attempt to engage more deeply with those working in the anthropology and sociology of technology who are, and always have been, my compass and point of reference. Somewhat ironically, my location at PARC and the marketing of the original text as a contribution in computer science have meant that the book contained in Chapters 2 through 10 of this edition received much greater visibility in computing – particularly HCI – and in cognitive science than in either anthropology or STS. Although I am deeply appreciative of that readership and the friends from whom I have learned within those communities, it is as a contribution to science and technology studies that the present volume is most deliberately designed.

1

Readings and Responses

This chapter provides a synopsis and some contextualization of the analvsis offered in the original edition of *Plans and Situated Actions* (*P&SA*), published in 1987, followed by my reflections on the reception and readings of that text. My engagement with the question of human-machine interaction, from which the book arose, began in 1979, when I arrived at PARC as a doctoral student interested in a critical anthropology of contemporary American institutions¹ and with a background as well in ethnomethodology and interaction analysis. My more specific interest in the question of interactivity at the interface began when I became intrigued by an effort among my colleagues to design an interactive interface to a particular machine. The project was initiated in response to a delegation of Xerox customer service managers, who traveled to PARC from Xerox's primary product development site in Rochester, New York, to report on a problem with the machine and to enlist research advice in its solution.² The machine was a relatively large, feature-rich photocopier that had just been "launched," mainly as a placeholder to establish the company's presence in a particular market niche that was under threat from other, competitor, companies. The machine was advertised with a figure dressed in the white lab coat of the scientist/engineer but reassuring the viewer that all that was required to activate the machine's extensive functionality was to "press the green [start] button" (see Fig. 1.1).

² The project is discussed at length in Suchman (2005).

¹ A defining text of what came to be known as "anthropology as cultural critique" is Marcus and Fischer (1986). See also Gupta and Ferguson (1997); Marcus (1999); Strathern (1999).



Readings and Responses

FIGURE 1.1. "Pressing the green button" Advertisement for the Xerox 8200 copier, circa 1983 © Xerox Corporation.

It seemed that customers were refuting this message, however, complaining instead that the machine was, as the customer service managers reported it to us, "too complicated." My interest turned to investigating just what specific experiences were glossed by that general complaint, a project that I followed up among other ways by convincing my colleagues that we should install one of the machines at PARC and invite our co-workers to try to use it. My analysis of the troubles evident in these videotaped encounters with the machine by actual scientists/engineers led me to the conclusion that its obscurity was not a function of any lack of general technological sophistication on the part of its users but rather of their lack of familiarity with this particular machine. I argued that the machine's complexity was tied less to its esoteric technical characteristics than to mundane difficulties of interpretation characteristic of any unfamiliar artifact. My point was that making sense of a new artifact is an inherently problematic activity. Moreover, I wanted to suggest that however improved the machine interface or instruction set might be, this would never eliminate the need for active sense-making on the part of prospective users. This in turn called into

10

Human–Machine Reconfigurations

question the viability of marketing the machine as "self-explanatory," or self-evidently easy to use.³

My colleagues, meanwhile, had set out on their own project: to design an "intelligent, interactive" computer-based interface to the machine that would serve as a kind of coach or expert advisor in its proper use. Their strategy was to take the planning model of human action and communication prevalent at the time within the AI research community as a basis for the design. More specifically, my colleagues were engaged with initiatives in "knowledge representation," which for them involved, among other things, representing "goals" and "plans" as computationally encoded control structures. When executed, these control structures should lead an artificially intelligent machine imbued with the requisite condition–action rules to take appropriate courses of action.

My project then became a close study of a second series of videotaped encounters by various people, including eminent computer scientists, attempting to operate the copier with the help of the prototype interactive interface. I took as my focus the question of interactivity and assumptions about human conversation within the field of AI, working those against findings that were emerging in sociological studies of faceto-face human conversation. The main observation of the latter was that human conversation does not follow the kind of message-passing or exchange model that formal, mathematical theories of communication posit. Rather, humans dynamically coconstruct the mutual intelligibility of a conversation through an extraordinarily rich array of embodied interactional competencies, strongly situated in the circumstances at hand (the bounds and relevance of which are, in turn, being constituted through that same interaction). I accordingly adopted the strategy of taking the premise of interaction seriously and applying a similar kind of analysis to people's encounters with the machine to those being

³ As Balsamo succinctly points out, "to design an interface to be 'idiot-proof' projects a very different level of technical acumen onto the intended users than do systems that are designed to be 'configurable'" (Balsamo in press: 29). It should be noted that this agument carried with it some substantial – and controversial – implications for technology marketing practices as well, insofar as it called into question the assertion that technology purchasers could invest in new equipment with no interruption to workers' productivity and with no collateral costs. On the contrary, this analysis suggests that however adequate the design, long-term gains through the purchase of new technology require near-term investments in the resources that workers need to appropriate new technologies effectively into their working practices. Needless to say, this is not a message that appears widely in promotional discourses.